Application of a *Drosophila melanogaster* model to study Familial Isolated Pituitary Adenomas syndrome pathogenesis *in vivo*

Thesis presented by:

Elena Daniela Aflorei, MD, MSc

This thesis is submitted for the Degree of

Doctor of Philosophy (PhD)

Under the supervision of:

Prof Márta Korbonits, MD, PhD

Prof Ralf Stanewsky, PhD

Centre for Endocrinology,

William Harvey Research Institute,

Barts and the London School of Medicine and Dentistry,

Queen Mary University of London

London, UK, March 2016

STATEMENT OF ORIGINALITY

I, Elena Daniela Aflorei, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated.

Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: March 2016

LIST OF COLLABORATION:

1. The library preparation and the RNA sequencing were performed at the Barts and the London School of Medicine genomics core facility (Genome Centre).
2. The multiplex qPCR method, performed with the help of collaborators from Royal Veterinary College (Dr. Rob Fowkes and Dr Samantha M. Mirczuk)
3. The transgenic *Drosophila* stocks were generated by BestGene Inc California USA.

LIST OF PUBLICATIONS

1. **Publications arising from the data presented in this thesis are under preparation.**

2. **Publications regarding familial pituitary adenomas (attached at the end of the thesis):**

   Aflorei, E. D. & Korbonits, M. **Epidemiology and etiopathogenesis of pituitary adenomas**. *Journal of Neuro-Oncology* 117, 379-394, (2014).

   Baciu, I., Radian, S., Capatina, C., Botusan, I., Aflorei, E. D., Stancu, C., Dumitrascu, A., Ciubotaru, V. & Coculescu, M. **The p.R16H (C.47G>A) AIP gene variant in a case with invasive non-functioning pituitary macroadenoma and screening of a control cohort**. *Acta Endocrinologica* IX, 12, (2013).

3. **Congress presentations regarding the data arising from this thesis**

   Aflorei, E.D. et al. **Functional homology between human and fruitfly AIP protein - an in vivo assay system to test the pathogenicity of *AIP* mutations**. *The Endocrine Society Annual Meeting ENDO 2015* (San Diego, CA, USA, March 2015); **P SAT-440**.

   Aflorei, E.D. et al. **Human *AIP* gene rescue lethality in a *Drosophila melanogaster* knockout model of AIP orthologue**. *16th European Congress of Endocrinology 2014* Vol. *Endocrine Abstracts* pg. 55 (Wrocław, Poland, May 2014) - **Oral Communication OC12.1.**

   Aflorei, E.D. et al. **Overexpression of human AIP gene can rescue the lethality of the knockout fruitfly model**; *XXII Romanian Congress of Endocrinology* (Targu Mures, Romania, October 2014) - **Oral Communication**.

   Aflorei, E.D. et al. **The role of AIP in the integrin pathway** in *Gondon Research Conference - Fibronectin, Integrins & Related Molecules* (Ventura, CA, February 2013). **P WED-THUR B1**

Aflorei, E.D. et al. ***Drosophila melanogaster*** **as a model organism to study aryl hydrocarbon receptor interacting protein (AIP) gene function**. *16th European Congress of Endocrinology, Copenhagen, Denmark* 27 April-1 May (2013), *Endocrine Abstracts* pg. 302, **P825**.

Aflorei, E.D. et al. **Danio rerio as a model organism for FIPA syndrome. Preliminary data**. *The Romanian Society of Endocrinology National Symposium with international participation (*2012), **Poster presentation**.

Aflorei, E.D. et al. **Aryl Hydrocarbon Receptor-Interacting Protein- (AIP-) homologue in the fruitfly**. *15th Congress of the European NeuroEndocrine Association Vol. Austrian Journal of Clinical Endocrinology and Metabolism* pg. 56 (Vienna, Austria, 2012). **P061**

Aflorei, E.D. et al. **AIP (aryl hydrocarbon receptor-interacting protein)-homologue in the fruitfly** in *William Harvey Research Institute Day* (London, UK, 2012), **Poster presentation**

Aflorei, E.D. et al. **Development of novel AIP (Aryl Hydrocarbon Receptor Interacting Protein) gene study models using the fruitfly and the zebrafish**. *15th International and 14th European Congress of Endocrinology* (Florence, Italy, 2012), **P1334**.

4. Other publications

Garcia, E. A., Trivellin, G., Aflorei, E. D., Powell, M., Grieve, J., Alusi, G., Pobereskin, L., Shariati, B., Cudlip, S., Roncaroli, F., Mendoza, N., Grossman, A. B., Harper, E. A. & Korbonits, M. **Characterization of SNARE proteins in human pituitary adenomas: targeted secretion inhibitors as a new strategy for the treatment of acromegaly?** *The Journal of clinical endocrinology and metabolism*, E1918-1926, (2013).

Musat, M., Stanescu, B., Grigorie, D., Danciulescu, R., Hortopan, D., Nicolicea, V., Aflorei, E. D. & Poiana, C. **Cushing Syndrome in pregnancy - review of the literature and case report**. *Gineco.ro* 5, 122-125, (2009).

.

ABSTRACT

A phenotypically distinct subgroup of familial isolated pituitary adenoma (FIPA) families has mutations in the aryl hydrocarbon receptor-interacting protein (AIP) gene leading to young-onset acromegaly in most patients. These patients typically develop invasive pituitary adenomas, but the mechanisms by which AIP inactivation promotes pituitary tumorigenesis and an aggressive behaviour remain unknown. To date, more than 70 different AIP variants have been reported and determining the pathogenicity of missense variants is a challenging problem.

The *Drosophila* AIP orthologue (*CG1847*) is located on the X chromosome and encodes a protein of similar size and structure to human protein (hAIP). I have generated CG1847 deficient flies via two methods: *in vivo* RNAi knockdown and imprecise excision of a transposable P-element, which generated a putative null allele of CG1847. Our data show that knockdown and knockout of CG1847 results in lethality confirming that *AIP* is an essential gene.

To reveal the potential underlying molecular mechanisms of loss of AIP, a whole transcriptome analysis was performed in mutant versus control male larvae. This allowed us to determine gene expression profiles and to identify key pathways that are significantly altered in the mutant, and that are related to embryonic development or survival.

To functionally test the homology between *hAIP* and *CG1847*, I used the Gal4/UAS system to perform rescue experiments. I subsequently tested whether wild-type hAIP, a truncated hAIP and four missense mutations identified in FIPA families could rescue the lethality of $CG1847^{exon1\_3}$ mutants by expressing hAIP during fly development.

In this thesis were identified novel *AIP* features. *CG1847* is a *Drosophila melanogaster AIP* orthologue and is essential for normal development. RNA sequencing revealed possible new underlying *CG1847* molecular mechanisms as the tumour suppressor function of AIP might involve the regulation of cytoskeletal organisation. *Drosophila* is a useful *in vivo* system to study human *AIP* missense variants to establish pathogenicity.

TABLE OF CONTENTS

STATEMENT OF ORIGINALITY....2
ABSTRACT....5
TABLE OF CONTENTS....6
LIST OF FIGURES....9
LIST OF TABLES....12
LIST OF MATING SCHEMES....13
LIST OF ABBREVIATIONS....14
ACKNOWLEDGEMENTS....16
CHAPTER 1: INTRODUCTION....18
1.1 Pituitary gland....18
1.2 Pituitary tumorigenesis....22
1.3 Familial isolated pituitary adenomas....36
1.4 *Drosophila melanogaster*: a model system....49
AIMS OF THE STUDY....56
Overview on the subject....56
The structure of this thesis....56
CHAPTER 2: MATERIALS AND METHODS....57
2.1 Materials....57
2.2 Standard fly techniques....60
2.3 General molecular biology techniques....76
2.4 Cloning....84
2.5 Illumina TruSeq stranded mRNA sample preparation - Low sample protocol....94
2.6 Sequence alignment to reference transcriptome....94
2.7 STRING analysis....98
2.8 Validation by multiplex- qPCR of selected transcripts....98
2.9 Statistical analysis....99
CHAPTER 3: ESTABLISHING A *DROSOPHILA MELANOGASTER* MODEL TO STUDY AIP FUNCTION....100
3.1 INTRODUCTION....100
3.2 BACKGROUND....101

3.3 OBJECTIVES 106

3.4 RESULTS 106

3.5 DISCUSSION 135

3.6 CONCLUSION 147

CHAPTER 4: THE *DROSOPHILA AIP* ORTHOLOG IS ESSENTIAL FOR ACTIN CYTOSKELETON STABILISATION AND CELL ADHESION 148

4.1 INTRODUCTION 148

4.2 BACKGROUND 149

4.3 OBJECTIVES 153

4.4 RESULTS 153

4.5 DISCUSSION 180

4.5 CONCLUSIONS 189

CHAPTER 5: AN *IN VIVO* SYSTEM TO TEST THE PATHOGENICITY OF AIP MISSENSE MUTATIONS 191

5.1 INTRODUCTION 191

5.2 BACKGROUND 193

5.3 OBJECTIVES 195

5.4 RESULTS 195

5.5 DISCUSSION 209

5.6 CONCLUSIONS 218

CHAPTER 6: GENERAL CONCLUSIONS AND FUTURE DIRECTIONS 219

6.1 *CG1847* is a *Drosophila melanogaster AIP* orthologue and is essential for normal development 219

6.2 RNA sequencing reveals possible new underlying CG1847 molecular mechanisms 220

6.3 The cytoskeletal disorganisation might be related to CG1847 loss of function, this being the mechanism for the tumour suppressor function of AIP 221

6.4 Human AIP variants have different capacities for compensating for CG1847 loss of function in an *in vivo* model 222

APPENDICES 223

Appendix 1 AIP sequence..... 223
Appendix 2 CG1847 sequence 224
Appendix 3: Fly food recipe 225
Appendix 4: Structure of Inverted Repeats (IR) for RNAi constructs..... 225
Appendix 5: Primers used in this study..... 227
Appendix 6: Immunostaining protocols..... 230
Appendix 7: Antibodies used in this study..... 233
Appendix 8: Preparing competent cells..... 234
Appendix 9: TruSeq Stranded mRNA Sample Preparation – Low Sample (LS) Protocol..... 236
Appendix 10: RNA-seq – Table..... 244
REFERENCE LIST 258

LIST OF FIGURES

Figure 1: Molecular regulation of anterior pituitary gland development.....20

Figure 2: Hormonal axis, positive and negatives feedback loops.....22

Figure 3: Cascade of factors involved in anterior pituitary development and tumorigenesis.....28

Figure 4: The pathogenesis of pituitary tumours due to germline mutations.....34

Figure 5: A schematic illustration of the alternative splicing of human *AIP* into four isoforms.....39

Figure 6: The AIP crystal structure.....40

Figure 7: Representation of the TPR motifs of AIP.....41

Figure 8: Life cycle of *Drosophila melanogaster*.....52

Figure 9: Mitotic chromosomes of *D. melanogaster*.....53

Figure 10: Common phenotypic markers used in this study.....54

Figure 11: A schematic illustration of the alternative splicing of CG1847 into two isoforms.....55

Figure 12: Shematic representation of UCSC Browser Blat alignment of IR mapping to *CG1847*.....60

Figure 13: Design for larval staging.....73

Figure 14: Schematic structure of the pGEM-T Easy Vector.....87

Figure 15: Schematic structure of genomic rescue construct pW@RpA+CG184.....89

Figure 16: Schematic structure of genomic rescue construct pUASk10attB-hAIPwt.....89

Figure 17: Schematic structure of genomic rescue construct pUASk10attB-hAIPwt.....91

Figure 18: FastQC quality control of cDNA libraries.....95

Figure 19: Overview of the bioinformatic workflow.....97

Figure 20: UAS-Gal4 system in *Drosophila*.....101

Figure 21: Schematic diagram of the phiC31 integration system into the *Drosophila* genome.....103

Figure 22: RNA-Seq workflow.....105

Figure 23: Sequence alignment of *Drosophila* CG1847 protein and human AIP.....107

Figure 24: Expression of *CG1847* during development in different tissues and organs.....108

Figure 25: Three dimensional theoretical model of *CG1847*.....109

Figure 26: CG1847 RNAi knockdown.....110

Figure 27: Quantification of *CG1847* RNAi knockdown efficiency.....111

Figure 28: Schematic representation for primer annealing.....112

Figure 29: Agarose gel for selected number of *CG1847* mutant alleles.....113

Figure 30: Creation of *CG1847* knockout fly line.....114

Figure 31: Schematic diagram of the CG1847 locus in wt and $CG1847^{exon1\_3}$ mutants.....114

Figure 32: Agarose gel of PCR results - selected stocks resulted from the P-element excison screen..115

Figure 33. Comparation between the *CG1847*wt sequence and the chromatograms of $CG1847^{2.39A}$. .115

Figure 34: BLAST alignment.....116

Figure 35: *CG1847* expression in mutant male larvae.....117

Figure 36: Survival rate between 24 and 96 h AEL.....119

Figure 37: Delayed larval development in $CG1847^{exon1\_3}$ mutants.....120

Figure 38: Results of rescue experiments. 121
Figure 39: Rescue experiment of the CG1847 lethal phenotype with a wt CG1847 construct. 122
Figure 40: Chi squared contingency test for the distribution of female's genotypes. 123
Figure 41: Rescue of the *CG1847* lethal phenotype by ubiquitous or tissue-targeted overexpression of wt hAIP. 124
Figure 42: RNA Quality Control. 126
Figure 43: RNA levels for heat shock protein genes are downregulated in *CG1847* mutants. 128
Figure 44: STRING analysis: the heat shock protein cluster. 129
Figure 45: STRING analysis: the Osiris, Tweedl and cuticule proteins cluster 130
Figure 46: RNA levels for Osiris and Tweedl transcripts are significantly changed in *CG1847* mutants. 130
Figure 47: Schematic representation of the 3 main stepts of GeXP Genetic amplification. 133
Figure 48: Representative electropherograms corresponding to gene expression profiles. 133
Figure 49: Multiplex qPCR results. 134
Figure 50: Early wing development. 150
Figure 51: A schematic summary of adhesion mechanism. 151
Figure 52: Schematic representation of the FLP/FRT system used to generate mosaic clones. 152
Figure 53: The CG1847 silencing in wing discs produces wing blisters. 154
Figure 54: Quantification of blister formation. 155
Figure 55: The CG1847 silencing with UAS-*CG1847-RNAi-T2* produces wing blisters. 156
Figure 56: Quantification of blister formation. 156
Figure 57: Homozygous *CG1847* mutant cell clones induced by the FLP/FRT system. 158
Figure 58: Blister formation in adult females with mitotic clones. 159
Figure 59: Quantification of the area size of of mitotic recombined clones. 160
Figure 60: Wing disc clones stained with DAPI and phalloidin. 161
Figure 61: Quantification of cell density. 162
Figure 62: CG1847 deficient mutants display normal muscle attachment sites pattern 164
Figure 63: Normal distribution of βPS integrin, integrin-associated proteins and actin at the basal surface. 166
Figure 64: CG1847 silencing resulted in actin disorganisation 167
Figure 65: Schematic representation of the orthogonal sections. 168
Figure 66: *CG1847* silencing specifically in the wing resulted in loss of adhesion. 169
Figure 67: Overexpression of *CG1847-RNAi* in the wing causes persistence of gaps in the actin layer at 28h APF. 170
Figure 68: Overexpression of *CG1847-RNAi* in the wings resulted in wider veins. 171
Figure 69: CG1847 controls actin and PINCH stability in pupal wings. 172
Figure 70: βPS and Parvin are normal in pupal wings lacking *CG1847*. 173

Figure 71: Significant difference in actin and PINCH staining intensity between CG1847 knockdown samples and control. 174
Figure 72: Overexpressed human AIP is localised in the cytoplasm of *Drosophila* wing cells. 176
Figure 73: Quantitative PCR analysis of transcripts with a possible involvement in blister formation in *Drosophila* adult wings. 179
Figure 74: Genetic hierarchy of integrin–actin linker complex assembly at muscle attachment sites and in wing epithelium. 184
Figure 75: Schematic representation of the rescue experiments. 194
Figure 76: Schematic diagram of the *UAS-hAIP* constructs. 198
Figure 77: hAIP chromatograms. 199
Figure 78: Results of rescue experiment with wild-type AIP. 200
Figure 79: *pUAS-hAIPtrunc* DNA to protein sequence translation. 201
Figure 80: Quantitative analysis of *in vivo* rescue experiments using hAIP missense variants. 202
Figure 81: Similar distribution of the males phenotypes 203
Figure 82: Genotyping of rescued males with hAIPwt. 205
Figure 83: Two types of F1 males due to nondisjunction. 206
Figure 84: Genotyping for the two aberrant male types. 206
Figure 85: Genotyping of non-Fm6 non-CyO males resulting from crosses with the hAIPtrunc construct. 207
Figure 86: Overexpression of pUAS-hAIP constructs driven by actin-Gal4. 208
Figure 87 Overview of normal meiosis in the *CG1847*$^{exon1\_3}$ mutant stock. 216
Figure 88: Primary nondisjunction in *CG1847*$^{exon1\_3}$ mutant stock. 217
Figure 89: Secondary non-disjunction in the female. 217
Figure 90: TruSeq Stranded mRNA Sample Preparation LS Workflow 236

LIST OF TABLES

Table 1: Population based studies. **25**

Table 2: Selected genes that may be involved in molecular pathogenesis of pituitary adenomas. **31**

Table 3: Interacting partners of aryl hydrocarbon receptor-interacting protein (AIP). **43**

Table 4: *Drosophila melanogaster* lines used in this study. **58**

Table 5: *Drosophila melanogaster* transgenic stocks generated and used in this study **59**

Table 6: The reaction mix used for semi-quantitative RT-PCR **79**

Table 7: The reaction mix used to identify large genomic deletions **80**

Table 8: The reaction mix used to identify small genomic deletions **80**

Table 9: The reaction mix for direct site mutagenesis **82**

Table 10: List of plasmids used in this study. **84**

Table 11: The reaction mix used to generate gene specific constructs **85**

Table 12: The typical Taq DNA polymerase reaction mix for adding A' overhangs: **86**

Table 13: Missense variants generated in the study **91**

Table 14: *E. coli* strains used in this study **92**

Table 15: Summary of P-element excision screen. **112**

Table 16: Determination of lethality stage during development. **118**

Table 17: Example of an excerpt from merged.gtf file **127**

Table 18: Human orthologues for *Drosophila* transcripts **131**

Table 19: modENCODE RNAi screen for CG1847 ..... **140**

Table 20: Quantification of blister formation in CG1847^{exon1_3}, FRT19A /FRT19A females. **159**

Table 21: Quantification of blister formation in control *w,Dm,FRT19A /FRT19A* females **160**

Table 22: List of genes with a putative involvement in blister formation **178**

LIST OF MATING SCHEMES

Mating scheme 1: Crossing scheme for generating CG1847 mutant fly 62

Mating scheme 2: Replacement of balancer chromosome in the *CG1847* mutant stock.....63

Mating scheme 3: Crossing scheme for generating revertant stock.....63

Mating scheme 4: Replacement of balancer chromosome in the control stock.....63

Mating scheme 5: Recombination of the *CG1847*$^{exon1\_3}$ mutant allele and the FRT19A site. 64

Mating scheme 6: Recombination of the *CG1847*$^{exon1\_3}$ mutant and *mys*$^{5.4}$ 65

Mating scheme 7: Combination of the *CG1847*$^{exon1\_3}$ mutant and actin-Gal4 driver 67

Mating scheme 8: Combination of the *CG1847*$^{exon1\_3}$ mutant and fat body driver *Cg*-Gal4 67

Mating scheme 9: Combination of the *CG1847*$^{exon1\_3}$ mutant and haemocytes driver Crq-Gal4.....68

Mating scheme 10: Combination of the *CG1847*$^{exon1\_3}$ mutant and insulin secreting cells driver dilp-Gal4 68

Mating scheme 11: Combination of the *CG1847*$^{exon1\_3}$ mutant and muscle cells driver Mef2-Gal4 69

Mating scheme 12: Combination of the *CG1847*$^{exon1\_3}$ mutant and nervous system driver elav-Gal4 ...69

Mating scheme 13: Combination of the *CG1847*$^{exon1\_3}$ mutant and glial cells driver repo-Gal4.....70

Mating scheme 14: Combination of the *CG1847*$^{exon1\_3}$ mutant and haemolymph driver HE-Gal4 70

Mating scheme 15: Combination of the *CG1847*$^{exon1\_3}$ mutant and gut driver drm-Gal4.....70

Mating scheme 16: Combination of the *CG1847*$^{exon1\_3}$ mutant and malpighian tubules driver c42-Gal4 71

Mating scheme 17: Combination of the *CG1847*$^{exon1\_3}$ mutant and fat heart driver tinC-Gal4 71

LIST OF ABBREVIATIONS

♂	male
♀	virgin female
A	adenine
aa	Amino acids
AC-L	Affinity Capture-Luminescence
AC-MS	Affinity Capture-MS
ACTH	Adrenocorticotrophic hormone
AEL	After egg laying
AFP	After puparium formation
AhR	Aryl hydrocarbon receptor
AIP	Aryl hydrocarbon receptor interacting protein
AIP	Aryl hydrocarbon receptor interacting protein gene
AIPmut	Mutation in the aryl hydrocarbon receptor interacting protein gene
AJs	Adherent junctions
ANKRs	Ankyrin repeats
ARNT	Aryl hydrocarbon receptor nuclear translocator
ATP	Adenosine triphosphate
bp	Base pairs
β-actin	Beta actin
BJs	Basal junctions
BSA	Bovine serum albumin
C	Cytosine
cDNA	Complementary deoxyribonucleic acid
CDK	Cyclin dependent kinase
cAMP	Cyclic adenosine monophosphate
CNC	Carney Complex
cGMP	cyclic guanosine monophosphate
cM	centi-Morgan
CNC	Carney complex
CDKN1B	cyclin-dependent kinase inhibitor 1B Gene
CNS	Central nervous system
CPR	Cuticular proteins
CRH	Corticotrophin-releasing hormone
DAPI	4',6-diamidino-2-phenylindole

DBS	Double-stranded break
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dNTP	Deoxyribonucleotide triphosphate
dTTP	2'-deoxythymidine 5'-triphosphate
E-Cad	E-Cadherin
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
ERK	Extracellular signal-regulated kinase
FCS	Foetal calf serum
FERM	band 4.1, ezrin, radixin, and moesin
FIPA	Familial isolated pituitary adenoma
FKBP	FK506 binding protein
FSH	Follicle-stimulating hormone
G	Guanine
Gsα	G protein α s subunit
Gb	Giga bits
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	green fluorescent protein
GH	Growth hormone
GHRH	Growth hormone-releasing hormone
GNAS	Guanine nucleotide-binding protein, alpha stimulating activity polypeptide
GPCR	G protein-coupled receptor
GPR101	G protein-coupled receptor 101 gene
hAIP	Human AIP
hh	hedgehog
hsp90	Heat shock protein 90
IFS	Isolated familial somatotrophinoma
IR	Inverted repeats

IGF-I	Insulin-like growth factor I
IHC	Immunohistochemistry
IPP	ILK-PINCH-Parvin complex
Kb	Kilobase pairs
KD	Knock down
LH	Luteinising hormone
LOH	Loss of heterozygosity
MASs	Muscle attachment sites
Mb	Mega bases
MAPK	Mitogen-activated protein kinase
MAS	McCune-Albright syndrome
Min	minutes
ml	Millilitres
MEN	Multiple endocrine neoplasia
mod ENCODE	model organism ENCyclopedia Of DNA Elements
mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
NF1	Neurofibromatosis 1
NFPA	Non-functioning pituitary adenoma
ns	not significant
P{EP}	P-element construct
PAP	Pituitary adenoma predisposition
PBS	Phosphate buffered saline
PBT	Phosphate buffered saline + Triton-X 100
PCR	Polymerase chain reaction
PDE	Phosphodiesterases
PIP	PINCH-ILK-Parvin complex
PKA	Protein kinase A
PG	Prothoracic Gland
PI	Pars intercerebralis
PL	Pars lateralis
PKA	Protein kinase A
PP	Prepupae
PRKAR1A	cAMP-dependent protein kinase A type 1-alpha regulatory subunit gene
PRL	Prolactin
PS integrins	position-specific (PS) integrins

PTTH	Prothoracicotropic hormone
q	long arm of a chromosome
Rb	Retinoblastoma gene
RC	Reconstituted Complex
RG	Ring Gland
RET	Rearranged during transfection proto-oncogene
RFP	Red fluorescent protein
RNase	Ribonuclease
RNA	Ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
Rpl32	Ribosomal protein 32
rpm	Rotations per minute
RT-PCR	Reverse transcription-polymerase chain reaction
RT	Room temperature
SE	Standard error
sec	Seconds
siRNAs	Short-interfering RNA molecules
SNP	Single nucleotide polymorphism
ss	Spineless
SSTR	Somatostatin receptor
T	Thymine
TAE	Tris-acetate-ethylenediaminetetraacetic acid
T3/T4	Thyroid hormones
TBS	Tris buffered saline
TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TRH	Thyrotrophin-releasing hormone
TPR	Tetratricopeptide repeat
TSG	Tumour suppressor gene
TSH	Thyroid-stimulating hormone
Tris	Tris(hydroxymethyl)aminomethane
USP8	Ubiquitin specific peptidase 8 gene
UTR	Untranslated region
VDRC	Vienna *Drosophila* RNAi Center
wt	wild-type
Zf	Zebrafish
WPP	White pre-pupae
Y2H	Two yeast hybrid

ACKNOWLEDGEMENTS

A PhD is an exhausting process that goes far beyond the hours spent in the lab during those very long 4 years. At its end I can appreciate that it was a highly rewarding experience that wouldn't have been possible without the support of so many people. I'm afraid I will inevitably forget some of them. To them I apologise in advance.

First, I especially thank Professor Márta Korbonits, who was not only my supervisor, but also a mentor and a trustful friend. Márta gave me the liberty to think independently and critically, and she taught me how to communicate and present effectively. She allowed me to use my creativity in research as she always encouraged and welcomed my new, maybe sometimes naive, ideas. I consider myself extremely fortunate to have encountered her in my early scientific career.

I enclose here a few words for my family:

> Suportul unei familii adevărate este neprețuit. Dragii mei părinți, voi ați fost sursa puterii mele. Vă mulțumesc pentru felul în care m-ați crescut şi educat, pentru omul puternic care am devenit astăzi. Vă mulțumesc că mi-ați oferit sprijin şi suport chiar şi atunci când poate nu ați înțeles foarte bine încotro mi se îndreaptau gândurile şi paşii. Vă mulțumesc din inimă.
>
> Câteva cuvinte de mulțumire mi se par insuficiente pentru cel care a fost cel mai puternic sprijin atunci când doctoratul mi s-a părut mai greu decat aş fi putut eu duce. Pentru tot suportul psihic îi mulțumesc celui mai minunat bărbat, cel care a fost lumina mea din întuneric. Cosmin, tu ai crezut în mine şi m-ai învățat că dragostea şi încrederea sunt posibile indiferent de distantă. Ți mulțumesc pentru libertatea şi suportul oferite.

In this amazing journey I had 2 special companions, two very amazing ladies from which I have learned that friendship is possible beyond cultural backgrounds. Thank you Laura and Ida. I find it fascinating the fact that life brought us together: three very different women, with different life values, but who understood each other more than anybody else. You will be greatly missed!

A huge THANK YOU goes to all my friends in the Endocrinology department and all the staff that made the lab a joyful place to work. I treasured every minute spent in your company. Special thanks for a few special people (more or less in order of appearance), who supported me in many situations: Sayka Barry, Edwin Garcia, Giampaolo Trivellin, Federico Martucci, Craig Stiles,

Judit Dénes, Maria Hérincs, Plamena Gabrovska, Karen Young, Francesco Ferraù, Donato Iacovazzo, Francisca Caimari and Mary Dang.

For the collaboration and helpful support, I thank Professor Ralf Stanewsky and his very patient postdoc, Chenghao, for guiding me at the beginning of my journey in *Drosophila melanogaster* world. Of course, special thanks to the rest of his team, some of the people I met during these years: Min, Werner, Joanna, Adam, Maite and Sanne.

I thank Nic Tapon and John Marshall for their advice, guidance, help and support.

I am infinitely grateful to Nick Brown for his support regarding antibodies, fly stocks, but most precious, the extremely useful comments on the project. A significant part of this thesis it would not have been possible without the help of Nick' amazing postdoc, Benjamin Klapholz, who was patient enough to answer to all my very, very, long emails.

Unfortunately I met Dr Paulo Ribeiro only towards the end of my PhD. However, without his help and without the support of his team, this thesis wouldn't have had the results that it has today. Thank you Paulo for "adopting me". Especially, thank you Nina for sharing so much of your knowledge with me. You will remain an inspiration for me. Thank you Cecilia and Alex for being so patient with me and for sharing the microscopes.

I am especially grateful to Rob Fowkes and Samantha Mirczuk (Royal Veterinary College, London) for all the help regarding Multiplex qPCR.

In the international fly research community there is a free exchange of reagents and I would like to thank to Professor Christos Zervas and Professor Mary Beckerle for sending me the necessary antibodies.

I also have to express my gratitude to my two great mentors and friends, who both encouraged me and guided my steps towards a scientific career: Dr Madalina Musat and Serban Radian.

There were also many things that made this PhD possible. Thanks PubMed, FlyBase, Ensemble, Fly pushing. Thanks Wikipedia for answering to all my questions. Thanks Facebook for keeping my old friends so close. Thank you London for SouthBank Centre, Barbican Centre and RichMix, for museums, movies and music.

As unfortunately research, even though fascinating, is not possible without financial support, I would like to acknowledge William Harvey Research Foundation for generously funding the 3 years of research. I'm also grateful to the Society for Endocrinology and the European Society of Endocrinology who funded my attendances to national and international conferences that strongly influenced some of the ideas presented here.

CHAPTER 1: INTRODUCTION

1.1 Pituitary gland

The pituitary gland, or hypophysis, is one of the most important glands of the human endocrine system as it is involved in controlling the normal function of other endocrine glands, such as the thyroid gland, adrenals, and gonads. Its role is very complex as the different hormones secreted at this level can directly or indirectly regulate numerous processes concerning normal growth and development, sexual development and reproduction, metabolism, thermoregulation, stress response, sleep, and adaptation to changes in the external environment.

Anatomy of pituitary gland

The hypophysis is a small endocrine gland, weighing around 600 mg. The size of a normal adult pituitary gland was found to vary between 2.7 and 6.7 mm in a series of 14 women without known sellar or parasellar lesions[1], but even sizes of around 9.7 mm were reported as normal by a study of direct coronal scans in 50 normal female volunteers[2]. The pituitary gland has two main parts: adenohypophysis (the anterior pituitary) and neurohypophysis (the posterior pituitary). The hypophysis is connected to the hypothalamus by the pituitary stalk, which consists of blood vessels and axons of the hypothalamic neuronal cell bodies. The hypophysis and the hypothalamus, compose the "hypothalamo-pituitary axis"[3].

This gland is located at the base of the skull, above the sphenoid sinus, inside a bony cavity called the "sella turcica". This area is adjacent to many vascular and neurologic structures, hence processes which cause enlargement of the gland, may lead to significant mass effects. On the lateral sides of pituitary gland are the cavernous sinuses, which contain important structures as the internal carotid artery, and the cranial nerves III, IV, 2 of the 3 branches of cranial nerve V and VI. A pituitary tumour that extends laterally and invades the cavernous sinus may cause oculomotor paralysis and can sometimes even result in cerebral ischemia if the internal carotid artery is compressed[4]. A major neurologic structure located immediately above the hypophysis is the optic chiasm[5]. The optic chiasm is separated from the pituitary only by the diaphragma sella, a recess of the dura mater. This neurologic structure can be easily compromised through a mass effect from a pituitary adenoma. The most frequent manifestation of chiasmal disfunction is bitemporal hemianopia. As the pituitary gland is localised very close to the optic nerve chiasma, a tumorous process might result in partial blindness in the outer half of both visual fields. Headache is another complaint frequently reported in cases of pituitary adenomas; however, the mechanism of its appearance is the same as in case of any other space-occupying

brain lesion[6], as traction on the pain-sensitive structures[7] and dural stretch[8,9]. For example large tumors with cavernous sinus invasion might lead to headache as the sinus does contain pain-producing structures, as the internal carotid artery and trigeminal nerve and ganglion. However, other factors apart from tumor size, such as family history of headache or the type of hormonal secretion might be important factors for the etiology of pituitary adenoma related headache[10].

The anterior pituitary gland secretes six major hormones: growth hormone (GH), prolactin (PRL), adrenocorticotropin hormone (ACTH), luteinizing hormone (LH), follicle-stimulating hormone (FSH), and thyroid-stimulating hormone (TSH). Around 50% of anterior lobe cells are GH secreting cells[11]. The main role of GH is in controlling body growth. This hormone acts either directly on multiple tissues or indirectly, by stimulating the tissue production of insulin-like growth factors (IGFs, mainly IGF-I). In men and nulliparous women, PRL secreting cells or lactotrophs may account for approximately 10% of the anterior pituitary cells. PRL stimulates breast enlargement and milk production during and after pregnancy and, as a result, in multiparous women the number of lactotrophs can be increased. ACTH is secreted together with pro-opiomelanocortin (POMC) derivatives by corticotroph cells (approximately 10-20% of anterior lobe). ACTH has a crucial role in the secretion of glucocorticoid hormone (cortisol) from the adrenal gland cortex. FSH and LH secreting cells, or gonadotroph cells, account also for around 10% of the anterior pituitary cells and are involved in the sex steroid hormone production and regulation of germ cell maturation. Thyrotroph cells (only 5% of the adenohypophyseal cells) secrete TSH, which stimulates thyroid hormone (T3/T4) production in the thyroid follicles. Thyroid hormone mainly controls metabolism and thermogenesis[12].

The posterior lobe consists mainly of the axons of neurons that are localised in the hypothalamus - neurons that secrete the antidiuretic hormone (ADH), involved in regulation of water retention, and oxytocin, a hormone with a role in social interactions and uterine contractions during delivery. As the posterior pituitary lobe is not the subject of this thesis, it will not be presented in more detail.

Pituitary development

The embryonic pituitary development was subject of numerous studies, and the maturation process of anterior pituitary cells is very well understood. From a developmental perspective, the anterior pituitary cells are embryologically derived from Rathke's pouch[13], a process that involves the complex interplay of lineage-specific transcription factors and locally-produced growth factors[14] (Figure 1).

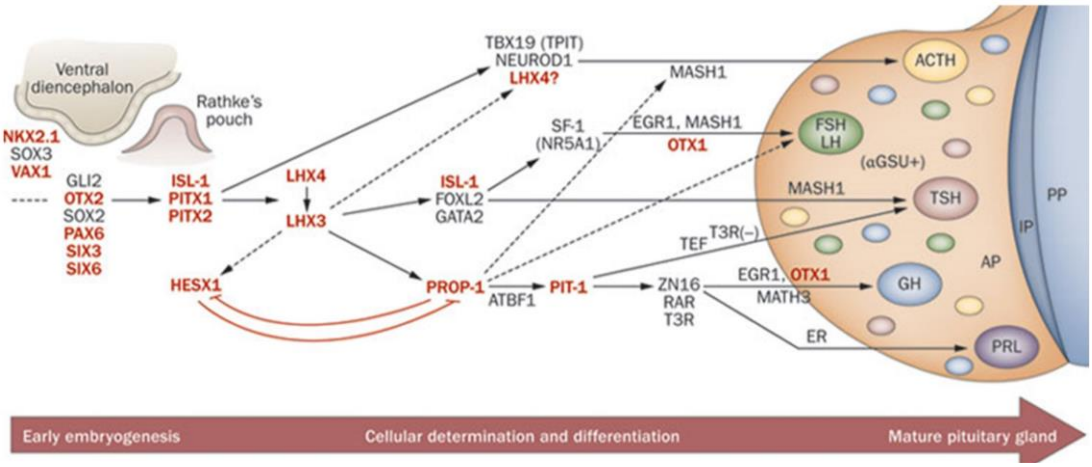

Figure 1: Molecular regulation of anterior pituitary gland development. Multiple transcription factors contribute to the establishment of the first structure of the developing pituitary gland, Rathke's pouch, and the subsequent differentiation of the five specialized, hormone-secreting cell types characteristic of the mature anterior pituitary gland: corticotrophs (ACTH), gonadotrophs (FSH and LH), thyrotrophs (TSH), somatotrophs (GH) and lactotrophs (PRL). Homeodomain-containing transcription factors critical to this process are highlighted in red, but the diagram includes other relevant transcription factors to give a broad picture of the cascade. Arrows indicate upstream relationships in molecular signalling pathways but do not necessarily imply direct activation. Flat arrowheads denote repressive relationships. The placement of specific cell types in the diagram does not reflect their actual location within the anterior pituitary gland. Adapted from Prince, K. L. *et al*. (2011)[15]

Rathke's pouch is an ectodermal invagination that is formed anterior to the roof of the oral cavity at a very early stage of development, between the fourth and fifth week of gestation. Later in development, the pouch becomes distinct from the oral cavity and nasopharynx but maintains the connections with the stalk and hypothalamic infundibulum[16]. In the initial stages, Rathke's pouch cells express several transcription factors as LHX3, LHX4, and ISL-1, which belong to the LIM homeodomain family. These factors have an essential role in the very early determination towards pituitary development[17]. Recent studies have shown that SOX2 and SOX3 play key roles in regulating early pituitary morphogenesis both in rodent and man[18]. PITX1 and PITX2 are other essential transcription factors which are initially expressed in the oral ectoderm, and subsequently these factors will have a major role in the normal development of all pituitary cell types[19,20].

In the anterior lobe, the development of different hormonal cell lineages is under the influence of other transcription factors. Maybe the most important transcription factor is PROP1 which directly or indirectly controls the development of all anterior pituitary cell lineages. Somatotrophs, lactotrophs and thyrotrophs originate from a common lineage, under the control of PROP1 and Pit-1 expression. Both factors are essential for GH, PRL, and TSH secretion[21]. PROP1 in conjunction with other transcription factors is important for corticotrophs, and gonadotrophs lineages. Gonadotroph cell development is influenced by the expression of two

nuclear receptors, SF-1 (steroidogenic factor) and DAX-1 (dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1)[22]. PROP1 also have an indirect influence on development of LF and FSH secreting cells. Among these numerous transcription factors involved in the terminal differentiation of pituitary cell types, the T-Pit transcription factor is important for the development of corticotrophs, which express the pro-opiomelanocortin[23]. Mutations in the genes encoding the transcription factors involved in the linear development of anterior pituitary gland, such as Pit-1, PROP1, SF-1, DAX-1, and T-Pit, result in different pathologies that involve selective or combined pituitary hormone deficits. Fortunately, these are rare abnormalities.

Regulation

All the pituitary hormones are secreted in a pulsatile manner, reflecting the fact that the pituitary gland is under the control of the nervous system through the hypothalamus. Various external stimuli, such as ambient temperature, level of physical exercise, physical or psychological stress, or supplied nutrients, lead to secretion of specific hypothalamic releasing or inhibitory factors. The hypothalamic factors act on the surface receptors of specific pituitary cells, and as a response, pituitary hormones will be synthesised, secreted or inhibited. The pituitary hormones elicit specific responses in peripheral target tissues[12].

The stimulated peripheral glands will be consequently followed by specific hormone production that, in turn, will act via a feedback loop to control anterior pituitary function. There are mainly two mechanisms by which the hormonal products of peripheral glands, in turn, control at the level of the hypothalamus and pituitary: positive and negative feedback (Figure 2).

The hypothalamus controls expression and secretion of anterior pituitary hormones through positive or negative regulation by releasing or inhibiting hypothalamic factors that act on their corresponding receptors on the adenohypophyseal cells. Through this mechanism, GH secretion is stimulated by GH-releasing hormone (GHRH); GH acts in the liver and other target tissues by mediating IGF-I production. Corticotrophin-releasing hormone (CRH) is secreted by hypothalamic neurons and induces positive regulation of ACTH secretion; the final result is glucocorticoid secretion from the adrenal glands' cortex. Pituitary TSH is stimulated by a hypothalamic hormone, thyrotropin-releasing hormone (TRH). TSH acts through the same positive regulation on thyroid cells to promote thyroid hormone secretion (T3/T4). Gonadotroph cells have specific receptors for gonadotropin-releasing hormone (GnRH) and, when stimulated, will produce variable amounts of LH and FSH, depending on the frequency and amplitude of GnRH pulses[24]. In turn, LH and FSH control the production of sex steroid hormones (oestradiol,

progesterone, and testosterone) secreted from the ovaries and testes. PRL is the only pituitary hormone that is not stimulated by hypothalamic releasing factors.

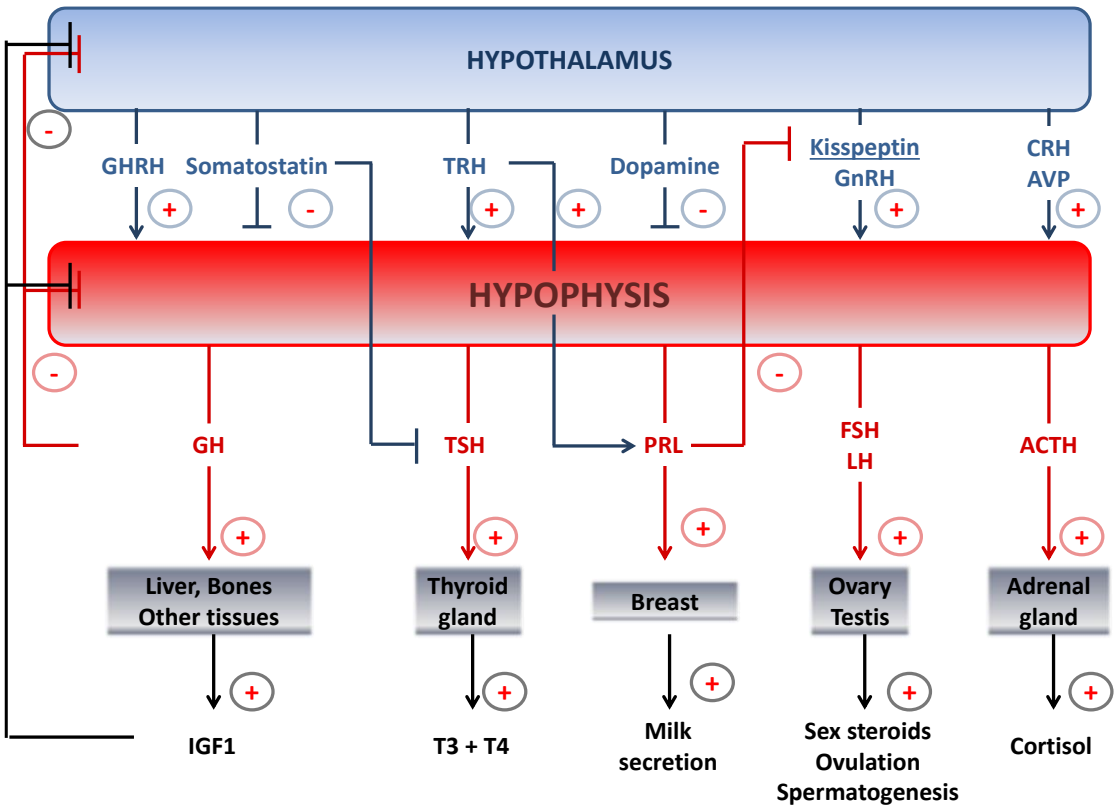

Figure 2: Hormonal axis, positive and negatives feedback loops. Schematic representation of the mechanisms regulating the anterior pituitary hormone secretion at the pituitary and at the hypothalamic level. Positive regulation is indicated by arrows, whereas negative regulation is shown by flat arrowheads.

The negative feedback mechanisms are exerted both by hypothalamic inhibitory factors and by hormones released from the target glands[25,26]. Pituitary GH secretion is inhibited both by hypothalamic somatostatin and by IGF-I. In the case of PRL, the negative influence is exerted by dopamine. The glucocorticoid hormones secreted by adrenal glands have an inhibitory effect on the secretion of both ACTH and CRH. TSH and TRH are under the negative control of thyroid hormones produced in the thyroid gland. FSH and LH (and the hypothalamic factor GnRH) are inhibited by a negative feedback of the sex steroid hormones[24].

1.2 Pituitary tumorigenesis

Definition

Pituitary adenomas can arise from each of the cell types of the anterior pituitary (except of folliculo stellate cells). These tumours are usually benign monoclonal tumours[27], presenting either due to hypersecretion of pituitary hormones, and/or due to local space occupying effects

and hyposecretion of some or all pituitary hormones. The main pituitary tumour types are prolactinomas, clinically non-functioning pituitary adenomas (NFPAs), GH-secreting adenomas and ACTH-secreting adenomas. TSH-secreting adenomas and gonadotroph adenomas with clinically relevant LH/FSH secretion are less frequent. A recently identified subgroup of pituitary adenomas detected as incidental findings during brain imaging are often referred as pituitary incidentalomas. The vast majority of pituitary tumours are benign, with very slow growing rates of the intrasellar masses, although dural invasion is present in almost 45% of the cases[28]. In these cases, the pituitary tumours present obvious extrasellar extension and invasion. Pituitary carcinomas are extremely rare cases, representing less than 0.2% of pituitary tumours; they are defined as pituitary tumours with a distant metastasis[29]. Pituitary adenomas are classified as microadenomas when they are less than 1 cm in diameter. They are usually restricted to the sella turcica, and usually do not have significant compressive effects on the surrounding tissue. Macroadenomas account for around 40% of all pituitary adenomas, can put pressure on the optic chiasm and the pituitary stalk, and can invade areas around the pituitary gland such as the cavernous sinus, the suprasellar area or the sphenoid sinus.

Epidemiology of pituitary adenomas

While previously pituitary adenomas were considered rare, recent studies have shown that the prevalence of pituitary tumours is higher than expected, probably due to better imaging techniques and improved diagnostic modalities. Pituitary adenomas account for 15% of all intracranial neoplasms, being the third most frequent tumour type after meningiomas and gliomas[30].

Data from national cancer registries, autopsy studies, imaging data, referrals to specialized centers, or population-based studies are the usual source for pituitary adenoma epidemiological statistics. Unfortunately, some of these sources are not entirely reliable. Cancer registries are extremely useful databases to monitor new disease cases every year at a national level. However, due to lack of mandatory reporting, the data available is actually an underestimation of the real prevalence. One of the first studies on the prevalence of pituitary adenomas based on autopsy samples was published by Costello et al. 1936[31]. In 22.5% of unselected autopsy cases a pituitary adenoma was identified. More recent studies have found similar detection frequencies depending on the thickness of sections made through the pituitary gland. Therefore, it is becoming increasingly clear that pituitary adenomas are frequently found during autopsy of individuals from the general population. Still, when assessing autopsy data it is necessary to take into account that these studies usually represent a relatively older population.

In 2004, Ezzat *et al*.[32] performed a meta-analysis of the prevalence of pituitary adenomas based on the English-language articles available at the time and identified 33 articles based on autopsy and radiological data. The results of the final statistical analysis suggest a wide interval in the prevalence of pituitary adenomas. The imaging studies were ranging from 1% to almost 40% (with an estimated mean prevalence rate of 22.5%). With regards to the postmortem studies, the prevalence of pituitary adenomas had a similar interval: 1% to 35% with an estimated mean of 14.4%. The overall estimated prevalence of pituitary adenomas in the general population was calculated by Ezzat *et al*. as 16.7%[32]. Community-based cross-sectional studies usually report either pituitary adenoma prevalence[33-35], or their incidence[36], although a recent article from Malta evaluated both of these parameters[37].

The first study that aimed to evaluate the prevalence of pituitary adenomas in a well-defined population was published by Daly *et al*. in 2006[33]. They evaluated only clinically relevant pituitary adenomas and demonstrated a prevalence of one case per 1,064. These results were replicated by Fernandez *et al*.[35] in the UK showing a prevalence of one case per 1,289. The results from both studies show significantly higher prevalence compared to previous studies based on data from tertiary referral centres (between 1:5,263 and 1:3,571[38]). This significant difference is probably due to different patient identification methods, and better imaging and diagnostic practices. These data were subsequently confirmed by other population-based studies (Table 1).

The most recent epidemiologic study regarding the prevalence and incidence of pituitary adenomas took place in Iceland. It was a retrospective observational study and included all pituitary adenomas diagnosed on this area over more than 40 years. An extensive clinical database was generated and used for calculating the rates of prevalence and incidence. The authors confirmed the increased overall prevalence and incidence rates. In this study it is claimed that these findings cannot be totally justified by a facilitated access to imaging investigations in recent years as most of patients were already symptomatic at diagnosis[39]. However, the use of performing imaging investigations in recent years has had a significant impact on the diagnosis of pituitary lesions with or, more often without, clinical relevance. As a result, a clinical guideline on the management of pituitary incidentalomas was necessary and has been recently published[40]. The Endocrine Society recommendations for patients with a pituitary incidentaloma is to have a complete evaluation which should include detailed personal and familial past medical history, physical examination, laboratory investigations for both hormone hypersecretion and hypopituitarism. The visual field examination should not be missed. The frequency of the investigations should be adjusted accordingly to the changes in size of adenomas[40].

Study/ reference	Size of population	Size		Gender		Mean age at diagnosis	Incidence	Prevalence	Types % (Female/Male ratio)				
		% Micro	% Macro	% Males	% Females				PRL	NFPA	GH	ACTH	TSH
Daly *et al*.[33]	71972	57.4	42.6	32.4	67.6	40.3 (range 12–86yr)	-	1/1064	66.2 (2.88)	14.7 (0.42)	13.2 (0.5)	5.9 (4)	0
Fontana *et al*.[34]	54607	N/A	N/A	27	73	-	-	1/1241	56	30	9	5	0
Fernandez *et al*.[35]	81498	58.7	41.3	33.3	66.7	37 (range 16–79yr)	-	1/1289	57.1 (8)	28 (0.5)	11.1 (0.75)	1.58 (0.1)	0
Raapana *et al*.[36]	722,000 to 733,000	54	46	19.04	80.99	40 (range 27.3–55yr)	4.00/100,000/ year	1/1470	51 (4.3)	37 (1.2)	8.5 (0.6)	3 (2.8)	1.2 (2)
Grupetta *et al.*[37]	394,640 to 417,608	56.6	43.4	30.4	69.6	40.6 (SD ±15.0yr)	4.27/100,000/ year	1/1321	46.2 (1.4)	34.2 (4.4)	16.5 (1.36)	2.2 (6)	0.94 (2)
Agustsson *et al.*[39]	210.912 to 321.857	41.2	54.8	40.3	(59.7	44 (range 4–94yr)	0.6/100,000/ year 1955–1972 5.8/100,000/ year 2003–2012	1/865	39.9 (3.0)	43.1 (0.95)	11.3 (0.65)	5.7 (2.85)	0

Table 1: Population based studies. A review of the literature regarding the prevalence of different types of pituitary adenomas (updated from Aflorei *et al* 2014[41])

Etiopathogenesis of pituitary adenomas

Tumorigenesis requires two overlapping steps: initiation and promotion[42]. Most of the data suggest that human pituitary adenomas are the result of an intrinsic pituitary defect due to inherited or acquired genetic or epigenetic changes that probably confer growth advantage[43], which represents the initiation step. In the promotion step, additional genetic changes, growth factors and environmental factors are involved, for example circulating hormones such as oestrogens or the microenvironment, with additional permissive effect on the behaviour of the altered cells. These defects lead to monoclonal expansion of a single modified cell, although heterogeneity in recurrences after treatment has been described[44].

Hormonal factors involved in pituitary tumour pathogenesis

Despite several animal models suggesting that hormonal effects, such as hypothalamic trophic or inhibitory hormones, or abnormal feedback regulation from peripheral hormones could lead to pituitary adenomas[45,46], there is little evidence for this in humans. Ectopic GHRH or CRH secretion result in pituitary hyperplasia, but an association with adenoma development has not been shown[47,48]. The lack of cortisol feedback in untreated Addison disease or congenital adrenal hyperplasia has not been consistently associated with corticotrophinomas, although individual case reports have been described[49]. The preponderance of microprolactinomas in females maybe linked to estradiol influence on the pituitary gland; however, there is no data suggesting a link between the use of the contraceptive pill and pituitary adenomas[50]. Factors involved in the physiological regulation of the pituitary gland, such as somatostatin analogues and dopamine agonists, are utilised therapeutically as analogues of inhibitory hypothalamic factors and have been used successfully in the treatment of acromegaly and prolactinomas.

Environmental factors involved in pituitary tumour pathogenesis

Most of the available evidence regarding the role of environmental factors as potentially causal factors in pituitary adenomas tumorigenesis comes from animal experiments. In mice and cats, environmental causal factors for pituitary adenomas have been found to include chemical and radiation exposure.

Some of the earliest experiments were performed by Gorbman[51], and demonstrated that mice exposed to radioactive iodine (I131) developed tumours of the pituitary gland. More recent experiments[52] confirmed that when mice are exposed to radiation they will develop pituitary adenomas. Regarding the role of environmental carcinogens on the development of pituitary adenomas, it has been shown that Fischer 344 rats receiving acrylamide in drinking water for 2 years developed pituitary tumours[53]. Acrylamide is known as an industry-related carcinogenic

substance, but recent discoveries have found that acrylamide is also formed in significant concentrations during high-temperature cooking of many common foods[54]. Recently published data reported that naturally occurring hypersomatotropic cats have increased plasma concentrations of organohalogenated contaminants as polyclorobyphenils and other substances with oestrogen-like activity compared to diabetic cats and controls. The mechanism of tumour formation seems to involve a link between AIP-induced xenobiotic metabolism, toxicity and cell cycle regulation[55].

However, very little evidence is available regarding the role of environmental factors in the aetiology of pituitary adenomas in humans. In 2010, Canavo *et al*.[56] investigated the epidemiological characteristics of patients with acromegaly in Sicily. The study focused on the relationship between the prevalence of acromegaly and environmental pollution. Area of Sicily were divided into 4 zones with different pollution levels. In the most polluted zone, there were high concentrations of toxic substances such as benzene, toluene, cyclohexene, and ethane. There was increased prevalence of the disease in the highly polluted area compared to less polluted areas. The risk ratio (RR), calculated for the most polluted area, assuming the least polluted area as a reference, showed that the population of the most polluted region had an 8.3-fold increased risk of developing acromegaly ($P<0.0014$).

Given the fact that AIP is a molecular chaperone to the receptor of the environmental toxin dioxin, the prevalence of pituitary adenomas was investigated in Seveso, Northern Italy, following an industrial accident where considerable dioxin pollution affected well circumscribed areas[57]. Although no significant increase was identified, this is an issue that warrants further study.

Genetic factors involved in pituitary tumour pathogenesis

Data on genetic predisposition and on acquired somatic changes (mutations or epigenetic alterations) associated with pituitary adenomas is steadily accumulating[58]. The differentiation of the pituitary gland is under the coordination of a series of very specific and temporally events[59], which are highly regulated by a various transcription factors. The succession of sequential activation of these factors has been shown to be also involved in every step of the tumorigenesis process. Figure 3 depicts an overview of pituitary development and intracellular tumour cell signalling associated with cell proliferation and tumour development.

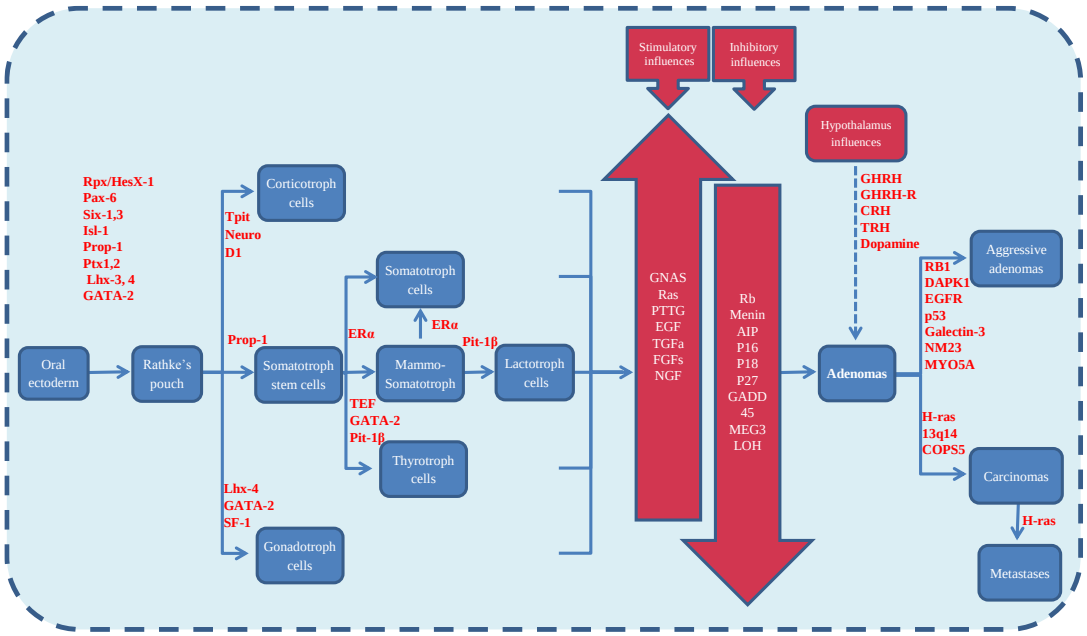

Figure 3: Cascade of factors involved in anterior pituitary development and tumorigenesis. Sequential activation of series of factors that are involved in tumorigenesis process, from progressive differentiation of mature pituitary gland cell types to pituitary adenomas and carcinomas. Oncogene activation, tumour suppression gene (TSG) inactivation and different factors probably contribute to transformation from normal pituitary to adenoma. The role of hypothalamic factors in humans has not been convincingly shown in the development of human pituitary adenomas (dashed line). Additional mutations may facilitate aggressive behaviour or malignant transformation (adapted from Aflorei et al 2014[41]).

Numerous genes have been suggested to be involved as oncogenes or tumour suppressor genes (TSG) in the pathogenesis of sporadic pituitary adenomas. Table 2 presents a brief summary of some of the most frequent genetic disruptions observed in pituitary adenomas).

Gene	Mechanism of normal function	Result of altered function	Oncogene /TSG
Somatotroph adenoma			
CCND1	Involved in progression through the G1-S phase of the cell cycle	Increased expression, can stimulate both cell proliferation and apoptosis in GH_3 cells	Oncogene
CREB	Phosphorylation-dependent transcriptional activator of cAMP response elements (CREs)	Constitutive activation by phosphorylation	Oncogene
GHR	Transmembrane receptor that mediates GH action	Loss-of-function somatic mutation	-
GHRH	Stimulates GH secretion	Increased expression	-
GHRH-R	Transmembrane receptor that mediates GHRH action	Truncated alternatively spliced nonfunctioning receptor	-
GNAS1	Alpha subunit of the stimulatory G protein that activates adenylate cyclase	Predominant maternal origin of GNAS1 transcripts; 40% of GH-secreting pituitary adenomas have somatic mutations	Oncogene

GPR101[60]	G protein-coupled receptor 101 gene	increased by a factor as high as 1000 in the pituitary tumours	Oncogene
SSTR2	Specific high-affinity G-coupled receptor for somatostatin	Decreased expression	-
Lactotroph adenoma			
BMP4	Involved in the control of the differentiation and proliferation of the different cell types in the anterior pituitary	Overexpressed in prolactinomas	TSG
DRD2	G protein-coupled receptor for dopamine	Decreased expression	-
FGFR4	Membrane-anchored receptor for fibroblast growth factor	Increased expression of a N-terminally truncated cytoplasmic isoform (ptd-FGFR4) by alternative transcription initiation	Oncogene
TGF-α[61]	Competes with EGF for binding to the EGF receptor and stimulates its phosphorylation in order to produce a mitogenic response	Overexpressed under the prolactin promoter influence	Oncogene
Corticotroph adenoma			
CCNE1	Promotes progression through the G1-S phase of the cell cycle	Increased expression	Oncogene
HDAC2	Enzyme that deacetylates lysine residues on the N-terminal region of the core histones	Decreased expression	Oncogene
NR3C1	Nuclear receptor for glucocorticoids	Loss-of-function somatic mutation	-
SmarcA4	Member of the SWI/SNF protein family with helicase and ATPase activities. Regulates gene transcription by altering chromatin structure	Decreased expression, altered subcellular localization	TSG
USP8[62,63]	Involved in regulation of the endosome morphology via protein ubiquitination	Gain-of-function somatic mutations in corticotropinomas	Oncogene
Nonfunctioning adenoma			
DKC1	Pseudouridine synthase that modifies rRNA and regulates telomerase activity	Loss-of-function somatic mutation	TSG
MEG3	Induces apoptosis and inhibits proliferation of tumour cells	Decreased expression	TSG
PITX2	Member of the bicoid-like homeobox transcription factor family, which is involved in the Wnt/Dvl/β-catenin pathway	Increased expression	-

PLAGL1	Zinc finger transcription factor that plays a role in pituitary development, differentiation, maturation and tumorigenesis	Decreased expression	TSG
PRKCA	Kinase that participates in growth factor- and hormone-mediated transmembrane signalling and cell proliferation	Increased expression, gain-of-function mutation	Oncogene
Most or all pituitary tumour types			
AKT1	Regulates many processes including metabolism, proliferation, cell survival, growth and angiogenesis	Increased expression, especially in NFPAs	Oncogene
AKT2	Regulates many processes including metabolism, proliferation, cell survival, growth and angiogenesis	Increased expression, especially in NFPAs	Oncogene
BAG1	Inhibits the chaperone activity of HSP70/HSC70 and the pro-apoptotic function of PPP1R15A	Increased expression	-
CCNA1, B1, B2	Involved in the control of the G2/M phases of the cell cycle	Increased expression	Oncogene
CDKN1A – p21	Regulator of cell cycle progression at G1	Decreased expression in NFPAs, Increased expression in hormone producing adenomas	TSG
CDKN2A	Induces cell cycle arrest in G1 and G2 phases	Decreased expression	TSG
PIT1[64]	Member of the POU transcription factor family; plays a key role in the specification, expansion and survival of somatotrophs, lactotrophs and thyrotrophs during development	Overexpressed in GH, PRL and TSH pituitary adenomas	Oncogene
PTTG	Cell cycle regulation and cell senescence	Increased expression, especially in corticotrophinomas	Oncogene
Invasive adenoma			
DAPK1	Positive mediator of programmed cell death induced by gamma-interferon	Decreased expression either by promoter methylation or by homozygous deletion of the promoter CpG island	TSG
EGFR	Transmembrane glycoprotein required for normal cellular proliferation, adhesion, migration and differentiation	Increased expression	Oncogene
Galectin-3[65]	Extracellular Gal-3 mediates cell migration, cell adhesion, and cell-to-cell interactions; intracellular Gal-3 inhibits apoptosis	Up-regulated during neoplastic progression	Oncogene
MYO5A[66]	Actin-dependent molecular motor, with roles in tumour cell migration, invasion, and metastasis	Increased expression	-

NM23[67]	N-terminal kinase domain could phosphorylate and downregulate cyclin B and could prevent the progression of cell from G2 to M phase of the cell cycle	Allelic loss results in reduced NM23 expression	TSG
RB1	Key regulator of entry into cell division	Decreased expression partly by promoter methylation	TSG
Pituitary carcinoma			
COPS5	Probable protease subunit of the COP9 signalosome complex, which is involved in various cellular and developmental processes	Increased expression	-
HRAS	GDP/GTP binding protein that regulates cell division in response to growth factor stimulation	Gain-of-function somatic mutations	Oncogene

Table 2: Selected genes that may be involved in molecular pathogenesis of pituitary adenomas. This list is not exhaustive and exemplifies some of the complex genetic disruptions observed in pituitary adenomas. Updated from Gadelha *et al.*, for references see original review article[68].

The key mechanisms that seem to be involved in the pituitary tumorigenic process are oncogene activation and TSG inactivation. These can occur either independently or in combination.

Gain-of-function mutations occur most often in genes affecting signal transduction pathways and lead to prolonged activation of the downstream pathway. These are most commonly dominant mutations so a single mutated allele is sufficient to induce a phenotype. The most frequently observed genetic change in pituitary adenomas is the somatic heterozygous activating mutation of the *GNAS* gene coding for the G protein α-subunit. This is called the *gsp* mutation, which can be present in up to 40% of GH-secreting pituitary adenomas[69,70]. The mutation abolishes the GTP-ase activity of the α subunit, which leads to a constitutively activated adenylate cyclase, increased cyclic adenosine monophosphate (cAMP) levels and protein kinase A (PKA) activation. In turn, the cAMP response element-binding protein (CREB) is phosphorylated and leads to sustained GH hypersecretion and cell proliferation. *H-ras* mutations have been described in a few cases of invasive prolactinomas or distant metastatic pituitary carcinomas[71-73]. As most of these mutations were found in rare carcinoma samples, it was suggested that they may be important in malignant transformation and metastasis rather than pituitary adenoma initiation.

Loss of TSGs on both alleles may initiate tumour cell growth (Knudson's two-hit theory[74]). In inherited conditions the first genetic alteration event can be an inherited germ-line mutation of one allele, followed by a second somatic alteration event affecting the TSG, leading to the gene

being completely turned off. The second event can either be a large deletion (which can usually be detected by testing for loss of heterozygosity (LOH) at the mutation locus or microsatellites around it), a somatic point mutation, a reduction in gene expression due to promoter methylation, or the presence of inhibitory microRNAs[75]. The retinoblastoma gene (*Rb*) was one of the first described TSGs in pituitary adenoma. Although *Rb* knockout mice develop intermediate lobe ACTH-secreting pituitary adenomas, and germline mutations cause retinoblastoma in humans, no *Rb* mutations were identified in human pituitary adenomas[43]. Methylation of the *Rb* promoter[76] has been described in human pituitary tumorigenesis. Areas of the genome typically lost in pituitary adenomas include the p and q arms of chromosome 11, locus 11q13, 13q12-14, 10q, and 1p, 9p, 13, 3, and 12 and X[77]. Many of these include TSGs, but no particular chromosomal loci in pituitary adenomas are free of allelic deletions.

Other tumour-initiating and promoting factors have been studied using animal models and human tissue samples. Molecules that have been characterized to be involved in the proliferative potential of pituitary cells and tumour growth include cell cycle regulators, disrupted growth factors, transcriptional regulators or pituitary miRNAs.

Cell cycle disruption: Proteins regulating the cell cycle can also be considered as oncogenes and TSGs in a tumorigenic process. Several transgenic mouse models have demonstrated that both inactivated[78,79] and overexpressed cell cycle regulators are sufficient to initiate pituitary tumorigenesis. Cyclins and stimulated cyclin-dependent kinases (CDKs) promote initiation of the cell cycle by phosphorylation and therefore inactivation of Rb. This confers to the cyclins the capacity to lead to tumorigenesis and several have been shown to be overexpressed in different types of pituitary adenomas[80].

Growth factors: Various growth factors and their cognate receptors are essential for regulating pituitary cell growth and for regulating normal hormone production[81]. Transforming growth factor (TGFα)[61] is a mitogenic protein which in pituitary tissue is expressed mainly in lactotrophs, where might be overexpressed under the influence of the PRL promoter. TGFα does not induce tumours of other pituitary cell subtypes, indicating a specific role in tumorigenesis of prolactinomas. Fibroblast growth factors (FGFs) are a very complex family of ligands that are involved in pituitary development and growth. In an article published in 2001, Ezzat *et al*.[82] demonstrated that the N-terminally truncated isoform of FGF receptor-4 (ptd-FGFR4), an altered growth factor receptor isoform, can be implicated in the neoplastic process of pituitary adenomas. Disruption of other growth factors involved in pituitary tumorigenesis include epidermal growth factor[83], nerve growth factor[84] and bone morphogenetic protein 4 (BMP4)[85].

BMP4 is required for early embryonic and pituitary development and has been shown to be highly expressed in pituitary adenomas in mouse models.

Transcriptional regulators: Anterior pituitary cell differentiation is a very complex process during which a series of transcription factors play key roles in each stage. PITX1 and PITX2 are some of the first transcription factors expressed in the Rathke's pouch and subsequently will have a key role in the normal development of all pituitary cell types[20,19]. PITX1 has been demonstrated in all normal anterior pituitary cell types and in the majority of pituitary adenoma subtypes[86]. One recent study reported reduced PITX1 mRNA expression in corticotroph tumours compared with other pituitary tumours. PITX2 (pituitary homeobox 2) is another pan-pituitary transcriptional activator involved in very early stages of pituitary cell differentiation. Pellegrini-Bouiller *et al*[87] described absent PITX2 mRNA expression in corticotroph adenomas, with high PITX2 expression in gonadotroph tumours. Interestingly, although Ptx2 expression was observed in PRL secreting tumours, no expression was observed in somatotroph adenomas. This might suggest that PITX2 has an involvement in the terminal differentiation of these types of pituitary cells. A third factor implicated in early pituitary development is the Prop-1, which is required for *Pit1* gene expression. Inactivating Prop-1 mutations occur in subjects with combined pituitary hormone deficiency. Regarding this very important transcription factor, RT–PCR analyses have demonstrated appropriate Prop1 expression in normal pituitary tissue and in all pituitary adenomas examined *et al*[88]. On the other hand Pit1 mRNA was increased up to fivefold in somatotropinomas and prolactinomas compared to normal pituitary tissue[64].

Therefore, the available data regarding the role of transcription factors in pituitary tumour pathogenesis are limited and quite old, and there are necessary more investigations to understand their involvement.

The overwhelming majority (~95%) of pituitary adenomas arise sporadically[89,90]. However, genetic factors may also play a role. It is well described that certain animal strains are more prone to pituitary adenoma development, primarily prolactinomas[91,92]. Racial differences have also been identified in humans, with black people having increased prevalence of pituitary adenomas. Racial differences in pituitary adenoma incidence were first reported in 1976[93] and showed the incidence in black women to be tripled when compared to white women, while incidence rates for black men were 4 times as high as for white men. A more recent article[94] published in 2011 also found that, in the USA, blacks had significantly higher incidence rates of pituitary tumours than Asian/Pacific islanders, American Indians/Alaska natives, and whites. There were no significant differences between the latter 3 groups. One of the few possible explanations for this high incidence is the fact that the racial differences in incident rates are

naturally occurring. Another possible reason is that black people may have a different clinical presentation of pituitary adenomas that draws attention to this condition. A third possibility is that the higher incidence rates are an incidental finding. However, the real cause for these higher incidence rates is still unclear. Despite the fact that in 1991 Goldstone *et al*[95]. raised the question that there might be a genetic basis for the differences in cancer incidence between Afro-Americans and Euro-Americans, no further experiments, involving ether whole genome sequencing or next generation sequencing, were performed. This might be due to the fact that the costing involved in order to reveal the genes/ transcripts significantly changed in Afro-Americans versus Euro-Americans or other populations can be significant.

Pituitary adenomas due to germline mutations

The detection of germline mutations in a patient has major implications for family members because the relatives are at risk of developing the disease. Familial genetic screening is an important step in these pathologies as the mutations often can be identified.

The main genetic causes of pituitary adenomas are summarised in Figure 4:

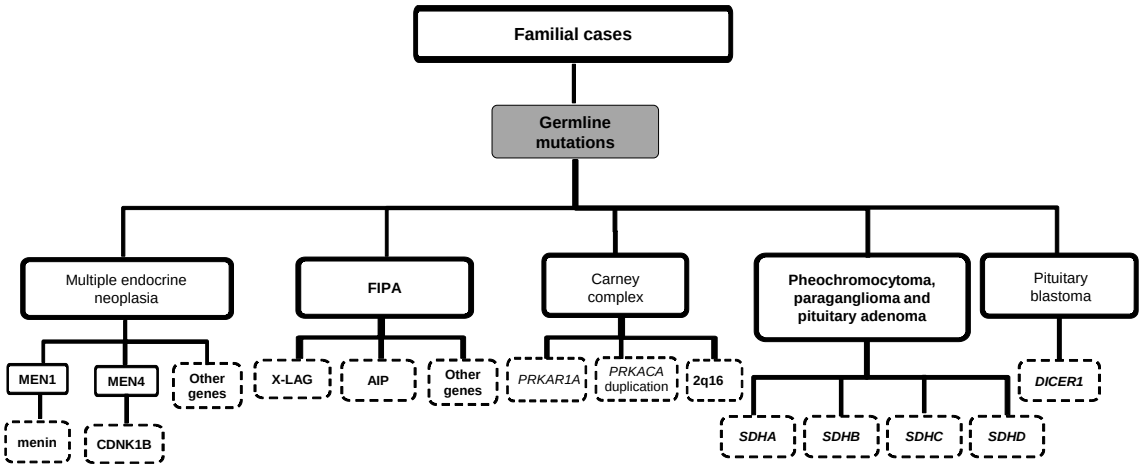

Figure 4: The pathogenesis of pituitary tumours due to germline mutations. A germline mutation in the *MEN1* gene, or rarely in the *CDKN1B* (coding for cell cycle regulator protein p27) gene causes MEN1 or MEN4; in a small proportion of affected cases, no gene abnormality can be found. A mutation in the protein kinase A regulatory subunit gene (*PRKAR1A*) is found in the majority (60%) of patients with the Carney complex; in the remainder, data suggest a causative gene in the 2q16 area. Patients have been described with an SDH mutation-related familial paraganglioma/phaeochromocytoma and familial pituitary adenomas[96-100]. Rare case reports of patients with neurofibromatosis type 1 and pituitary adenoma have been described. A mutation in the *DICER1* gene, a gene which regulates microRNAs, may cause an ACTH-secreting pituitary blastoma of childhood onset [101]. A fifth of families with familial isolated pituitary adenoma cases show a mutation in the *AIP* gene; in the majority of FIPA families, the causative gene has not yet been identified. Adapted from Korbonits et al [102].

Classical familial pituitary tumour syndromes have been described in around 5% of pituitary patients. The classical familial syndromes that predispose to pituitary adenomas are multiple endocrine neoplasia type 1 (MEN1) and type 4 (MEN4), Carney complex, and familial isolated pituitary adenomas (FIPA). More recent data raised the possibility that pituitary tumours can be caused by mutations in the *DICER1*[101] and succinate dehydrogenase *(SDH)* genes[98-100,97] among others. Isolated case reports have described pituitary adenomas in patients with neurofibromatosis type 1[103-108]. A very new syndrome of pituitary gigantism is called X-linked acrogigantism (X-LAG) and was published for the first time in 2014 by Trivellin *et al*.[60]. This pathogenesis is caused by microduplications on chromosome Xq26.3, affecting a gene named *GPR101*. GPR101 expression was found highly upregulated (up to 1000 times) in pituitary tumours of patients with Xq26.3 microduplications, compared with normal pituitary and tumours from patients who tested negative for microduplications[60].

MEN1 is an autosomal dominant syndrome that is caused by an inactivating mutation in the *MEN1* gene. The main manifestation involves pancreatic, pituitary, and parathyroid gland tumours, and there is typically a high penetrance, with over 95% of the patients manifesting the disease by the age of 50. The incidence of MEN1 is estimated to be around 0.25% from post-mortem studies and the prevalence estimated to be 1 in 20,000 to 40,000 individuals[109,110]. MEN1 affects both sexes equally but the pituitary manifestation, most commonly prolactinomas, has a female preponderance and sometimes paediatric onset. Pituitary adenomas can be the first presentation of MEN1 syndrome in 14% of cases.

MEN4 is a novel and rare familial syndrome seen in patients with MEN1-like features, but no *MEN1* gene mutations. Germline mutation in the CDKN1B has been described in a few cases, and single patients have been described with mutations in the genes coding for the CDKIs p15, p18, and $p21^{cip1}$ [111,112]. $p27^{Kip1}$ is known to interact with the menin protein which may explain the similar phenotype to MEN1.

Carney complex (CNC) is an autosomal dominant disorder with numerous manifestations including myxomas (benign tumours of the skin, breast, and other sites), testis and adrenal tumours as well as somatotroph hyperplasia or adenomas. The majority of the cases are caused by an inactivating mutation in the regulatory subunit of protein kinase A *PRKAR1A*[113].

X-LAG syndrome is a very recently described paediatric disorder characterised by a microduplication on chromosome Xq26.3, and increased growth hormone secretion. At this locus 4 genes were identified. Of these 4 genes, *GPR101* was found upregulated in pituitary lesions from patients with X-LAG. A variant was identified at position 924 (substitution of a

guanine with a cytosine (c.924G→C; p.E308D); however, the role of this change is currently under investigation. *GPR101* encodes a G-protein–coupled receptor, and pituitary-specific *GPR101* overexpression may be due to a gene-dose effect. Genetic analyses revealed that the phenotype associated with the identified *GPR101* mutation is associated with an extremely early-childhood onset form of gigantism. In contrast, the patients with gigantism of other cause develop the disease during adolescence[60].

FIPA: Case reports about families with pituitary adenomas but no other syndromic features have been described previously, but the first comprehensive report establishing this new diagnosis was only published in 2006[114]. In about 20% of these families, a mutation in the Aryl Hydrocarbon Receptor-Interacting Protein (AIP) gene was described[115], while in others the disease causing gene or genes have not been identified. More details on the section 1.3.

1.3 Familial isolated pituitary adenomas

Despite numerous studies that aim to understand the aetiology and pathogenesis of pituitary adenomas, many questions remain unanswered. One new puzzle has been the finding that heterozygous germline mutations of the *AIP* gene may lead to FIPA[116].

FIPA (OMIM #102200) is defined as the occurrence of a pituitary adenoma in two or more members of the same family in the absence of other syndromic clinical features, such as those characteristic of MEN1 and MEN4, CNC. This pathology is inherited in an autosomal dominant manner. However, the penetrance of the disease in FIPA families is highly variable within a range of 15-33%. Up to 48.6% of the families with *AIP* mutations had three or more pituitary adenoma patients per family. The maximum number of affected individuals within the same family was eight (six of them prospectively diagnosed) in a family carrying the p.R304*[117]. Owing to this incomplete penetrance, some of the carriers fail to express the phenotype. As the trait may not be expressed in one or more generations, it poses a challenge in detecting the inheritance in subsequent generations of the same kindred[118].

AIP mutations have been described as the cause of pituitary adenomas in about 15-20% of FIPA families, with a higher prevalence, up to 36-50%, in cases of families with only somatotropinomas[89,115,116,119,120]. Despite the fact that *AIP* germline mutations were not identified in patients with sporadic pituitary adenomas in some of the initial studies[121], recent publications have shown that about 4% of the cases of apparently sporadic pituitary adenomas have *AIP* mutations[115,122,123].

To date, more than 70 different *AIP* variants have been reported[120]. Various changes have been described throughout the entire region of the gene: missense, nonsense, deletions and insertions (in-frame or causing frame-shift), splice-site and promoter mutations as well as large genomic deletions. All these are germline mutations, and no *AIP* somatic mutations have been identified. The observed loss of heterozygosity (LOH) of the wild-type *AIP* allele in the tumour tissue of affected individuals supports the role of *AIP* as a TSG[124-126].

FIPA history

Until 2006 the literature on families with two or more members diagnosed with different types of pituitary adenomas, without any other manifestations of MEN1 or CNC was scarce. In modern literature, the first prolactinoma family was described in 1967[127] and was followed by a description of two acromegaly families[128,129]. A few years later, this was followed by the description of a corticotroph adenoma family[130]. Another familial prolactinoma case was again reported by Berezin *et al*. in 1995[131]. In the same year, a family with 3 patients affected by gigantism or acromegaly was described. It was the first time that association with the MEN1 gene locus had been ruled out[132]. The first article that reported loss of heterozygosity (LOH) on chromosome 11q13 in pituitary adenomas of 2 siblings with familial acromegaly was published in 1997[133]. The MEN1 gene is localized at this locus. Two hypotheses were suggested: the first one stated that familial acromegaly was an alternative form of the MEN1 syndrome; the second hypothesis was that an independent gene located in the immediate vicinity is the one involved the pathogenesis of familial acromegaly[133]. The second hypothesis was convincingly confirmed 8 years later by Soares *et al*. in an article published in 2005 on 8 novel families[126].

With every article that was published it became more obvious to the scientific world that pituitary adenomas could occur in a familial setting. In order to characterize these families with isolated pituitary adenomas as a distinct clinical entity different than MEN1 and CNC, in 2005 Daly *et al*. proposed a new syndrome: familial isolated pituitary adenomas (FIPA) (OMIM 102200)[134]. Since then, international collaborative efforts resulted in a collection of hundreds of families that exhibited different patterns of pituitary adenomas. Families with up to five affected cases were reported; these kindreds were characterized either as “homogeneous” (same adenoma phenotype among affected cases) or “heterogeneous” (different adenoma phenotypes among affected members)[135,136]. When these cases were investigated in more detail, a first-degree relationship between affected members was observed in the majority of the FIPA kindreds and an autosomal dominant inheritance pattern with incomplete penetrance was suggested for FIPA, based on pedigree analysis[90].

However, the exact genetic cause remained elusive until 2006 when Vierimaa *et al*[116] performed a study on two families with pituitary adenomas from Northern Finland. The Finnish group used high-stringency criteria and performed a genome-wide single-nucleotide polymorphism study on 16 patients in order to identify the gene locus. A linkage was identified at chromosome 11q12–11q13, but the *MEN1* gene locus was included. As none of the patients included in this cohort carried MEN1 mutations, a further screen was performed in both families included in the study. Further investigations on the *AIP* status in FIPA families showed that this pathology is an autosomal dominant disease with variable penetrance. The fact that *AIP* might be a TSG was suggested by the discovery that the majority of known affected individuals have LOH at the *AIP* locus, losing this way the wild-type allele, in concordance with the Knudson's two-hit hypothesis[137]. In one particular case described in 2007 by Iwata *et al*. a missense mutation (V49M) was identified in a patient with gigantism. However, the wt allele of the AIP gene was retained in the GH-secreting adenoma, but the LOH cannot be completely ruled out possibly due to contamination with normal pituitary tissue[138].

One important problem that remains to be solved is that mutations in the *AIP* gene were identified only in about 20% of the families. It remains unknown which genes are involved in the pituitary tumorigenesis of the remaining FIPA families.

AIP gene

AIP gene has 6 exons, 8,075 bp, and is localized on chromosome 11: 67,250,505-67,258,579 on the forward strand (UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly). All 6 exons are part of the coding region (total size for the coding region: 7,835).

Cloning

Kuzhandaivelu *et al*.[139] were the first to clone the AHR-interacting protein (AIP) gene in 1996. At the time, they designated it as XAP2 due to the fact that they used a yeast 2-hybrid system to identify proteins that interact with the hepatitis B virus (HBV) X protein. One year later, Carver and Bradfield[140] showed that AIP (named ARA9 in this paper) interacts with AHR in a ligand-dependent manner, also by using a yeast 2-hybrid assay.

Mapping

The mapping of *AIP* gene was a long process which started in 1993 when Thakker *et al*.[141] published the cases of 4 non-MEN1 patients with acromegaly and LOH for chromosome 11q13. In 1999, Gadelha *et al*. narrowed the region of LOH to 11q13, as they found LOH in all pituitary adenoma tissues from affected members of 2 unrelated families with acromegaly. As none of the patients had germline mutations in the *MEN1* gene, they concluded that the pituitary lesions

might be due to another mutated gene located in the 11q13 region[124]. One year later, the same group[125] managed to further narrow the LOH region to 8.6 cM: 11q13.1-11q13.3. The 8.6 cM region was reduced in 2005 to an only 2.21 Mb region as Soares *et al*.[126] used haplotyping and allelotyping techniques to evaluate 8 familial isolated somatotropinoma cases which were compared to 15 sporadic somatotropinomas. Only three genes from this region were considered potential candidates and were further submitted to sequencing, but no mutations were actually identified. Only in 2006 Vierimaa *et al*. found linkage to chromosome 11q12-q13 by genotyping of a large Finnish family with pituitary adenoma predisposition[116].

AIP protein

AIP belongs to the group of proteins harbouring conserved C-terminal tetratricopeptide-repeat (TPR) domains of 34 amino acids residues forming two palindrome alpha helices. The AIP protein has 3 TPR domains and a final 7th helix[142]. The N-terminal of AIP has a peptidyl-prolyl cis-trans isomerases (PPIase)-like domain and due to this structure AIP protein has a high degree of homology with proteins that belong to immunophilin family, such as FKBP52 (52 kDa FK506-binding protein).

The human *AIP* gene has four alternative mRNA transcripts (Figure 5) reported by the Ensembl database. Three of the four transcripts encode for a protein as they have the PPIase-FKBP domain and some of the TPR domains. Two of the isoforms are predicted to encode for shorter proteins that are also reported in GenBank, but their biological relevance is still unknown[143]. However, only the longest isoform has all three TPR domains, encoding for a 330-amino acid protein with a molecular mass of approximately 37 kDa (UniProt: O00170)[144].

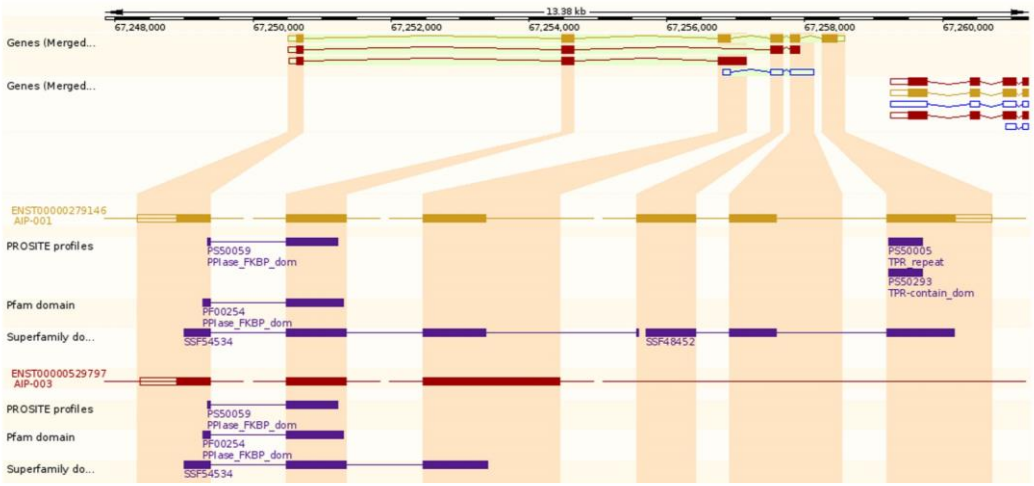

Figure 5: A schematic illustration of the alternative splicing of human *AIP* into four isoforms. (http://www.ensembl.org/Homo_sapiens/Gene/Splice?db=core;g=ENSG00000110711;r=11:67250512-67258574)

The crystal structure of human AIP protein was solved by two groups. In 2012, Linnert *et al*. published the N-terminal part of the protein[145] while the C-terminal part was crystalized in 2013 by Morgan *et al*[142] (Figure 6). The main motifs of the protein are, according to UniProt (accessed on 22.07.2015): PPIase domain: amino acids 31-121; TPR1 domain: amino acids 179-212; TPR2 domain: amino acids 231-264; TPR3 domain: amino acids 265-298 and the C-terminal α-7 helix: amino acids 300-330[144].

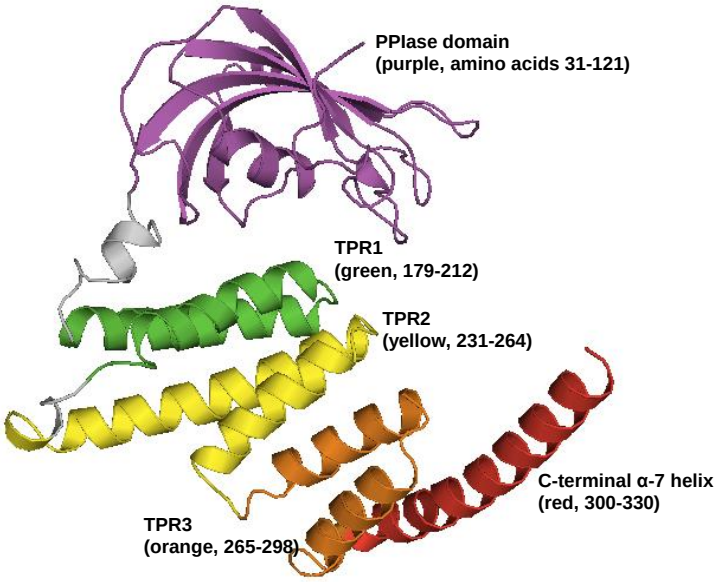

Figure 6: The AIP crystal structure. The most highly conserved residues are in the TPR domains, three antiparallel double helices and in the final alpha helix.

In 1998, Das *et al*.[146] published the crystal structure of the TPR domains of protein phosphatase, PP5. They showed that pairs of antiparallel α-helices, consisting of 34 amino acids, are packed together in an arrangement so that the protein folds into a right-handed super-helical structure necessary for the recognition of target proteins. Based on this three dimensional protein structure, Russell *et al*.[147] predicted that some of the amino acid residues in TPR regions are conserved (Figure 7), and are likely to mediate protein-protein interaction. Their focus was mainly on the residues that may be important for interactions with Hsp90. Only a few years later, different groups introduced point mutations in AIP (Bell & Poland 2000[148], Meyer *et al*. 2000[149], Laenger *et al*. 2009[150]), confirming the importance of the specific conserved amino acids for AIP–AhR or AIP–Hsp90 interactions. These changes were also identified in patients with pituitary adenomas. The first patient with a mutation in one of these crucial amino acids was published by Daly *et al*. in 2007[151]. The detected change was in lysine (K) at position 11: p.K241E. Later, Leontiou *et al*. 2008[89] identified a change in the cysteine (C) at position 8 (p.C238Y)), while the isoleucine (I) at position 27 (p.I257V) was the last one reported and it was detected in a patient with a TSH-secreting pituitary adenoma (Montanana *et al*. 2009[152]).

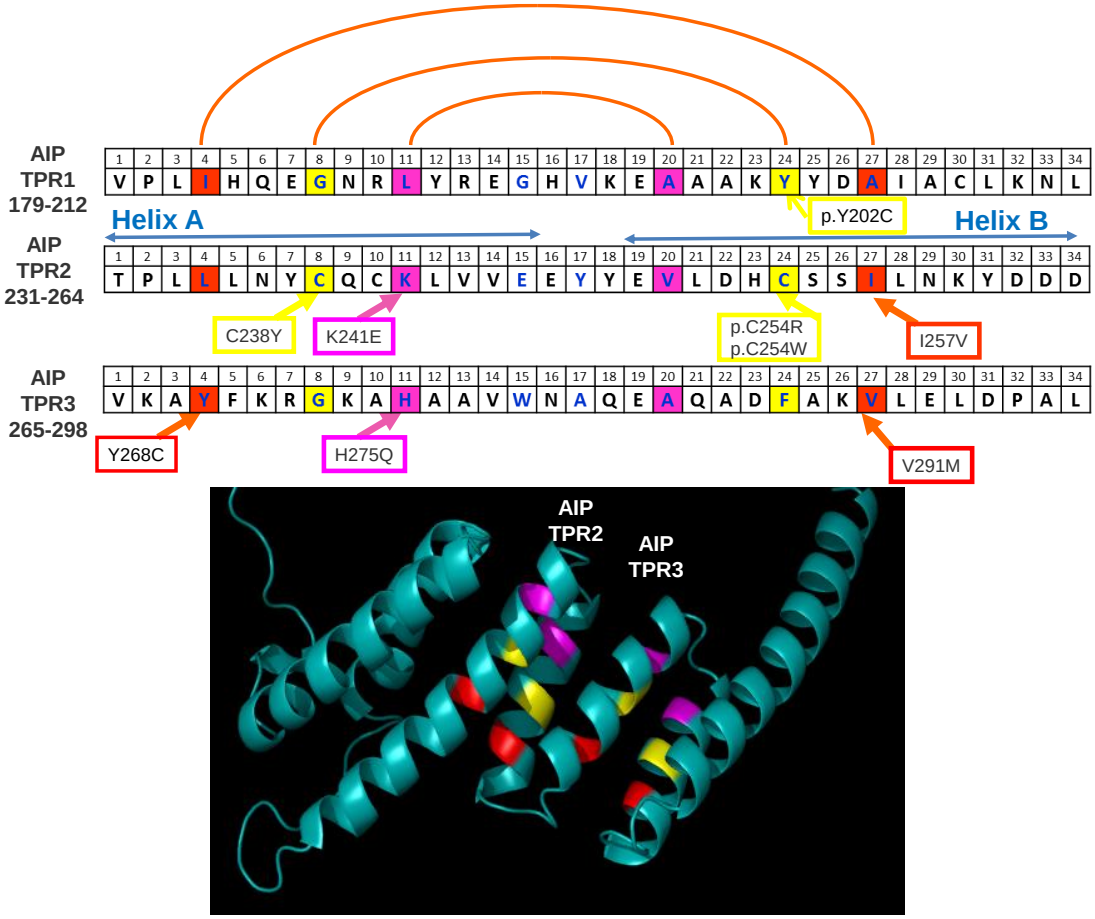

Figure 7: Representation of the TPR motifs of AIP. Top panel: A TPR motifs are composed of a pair of antiparallel helices, A and B. Consensus amino acids are located at positions 4, 7, 8 and 11 in helix A and at positions 20, 24, 27 and 32 in helix B). Residues 8 and 20 are located at the position of closest contact between the A and the B helices of a TPR, whereas residue 27 on helix B is located at the interface of three helices (A, B and the A helix of the next TPR motif) within a three-helix bundle. Patients with familial isolated pituitary adenoma have been identified with mutations affecting these crucial amino acids, such as the cysteine (C) at position 8 (p.C238Y), the lysine (K) at position 11 (p.K241E) and the isoleucine (I) at position 27 (p.I257V). Adapted from Trivellin *et al*. 2011[153].

AIP protein is ubiquitously expressed in both developmental and adult stages. AIP expression is substantial and has been confirmed in human and murine multiple tissues including heart, brain, skeletal, liver, muscle, kidney, testis, ovary and pituitary etc. At the cellular level AIP is predominantly cytoplasmic[139,154] but nuclear expression was also reported[149]. High expression of AIP is apparent in two types of adult pituitary cells: growth hormone (somatotrophs) and prolactin (lactotrophs)[89].

AIP interacts with a very wide repertoire of molecular partners[155] (Table 3). The TPR domains mediate AIP's interactions with most of its known partners, such as AHR, Hsp-70/90, survivin, phosphodiesterase-2A&4A5, TOMM20 and $G_{\alpha 12/13}$[153].

Partner	Full name	AC-MS	co-IP	Y2H	RC	Other	Confirmed interaction	Function	Organism and/or cell type(s)
ACTB[156]	Actin, cytoplasmic 1	✓					N	Cytoskeletal component	COS-7 cells
AGO1[157]	Argonaute RISC catalytic component 1	✓	✓				N	RNA interference	HEK 293, HeLa, and A549 cells
AhR[154,158-160]	Aryl hydrocarbon receptor		✓	✓	✓		Y	Adaptive and toxic responses, development	HeLa, Hepa1c1c7, COS-1, B-cells
ARNT[140,161]	Aryl hydrocarbon receptor nuclear translocator		✓	✓	✓		Y	Transcription factor activity, sequence-specific DNA binding	HeLa cells, HepG2, HeLa, COS7 cells
BIRC5[162]	Baculoviral IAP repeat containing 5 (Survivin)	✓	✓			✓	Y	Suppression of apoptosis	HeLa, MCF-7, Raji cells
CDC37[163]	Hsp90 co-chaperone cell division cycle 37	✓				AC-L	N	cell division, regulation of cell cycle	Stable polyclonal 293T cell lines
CDK9[164,165]	Cyclin-dependent kinase 9	✓					Y	regulation of transcription	HEK293T cells
EGFR[166]	Epidermal growth factor receptor					PCA	N	Cellular proliferation, survival, adhesion, migration, differentiation	Human fetal brain
EIF2S3[167]	Eukaryotic translation initiation factor 2 subunit 3					Co-F	N	GTPase activity, GTP binding	3 populations of cells
Gα13	Guanine nucleotide binding protein (G protein), alpha 13			✓	✓		N	Mediates receptor-stimulated signalling pathways	HEK293T, epa1c1c7, COS-7 cells
GNAQ[168]	Guanine nucleotide binding protein (G protein), q polypeptide				✓		N	modulators or transducers in transmembrane signalling systems	HEK293T, COS-7 and Hepa1c1c7 cells
HSP90AA1[161,163,169,170]	Heat shock protein 90kDa alpha (cytosolic), class A member 1	✓	✓		✓	AC-L	Y	protein folding; response to stress	293T, HeLa, COS7 and SK-N-MC, HEK-293 cells; Mouse embryonic fibroblasts,
HSP90AB1[163]	Heat shock protein 90kDa alpha (cytosolic), class B member 1		✓			AC-L	N	DNA binding, double-stranded RNA binding,	293T cell lines
HSPA8[171]	Heat shock cognate 71 kDa protein		✓				N	Repressor of transcriptional activation	COS-7 cells HeLa cells
HSPA9[167]	Heat shock 70kDa protein 9 (mortalin)					Co-F	N	ATP binding; poly(A) RNA binding	3 populations of cells

HSC70	Heat shock cognate 70	✓					N	Protein folding, Mitochondrial protein import,	HeLa cells
IRF7[157]	Interferon regulatory factor 7	✓			✓		N	transcription factor activity, sequence-specific DNA binding	HEK 293, HeLa, and A549 cells
NADSYN1[163]	Glutamine-dependent NAD(+) synthetase	✓				AC-L	N	glutaminase activity; NAD+ synthase activity	293T cell lines
NR3C2[169]	Nuclear receptor subfamily 3, group C, member 2		✓				N	Mineralocorticoid receptor	Mouse embryonic fibroblasts, SK-N-MC, HEK-293
PDE2A3[172]	cGMP-dependent 3',5'-cyclic phosphodiesterase 2A		✓	✓		Co-loc	Y	cAMP and cGMP degradation	Human brain, COS-1, HeLa cells
PDE4A5[89,173]	cAMP-specific 3',5'-cyclic phosphodiesterase 4A		✓	✓			Y	cAMP degradation	Rat brain COS-7 cells
PPARα[174]	Peroxisome proliferator-activated receptor alpha		✓		✓		N	Regulation of energy homeostasis	Mouse liver
PTGES3[163,169]	Prostaglandin E synthase 3	✓	✓			AC-L	Y	prostaglandin-E synthase activity; telomerase activity	293T cell lines, Mouse embryonic fibroblasts, SK-N-MC, HEK-293
RET[162]	Rearranged during transfection tyrosine-kinase receptor		✓	✓		PCA	N	Development, maturation, survival	Human fetal brain, rat pituitary, neuroblastoma and HEK293 cells
TNNI3K[175]	TNNI3 interacting kinase			✓			N	Promotes cardiomyogenesis, enhances cardiac performance	Human heart
TOMM20[142,171]	Mitochondrial import receptor subunit TOM20 homolog				✓		N	Mitochondrial import receptor	Human fetal liver, COS-7, HeLa cells
USP19[163]	Ubiquitin specific peptidase 19		✓			AC-L	N	ubiquitin-specific protease activity	293T cell lines,
YES1[165]	YES proto-oncogene 1, Src family tyrosine kinase	✓					N	non-membrane spanning protein tyrosine kinase activity	HEK293T cells

Table 3: Interacting partners of aryl hydrocarbon receptor-interacting protein (AIP). The techniques used to identify the various interactions, the functions of the different AIP partners, cell types where the interactions have been examined are reported. Adapted from Trivellin et al.[153] and updated from UNIPROT database[144]

Previous data suggested that the ubiquitously expressed AIP protein acts as a co-chaperone to heat-shock protein-90 (Hsp90) and regulates nuclear-cytoplasmic shuttling of the nuclear receptor AhR, whose ligands include the environmental toxin dioxin[140].

AHR

Over the last few years there was an increased interest in understanding the function of AIP in the AHR signalling pathway. However, the vast majority of studies have focused on toxicology rather than mechanistic aspects of AHR function[149,176,177]. As a consequence, little is known with regards to the role of AHR in the regulation of tumorigenesis. AHR is chronically activated in tumour cells and facilitates all the major stages of tumorigenesis, from initiation and progression to metastasis[178]. *In vitro* studies have shown that AIP stabilizes AHR in the cytoplasm by forming an AIP/AHR/Hsp90/p23 complex. AHR is mainly a ligand-activated transcription factor, which binds to the environmental toxin 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). Upon TCDD binding, the AHR/AIP/Hsp90/p23 complex translocates into the nucleus, where AHR undergoes conformational changes that enable it to detach from the complex. AHR then binds to DNA at xenobiotic response element sequences, inducing the transcription of xenobiotic-metabolizing enzymes[179]. Although TCDD is a known carcinogen that induces toxicity in the liver, endocrine glands or skin, it is still unclear whether this is involved in pituitary tumorigenesis. Epidemiology studies on the relationship between TCDD exposure and pituitary tumorigenesis in humans have not been conclusive[56,57].

Although no significant increase was identified, this is an issue that warrants further study. A mouse model of AHR deficiency is available. However, *AHR* mutant mice do not develop pituitary adenomas[180,181] and consequently the role of AHR on pituitary tumorigenesis remains elusive.

Site-directed mutagenesis experiments were used for finding the putative phosphorylation sites involved in AIP-AhR interaction. These were determined by a two-dimensional phosphopeptide mapping analysis. The findings revealed that changes in any or all of serine residues 43, 53, 131, 132, and 329, did not interfere with AIP ability to bind AHR. However, serine residue S53 could be essential for the nuclear translocation of AIP[182]. Mass spectrometry studies revealed possible other putative phosphorylation sites[183].

cAMP pathway

While the vast majority of studies involving AIP partners have focused on the AhR pathway, other AIP partners have been identified that could mediate disease pathogenesis[153]. Recently, two studies have highlighted a potential role of G proteins (Gαs and Gαi) in this process.

According to these reports, *AIP* inactivation is associated with low cAMP-Gαi signalling[184,185] enabling the development of somatotropinomas in AIP positive patients[185]

It was previously shown that cAMP/PKA signalling is involved on the synthesis and secretion of GH and PRL[186,187]. Moreover, as cAMP production is regulated by G proteins, the cAMP pathway plays a crucial role in multiple endocrine tissues, and many hormone receptors are transmembrane G protein coupled receptors (GPCRs). G proteins have three different subunits, α, β and γ. Up to 40% of sporadic GH-producing tumours are caused by mutations in the *GNAS1*[188] gene, which encodes the Gsα subunit. Pathogenic changes in this gene are also responsible for somatotropinomas and prolactinomas in McCune-Albright syndrome[189]. Deregulated cAMP signalling may also arise due to mutations in the gene which encodes for the protein kinase A type 1A regulatory subunit (*PRKAR1A*). Mutations in *PRKAR1A* have been identified in patients with GH/ PRL-producing tumours in the context of Carney complex[190-192].

Therefore, a potential role of AIP in cAMP signalling remains an attractive hypothesis to explain the effect of *AIP* loss in pituitary tumours. However, it is still unclear whether G proteins directly interact with AIP or if there are other specific molecular interactions necessary for the described functional changes to occur. Nevertheless, the AIP-cAMP link is a promising candidate pathway in the development of AIP positive tumours.

Cytoskeletal proteins

In sharp contrast to AHR and cAMP signalling, some of the proposed AIP partners have not been confirmed by independent studies. These include the cytoskeletal proteins actin and dynein, the epidermal growth factor receptor (EGFR), and the cardiac-specific kinase TNNI3K[166,175].

The putative interactions of AIP with cytoskeletal proteins are still controversial. One study reported no or very weak binding of the AIP PPIase-like domain to cytoplasmic dynein[193]. A different group[156] reported that AIP-mediated cytoplasmic retention of AHR[149,194,195] requires anchoring of the complex to actin filaments and proposed the existence of a direct AIP-actin interaction. However, in subsequent studies by Petrulis *et al.*[177], this direct interaction was not validated. This discrepancy may be due to the fact that different cell types were used in the two studies. These results highlight the need for further studies regarding the role of the actin cytoskeleton in the regulation of AIP function.

Based on the function of the actin cytoskeleton, it remains an interesting candidate for both the pituitary tumorigenesis process and the aggressive behaviour of pituitary adenomas. Actin is ubiquitously and highly expressed and has key roles in cell motility. Cell migration is a central process essential for normal development[196], as well as a major player in pathologic processes,

especially tumorigenesis and metastasis[197,198]. The core molecular mechanism of cell migration involves a combination of actin polymerisation, integrin receptor ligand binding and actomyosin contraction. The actin cytoskeleton dictates how adhesions are assembled at the leading edge of a cell, by serving as a dynamic scaffolding network. This is a two-way process, as actin polymerisation and organisation is mediated by migration-related signals via integrin receptors[199]. It was shown that actin polymerisation is more intense at the leading edge of cells, where integrin receptors bind to their extracellular ligands and promote cell motility and migration[200].

Bearing in mind that AIP is a co-chaperone protein, with a wide range of interacting partners and with ubiquitous expression, it is likely that the effect of *AIP* mutation is due to poorly understood molecular interactions or disturbed signalling pathways, and not due to AIP expression levels or pattern.

The adenohypophysis is a tissue with epithelial structure and most of this gland is formed of sinuous strings of epithelial cells in close contact with abundant vascular tissue. Consequently, to better understand the involvement of AIP in tumorigenesis and invasion of pituitary adenomas, it is essential to investigate how loss of this protein affects its interacting partners, the structure of the epithelial tissue and the connection with surrounding tissues. The developing *Drosophila* wing is a very simple *in vivo* model to investigate the molecular mechanisms involved in cell shape changes associated with rearrangement of epithelial sheets during tumorigenesis.

Clinical features

There is considerable phenotypic difference between families that exhibit or not *AIP* mutations, with childhood- and young-onset somatotroph adenomas often leading to gigantism, predominating in the families with *AIP* mutations. Numerous studies have revealed that both *AIPmut* positive tumours and *AIPmut* negative sporadic somatotrophinomas with low AIP protein expression show an invasive phenotype[89,119,201,202]. It was also published that up to 70% of *AIP* mutations are associated with somatotroph or somatolactotroph adenomas[102,120,153].

A very recent study conducted by Hernandez-Ramirez *et al*. in 2015[117] revealed a number of novel FIPA aspects. The majority (71.7%) of FIPA *AIPmut* positive patients are young at the onset of disease as they had the first signs of pituitary adenoma during the second and third decades of life. Contrary, only 39.2% the *AIPmut* negative FIPA patients, developed the pituitary adenomas around the same stage of life. Most commonly, in both *AIPmut* positive and negative families, the patients were diagnosed with GH excess; however, a significantly higher number of

cases were diagnosed in the *AIPmut* positive. Regarding the size and extension of pituitary adenomas between these two types of families *AIPmut* positive FIPA patients were found to have significantly larger tumours, as revealed by evaluating the maximum diameter. Although the group published that they found a higher prevalence of macroadenomas in *AIPmut* positive FIPA, overall, not all the *AIPmut* patients had tumours with an aggressive behaviour. Interestingly, a phenotype-genotype correlation was also found as patients with truncating AIP mutations usually have a younger age at the onset of the disease. Consequently, a higher percentage of gigantism was found among patients with truncating *AIPmuts* compared to those with nontruncating *AIP* variants[117].

Previous studies have published that there an increase ratio of males to females in AIPmut positive familial cases; however, this finding was not confirmed by this study. Even more, it was found an equal number of male and female among unaffected carriers.

Animal models for pituitary tumorigenesis due to loss of AIP

The AIP protein sequence is evolutionarily conserved among species. The percentage of homology between species varies from 100% of protein identity in primates (*P. troglodytes*), to around 94% in rodents (*R. norvegicus:* 93%, *M. musculus:* 94%). The homology levels are much lower in more inferior organisms such as *D. rerio* (66%), or *D. melanogaster* (40%); however, in all the species the protein has the same PPIase-like domain and TPR repeats[153]. The fact that AIP is highly conserved among the species might be related to the fact that this is a disease-associated protein[203].

For many years, animals have been used as models for understanding the mechanisms behind a wide variety of neoplasms that occur in humans. The initial studies used rodents that spontaneously develop cancers, which are also prevalent in humans, such as lymphoma, bladder cancer, and melanoma. The increased performance in the field of genetic engineering technologies led to a boost in recent years in the use of genetically-modified rodents that manifest a wide variety of neoplastic conditions. These changes have served to radically change the landscape of the studies that are focused on the mechanisms of pituitary adenomas tumorigenesis.

Mouse models

A first *Aip* KO mouse model was generated by Lin *et al*. in 2007 and demonstrated an essential role for AIP during embryonic development as $Aip^{-/-}$ mice die at E10.5-14.5[204]. Most of the abnormalities were present in the cardiovascular system: double-outlet right ventricle,

ventricular septal defects and pericardial oedema. In the same time, the heterozygous *Aip*$^{+/-}$ mice were apparently normal, with no obvious phenotype.

A second attempt to generate the mouse model came from the same group and resulted in a hypomorphic model of *Aip* deficiency. The penetrance of the phenotype (a patent *ductus venosus*), was variable depending on the genotype; up to 83% of the *Ara9*$^{(fxneo/fxneo)}$ mice[181]. The same phenotype was described in a model of *Ahr* KO mouse[205], and the fact that the levels of ARA9 expression were perfectly correlated with the frequency of the phenocopy of the *Ahr*-null allele led the authors to the conclusion that ARA9 is an essential protein for AHR signalling during development.

As there was still necessary a mouse model for pituitary tumorigenesis, in 2010 Raitila *et al*. published a heterozygous partial deletion of *Aip*. In contrast to previous models, the heterozygous mice developed normally. However, as expected, this model developed pituitary adenoma, in particular somatotropinomas, more frequently of the pars distalis of the anterior pituitary. Only a small percentage of these mice developed prolactinomas and corticotrophinomas. LOH was also detected in the tumour tissue. The adenomas were first detected at six months, and reached full penetrance by 15 months. Ki-67 analysis of the tumour suggested a more aggressive disease, as the *Aip*-deficient tumours have a higher proliferation index[206]. Additionally, IGF-1 concentrations were significantly elevated, similar to the clinical presentation in the human pathology, thereby providing increased evidence for AIP involvement in pituitary tumorigenesis. Although this mouse model strongly resembles the human disease, it has two important limitations. In human pathology, almost 25% of *AIP* mutation-positive somatotropinomas developed during childhood and adolescent stages[207-209]; however, the mouse model developed the tumour at adulthood with full penetrance at 15 months. Secondly, low penetrance is a characteristic of AIP-associated which is also different than the described model.

Zebrafish models

The interest of using zebrafish (*Danio rerio*) in genetic studies comes from its ease of use, relying on the rapid development cycle, transparent embryos, fully sequenced genome and availability of mutant strains[210]. The zebrafish (Zf) model offers high anatomical and functional similarities to human neuroendocrine system. Zebrafish is an excellent model for the study of the human endocrine system[211], displaying high anatomical and functional similarities with humans: the hypothalamus regulates pituitary function, producing oxytocin and vasopressin that are released from the posterior pituitary, and also produces releasing factors (at least six), which in

turn control the synthesis and secretion of the anterior pituitary hormones. Our laboratory has pioneered the use of *Danio rerio* and we have developed research tools for studying the effect of loss of *AIP*[212]. We generated *Aip* knockdown Zf embryos with antisense morpholino oligonucleotides injected to randomly-selected embryos at one-cell stage, using injected 5-base mispaired oligonucleotide as control morpholino. Aip Morpholino KD Zf embryos demonstrate brain, pericardium, eye, and swim bladder anomalies along with general developmental delay, pointing to wide developmental role of the *AIP* gene. *Aip* morphant embryos exhibit stronger PRL immunostaining in the pituitary compared to controls suggesting a possible increase in proliferative activity (hyperplasia or tumour) at the pituitary level in the absence of *AIP* gene function[213]. This attractive model can be proposed as an intermediary stage between cell culture/*Drosophila* and mammalian experimentation ultimately refining analysis and reducing costs. Nevertheless, despite all the efforts in understanding the physiological function of AIP regarding cell proliferation, the actual mechanism and the identity of its interacting partners relevant to pituitary tumorigenesis are still unknown.

1.4 *Drosophila melanogaster*: a model system

Drosophila as a model for cancer studies

One main question that is always asked is: how relevant is *Drosophila melanogaster* for the study of human cancer mechanisms? The signalling pathways involved in tissue and organ development, cell proliferation, cell survival, and cell migration are highly conserved in *Drosophila*[214,215]. Different studies in flies have been extremely informative: *Drosophila* tumours successfully model many of the features of mammalian cancers including an unlimited proliferative potential. Also, in malignant neoplasia, *Drosophila* tumours mimic the disruption of tissue architecture and the proprieties to invade and to metastasize causing host lethality. The results of fruit fly studies were shown to be transferable to humans: more than half of the known human disease genes, have homologues in *Drosophila melanogaster*[216].

As most of the signalling pathways[214,217] and most of molecular mechanisms involved in the control of growth and the cell cycle are well-conserved[218-220], in this project I propose to utilise the experimental advantages of the fruit fly to discover the conserved functions of AIP.

The first studies using *Drosophila* as an animal model were conducted by William Castle at Harvard University in 1901. This research was considerably improved by Thomas Hunt Morgan who first isolated a naturally occurring *Drosophila* mutation – *white.* This mutation causes a change in the eye colour, from red to white. Morgan and three of his students (Sturtevant, Bridges and Muller) demonstrated that genes are carried on chromosomes[221,222]. Their work on

multiple mutations allowed them to discover the notions of "crossing over" and genetic linkage. Another major step was introduced by Muller, who used balancer chromosomes in order to be able to maintain homozygous lethal mutations.

But the moment that established the fruit fly as a valuable model system in research was the Nobel Prize awarded to Lewis, Nüsslein-Volhard and Wieschaus in 1995 due to their work in genes controlling development. More recent studies focused on tumorigenesis in *Drosophila* and understanding of some of the essential characteristics of human neoplasia.

Today, it is consider that around 50% of the proteins that are involved in human diseases, including cancer, have a conserved orthologue in fruit fly[223,224] . In some cases, it was already shown that the degree of conservation is so high that the human genes can rescue the loss of corresponding *Drosophila* orthologues[225]. As a result, it should not be a surprise anymore that very important findings derived from fly research are extremely relevant to human medicine.

There are *Drosophila* studies that led to the identification of a protein and its molecular function long before its implication in cancer of the corresponding human homologue was discovered. An important example is *Notch*, which was first identified in fruit fly, and later genetic and molecular fly studies have unveiled some of the *Notch* targets and partners. Many years later aberrant expression of human *NOTCH1* was identified as being involved in T cell acute lymphoblastic leukaemia[226]. Another important cancer gene that was first identified in fruit fly is *hedgehog* (*hh*) – a segment polarity gene. The Hedgehog signalling pathways were first understood with the help of *Drosophila* studies and a few years later human mutations affecting *hh* signalling were identified as a causing factor in basal cell carcinoma and medulloblastoma[227].

There are numerous studies on *D. melanogaster* that helped to understand the molecular basis of human cancer, I will mention a few below.

Drosophila as a model for BRCA2 mutations

Heterozygous mutations in *BRCA2* confer a high risk of breast cancers in humans, but the BRCA2 gene is the first example on how studies involving *Drosophila melanogaster* led to the discovery of the cancer-relevant proteins in humans. Studies involving genetic analysis of BRCA2 in fruit flies have shown that this gene plays a major role in homologous recombination by having a protective role against large deletions[228]. This gene is also an important example on how a protein that is involved in a human cancer involving a tissue that has no equivalent in flies (the mammary gland in this case) may be better understood due to basic research carried out in *Drosophila*.

Drosophila as a model for glioblastoma

Another example includes fruit fly models of glioblastoma, one of the most malignant human brain tumours. This type of cancer is due to mutations that activate the epidermal growth factor receptor (EGFR) and phosphatidylinositol-3 kinase (PI3K) signalling pathways. The authors showed that loss of *Rbf1*, one of the two retinoblastoma (RB) genes in flies, led to constitutive co-activation of EGFR-Ras and PI3K pathways in *Drosophila* glia and to the development of invasive glial cells that create transplantable tumour-like growths, mimicking human gliomas[229].

Drosophila as a model for Neurofibromatosis 1

Neurofibromatosis 1 (NF1) is typically a childhood onset multisystem genetic disorder characterized by café-au-lait spots, skeletal dysplasias, and by the benign and malignant brain tumours of the peripheral nervous system. Although most cases do not develop metastasis, the health burden is still significant as the majority of affected children/adolescents suffer from debilitating skeletal defects and learning disabilities. The genetic cause is mutation of the *NF1* gene[230]. An important step forward in understanding this disease pathology was the use of *Drosophila*. The first significant results were published in 2001 by Williams *et al*. and they successful provided evidence that NF1 regulates Ras signalling[231]. The end point of these studies was the development of Ras pathway inhibitors used in the treatment of neurofibromatosis patients[232].

Overall, the record of *Drosophila* use as a model for tumorigenesis is impressive. The tools developed in this field offer the advantage of detailed *in vivo* exploration of interactions between cells, tissues, and genes. There is a constant increase in the number of laboratories that choose to take advantage of this valuable animal model to explore cancer mechanisms and even therapeutics[233].

Advantages of *Drosophila melanogaster* as a model system

As an animal model, *Drosophila* has important advantages that make it extremely valuable in research. Males are easily distinguishable from females. These insects are very small and as such require a limited space in the lab, are cheap, clean, harmless (some people might present occasional allergies), and easy to maintain in stable stocks. They have very short generation time, about 10 days at 25°C, and this allows multiple and parallel independent experiments to be performed and repeated within a relatively short time period. Even more, as the length of the life cycle is temperature-dependent, this enables researchers to increase the fruit fly life span by raising them at 18°C, or to shorten it at 29°C, based on the experimental requirements

The life cycle includes different phases that are easily identified (Figure 8).

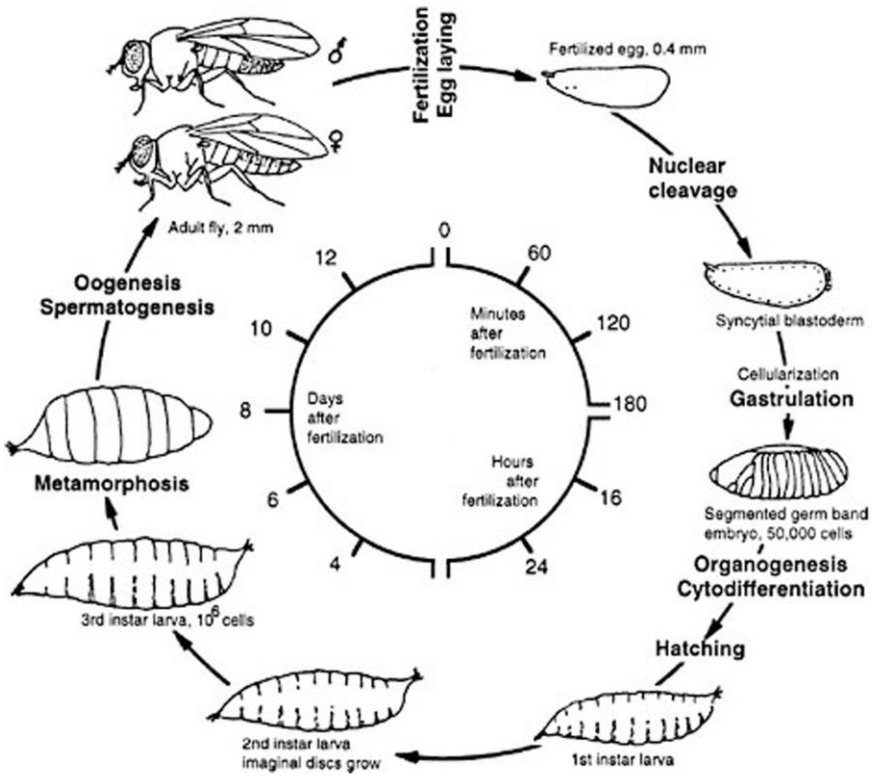

Figure 8: Life cycle of *Drosophila melanogaster*. The larva hatches 1 day after the egg is fertilised. First, second, and third instar are larval stages, each ending with a molt. During pupation most of the larval tissues are destroyed and replaced by adult tissues derived from the imaginal discs that were growing in the larva. Times are given for the life cycle at 25°C[234]. Source: http://www.nap.edu/openbook.php?record_id=9871&page=162

One of the most important steps in *Drosophila* research was the sequencing of the entire genome[223], a crucially important source of information that allowed the development of new bioinformatics studies. The entire genome sequencing revealed that more than 90% of the fruit fly genes are similar to the human and mouse genes. However, the *Drosophila* annotated genome, with its very small number of chromosomes, 180 million bases and ~13,600 estimated genes, offers a much simpler system than mammalian models[223].

Drosophila has only four pairs of chromosomes designated as 2, 3, 4 and X/Y (Figure 9). The Y chromosome has only a few genes that are important for spermatogenesis, but this chromosome is not essential for fly viability.

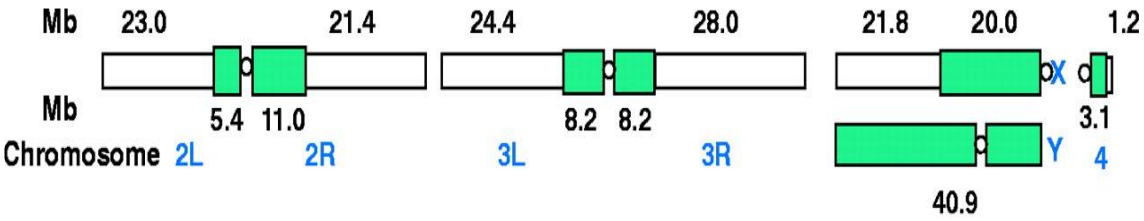

Figure 9: Mitotic chromosomes of *D. melanogaster.* The euchromatic regions, heterochromatic regions, and centromeres are displayed. Arms of the autosomes are designated 2L, 2R, 3L, 3R, and 4. The euchromatic length in megabases was derived from the sequence analysis. The Y chromosome is nearly entirely heterochromatic. Adapted from Adams et al. 2000[223]

Apart from the fast development, another advantage of *Drosophila* is the fact that for each chromosome there were created balancer chromosomes. The first records of balancer chromosome use in *Drosophila* studies date back to 1975[235]. These are special modified chromosomes, which were designed via multiple, nested chromosomal inversions. The balancer chromosomes are able to suppress crossing over between homologous chromosomes during meiosis. If, by chance, the crossing over involving a balancer chromosomes occurs, the resulting chromatids either lacking some genes or have duplicated genes. As a consequence, the progeny is not viable.

Typically, a balancer chromosome contains one or more dominant mutations that are visible phenotypic markers that enable researchers to follow the balancer through crosses. Furthermore, the main advantage of using balancer chromosomes is the ability to maintain lethal mutations in heterozygous stable stocks. Without balancer chromosomes lethal mutations would otherwise impair the survival of homozygotes[216]. The names of the balancer chromosomes are standardized based on the chromosome they serve to stabilize. The first letter of the chromosome's name represents the number of the chromosome it stabilizes. F stands for the first chromosome, S stands for second, and T stands for third. The second letter is an M, which stands for "multiply inverted". The M is followed by a number to distinguish balancers of the same chromosome: FM6, FM7s, and TM3. They also carry the name of the phenotypic marker the balancer carries. Additionally, the genetic marker or markers of the balancer are listed after the name and separated by a comma[236]. Below are a few examples of some markers used in this study, at least one for each chromosome (Figure 10).

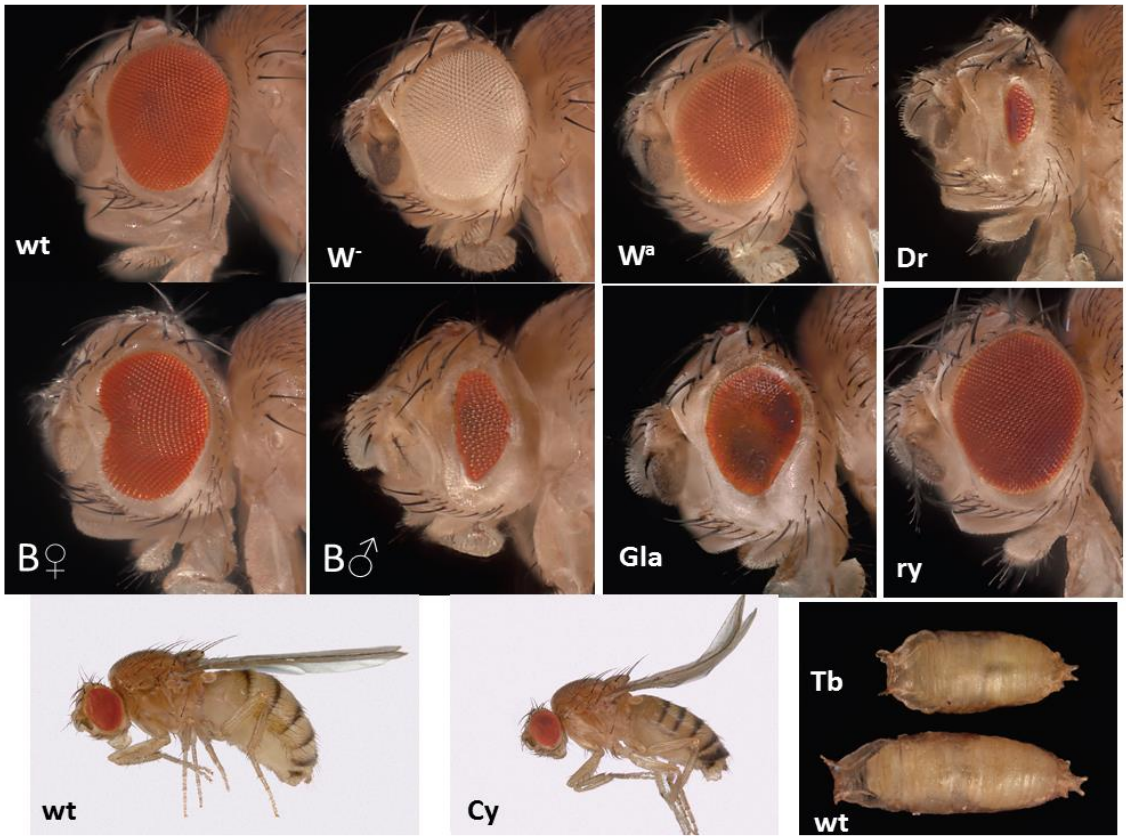

Figure 10: Common phenotypic markers used in this study. Due to phenotypic markers the genetic crosses are totally predictable and verifiable, as the markers aids on tracing the inheritance of alleles (are clearly very distinct from wild-type)[237].

In the fly research community there is a free exchange of reagents and information (via databases such as Flybase, Virtual fly, etc.). Also part of the advantages is the wide spectrum of methods that can be used. The adaptation of the Gal4/UAS system[238] from yeast (detailed in section 3.2.1), together with immunostaining and *in situ* hybridisation, mutants and rescue experiments, made it possible to achieve an increasingly more detailed analysis of tumorigenesis mechanisms.

In addition to the advantages mentioned above, *Drosophila* provides the ability to perform forward genetic screens to generate mutant stocks that allow investigation of fundamental questions regarding tumour development during embryogenesis and in the adults.

Disadvantages of *Drosophila* as model system

All animal models have limitations, and *Drosophila* is no exception. This animal model organism has as a main disadvantage the issue of care and intensive handling as fly stocks have to be maintained alive and cannot be frozen. Another negative aspect of is that its small size precludes detailed study of pathology and cause of death. Together with some lacking information

regarding detailed anatomy during development, this poses a problem for developmental studies.

There are also a few limitations of the fruit fly as an animal model for tumorigenesis. Flies do not have an adaptive immune system. Although flies possess a rudimentary heart with internal organs, they also do not have a closed circulatory system of endothelial cell-derived blood vessels. Flies also lack telomerase and use a different mechanism in order to maintain the length of the telomere. As a result of these differences, *Drosophila* is not a perfect cancer model to model mammalian carcinogenesis, but the fruit fly might be extremely informative about the essential factors that initiate tumour formation.

Drosophila's AIP orthologue

The *Drosophila melanogaster* gene *CG1847* (NM_132530.4)[239] is the fruit fly's single orthologue of human AIP (40% amino-acid identity). The gene is located on chromosome X at position 10F2, from base pair 11,869,170 to base pair 11,871,168 on the positive strand. It has 3 exons. As it is located on X chromosome, males will have only one copy of this gene.

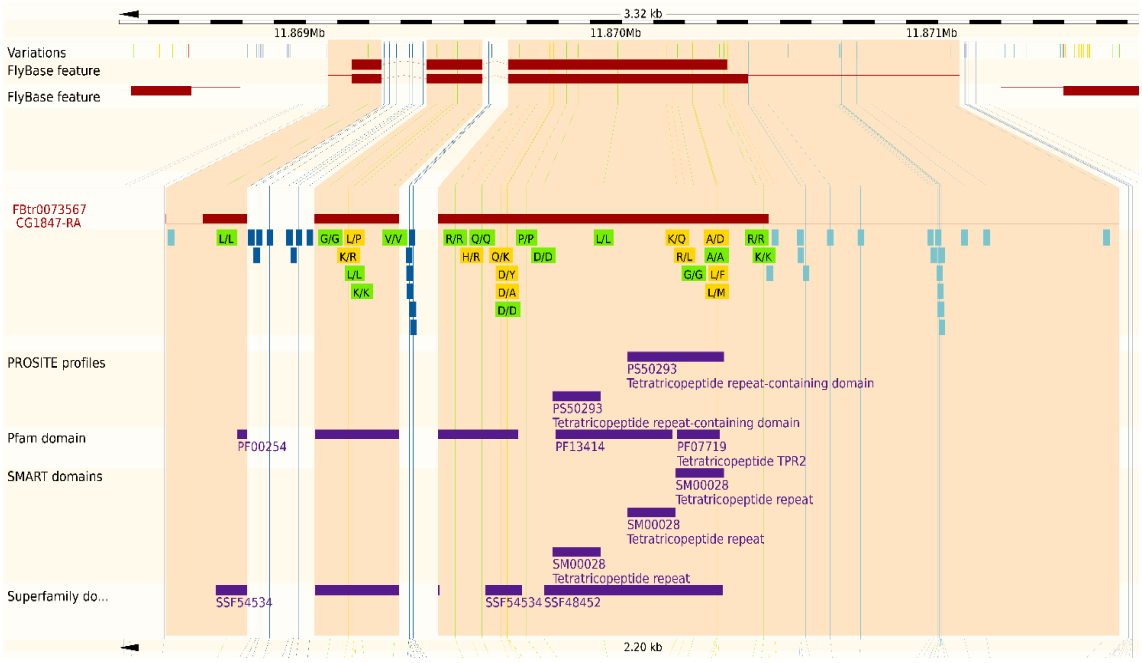

Figure 11: A schematic illustration of the alternative splicing of CG1847 into two isoforms. (http://www.ensembl.org/Drosophila_melanogaster/Gene/Splice?db=core;g=FBgn0030345;r=X:11762233-11766188;t=FBtr0073567)

Structurally, CG1847 is a 320 aa protein and shares a significant degree of identity and homology with its human orthologue as it has one peptidyl-prolyl *cistrans* isomerases (PPIase)-like domain and three TPR repeats (Figure 11). However, currently there is no published data focusing on the function of this *Drosophila* gene.

AIMS OF THE STUDY

Overview on the subject

The main aim of this study was to understand the role of *AIP* orthologue during *Drosophila* development and to reveal the signalling pathways and the AIP orthologue partners that enable cells to proliferate, migrate, and invade into surrounding tissues when AIP protein structure is changed.

The structure of this thesis

The aims of the work presented in this thesis are to understand the results of loss of CG1847 in an *in vivo* model. More specifically;

- First (Chapter 3):
    - To establish *Drosophila melanogaster* as a model to study AIP/CG1847 silencing.
    - To characterise CG1847 function during fly development by silencing CG1847 (by knockdown and knockout).
    - To reveal the potential underlying molecular mechanisms of loss of AIP by performing a whole transcriptome analysis using Illumina Next Generation Sequencing. RNA-seq was performed in 48 h old mutant (*CG1847*$^{exon1\_3}$) versus control male larvae isolated using fluorescent markers. This allowed us to determine gene expression profiles using an established analysis pipeline and to identify key pathways that are significantly altered in the mutant and are related to embryonic development or survival.
- Second (Chapter 4):
    - To reveal the molecular partners of *Drosophila* AIP involved in cell-cell adhesion and to gain further insights on how CG1847 silencing leads to cell detachment, with a possible impact on tumorigenesis.
- Third (Chapter 5):
    - To test the functional conservation between human and fly protein by rescuing the lethality of *CG1847*$^{exon1\_3}$ mutants via expressing human AIP cDNA under the control of a ubiquitous promoter during fly development.
    - To discriminate between pathogenic and non-pathogenic *AIP* mutations as this is a challenging problem in the management of patients carrying a missense *AIP* variant.

CHAPTER 2: MATERIALS AND METHODS

2.1 Materials

2.1.1 *Drosophila* stocks

The *Drosophila melanogaster* strains used in this study and their purpose and references are listed below, in Table 4.

Control stocks	
w^{iso}	Gift from Nic Tapon
Df(1) yw	Lindsley and Zimm (1992)[240]
yw,CG1847$^{2.39A}$/FM6	Precise excision generated in this project
yw,CG1847$^{2.39A}$/FM7c,Dfd-YFP	Precise excision generated in this project
Stocks containing balancer chromosomes	
yw ; Bl/CyO ;	Lindsley and Zimm (1992)[240]
yw ; ; Dr/TM3	Lindsley and Zimm (1992)
TrxR-1 D2/FM6	Gift from Fanis Missirlis[241]
w^{+} Baz/FM7c,Dfd-YFP	Bloomington *Drosophila* Stock Center: Stock ID: 23229
Stocks containing GAL4 driver transgenes	
yw; Act-Gal4/CyO	Luo et al. (1994)[242]
elav-Gal4	Luo et al. (1994)
Hs-hid;hh-Gal4, UAS-DIAB/TM6B	Gift from Nic Tapon
nub-Gal4/CyO	Gift from Nic Tapon
Cg-Gal4	Bloomington *Drosophila* Stock Center Stock ID 7011
Crq-Gal4	Gift from P. Ribeiro
Dilp-Gal4/CyO	Ikeya et all 2002[243]
Mef2-Gal4	Bloomington *Drosophila* Stock Center Stock ID 25756
repo-Gal4	Glial cells specific driver
drm-GAL4	Bloomington *Drosophila* Stock Center Stock ID 7098
HE-GAL4	Bloomington *Drosophila* Stock Center Stock ID 8699
c42-GAL4	Gift from P. Ribeiro[244]
tinC-GAL4	Gift from M. Frasch (2001)[245]
Stocks containing UAS-RNAi transgenes	
UAS-CG1847R-1	National Institute of Genetics - Fly Stock Center: Stock ID: 1847R-1
UAS-CG1847R-2	National Institute of Genetics - Fly Stock Center: Stock ID 1847R-2

UAS-CG1847-T1: w1118; P{GD9582}v43701	Vienna *Drosophila* RNAi Center: Stock ID: v43701
UAS-CG1847 T2: w1118; P{GD9582}v43702/TM3	Vienna *Drosophila* RNAi Center: Stock ID: v43702
UAS-Dcr-2; ; UAS-CG1847T1 RNAi	
UAS-Dcr-2; ; UAS-CG1847T2 RNAi	
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GL00168}attP2/TM3, Sb[1]	Bloomington *Drosophila* Stock Centre Stock ID: 35270
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00277}attP2	Bloomington *Drosophila* Stock Centre Stock ID: 33399
Mutant *Drosophila AIP* stocks	
w P{EP}CG1847^{G1839}*	Bloomington *Drosophila* Stock Center: Stock ID: 32600
yw,CG1847^{exon1_3}/FM6	
yw,CG1847^{exon1_3}/FM7c,Dfd-YFP	
Stocks used for generation of mitotic recombination clones (Flp-FRT lines)	
w,Dm,FRT19A^{NeoR}	Gift from Nick Brown
P{w[+mC]=Ubi-mRFP.nls}1, w[*], P{ry[+t7.2]= hsFLP}122 P{ry[+t7.2]=neoFRT}19A	Bloomington *Drosophila* Stock Center Stock ID: 31418
*w*ovo[D1]v^{24} P{FRT(whs)}101/C(1)DX, y1 f1/Y; P{hsFLP}38*	Bloomington *Drosophila* Stock Center Stock ID: 1813
Stocks containing GFP-tagged proteins	
mys$^{5.4}$(βPS-GFP)	Gift from Nick Brown[246]
Recombinant lines	
CG1847^{exon1_3} FRT19A^{NeoR} / FM7c, Dfd::YFP	
CG1847^{exon1_3}/FM6; Actin-Gal4/CyO	Ubiquitously driver
CG1847^{exon1_3}/FM6; Cg-Gal4/CyO	Fat body specific driver
CG1847^{exon1_3}/FM6; Crq-Gal4/CyO	Haemocytes specific driver
CG1847^{exon1_3}/FM6; Dilp-Gal4-Gal4/CyO	Insulin secreting cells - specific driver
CG1847^{exon1_3}/FM6; Mef2-Gal4/TM3	Muscle specific driver
CG1847^{exon1_3}/FM6; elav-Gal4/TM3	Neurons specific driver
CG1847^{exon1_3}/FM6; repo-Gal4/TM3	Glial cells specific driver
CG1847^{exon1_3}/FM6; drm-GAL4/TM3	Gut specific driver
CG1847^{exon1_3}/FM6; HE-GAL4/TM3	Haemolymph specific driver
CG1847^{exon1_3}/FM6; c42-GAL4/TM3	Malpighian tubules specific driver
CG1847^{exon1_3}/FM6; tinc-GAL4/TM3	Heart specific driver
Other lines	
y w ; ; Ki, pp, Δ2-3	

Table 4: *Drosophila melanogaster* lines used in this study. Lines in white boxes were obtained from the Vienna *Drosophila* RNAi Center and the National Institute of Genetics - Fly Stock Center (Japan, http://www.shigen.nig.ac.jp/fly/nigfly/), as indicated in the text. Those in light grey boxes were existing lab stocks or gifts from other labs. The lines in dark grey boxes were generated during this study.

Transgenic lines carrying either wt CG1847 under the expression of its own promoter or human *AIP* cDNA variants generated by BestGene and used in this study are listed in Table 5:

Transgenic lines	
CG1847/CyO 1M	BestGene Stock ID: 12135-1-1M-Ch2
CG1847/CyO 3M	BestGene Stock ID: 12135-1-3M-Ch2
UAS_hAIPwt/CyO 1M	BestGene Stock ID: 14089-1-1M-Ch2
UAS_hAIPwt/CyO 2M	BestGene Stock ID: 14089-1-2M-Ch2
UAS_hAIPtrunc/CyO 1M	BestGene Stock ID: 12855-1-1M-Ch2
UAS_hAIPtrunc/CyO 2M	BestGene Stock ID: 12855-1-2M-Ch2
UAS_hAIP-R16H/CyO 1M	BestGene Stock ID: 14335-1-1M-Ch2
UAS_hAIP-R16H/CyO 2M	BestGene Stock ID: 14335-1-2M-Ch2
UAS_hAIP-C238Y/CyO 1M	BestGene Stock ID: 14335-2-1M-Ch2
UAS_hAIP-C238Y/CyO 2M	BestGene Stock ID: 14335-2-2M-Ch2
UAS_hAIP-A299V/CyO 1M	BestGene Stock ID: 14335-3-1M-Ch2
UAS_hAIP-A299V/CyO 2M	BestGene Stock ID: 14335-3-2M-Ch2
UAS_hAIP-R304Q/CyO 1M	BestGene Stock ID: 14335-4-1M-Ch2
UAS_hAIP-R304Q/CyO 2M	BestGene Stock ID: 14335-4-2M-Ch2

Table 5: *Drosophila melanogaster* transgenic stocks generated and used in this study

2.1.2 *Drosophila* husbandry

Fly strains were kept and raised in vials containing standard food medium (recipe in Appendix 3) and dry yeast. Two copies of each stock were kept in plastic vials (82x25 mm, B.T.P DREWITT) containing fly food and stored at 18˚C 65% humidity under 12 h: 12 h light: dark cycles. The stocks were flipped into fresh food vials every 4 weeks. For the working stocks or crosses, the ambient conditions were set to a 12 h: 12 h light: dark cycle, constant 25˚C, and 65% humidity. If required, crosses were flipped every 8-10 days to prevent the mix of individual flies from different generations.

2.1.3 Collection of Adult Flies

Unless otherwise indicated, fly crosses were maintained at 25°C. Recently eclosed males and females of the desired strains were collected using CO_2 anaesthetisation. Male and female flies were separated during the first 4 hours after eclosion in order to select virgin females. Each food vial contained a maximum of 25-30 flies (depending on the purpose) and was kept at 18°C until the flies were crossed with the desired genotype.

2.2 Standard fly techniques

2.2.1 RNAi-mediated silencing of CG1847

CG1847 was silenced in *Drosophila melanogaster* using RNA-interference (a detailed description of the UAS-GAL4 system is described in section 3.2.1). Four *UAS-CG1847-RNAi* lines were tested (Table 4). *UAS-CG1847-RNAi R1* and *R2* lines were purchased from NIG-Fly Stocks (the RNAi constructs were inserted on chromosome 2 and 3, respectively). To confirm that the results were not due to an RNAi off-target effect, two additional RNAi lines, *UAS-CG1847-RNAi T1* and *T2* were purchased from the Vienna Stock Center. These lines carry a different RNAi construct inserted on the 3rd chromosome. (Table 4). For more details regarding how these stocks were generated, the sequence of the inverted repeats, and the cloning strategies (Appendix 4 and Supporting Table 1). Figure 12 (below) depicts a schematic representation of the inverted repeats (IR) alignment to *CG1847* showing non-overlapping of the IR constructs used by the two providers.

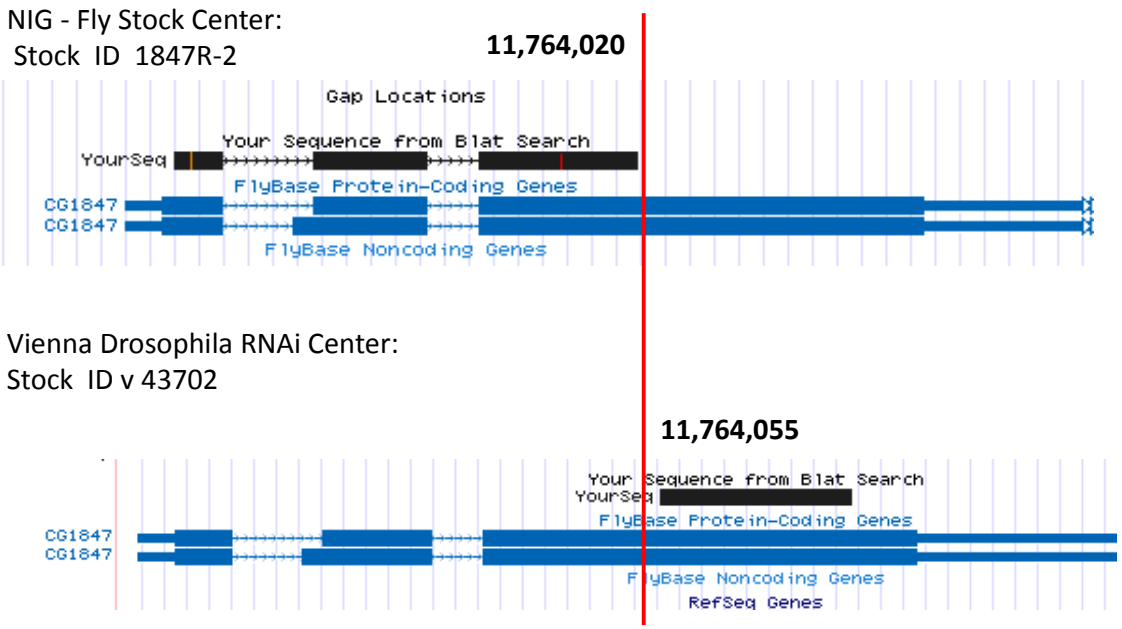

Figure 12: Shematic representation of UCSC Browser Blat alignment of IR mapping to *CG1847.* The 2 IR constructs (in black) are targeting different areas of the gene.

The Vienna stocks were combined with flies carrying the *UAS-Dicer-2* transgene (*Dcr-2)* on the first chromosome in order to enhance RNAi potency[247]. This system facilitates the cleaving of the double-stranded RNA and the formation of the RNA-induced silencing complex (RISC), leading to a greater silencing of *CG1847* (Table 4).

Using 3 different drivers, I ectopically expressed specific *CG1847* RNAi in all *Drosophila* cells, or specific subsets. The drivers used were *elav-Gal4* (inserted in the X chromosome), *Act-Gal4, nub-Gal4* and *hh-Gal4*, inserted in 2nd and 3rd chromosomes, respectively (Table 4). Six to eight females carrying the universal or tissue-specific Gal4 driver were crossed with 3-5 males of the UAS-*CG1847* RNAi line and maintained at 25°C.

As stronger phenotypes resulted with *UAS-CG1847-RNAi*-R1, *UAS-CG1847-RNAi*-R2 and *UAS-CG1847-RNAi*-T2 compared with *UAS-CG1847-RNAi*-T1, I will further present and discuss only these data.

For *UAS-CG1847-RNAi R1,* the expected genotypes were:

- *Act-Gal4 / UAS-CG1847RNAi-R1*
- *elav-Gal4 /+; UAS-CG1847RNAi-R1 / +*
- *UAS-CG1847 R1 / +; + / hh-Gal4*

For UAS-*CG1847-RNAi* R2, the expected genotypes were:

- *Act-Gal4 / CyO; UAS-CG1847 R2 / +*
- *elav-Gal4 / +;; UAS-CG1847 R2 / +*
- *UAS-CG1847 R2 / hh-Gal4.*
- *nub-Gal4 / CyO; UAS-CG1847 R2 / +*

For the third RNAi line, the combinations had the following genotypes:

- *Dcr-2 / +; Act-Gal4 / CyO ; UAS-CG1847 T2 / +*
- *Dcr-2 / elav-Gal4; UAS-CG1847 T2 / +*
- *Dcr-2 / +;; UAS-CG1847 T2 / hh-Gal4.*
- *Dcr-2 / +; nub-Gal4 / CyO ; UAS-CG1847 T2 / +*

The RNAi efficiency was further tested by measuring the specific mRNA levels through semi-quantitative RT-PCR (details in Chapter 3.2). Primer sequences are listed in Supporting Table 2.

2.2.2 Generation of mutant CG1847 flies - Imprecise Excision screen

To investigate the possible roles of the *CG1847* gene, located on the X chromosome, a CG1847 mutant line was created. The *CG1847* gene was mutated by P-element transposase-mediated deletion of genomic DNA. For this, a fly line was obtained, in which a P-element is inserted within the 5'UTR of *CG1847*: w*P{EP}CG1847^{G1839} (Bloomington *Drosophila* Stock Center). This is a transgenic insertion stock generated through mobilization of a P-element construct P{EP} which carries a w[+mC] mini-white visible marker and UAS binding sites for GAL4 transcriptional

regulation[248]. The P-element was mobilised by crossing females homozygous for the P element with transposase carrying males of the strain: *yw;+;Ki,pp,Δ2-3* (Mating scheme 1).

F0: ♀ *w P[w+] / w P[w+]* **X** ♂ *y w; ; Ki, pp, Δ2-3 / Y Ki, pp, Δ2-3*

Set up about 200 individual crosses with females that have a lethal mutation on the X chromosome (allows easier selection in the next generation)

F1: ♀ *w P[w+] / Y ; ; Ki, pp, Δ 2-3 / +* **X** ♂ *TrxR-1 D2 / FM6*

In the 3rd generation, screen for virgins that have w eyes, i.e. those who lost the P-element

F2: ♀ *w [E{P}] / FM6* **X** ♂ *FM6 / Y*

In the next generation, the stable stock is obtained. Putative excision alleles were identified based on the lethality of male flies

F3: ♀ *w [E{P}] / FM6* ♀ *FM6 / FM6* Infertile ♂ *FM6 / Y* ♂ *w [E{P}] / Y* LETHAL

Mating scheme 1: Crossing scheme for generating CG1847 mutant fly

Individual F1 males (which carry both the P element and the transposase) were crossed with 3 first chromosome balancer *w*, FM6 females. These females also carried a lethal mutation on the other X chromosome TrxR-1D2 (Thioredoxin reductase1)[241] which allowed selection of the desired alleles in the next generation as both TrxR-1D2/Y males and TrxR-1D2/ w [E{P}] (possible mutated) were lethal. In F2 the recently eclosed adults were selected for the presence of the dominant *Bar* marker and possibly one copy of the mutated *CG1847* identifiable by *Bar*$^{+}$ eye. Using the eye shape phenotype as a marker, the offspring with the desired genotype was selected. Single white-eyed female progeny from the F2 generation were back-crossed to *FM6* males to create a stable stock. Putative excision alleles were identified based on lethality of male flies. Females homozygous for the *CG1847* mutation are not viable, while the hemizygous males do not emerge as adult flies.

Mapping: Genomic DNA was isolated from heterozygous mutant females. The resulting stable stocks were screened by PCR to identify the shorter amplicons using the primers described in Appendix 5. The putative mutants were then sequenced to confirm the imprecise excision. In stock Δ25A a deletion of exons 1 and 2, and 3rd exon of *CG1847* was identified, generating the desired *CG1847* mutant (henceforth designated as *CG1847*$^{exon1\_3}$).

For collection of male larvae carrying the *CG1847* mutation, the FM6 balancer chromosome was replaced with the FM7c balancer, which also carries a fluorescent marker Dfd-YFP[249].

F0:	♀	*CG1847^{exon1_3}* / *FM6*	X	♂	*FM7c, Dfd-YFP* / *Y*
F1:	♀	*CG1847^{exon1_3}* / *FM7c, Dfd-YFP*	X	♂	*FM7c, Dfd-YFP* / *Y*

Mating scheme 2: Replacement of balancer chromosome in the *CG1847* mutant stock

2.2.3 Generation of control stock for mutant *CG1847* (revertant) - Precise Excision screen

The type of mutagenesis screen described above (section 2.2.2) has the advantage that it could generate both the desired mutant stock via imprecise excision of the P element, and, at the same time, a control stock (revertant) through precise excision[250]. The revertant can therefore be used as a control as it has the same genetic background as the mutant line. The P-element was mobilised in the same screening experiment as in section 2.2.2 (see below).

F0:	♀	*w P[w+]* / *w P[w+]*	X	♂	*y w; ; Ki, pp, Δ2-3* / *Y Ki, pp, Δ2-3*

Set up about 200 individual crosses

F1:	♀	*w P[w+]* / *Y* ; ; *Ki, pp, Δ 2-3* / *+*	X	♂	*w FM6* / *l(1)*

In the 3rd generation screen for virgins that have w eyes, i.e. those who lost the P-element

F2:	♀	*w [E{P}]* / *FM6*	X	♂	*FM6* / *Y*

In the next generation I obtained the stable stock. Putative excision alleles were identified based upon male's lethality

F3:	♀ *w [E{P}]* / *FM6*	♀ *FM6* / *FM6* Infertile	♂ *FM6* / *Y*	♂ *w [E{P}]* / *Y* **VIABLE**

Mating scheme 3: Crossing scheme for generating revertant stock

The stocks with viable males were screened by PCR, and in stock 2.39A, I confirmed by sequencing that the excision deleted almost the entire P element, with the exception of a 12bp region (footprint of the P element) generating the *CG1847* control line. For further reference, this stock will be named *CG1847$^{2.39A}$*.

As above, the FM6 balancer chromosome was also replaced with the FM7c,Dfd-YFP.

F0:	♀	*CG1847$^{2.39A}$* / *FM6*	X	♂	*FM7c, Dfd-YFP* / *Y*
F1:	♀	*CG1847$^{2.39A}$* / *FM7c, Dfd-YFP*	X	♂	*FM7c, Dfd-YFP* / *Y*

Mating scheme 4: Replacement of balancer chromosome in the control stock

2.2.4 Recombination crosses

Recombination of CG1847^{exon1_3} mutant allele with FRT19A^{NeoR} site

The gene *CG1847* is located at position 10F2 on the X chromosome. To generate mitotic clones mutant for *CG1847*, we recombined our *CG1847* mutant allele with a transgenic stock carrying FRT recombination sites in the X chromosome. The FRT site is inserted on the X chromosome at 19A2. For recombination, flies harbouring *CG1847^{exon1_2}* were crossed with those carrying the *FRT19A^{NeoR}* site [*w*, *Dm*, *FRT19A^{NeoR}* / *w*, *Dm*, *FRT19A^{NeoR}*], a gift from Nick Brown (Table 4).

F0: ♀ $\frac{CG1847^{exon1\_3}}{FM6}$ X ♂ $\frac{w,Dm,FRT19A^{NeoR}}{Y}$

F1: ♀ $\frac{CG1847^{exon1\_3}}{w,Dm,FRT19A^{NeoR}}$ X ♂ $\frac{FM6}{Y}$

Set up 30 individual crosses

F2: ♀ $\frac{CG1847^{exon1\_3}\,FRT19A^{NeoR}}{FM6}$ X ♂ $\frac{FM6}{Y}$

Mating scheme 5: Recombination of the *CG1847^{exon1_3}* mutant allele and the FRT19A site.

From the F1 generation females with both *CG1847^{exon1_3}* and the FRT19A^{NeoR} on different X chromosomes were selected among the progeny. These females were selected on neomycin food, and were mated with first chromosome balancer males (FM6). In the last generation, individual recombined females were crossed back to three FM6 males.

The identified recombinants were screened by PCR for the *CG1847^{exon1_3}* allele and for the presence of the neomycin the resistance gene, in order to confirm these combination event (for primers sequences see Appendix 5).

Recombination of CG1847^{exon1_3} with mys5.4

The mys gene is located at 7D5. For recombination, flies harbouring *CG1847^{exon1_3}* were crossed with those carrying mys$^{5.4}$ (a βPS-GFP). In the next generation, females with both *CG1847^{exon1_3}* and *mys$^{5.4}$* on different X chromosomes were selected from the progeny. These females were selected based on the lack of a FM6 chromosome (females without Bar^{+}) and were mated with first chromosome balancer males (FM6). In the last generation individual recombined females were crossed back to three FM6 males (Mating scheme 6).

The identified recombinants were screened for the presence of GFP at the site of muscle attachment.

F0: ♀ $\frac{CG1847^{exon1\_3}}{FM6}$ X ♂ $\frac{mys^{5.4}}{Y}$

F1: ♀ $\frac{CG1847^{exon1\_3}}{mys^{5.4}}$ X ♂ $\frac{FM6}{Y}$

Set up 30 individual crosses

F2: ♀ $\frac{CG1847^{exon1\_3}mys^{5.4}}{FM6}$ X ♂ $\frac{FM6}{Y}$

Mating scheme 6: Recombination of the *CG1847^{exon1_3}* mutant and *mys$^{5.4}$*

2.2.5 Generation of homozygous mutant clones

Mitotic recombination was induced using the FLP/FRT system[251]. Virgin females of the genotype: *CG1847^{exon1_3}*, *FRT19A* / FM6 were mated with *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}* males. Eggs were collected for 4 hours and left at 25°C for 48 hours, at which stage they were subjected to a heat shock of one and a half hours at 37°C in a water bath on 2 subsequent days. The larvae were allowed to develop at 25°C until the 3rd instar stage and the desired genotype *CG1847^{exon1_3}*, *FRT19A* / *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}* was selected based on fluorescence. The larval imaginal wing discs were prepared at late 3rd instar. *CG1847^{exon1_3}* homozygous mutant clones were identified based on the absence of RFP expression.

To assess the role of the *CG1847^{exon1_3}* mutant in the wing homozygous clones of adult flies, larvae from a cross between *CG1847^{exon1_3}, FRT19A* / FM6 females and *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}* males were subjected to heat shock as described above and then allowed to reach maturity. As control, I used the progeny from the cross between *w*, *Dm*, *FRT19A^{NeoR}* females and *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}* males. Larvae from control crosses were subjected to the same heat-shock treatment described above.

The generation of somatic clones was investigated in wing imaginal discs dissected from 3rd instar larvae and the size of the clones was compared between heterozygous mutant progeny with somatic wing clones (females *CG1847^{exon1_3}, FRT19A^{neoR}* / *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}*) and control progeny (females *w*, *Dm*, *FRT19A^{NeoR}* / *Ubi-mRFPnls*, *hsFLP*, *FRT19A^{neoR}*).

Adult wing phenotype was compared between adult progeny.

2.2.6 Rescue of *CG1847* function

To determine whether deletion of *CG1847* was responsible for male lethality, I tested whether it could be rescued by the introduction of a genomic rescue-construct containing wt *CG1847.* To this end, a wt *CG1847* genomic rescue construct was prepared by subcloning a genomic DNA fragment that contained a 2227 bp region located upstream of the start codon of the *CG1847* DNA, along with a 412 bp fragment located downstream of the start codon, into the P-element transformation vector pWhiteRabbit containing an *attB* site[252]. This construct was sent to Best Gene Inc. (USA) where it was microinjected into 200 embryos which harbour an *attP40* landing site (*attP40-y*^{1} *w*^{67c23}; *P{CaryP}attP40*), following a standard protocol to create transgenics[253]. The advantage of the PhiC31 integrase system is the integration of the plasmid DNA at precise landing sites as opposed to random integration. The PhiC31 integrase (also known as ΦC31 integrase) encodes a serine-type recombinase that mediates the sequence-specific recombination between two different attachment sites, *attB* and *attP*. However, these sites share a small 3 bp central region, where the crossover occurs[254].

Five lines were produced and balanced over *CyO*. Two of the five transgenic lines survived during transportation. Males from these two lines were individually crossed with heterozygous females carrying the mutant *CG1847*$^{exon1\_3}$ allele. In the F1 generation the ability to rescue developing mutant hemizygous males was evaluated.

2.2.7 Rescue of mutant lethality by expressing wt human AIP (hAIPwt) under the control of a ubiquitous promoter (*actin*)

In order to test *in vivo* the degree of functional conservation between fly and human AIP (hAIP), I generated an UAS::hAIPwt construct by inserting the h*AIP* coding sequence, downstream of the GAL4-dependent UAS promoter into the *pUASK10attB* vector derived from pUAST[238]. This construct was sent to BestGene Inc. where it was then microinjected into 200 embryos (section 2.2.6). Five lines were produced and balanced over CyO. All five transgenic lines survived during transportation, but only 2 lines were used in rescue experiments.

The Gal4-UAS system was used for rescue experiments at 25°C, as the UAS-hAIP^{wt} was overexpressed ubiquitously in the respective genetic mutant background.

To rescue *CG1847* mutant lethality, heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a ubiquitous driver (*Act-Gal4*, located on the 2nd chromosome) were required. These flies were generated as detailed below (Mating scheme 7).

F0: ♀ $\frac{CG1847^{exon1\_3}}{FM6}$ X ♂ $\frac{Bl}{CyO}$

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ $\frac{CG1847^{exon1\_3}}{FM6}$ X ♂ $\frac{FM6;\ +}{Y\quad CyO}$

F1b: ♀ $\frac{x\ ;\ +\ .}{FM6\quad Bl}$ X ♂ $\frac{x\ ;\ Act\text{-}Gal4}{Y\qquad CyO}$

F2: ♀ $\frac{CG1847^{exon1\_3};\ +}{FM6\qquad CyO}$ X ♂ $\frac{FM6;\ Act\text{-}Gal4}{Y\qquad Bl}$

F3: ♀ $\frac{CG1847^{exon1\_3};\ Act\text{-}Gal4}{FM6\qquad\quad CyO}$ X ♂ $\frac{FM6;\ Act\text{-}Gal4}{Y\qquad CyO}$

Mating scheme 7: C**ombination of the *CG1847*$^{exon1\_3}$ mutant and actin-Gal4 driver**

To ubiquitously express the hAIP cDNA, males from two transgenic lines were individually crossed with females *CG1847*$^{exon1\_3}$ */ FM6; Act-Gal4 / CyO*. In F1 generation, I evaluated their ability to rescue developing mutant males (*CG1847*$^{exon1\_3}$ / Y; UAS::hAIPwt / *Act-Gal4*).

2.2.8 Rescue of mutant lethality by expressing h*AIP* under the control of specific tissues promoters

The UAS-hAIP construct allowed us to express the hAIP not only ubiquitously during fly development, but also in specific fly somatic cells or tissues, under the control of tissue-specific GAL4 drivers. To this end, I combined our heterozygous *CG1847*$^{exon1\_3}$ mutant females with 10 different tissue-specific promoters: fat body, haemocytes, insulin-secreting cells, muscle, neurons, glial cells, gut, haemolymph, malpighian tubules, and heart (all the crosses below).

For fat body targeted rescue of lethality, UAS-hAIPwt was overexpressed under *Cg*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a fat body driver (*Cg*-GAL4) (Mating scheme 8).

F0: ♀ $\frac{CG1847^{exon1\_3}}{FM6}$ X ♂ $\frac{X\ ;\ Bl}{Y\ \ CyO}$

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ $\frac{FM6\ ;\ +\ .}{Y\quad CyO}$ X ♂ $\frac{CG1847^{exon1\_3}\ ;\ +\ .}{FM6\qquad Bl}$

F1b: ♀ $\frac{X\ \ ;\ +\ .}{FM6\ \ CyO}$ X ♂ $\frac{X\ ;\ Cg\text{-}GAL4}{Y\quad Cg\text{-}GAL4}$

F2: ♀ $\frac{CG1847^{exon1\_3}\ ;\ Bl\ .}{FM6\qquad\quad CyO}$ X ♂ $\frac{FM6\ ;\ Cg\text{-}GAL4}{Y\qquad CyO}$

F3: ♀ $\frac{CG1847^{exon1\_3}\ ;\ Cg\text{-}GAL4}{FM6\qquad\quad CyO}$ X ♂ $\frac{FM6\ ;\ Cg\text{-}GAL4}{Y\qquad CyO}$

Mating scheme 8: Combination of the *CG1847*$^{exon1\_3}$ mutant and fat body driver *Cg*-Gal4

For haemocyte-targeted rescue of lethality, UAS-hAIPwt was overexpressed under *Crq*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a haemocyte driver (*Crq*-GAL4), which was introduced into the second chromosome of mutant stock (Mating scheme 9).

F0: ♀ *CG1847*$^{exon1\_3}$ / *FM6* X ♂ *X* ; *Bl* / *Y* *CyO*

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ *FM6* ; *+* / *Y* *CyO* X ♂ *CG1847*$^{exon1\_3}$; *+* / *FM6* *Bl*

F1b: ♀ *X* ; *+* / *FM6* *CyO* X ♂ *X* ; *Crq-GAL4* / *Y* *Crq-GAL4*

F2: ♀ *CG1847*$^{exon1\_3}$; *Bl* / *FM6* *CyO* X ♂ *FM6* ; *Crq-GAL4* / *Y* *CyO*

F3: ♀ *CG1847*$^{exon1\_3}$; *Crq-GAL4* / *FM6* *CyO* X ♂ *FM6* ; *Crq-GAL4* / *Y* *CyO*

Mating scheme 9: Combination of the *CG1847*$^{exon1\_3}$ mutant and haemocytes driver Crq-Gal4

For the insulin-secreting cell targeted rescue of lethality, UAS-hAIPwt was overexpressed under *dilp*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and an insulin secreting cells driver (*dilp*-GAL4), which was introduced into the second chromosome of mutant stock (Mating scheme 10).

F0: ♀ *CG1847*$^{exon1\_3}$ / *FM6* X ♂ *X* ; *Bl* / *Y* *CyO*

In the next generation I performed 2 different crosses at the same time:

F1a: ♀ *FM6* ; *+* / *Y* *CyO* X ♂ *CG1847*$^{exon1\_3}$; *+* / *FM6* *Bl*

F1b: ♀ *X* ; *+* / *FM6* *CyO* X ♂ *X* ; *dilp-GAL4* / *Y* *dilp-GAL4*

F2: ♀ *CG1847*$^{exon1\_3}$; *Bl* / *FM6* *CyO* X ♂ *FM6* ; *dilp-GAL4* / *Y* *CyO*

F3: ♀ *CG1847*$^{exon1\_3}$; *dilp-GAL4* / *FM6* *CyO* X ♂ *FM6* ; *dilp-GAL4* / *Y* *CyO*

Mating scheme 10: Combination of the *CG1847*$^{exon1\_3}$ mutant and insulin secreting cells driver dilp-Gal4

For muscle cell targeted rescue of lethality, UAS-hAIPwt was overexpressed under *MEF2*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a muscle cells driver (*MEF2*-GAL4), which was introduced into the third chromosome of mutant stock (Mating scheme 11).

F0: ♀ *CG1847^{exon1_3} / FM6* X ♂ *X ; Dr / Y TM3*

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ *FM6 ; + . / Y TM3* X ♂ *CG1847^{exon1_3} ; + . / FM6 Dr*

F1b: ♀ *X ; + . / FM6 Dr* X ♂ *X ; Mef2-GAL4 / Y Mef2-GAL4*

F2: ♀ *CG1847^{exon1_3} ; Dr . / FM6 TM3* X ♂ *FM6 ; Mef2-GAL4 / Y TM3*

F3: ♀ *CG1847^{exon1_3} ; Mef2-GAL4 / FM6 TM3* X ♂ *FM6 ; Mef2-GAL4. / Y TM3*

Mating scheme 11: Combination of the *CG1847^{exon1_3}* mutant and muscle cells driver Mef2-Gal4

For nervous system targeted rescue of lethality, UAS-hAIPwt was overexpressed under *elav*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847^{exon1_3}* allele and a pan-neural driver (*elav*-GAL4), which was introduced into the third chromosome of mutant stock (Mating scheme 12).

F0: ♀ *CG1847^{exon1_3} / FM6* X ♂ *X ; Dr / Y TM3*

In the next generation I performed 2 different crosses at the same time:

F1a: ♀ *FM6 ; + . / Y TM3* X ♂ *CG1847^{exon1_3} ; + . / FM6 Dr*

F1b: ♀ *X ; + . / FM6 Dr* X ♂ *X ; elav-GAL4 / Y elav-GAL4*

F2: ♀ *CG1847^{exon1_3} ; Dr . / FM6 TM3* X ♂ *FM6 ; elav-GAL4 / Y TM3*

F3: ♀ *CG1847^{exon1_3} ; elav-GAL4 / FM6 TM3* X ♂ *FM6 ; elav-GAL4. / Y TM3*

Mating scheme 12: Combination of the *CG1847^{exon1_3}* mutant and nervous system driver elav-Gal4

For glial cell targeted rescue of lethality, UAS-hAIPwt was overexpressed under *repo*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847^{exon1_3}* allele and a glial cell driver (*repo*-GAL4), which was introduced into the third chromosome of mutant stock (Mating scheme 13).

F0: ♀ *CG1847^{exon1_3} / FM6* X ♂ *X ; Dr / Y TM3*

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ *FM6 ; + . / Y TM3* X ♂ *CG1847^{exon1_3} ; + . / FM6 Dr*

F1b: ♀ *X ; + . / FM6 Dr* X ♂ *X ; repo-GAL4 / Y repo-GAL4*

F2: ♀ *CG1847^{exon1_3} ; Dr . / FM6 TM3* X ♂ *FM6 ; repo-GAL4 / Y TM3*

F3: ♀ *CG1847^{exon1_3} ; repo-GAL4 / FM6 TM3* X ♂ *FM6 ; repo-GAL4 / Y TM3*

Mating scheme 13: Combination of the *CG1847^{exon1_3}* mutant and glial cells driver repo-Gal4

For haemolymph targeted rescue of lethality, UAS-hAIPwt was overexpressed under *HE*-GAL4 driver in our genetic mutant background. Heterozygous females carrying both the mutant *CG1847^{exon1_3}* allele and hemocytes driver (*HE*-GAL4), introduced into the third chromosome of mutant stock, were generated (Mating scheme 14).

F0: ♀ *CG1847^{exon1_3} / FM6* X ♂ *X ; Dr / Y TM3*

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ *FM6 ; + . / Y TM3* X ♂ *CG1847^{exon1_3} ; + . / FM6 Dr*

F1b: ♀ *X ; + . / FM6 Dr* X ♂ *X ; HE-GAL4 / Y HE-GAL4*

F2: ♀ *CG1847^{exon1_3} ; Dr . / FM6 TM3* X ♂ *FM6 ; HE-GAL4 / Y TM3*

F3: ♀ *CG1847^{exon1_3} ; HE-GAL4 / FM6 TM3* X ♂ *FM6 ; HE-GAL4 / Y TM3*

Mating scheme 14: Combination of the *CG1847^{exon1_3}* mutant and haemolymph driver HE-Gal4

For gut targeted rescue of lethality, UAS-hAIPwt was overexpressed under *drm*-GAL4 driver in our genetic mutant background. Heterozygous females carrying both the mutant *CG1847^{exon1_3}* allele the gut driver (*drm*-GAL4) were generated (Mating scheme 15).

F0: ♀ *CG1847^{exon1_3} / FM6* X ♂ *X ; Dr / Y TM3*

In the next generation, I simultaneously performed 2 different crosses:

F1a: ♀ *FM6 ; + . / Y TM3* X ♂ *CG1847^{exon1_3} ; + . / FM6 Dr*

F1b: ♀ *X ; + . / FM6 Dr* X ♂ *X ; drm-GAL4 / Y drm-GAL4*

F2: ♀ *CG1847^{exon1_3} ; Dr . / FM6 TM3* X ♂ *FM6 ; drm-GAL4 / Y TM3*

F3: ♀ *CG1847^{exon1_3} ; drm-GAL4 / FM6 TM3* X ♂ *FM6 ; drm-GAL4 / Y TM3*

Mating scheme 15: Combination of the *CG1847^{exon1_3}* mutant and gut driver drm-Gal4

For Malpighian tubule targeted rescue of lethality, UAS-hAIPwt was overexpressed under *c42*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a Malpighian tubule driver (*c42*-GAL4), which was introduced into the third chromosome of mutant stock (Mating scheme 16).

F0:	♀	*CG1847*$^{exon1\_3}$ / *FM6*	X	♂	*X* ; *Dr* / *Y* *TM3*

In the next generation I performed 2 different crosses at the same time:

F1a:	♀	*CG1847*$^{exon1\_3}$ / *FM6*	X	♂	*FM6* ; *+* . / *Y* *Dr*
F1b:	♀	*X* ; *+* . / *FM6* *TM3*	X	♂	*X* ; *c42-GAL4* / *Y* *c42-GAL4*
F2:	♀	*CG1847*$^{exon1\_3}$; *+* . / *FM6* *Dr*	X	♂	*FM6* ; *c42-GAL4* / *Y* *TM3*
F3:	♀	*CG1847*$^{exon1\_3}$; *Dr* . / *FM6* *TM3*	X	♂	*FM6* ; *c42-GAL4* / *Y* *TM3*
F4:	♀	*CG1847*$^{exon1\_3}$; *c42-GAL4.* / *FM6* *TM3*	X	♂	*FM6* ; *c42-GAL4* / *Y* *TM3*

Mating scheme 16: Combination of the *CG1847*$^{exon1\_3}$ mutant and malpighian tubules driver c42-Gal4

For heart targeted rescue of lethality, UAS-hAIPwt was overexpressed under *tinC*-GAL4 driver in our genetic mutant background. For this purpose I generated heterozygous females carrying both the mutant *CG1847*$^{exon1\_3}$ allele and a heart driver (*tinC*-GAL4), which was introduced into the third chromosome of mutant stock (Mating scheme 17)

F0:	♀	*CG1847*$^{exon1\_3}$ / *FM6*	X	♂	*X* ; *Bl* / *Y* *CyO*

In the next generation, I simultaneously performed 2 different crosses:

F1a:	♀	*CG1847*$^{exon1\_3}$ / *FM6*	X	♂	*FM6* ; *+* . / *Y* *CyO*
F1b:	♀	*X* ; *+* . / *FM6* *Bl*	X	♂	*X* ; *tinC-GAL4* / *Y* *tinC-GAL4*
F2:	♀	*CG1847*$^{exon1\_3}$; *+* / *FM6* *CyO*	X	♂	*FM6* ; *tinC-GAL4* / *Y* *Bl*
F3:	♀	*CG1847*$^{exon1\_3}$; *tinC-GAL4* / *FM6* *CyO*	X	♂	*FM6* ; *tinC-GAL4* / *Y* *CyO*

Mating scheme 17: Combination of the *CG1847*$^{exon1\_3}$ mutant and fat heart driver tinC-Gal4

Virgin females from each of these stocks were then mated with transgenic males carrying the hAIPwt and in the next generation their ability to rescue the hemizygous mutant males was evaluated.

2.2.9 Rescue of mutant lethality by expressing truncated or missense hAIP variants under the control of a ubiquitous promoter (actin)

For rescuing the *CG1847*$^{exon1\_3}$ mutant with different hAIP variants I generated five UAS::hAIP constructs by inserting either a truncated version of hAIP or four different missense variants that were identified in FIPA families (sections 2.4.5 and 2.4.6). These constructs were sent to BestGene Inc. where they were microinjected into embryos harbouring *attP40* landing sites, according to the standard protocol. Five lines were produced for each missense variant, and then each line was balanced over *CyO*. Not all the transgenic lines survived during transportation. However, for each of the rescue experiment I used 2 of the generated stocks (Table 5).

I tested whether these UAS::hAIP transgenes (truncated or missense variants) were able to rescue *CG1847*$^{exon1\_3}$ mutants by expressing them with a ubiquitous driver (actin) during fly development.

2.2.10 Larvae collection

Egg laying plates

To obtain developmentally staged egg/larvae collections precisely, I prepared embryo collection cages covering a 60 mm agar plate containing medium comprised of apple juice hardened with agar. Egg laying plates were produced according to the following protocol: For 1L of medium I used 750 ml of water, 21.5 gr agar, 250 ml of apple juice concentrate, 25 gr of sucrose, and 5 ml 9:1 (propionic acid : phosphoric acid). The water and agar mix was autoclaved. Separately was prepared the second mix from sugar, apple juice and 5 ml of 9:1 propionic acid: phosphoric acid. When the water-agar mix cool down till around 60°C it was added the second mix, and carefully pour into 60 mm plates, without generating bubbles.

Stocks

In stocks *CG1847*$^{exon1\_3}$ */ FM6* and *CG1847*$^{2.39A}$ */ FM6* stocks obtained during the P-element excision screen, the FM6 balancer chromosome was substituted with a *FM7c, Dfd::YFP* chromosome. As a result, I was able to differentiate mutant and control males (respectively, *CG1847*$^{exon1\_3}$ */ Y* and *CG1847*$^{2.39A}$ */ Y*) based on lack of fluorescence.

Staging to determine the lethality stage

Stocks were grown at low density at 25°C. 50 mated females were allowed to lay eggs for 2 hours in food bottles, in order to minimize the variations associated with parental rearing conditions.

In the F1 generation, an additional 50 females from the mutant stock (*CG1847*$^{exon1\_3}$ / *FM7::DfdYFP*) and 50 females from the revertant stock (*CG1847*$^{2.39A}$ / *FM7::DfdYFP)* were collected. They were allowed to mate with *FM7::DfdYFP/Y* males for 3 days. These parental flies were 4-10 days post-emergence before the start of the egg collections. Adults were set for 4 h of egg laying on apple juice plates. Mutant (*CG1847*$^{exon1\_3}$ / *FM7::DfdYFP*) and the revertant (*CG1847*$^{2.39A}$ / *FM7::DfdYFP)* larvae were reared at 25°C. Embryos were allowed to age until the desired developmental stage as described in Figure 13.

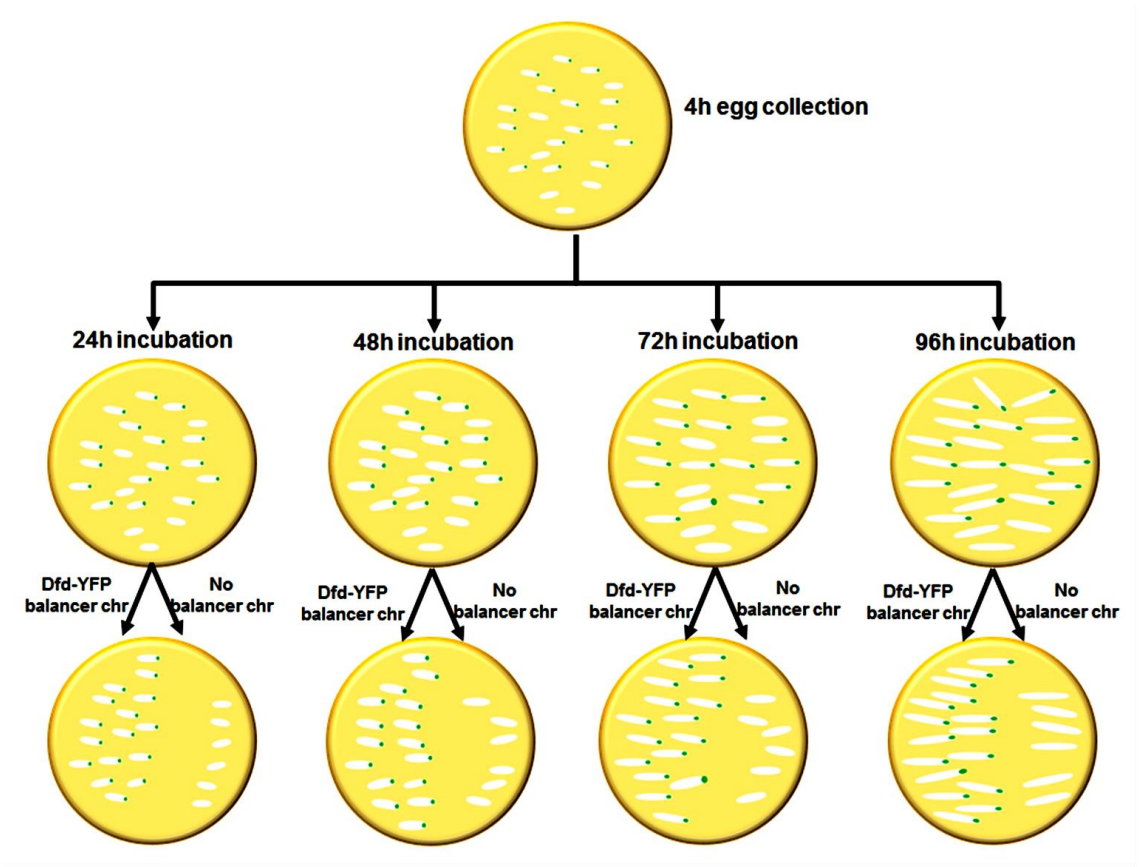

Figure 13: Design for larval staging. For determining the lethality stage for *CG1847*$^{exon1\_3}$ mutant males, larvae were selected for the presence or absence of the fluorescent marker

Staging to collect larvae for RNA extraction

For RNA extraction, mutant and control eggs were collected as described above and allowed to age for an additional 48 h, resulting in larvae from 46 to 50 h in age. The wandering non-fluorescent larvae (*CG1847*$^{exon1\_3}$ / *Y* or *CG1847*$^{2.39A}$ / *Y*) were collected using the fluorescent microscope. RNA was extracted using the technique described in section 2.3.2 B. The quality of the RNA was assessed using the 2100 Bioanalyzer (Barts Genome Centre).

2.2.11 Collection of pupae

Food bottles

To obtain large pupae collections, flies were raised in bottles containing 15 ml of standard culture medium (Appendix 3), in an incubator with constant temperature of 25°C.

Stocks

The homozygous stock *UAS-Dcr2*; + ; *UAS-CG1847-T2* was used both as control and to knockdown (KD) *CG1847* expression in wing tissues, using the *nubbin*-GAL4 driver (*UAS-Dcr-2*; *UAS-CG1847-T2>nub*-Gal4). In F0 generation, I collected 100 virgin females from the control stock (*UAS-Dcr2; UAS-CG1847-T2*). Fifty virgins were mated with *UAS-Dcr2* / *Y*; *UAS-CG1847-T2* males for control stock, while the other 50 females were crossed with *nub-Gal4* males for KD purposes. They were allowed to mate for 3 days at 25°C. These parental flies were 4-10 days old when I started the egg collections. Adults were set for 4 h egg collection which were further reared at 25°C until they reached the pupal stage.

Staging for pupae dissection

For staging purposes, pupae were picked with a wet brush at the white pupa stage. White pupae are very easy to recognize and consistency was crucial to pick white pupae at the same stage. They were transferred to an empty Petri dish containing double-sided tape. The genotype and the time of collection were carefully marked. Pupae were dissected at the desired stage: 24 h and 28 h after puparium formation (APF).

2.2.12 Dissection of *Drosophila* tissues

Imaginal wing disc dissection

To obtain imaginal wing discs for subsequent immunostaining studies, the head section of the 3rd instar larvae was inverted. After removing the gut and fat, imaginal discs were fixed in 4% PFA in PBS at room temperature for 30 minutes, followed by the steps detailed in Appendix 6.

Pupal wing dissection

24 or 28 h AFP pupae were dissected in double-sided tape at the bottom of a Petri dish. Using a pair of forceps, the operculum was removed and a small incision was made in the head of the pupa to release the pressure and prevent damage to the wings by excess fat. I used a syringe needle to gently make an incision in the cuticle along the back of the pupa and separate it in two halves, which were stuck on their side on the double-sided tape. Dissected pupae were

immediately placed in 0.5 ml 1X PBS in the wells of a 96 well plate, which was the kept at 4°C until all pupae were dissected. Pupae were fixed in 4% formaldehyde in PBT at 4°C overnight. After fixing, pupal wings were dissected. Pupae were transferred one by one to a watchmaker's glass dish with PBS. The wing epithelium is visible inside a translucent cuticle sac. While the pupa was held in position with one of the forceps, the cuticle was removed, starting near the wing hinge region. Wings were removed, with some excess hinge tissue and placed into 60 Well HLA Terasaki Microplates. For complete immunostaining protocol, see Appendix 6.

Dissection of adult wings

For adult wing imaging, flies were collected in 70% ethanol and kept at 4°C for at least 24h. For wing dissection, flies were placed in isopropanol and the wings were removed with a pair of Dumostar #5 forceps. For mounting, see section 2.2.15.

Dissection of adult fly heads

Total RNA or protein were extracted from adult fly heads of flies stored at -80°C before use. To separate the fly heads (around 25-30 heads), flies were kept on dry ice, vortexed vigorously, and placed on a piece of glace kept on top of dry-ice. Using a pair of forceps, heads were removed and placed in cooled 1.5 ml Eppendorf tubes, which were stored at -80°C prior to RNA/protein extraction.

2.2.13 Immunostaining

Dissected fly tissues were fixed in 4% PFA solution (4 % (w/v) paraformaldehyde in PBS for 20 min. Three washes in PBT 0.3%, each for 5 min, were perform. In the next step the tissues were permeabilised in PBT 0.3% for 1 h at RT on the rotor, blocked with PBT 0.3% +3%BSA 1h, and incubated overnight with primary antibodies at 4°C on the rotor. On the second day, the primary antibodies were removed and tissues were washed 3 times in 0.3% PBT. The samples were incubated with fluorescent secondary antibodies diluted in PBT 0.3% for 1h, washed again 3 times, each wash for 20 min in 0.3% PBT. In the final step the tissues were counterstained with 4',6-diamidino-2-phenylindole (DAPI) and mounted for confocal microscopy analysis. For detailed protocols for immunostaining see Appendix 6.

2.2.14 Antibodies

The complete list of the antibodies used during the course of this study, the dilutions, the use and the provider can be found in Appendix 7.

2.2.15 Mounting

Imaginal wing discs or pupal wings were mounted in 24 µl Vectashield (Vector Laboratories) using 22X22mm cover slips. Adult *Drosophila* wings were mounted on the slide with a drop of isopropanol. After waiting a few seconds for the isopropanol to dry, Euparal mounting medium (DS31 - Anglian Lepidopterist Supplies (ALS)) was added in small drops between the wings. A 22 X 40 mm coverslip was added and the slides were dried at 65°C for 4 hours.

2.2.16 Image Acquisition and Processing

Confocal images of imaginal wing discs or pupal wings were taken using an LSM 510 laser scanning microscope (Zeiss; release version 5.0 SP1.1. using the License Basic Software ZEN 2008 version 500267 configuration 4.02.00). Stacks of confocal images were collected at different focal plane spacing, depending on experimental needs. Images were then processed using Image J freeware[255] (http://rsbweb.nih.gov/ij/). Images were readjusted for each colour independently but always on the whole picture.

Adult wing imaging was performed with Panoramic 250 High Throughput Scanner (Barts Cancer Institute, QMUL).

2.3 General molecular biology techniques

2.3.1 Oligonucleotide design

For primer design, the publicly available software Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) was used. A list of primers used, their sequence, and their applications can be found in Appendix 6. Genomic sequences were retrieved from the University of California (UCSC) Genome Browser http://genome-euro.ucsc.edu/cgi-bin/hgGateway?redirect=auto&source=genome.ucsc.edu, version February 2009 (GRCh37).

Primers were ordered from Sigma-Aldrich (http://www.sigmaaldrich.com/united-kingdom.html) and they were delivered as a lyophilised pellet. To generate a 100 µM stock solution, the primers were re-suspended in distilled H_2O and then stored at -20˚C. Further dilution with distilled water was performed to obtain a 10 µM working solution.

2.3.2 Deoxyribonucleic acid extraction

Adult fly DNA isolation

For extracting *Drosophila melanogaster* DNA, I used the following buffers:

Buffer A (for 10 samples): 100 µl of 1 M Tris-HCl (pH7.5), 200 µl of 0.5 M EDTA (pH8), 20 µl of 5 M NaCl, 50 µl of 10% SDS, 630 µl of ddH2O.

Buffer B (for 10 samples) was prepared from 1 ml of 5 M KAc, and 2.5 ml of 6 M LiCl (no water).

One adult fly was placed in a 1.5 ml reaction tube, which was then frozen at -20 °C for ten minutes. Then, frozen flies were physically disrupted in 200 µl of buffer A with a plastic pestle and incubated at 65°C for 30 minutes. 100 µl of buffer B was added and mixed by inverting the tube, this was then incubated for ten minutes at room temperature and centrifuged at 14000 rpm for 15 minutes at room temperature. In the next step, 250 µl of the supernatant was transferred into a new 1.5 ml reaction tube and 200 µl of ice-cold isopropanol was added. The solution was inverted for mixing and then centrifuged again at 14000 rpm for 15 minutes. The genomic DNA pellet was re-suspended in 200 µl ice-cold 75% ethanol, air dried and dissolved in 30 µl ddH_2O. The samples were stored at -20°C until further use.

RNA isolation from adult flies

mRNA from adult flies was extracted from 15 fly heads. Vials containing previously collected flies that were stored at -80°C were placed into liquid nitrogen and left to cool. The samples were then vortexed in order to isolate the heads (section 2.2.12). The heads were placed in new 1.5 ml Eppendorf tubes and the RNA extraction was performed using 500 µl of TRI reagent. Tissues were homogenised with a pestle, followed by brief vortexing and incubation at room temperature for 10 minutes. The tubes were then centrifuged at 4˚C at 15000 x g for 15 minutes; the resulting aqueous phase was transferred to a new 1.5 ml Eppendorf. 250 µl of isopropanol was added and this mixture was vortexed to ensure mixing had occurred. This was then incubated at room temperature for a further 10 minutes. The mix was centrifuged at 12000 x g for 8 minutes at 4˚C. The resulting supernatant was discarded. 500 µl of 75% ethanol (prepared with nuclease free water) was added and a new centrifugation step was performed at 7500 x g for 6 min at 4˚C. The ethanol was removed by air drying the RNA pellet. RNA was eluted in 30 µl of nuclease free water and it was incubated at -20˚C for the same day analysis. The samples were stored at -80˚C.

RNA isolation from Drosophila melanogaster larvae

mRNA from *Drosophila* larvae was extracted using the RNeasy Micro kit from Qiagen, according to manufacturer's specifications. The collected larvae (section 2.2.10) were homogenized in 350 µl Buffer RLT using a pestle. The lysate was centrifuged 3 min at 12000 x g. The supernatant was carefully transferred to a new Eppendorf. In the next step, 1 volume of 70% ethanol was added to the lysate, and mixed well. The sample was transferred to an RNeasy MinElute spin column

in a 2 ml collection tube, which was centrifuged for 15 s at ≥8000 x g. The flow-through was discarded and 350 µl Buffer RW1 was added to the RNeasy MinElute spin column. A new centrifugation step took place for 15 s at ≥8000 x g. The flow-through was again discarded and a 10 µl DNase I stock solution mixed with 70 µl Buffer RDD was added to the column. The sample was incubated at room temperature for 15 min. Additional 350 µl Buffer RW1 were added to the RNeasy MinElute spin column and centrifuged for 15 s at ≥8000 x g. The RNeasy MinElute spin column was placed in a new 2 ml collection tube. The RNA sample was washed with 500 µl Buffer RPE followed by centrifugation for 15 s at ≥8000 x g. The flow-through was discarded and added 500 µl of 80% ethanol. The next centrifugation step was for 2 min at ≥8000 x g. The RNeasy MinElute spin column was placed in a new 2 ml collection tube and centrifuged at full speed for 5 min to dry the membrane. RNA was eluted by adding 14 µl RNase-free water. The last centrifugation step was for 1 min at full speed to elute and recover the RNA.

2.3.3 Nucleic acids quantification

DNA and RNA sample purity and concentration was assessed using a ThermoScientific Nanodrop 1000 Spectrophotometer. Purity was measured based on the ratio of OD260:OD280. Readings of >1.8 for DNA and >2.0 for RNA indicated acceptable levels of purity.

2.3.4 First strand cDNA synthesis

Complementary DNA (cDNA) was produced from RNA using a M-MLV Reverse Transcriptase Kit (Invitrogen). Before performing the reverse transcription, RNA samples were diluted in RNA free water and standardized to the same concentration (500ng RNA in 16.95 µl of H_2O). In the next step the cDNA dilutions were incubated in the thermocycler (G-Storm GT-12061) for 10 min at 65^{o}C and added to each individual reaction. For each reaction the master mix consisted of 0.5 µl M-MLV (100U), 5 µl M-MLV RT 5x buffer M531A, 1.25 µl dNTPs (20mM), 0.25 µl Random Hexamers (250 ng/µl), 0.05 µl RnaseOUT 40 U/µl, 1µl DTT 100 mM, and 1 µg RNA. A volume of 8.05 µl of the mix was added to each RNA sample (to make a final volume of 25 µl). The incubation was performed in the thermocycler under the following conditions: 10 min at 26^{o}C, 60 min at 37^{o}C and 10 min at 92^{o}C.

The cDNA was then frozen at -20˚C if not used immediately.

RT-PCR was performed according to the standard PCR protocol using the cDNA as the template and primers that spanned intron/exon boundaries to act as a control against contamination with genomic DNA. The integrity of the cDNA was usually verified by PCR of the housekeeping gene

Ribosomal protein L32 (RpL32). For PCR conditions refer to section 2.3.5. PCR products were analysed on 2% agarose gels.

2.3.5 Polymerase Chain Reaction (PCR)

PCRs were performed at optimized conditions, according to standard procedures.

The basic protocol was modified depending on the DNA/RNA sample and the primer annealing temperatures. The annealing temperature was set to the calculated temperatures seen in Supporting Table 2. Temperature gradients were used to identify the optimal annealing temperature for the various primer pairs.

To test the efficiency of RNAi-mediated gene silencing, the specific mRNA levels were measured via semi-quantitative RT-PCR (for the primer sequences and annealing temperatures are described in Appendix 6) using the following protocol:

Component	12.5 µl Reaction	Final concentration
5X Green GoTaq reaction buffer	2.5 µl	1X
$MgCl_2$	0.75 µl	1 mM
10 mM dNTPs	0.25 µl	0.2 mM
10 µM Forward Primer	0.5 µl	5 µM
10 µM Reverse Primer	0.5 µl	5 µM
Template DNA	variable	<1.000 ng
Go-Taq DNA Polymerase	0.06 µl	5 u/µl
Nuclease-Free Water	to 12.5 µl	

Table 6: The reaction mix used for semi-quantitative RT-PCR

Cycling was performed in 0.2 ml thin wall PCR tubes using a G-Storm GT-12061 thermocycler. Cycling procedures were typically:

- **95˚C 5 min**
- **94˚C 30 s**
- **Primer specific ˚C 30sec 31 cycles**
- **72˚C 30 sec**
- **72˚C 10 min**
- **Samples were then cooled to 4˚C**

PCR techniques were used to identify *Drosophila* mutants generated during the imprecise excision screen. The GoTaq polymerase (Promega) was used to identify large genomic deletions, according to the supplier's recommendations.

Component	12.5 µl Reaction	Final concentration
5X Green GoTaq reaction buffer	2.5 µl	1X
MgCl2	0.75 µl	1 mM
10 mM dNTPs	0.25 µl	0.2 mM
10 µM Forward Primer	0.5 µl	5 µM
10 µM Reverse Primer	0.5 µl	5 µM
Template DNA	variable	<1.000 ng
Go-Taq DNA Polymerase	0.06 µl	5 u/µl
Nuclease-Free Water	to 12.5 µl	

Table 7: The reaction mix used to identify large genomic deletions

Cycling was performed in thin wall 0.2 ml PCR tubes using a G-Storm GT-12061 thermocycler. Cycling procedures were typically:

- **95˚C 5 min**
- **Touch-down 71-62˚C 10 cycles**
- **94˚C 45 sec**
- **60.9˚C 45 sec 28 cycles**
- **72˚C 2 min and 30 sec**
- **72˚C 10 min**
- **Samples were then cooled to 4˚C**

The small genomic deletions from the screen were identified using Taq polymerase (NEB) according to the supplier's recommendations.

Component	25 µl Reaction	Final concentration
5XQ5 reaction buffer	5 µl	1X
10mM dNTPs	0.5 µl	100 µM
10µM Forward Primer	1.25 µl	0.5 µM
10µM Reverse Primer	1.25 µl	0.5 µM
Template DNA	variable	<1.000 ng
Q5 DNA Polymerase	0.25 µl	2,000 u/ml
Nuclease-Free Water	to 25 µl	

Table 8: The reaction mix used to identify small genomic deletions

Cycling was performed in 0.2 ml thin wall PCR tubes using a G-Storm GT-12061 thermocycler. Cycling procedures were typically:

- **94˚C 5 min**
- **94˚C 30 sec**
- **54.2˚C 30 sec 30 cycles**
- **72˚C 45 sec**
- **72˚C 10 min**
- **Samples were then cooled to 4˚C**

2.3.6 Agarose gel electrophoresis of DNA

PCR reaction products were loaded on agarose gels to verify their size. The bands were separated by electrophoresis for 60 minutes at 120 V on 2% agarose gels. The 2% agarose gels were prepared by mixing 2.4g of agarose (Sigma-Aldrich, UK) with 120 ml of 1x TAE buffer (Tris/Acetat/ethylenediaminetetraacetic acid (EDTA)) (National Diagnostics, UK). Gels were stained with 12 µl of 1000X Red Nucleic dye (Life Technologies). A DNA marker (GeneRuler^{TM} DNA Ladder Mix, 0.5 mg DNA/ml, Fermentas, UK) was run alongside the samples to determine PCR product sizes. 6x loading dye was added to DNA samples prior to gel loading.

2.3.7 DNA gel extraction

When necessary, DNA bands were visualised on UV lamp and excised from the agarose gel using a sterile scalpel. The extraction and purification of DNA from the agarose gel was performed using a QIAquick Gel Extraction Kit (Qiagen).The extracted DNA was finally eluted with 30 µl ddH_20. Following this, the PCR buffer components were removed and the purified fragments were then used for sequencing or ligation reactions.

2.3.8 Site-directed mutagenesis

In vitro site-directed mutagenesis allows the introduction of site-specific mutations in double stranded plasmids. Mutagenic primers were designed using the Stratagene's QuickChange primer design program at www.stratagene.com/qcprimerdesign. Two oligonucleotides flanking the nucleotide to be changed were designed. The primers utilised are listed in the Appendix 5.

The QuickChange XL Site-Directed Mutagenesis kit is performed in three steps: synthesis of the mutant strand, digestion of the parental strand with *Dpn*I and transformation of competent cells with plasmid DNA.

Mutant strand synthesis: a PCR was performed for denaturation of the DNA template, annealing of the mutagenic primers containing the desired mutation. The primer extension was performed with PfuUltra HF DNA polymerase (2.5 U/µl). Sample reactions were prepared as indicated:

Component	50 µl Reaction
10x reaction buffer	5 µl
10mM dNTPs	1 µl
125ng sense primer	1.25 µl
125ng reverse primer	1.25 µl
10ng/µl Template dsDNA	1 µl
PfuUltra DNA Polymerase	1 µl
Nuclease-Free Water	38.20 µl

Table 9: The reaction mix for direct site mutagenesis

The thermal cycling conditions were the following:

- **95°C 1 min**
- **95°C 50 sec 18 cycles**
- **60°C 50 sec**
- **68°C 5 min (1 min/kb of plasmid length)**
- **68°C 7 min**

Following PCR samples were placed on ice for 2 min followed by digestion with *Dpn*I.

***DpnI* digestion of template:** methylated parental and hemimethylated DNA were digested with *Dpn*I. To each PCR reaction, 1 µl of *Dpn*I restriction enzyme (10 U/µl) was added and the reaction was gently mixed several times by pipetting up and down. The reaction was centrifuged for 1 min and then incubated at 37°C for 1 h.

Transformation: mutated plasmids were transformed into XL10-Gold Ultracompetent cells for nick repair following the protocol recommended by the producer.

2.3.9 Automated DNA sequencing

Sequencing was carried out at the Genome Centre (Barts and The London, Queen Mary, University of London). The Genome Centre uses BigDye 3.1 chemistry with visualization on the ABI 3730 capillary sequencer. This is an automated capillary gel electrophoresis that generates read lengths of up to 850 base pairs with a quality (Phred) score of over 20. Each PCR product was sequenced using both forward and reverse primers. Sequence chromatograms were

visualised and analysed using the BioEdit Sequence Alignment Editor software (http://www.mbio.ncsu.edu/bioedit/bioedit.html) (Ibis Biosciences, Carlsbad, CA).

2.3.10 Protein extraction

For western blot analysis, 25 fly heads were collected as described in section 2.2.12. Proteins were extracted in 35 µl of protein extraction buffer. For 1000 µl were used 857 µl protein extraction buffer, 1 µl DTT 1M, and 142 µl of protease inhibitor 7X. Fly heads were homogenised with a pestle. Following this, the tissues were subjected centrifugation at 4˚C at 15000 x g for 15 minutes; the resulting aqueous phase was transferred to a new 1.5 ml Eppendorf prior to quantification.

2.3.11 Protein quantification according to the Bradford method

To prepare concentration standards, 12 µl of 10 mg/ml BSA (Bovine Serum Albumin Promega R3961) and 18 µl of ddH_2O were mixed in a PCR tube to obtain the Standard 6 (4000 µg/ml). In the next step, 15 µl of Standard 5 were mixed with 15 µl of ddH_2O to obtain the Standard 2 solution (2000 µg/ml). Each dilution step resulted in a further 2-fold change in the concentration from the previous one until I obtain the Standard 1 solution (125 µg/ml). For the standard 0 was used ddH_2O only. For triplicate readings, 4 µl of each standard or sample were mixed with 196 µl of 1X Bradford reagent and added to three contiguous wells of a 96-well plate. The absorbance was read with the Wallac VICTOR plate reader at 595nm. The results were exported to an Excel file for further analysis.

2.3.12 Western Blotting

Buffers:

6X SDS loading buffer: 0.3M Tris-Cl pH 6.8, 0.6M DTT, 12% SDS, 0.6% Bromophenol blue, 60% glycerol and ddH_2O till 50 ml.

*Running buffer***:** NuPAGE MES SDS Running Buffer (20X), Invitrogen (Product code: NP0002). Adjust with 1X running buffer to a final volume of 1 l.

Transfer buffer: 25mM Tris base, 190 mM Glycine, 20% (v/v) Methanol. Adjust with ddH_2O to a final volume of 1 l.

Phosphate buffered saline (PBS): Oxoid Products (Code BR0014). Dissolve 10 tablets in 1 l ddH_2O and autoclave.

Washing buffer (PBS-T): 0.1% Tween 20. Adjust with 1X sterile PBS to a volume of 1 l.

Blocking buffer: 5% (w/v) Semi-skimmed milk in PBT (2.5g of non-fat milk powder in 50 ml of 1 X 0.1% PBS-Tween).

Protocol:

After quantification, protein samples were normalised to the same concentration, 20 µg/ml protein, and the dilutions were kept on ice in 0.6 ml tubes. Equal volumes of 6X SDS loading buffer were added to the samples. Samples were boiled for 5 min at 95°C on a thermal cycler to denature proteins' quaternary and tertiary structure, allowing them to run through the gel.

The protein separation was carried out using pre-cast 12 wells NuPAGE Novex 4-12% Tris-Acetate Protein Gels (Life Technologies) at 120 V, for 1.5 hours in 1x Running Buffer or until the samples reached the bottom of the gel chamber. The separated proteins were transferred into a Protran BA85 nitrocellulose membrane (GE Healthcare) by a "semi-dry blotter" for 30 min at 15 V/500 mA. The membranes were then blocked for 1.5 h at room temperature in 5% milk in PBS-T. The membranes were incubated with the primary antibodies (Appendix 7) diluted in blocking buffer, overnight at 4°C. On the second day, after removal of primary antibodies, the membranes were washed with PBT 3X for 10 minutes at RT and incubated with secondary antibodies, diluted in blocking buffer, for 1.5 h at room temperature. Odyssey Infrared Imaging System (LI-COR) was used for image acquisition.

2.4 Cloning

2.4.1 Plasmids

Table 10 lists the plasmids used in the course of this study, and their purpose.

Plasmids used	Purpose	Reference
pGEM-T easy	vector used for sub-cloning	Promega
pW@RpA	Transformation vector	Kind gift from Nick Brown[256]
pW@RpA CG1847	Transformation vector	
pUAS-k10_attP	Transformation vector	Kind gift from Nick Brown
pcDNA3-Myc-AIP		
pUAS-k10_attP_AIPwt	Transformation vector	
pUAS-k10_attP_Myc-AIPtrunc	Transformation vector	
pUAS-k10_attP_AIP-R16H	Transformation vector	
pUAS-k10_attP_AIP-C238Y	Transformation vector	
pUAS-k10_attP_AIP-A299V	Transformation vector	
pUAS-k10_attP_AIPR-304Q	Transformation vector	

Table 10: List of plasmids used in this study. Lines in the white boxes were obtained from specific supplier or were gifts from other labs. The lines in the light grey boxes were existing lab stocks generated during previous studies. The lines in the dark grey boxes are genomic rescue constructs cloned during this study for generating the transgenic fly stocks.

2.4.2 Subcloning

PCR amplification

PCR was used to generate the gene specific constructs, which were used to generate transgenic flies. Q5 (NEB) High-Fidelity DNA Polymerase is a high-fidelity, thermostable DNA polymerase with 3´→ 5´ exonuclease activity, fused to a processivity-enhancing Sso7d domain to support robust DNA amplification. This enzyme was used to ensure high fidelity DNA amplification. The reaction mix was set up as described in the manufacturer's protocol as follows:

Component	25 µl Reaction	Final concentration
5X Q5 reaction buffer	5 µl	1X
10mM dNTPs	0.5 µl	100 µM
10µM Forward Primer	1.25 µl	0.5 µM
10µM Reverse Primer	1.25 µl	0.5 µM
Template DNA	variable	<1.000 ng
Q5 DNA Polymerase	0.25 µl	2,000 u/ml
Nuclease-Free Water	to 25 µl	

Table 11: The reaction mix used to generate gene specific constructs

Cycling was performed in 0.2 ml thin wall PCR tubes using a G-Storm GT-12061 thermocycler. Cycling procedures were typically:

92°C 30 sec
92°C 30 sec
50-72°C 30 sec 25-30 cycles
72°C 30 sec/kb
72°C 2 min
Hold 4°C
Samples were then cooled to 4˚C

Each PCR reaction consisted of 35 cycles and the annealing temperatures were set using temperature gradients to identify the optimal annealing temperature for each primer pair.

After the PCR reaction, the products were loaded on an agarose gel to verify the size of the products. Afterwards the PCR buffer components were removed from the fragments according to the protocol described in section 2.3.7). The purified fragments were then used for ligation reactions.

Addition of 3' A overhangs to PCR products

Because the PCR amplification was performed using a proofreading DNA polymerase, such as Q5, the resulted products have blunt ends. As a consequence, to make the DNA amplicons purified and extracted from the gel suitable for TA cloning onto pGEM-T easy backbone (below) the Taq DNA polymerase (NEB) was used in the next step to add an adenine residue to the 3´-end of both strands of the amplicons.

For the next step is critical to remove all the Proofreading DNA Polymerase by purifying the PCR product carefully (with QIAquick Gel Extraction Kit from Qiagen –section 2.3.7). It might be possible that the proofreading activity of any DNA Polymerase remains will degrade the freshly added A' overhangs.

The typical Taq DNA polymerase reaction mix for a typical 25 µl reaction:

	Final Concentration	Volume(µl)
Purified PCR product	0.15 to 1.5 pmol	Variable*
dATP (10 mM)	0.2 mM	0.5
10X Taq Buffer with Mg	1x (1.5 mM $MgCl_2$)	2.5
Taq DNA Polymerase (5 U/µl)	1U	0.1
ddH_2O	-	up to 25 µl

Table 12: The typical Taq DNA polymerase reaction mix for adding A' overhangs:

The specific amount of purified PCR products was calculated based of the size of the amplicons. The recommended amount is 10–100 ng PCR product for each 100 bp length of the PCR product.

The mix was incubated at 72 °C for 1 hour, immediately followed by the TA cloning protocol, as for increased efficiency, it is recommended to use fresh PCR products.

pGEM-T Easy Vector

Both the human gene (*hAIP*) and fly orthologue gene (*CG1847*) were amplified by PCR (previously described in section 2.4.2 in order to be subcloned into the pGEM-T easy vector. The pGEM-T easy vector system (Promega, UK) (Figure 14) is a useful system that was developed for the cloning of PCR products. The vector is linearized and has T overhangs at both ends (a 3' terminal thymidine). These T overhangs at the ends of the insertion site improve the efficiency of the PCR product ligation into the plasmids, as well as prevent vector re-circularisation.

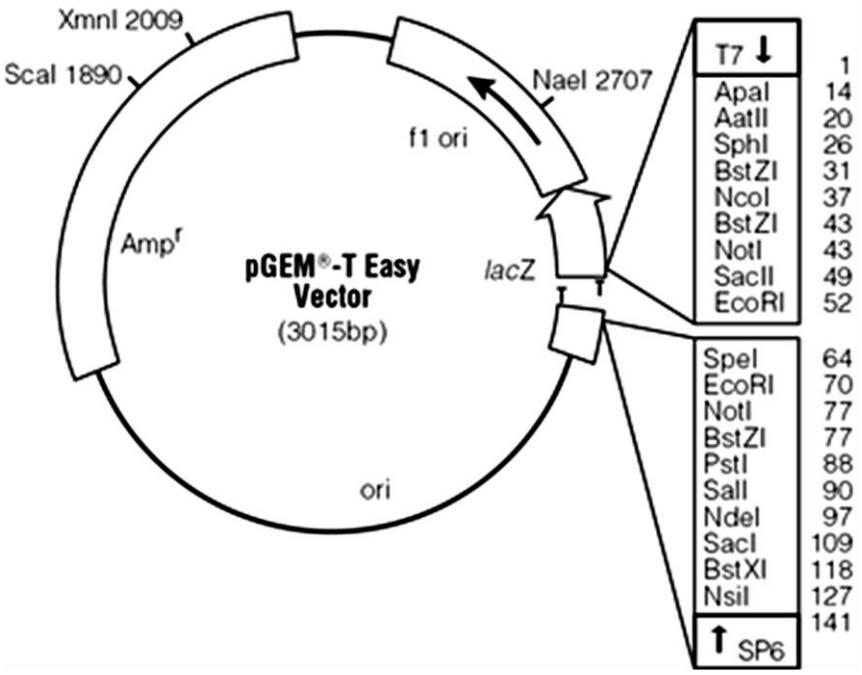

Figure 14: Schematic structure of the pGEM-T Easy Vector. The vector contains a multiple cloning site (MCS) flanked by T7 and SP6 RNA polymerase promoters. The ampicillin (Amp^r) resistance gene is useful for selection of colonies positive for the presence of the desired insert (source Promega protocols)

DNA restriction digest

The PCR products inserted into the pGEM-T vector were retrieved with either NEB or Promega restriction endonucleases, according to the producer's recommended protocol. The destination plasmids were also digested with the same enzymes to allow for the sub-cloning of the insert. The restriction digestion reactions were performed at 37°C in a water bath.

For directional cloning of the insert into the plasmid, inserts were cloned using two different restriction sites at the 5' and 3' ends, where possible. If this was not possible, only one restriction site was used and plasmids obtained from different colonies were screened for the right orientation of the insert. Where double digests were performed, digests were performed in a single reaction using the appropriate buffer. To confirm the presence of the insert and vector backbone I used restriction digestions of mini-preps.

Dephosphorylation of DNA fragments:

In order to avoid self-ligation of the plasmid backbones in cloning strategies, I used Antarctic Phosphatase (NEB), which removes the 5´ phosphates from DNA and RNA required by ligases.

1/10 volume of 10X Antarctic Phosphatase Reaction Buffer was added to 1 µg of DNA (plasmid was cut with restriction endonucleases and cleaned by gel extraction). I then added 1 µl of

Antarctic Phosphatase (5 units) and the mix was incubated for 60 minutes at 37°C. The Antarctic Phosphatase was heat inactivated for 15 minutes at 70°C.

DNA ligation

Ligations were performed using the DNA ligation kit (Promega). Before proceeding with the ligation of the PCR/digested fragments into the pGEM-T easy vector, the amount of the insert to be ligated was calculated. The amount of vector used was 100 ng. The following formula was applied:

[ng of vector × size of insert (kb)] ÷ size of vector (kb) × molar amount of (insert ÷ vector) = ng of insert

The ratio between the vector and the insert was calculated as followed:

a) If the insert was smaller than vector, then a 3:1 ratio (insert:vector) was used;
b) If the insert was almost equal to vector, then a 1:1 ratio (insert:vector) was used

As a control for self-ligation of the plasmid backbone I used the same calculations, but nuclease-free water was added to the reaction instead of the PCR product.

Reactions were incubated overnight at 4°C and subsequently used for transformation of *E.coli* competent bacteria as described in 2.4.7.

2.4.3 Cloning of genomic rescue construct pWhiteRabbitpolyA (pW@RpA)+CG1847

To obtain the genomic rescue construct pW@RpA+CG1847 (Figure 15), 2672 bp *CG1847* insert (*CG1847* with upstream and downstream sequences) was retrieved from pGEM-T easy using the restriction enzyme Not1 (NEB) and subsequently cloned into the pWhiteRabbitpolyA transformation vector. After the independent digestion reactions were carried out for the CG1847 inserts and the pW@RpA backbone, the samples were separated on 2% agarose gels. The fragments corresponding to the desired products sizes were cut from the gel and purified according to the purification protocol in section 2.3.7. The purified backbone was treated with Antarctic Phosphatase (section 2.4.2).

As described above, 100 ng of vector was used for DNA ligations, with an insert: vector ratio of 3:1. The reaction was incubated overnight at 4°C.

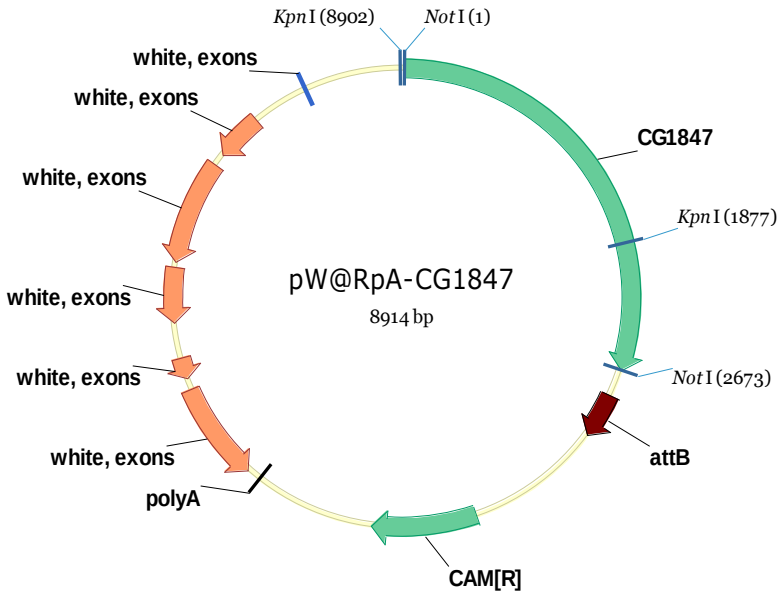

Figure 15: Schematic structure of genomic rescue construct pW@RpA+CG184

Afterwards I proceeded to transform the vector into competent *E.coli* as described in section 2.4.7.

2.4.4 Cloning of genomic rescue construct pUASK10attB+ full-length hAIPwt

To obtain the genomic rescue construct pUASk10attB-AIPwt (Figure 16) hAIP insert (1001 bp) was amplified from a pcDNA3-Myc-AIPwt vector existing in our lab (primers sequence in the Appendix 5).

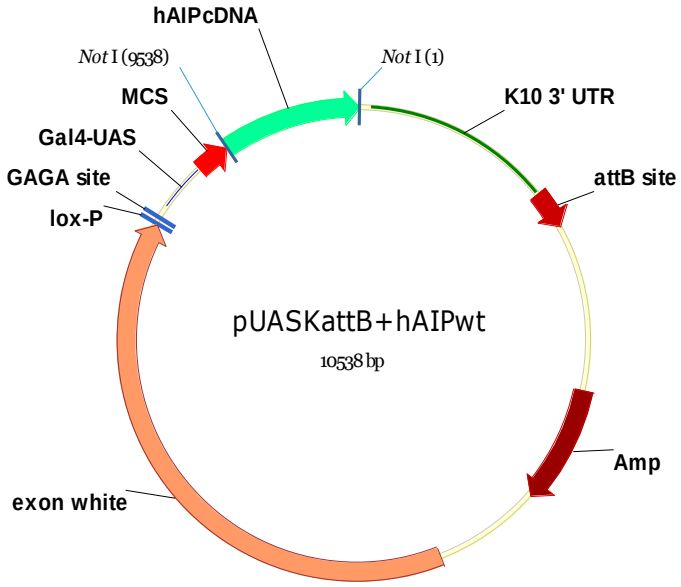

Figure 16: Schematic structure of genomic rescue construct pUASk10attB-hAIPwt

The PCR amplified product was subcloned into a pGEM-T easy vector. This construct was sequenced, and revealed that the last 7 bases of AIP together with the enzyme restriction site were not amplified. The QuickChange XL site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to introduce the missing 7bp to repair the pGEM-T easy+hAIPwt construct. The site-directed mutagenesis protocol is detailed in section 2.3.8.

The repaired hAIPwt insert was released from pGEM-T easy vector using the restriction enzyme Not1 (as one of the desired enzyme restriction sites was lost). The pUASK10attB backbone was also digested with Not1. The digestion products were separated on 2% agarose gels. The fragments corresponding to the desired products sizes were cut from the gel and purified according to the purification protocol in section 2.3.7.

The purified backbone was treated with Antarctic Phosphatase (section 2.4.2). As described above, 100 ng of vector was used in DNA ligations, with an insert: vector ratio of 3:1. The reaction was incubated overnight at 4°C. Afterwards I proceeded to transform competent *E.coli* as described in section 2.4.7. In further experiments, I generated a set of 4 constructs carrying *AIP* missense variants downstream of the attB site (for detailed description of the phiC31 system see section 3.2.3).

2.4.5 Cloning of genomic rescue constructs pUASK10attB+ truncated hAIP (hAIPtrunc)

To obtain the genomic rescue construct pUASk10attB-AIPtrunc (Figure 17) 907 bp of hAIP together with the upstream Myc tag were amplified from a pcDNA3-Myc-AIPwt vector. The last 86 bp of AIP, the ones encoding for the 7th alpha helix were deleted, generating a truncated hAIP variant (hAIPtrunc). This fragment was cloned into a pGEM-T easy vector after adding A' overhangs (detailed protocol in section 2.4.2), from which was released using the restriction enzyme Not1 (NEB). The pUASk10attB backbone was also digested with Not1.

After the independent digestion reactions were carried out at 37°C, the samples were separated on 2% agarose gels. The fragments corresponding to the desired products sizes were cut from the gel and purified according to the purification protocol in section 2.3.7.

The purified backbone was treated with Antarctic Phosphatase (section 2.4.2) to avoid re-circularization. As described above, 100 ng of vector was used for DNA ligations, with an insert: vector ratio of 3:1. The reaction was incubated overnight at 4°C.

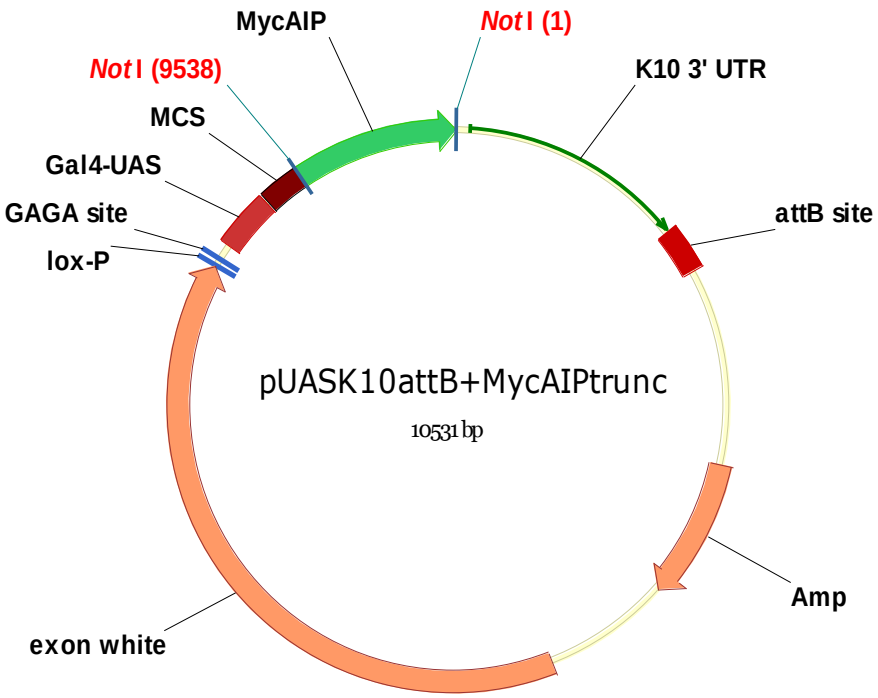

Figure 17: Schematic structure of genomic rescue construct pUASk10attB-hAIPwt

Afterwards I proceeded to transform competent *E.coli* as described in section 2.4.7.

2.4.6 Cloning of genomic rescue constructs pUASK10attB+ hAIP missense variants

The genomic rescue construct pUASK10attB+hAIP was used to introduce missense variants via site-directed mutagenesis (section 2.3.9). The primers for mutagenesis were designed using the QuikChange Primer Design tool:

Variant	Location	Pathogenic	Primers (Sequence 5' to 3')
c.47G>A (p.R16H)	N-terminus	No	ggacgggatccaaaaacatgtgatacaggaaggcc
			ggccttcctgtatcacatgtttttggatcccgtcc
c.713G>A (p.C238Y)	TPR2 domain	Yes	tgctgctcaactactgaccagtgcaagctggt
			accagcttgcactggtcagtagttgagcagca
c.896C>T (p.A299V)	TPR3 domain	Unlikely	acccagccctggtgcctgtggtgag
			ctcaccacaggcaccagggctgggt
c.911G>A (p.R304Q)	TPR3 domain	Yes	ctgtggtgagccaagagctgcgggc
			ggcccgcagctctcagctcaccacaggc

Table 13: Missense variants generated in the study

Site-directed mutagenesis was carried out using the QuikChange II XL Site-Directed Mutagenesis Kit according to the protocol suggested by the manufacturer (detailed protocol 2.4.4). Afterwards I proceeded to transform competent *E.coli* as described in section 2.4.7. Isolated

colonies were selected and grown overnight at 225 rpm/37°C in 5 ml LB broth supplied with 100 µg/ml ampicillin.

2.4.7 Transformation of competent cells with plasmid DNA

E. coli strains

The bacterial hosts used in this study are seen in Table 14

Strain	Genotype
JM109 (Promega)	endA1, recA1, gyrA96, thi, hsdR17 (rk–, mk+), relA1, supE44, Δ(lac-proAB), [F´ traD36, proAB, laqIqZΔM15]
JM109	Home-made –protocol at Appendix 8
XL10-Gold Ultracompetent cells	Tet^rΔ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 relA1 gyrA96 relA1 lac Hte [F´ proAB $lacI^q$ZDM15 Tn10 (Tet^r) Amy Cam^r].

Table 14: *E. coli* strains used in this study

To prepare competent cells for plasmid DNA transformation, I used the method detailed in Appendix 9.

Antibiotic usage

Selection for ampicillin resistance on L-Agar or in L-Broth was performed using 10 µg/ml ampicillin, from a 100 mg/ml stock solution (w/v) which was stored at – 20°C.

Preparation of LB/ampicillin plates

15g of LB (Luria Bertani) (Sigma-Aldrich, UK) media was mixed with 6g Agar (Sigma-Aldrich, UK) and 500 ml of ddH_2O was added. The medium was the autoclaved. After sterilized, the medium was melted and allowed to cool before adding 100 µg/ml of ampicillin. Then, 25 ml of medium were poured into petri dishes.

Transformation of competent cells with plasmid DNA

Cells were thawed on ice, and aliquoted into pre-chilled 1.5 ml falcon tubes to a volume of 50 µl. 25-50 ng of DNA of interest or up to 10 µl of ligation reaction were added under sterile conditions, and gently mixed. A control for ligation tube with 10 µl of control for ligation reaction was also prepared as well. The cells were incubated on ice for 30 min, and heat shocked for 45 sec in a 42°C water bath. The tube was immediately transferred to ice, and following a 2 min recovery step, 900 µl of pre-warmed SOC broth (Invitrogen) was added. The tube was shaken at 250 rpm at 37°C for one hour, and spread on a LB-agar plate containing the appropriate antibiotic. The plates were incubated overnight, inverted, at 37°C.

2.4.8 Plasmid DNA purification- Miniprep

In order to purify plasmid DNA QIAprep Miniprep kit (Qiagen, UK) was used according to the producer's specifications. A single isolated white colony from a plate was inoculated in 5 ml of LB medium containing 5 µl of 100mg/ml ampicillin and incubated overnight at 37°C at 250 rpm. After incubation, bacterial cells were harvested by centrifugation for 3 minutes at 8000 rpm at room temperature. The supernatant was then removed and the pellet re-suspended in 250 µl of buffer P1, followed by 250 µl of buffer P2 and mixed thoroughly by inverting the tube 4-6 times. 350 µl of buffer N3 were also added and mixed immediately by inverting the tube 4-6 times. The solution was centrifuged for 10 min at 13000 rpm. Supernatants were applied to a column and centrifuged for 30-60 seconds. The column was further washed with 0.75 ml of buffer PE and centrifuged for 30-60 seconds. To remove any residual buffer, the column was centrifuged for an additional 1 min. The plasmid DNA was eluted in 50 µl of H_2O. The concentration of each sample was determined using the Nanodrop. To confirm the integrity of the isolated plasmids a digestion with appropriate restriction enzymes was performed, separated on an agarose gel and sequenced.

2.4.9 Preparation Maxi prep for plasmidic DNA extraction

In order to purify higher volumes of plasmid DNA GenElute™ HP Plasmid Maxiprep Kit was used according to the producer's specifications.

The plasmid DNA was isolated from a 150 ml culture by centrifugation at 5000 X g for 10 minutes to pellet the cells. The cells were resuspended by adding 12 ml of Resuspension/RNaseA Solution and completely mixed by pipetting up and down, or vortexing. In the next step, 12 ml of Lysis Solution was added and the tube was immediately gently inverted 6 to 8 times. The mixture was allowed to sit and clear for 3 to 5 minutes. 12 ml of chilled Neutralization Solution were added to the mixture and gently invert 4 to 6 times. The tubes were allowed to sit until a white aggregate (cell debris, proteins, lipids, SDS, and chromosomal DNA) separated. In the next step were added 9 ml of Binding Solution and invert 1 to 2 times. Samples were immediately poured into the barrel of a filter syringe. The cell lysate was allow to sit for 5 minutes until the white aggregate should float to the top. GenElute HP Maxiprep Binding Columns were placed into 50 ml collection tube and prepared by adding 12 ml of the Column Preparation Solution. The columns were then spin at 3000 X g for 2 minutes. The eluate was discarded and the lysate was filtered in order for plasmid DNA to bind to the column. The columns were centrifuged at 3000 X g for 2 minutes and the step was repeated until the rest of the cleared lysate was filtered. 12 ml of Wash Solution 1 were added to the column, which was spun at 3000 X g for 2 minutes.

After discarding the eluate, 12 ml of Wash Solution 2 were added to the column and centrifuged at 3000 X g for 5 minutes. The plasmid DNA was eluted by adding 3 ml of ultrapure water and, for maximum recovery of plasmid, collection tubes were centrifuged at 3000 X g for 5 minutes. Sample concentration was determined using the Nanodrop.

2.5 Illumina TruSeq stranded mRNA sample preparation - Low sample protocol

Total RNA was isolated from eight larval collection (above, section 2.3.2) using Qiagen RNeasy MicroKits. RNA samples were purified using the DNase I as suggested in the Qiagen protocol and resuspended in 14 µl of RNAse free water. RNA samples were measured using Nanodrop and subjected to Agilent 2100 bioanalyzer for checking the integrity of the extracted RNA. The samples were normalized to 500 ng/µl. Equal volumes of RNA from four control and four mutant samples were used for generating the cDNA libraries according to the Illumina protocol (detailed protocol in Appendix 10). Samples were prepared using four replicates for controls and four replicates for mutants.

All libraries were sequenced in one lane, 30 million reads per sample on an Illumina HiSeq 1500 instrument.

Acknowledgement: The library preparation and the RNA sequencing were performed at the Barts and the London School of Medicine genomics core facility (Genome Centre).

2.6 Sequence alignment to reference transcriptome

Analysis pipeline short workflow

To identify the genes whose expression is affected in *CG1847* mutants a high-throughput RNA sequencing method was employed, which was performed with the help of colleagues at our Genome Centre (Dr. Charles Mein, and Dr Anna Terry Systems Administrator Bioinformatician).

The raw results from the Illumina Hi-Seq were cleaned for removing the adaptors, then the FASTQ files were exported for analysis in Bowtie[257] and TopHat[258] to generate the database of transcripts.

For the next step, assemble of the possible transcripts and their annotation, was used Cufflinks. Cufflinks (version 2.2.0) was run for each sample separately and the final results were assembled into a single merged.gtf file.

1. Pre-alignment

The results from the sequencer platform were submitted to a primary analysis. After a pre-alignment quality control, the unaligned reads were cleaned by trimming the bases from both 3' and 5' ends. The reads that were too short were removed after trimming.

The sequence alignment is a key step on this type of experiment. As a consequence, the read alignment program should be carefully chosen. One of the most efficient programs developed in the last few years – Bowtie[257] (http://bowtie-bio.sourceforge.net/index.shtml) – was the best option to align our data to the reference transcriptome for *Drosophila melanogaster*. Both the forward and reverse strands were considered for alignment. The unaligned reads were split and realigned by TopHat (http://ccb.jhu.edu/software/tophat/index.shtml), a junction aware pipeline[258].

2. Checking quality of data using FastQC

FastQC is a key control step that offers a simple quality control checks which provides a quality control report which is necessary to reveal the possible problems. This analysis presents the results in Per Base Sequence Quality mode using a Box Whisker type of plot. The quality scores, are also known as Phred quality scores, as they were developed initially for the Phred program for automation of DNA sequencing during the Human Genome Project[259,260].

All the samples were submitted for quality control (Figure 18).

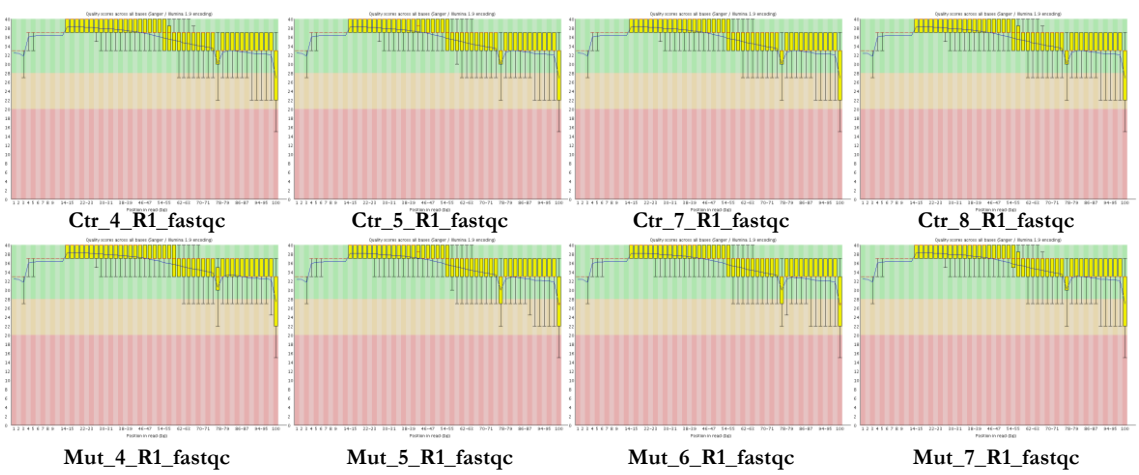

Figure 18: FastQC quality control of cDNA libraries. The background of the graph is divided along the y axis into three fields: green, which represents very good quality calls; orange, denoting calls of acceptable quality; and red, indicating poor quality calls. Control and mutant samples are shown in the top and bottom row, respectively. The yellow boxes represent the 25-75% quartile range, while the upper and lower whiskers represent the 10% and 90% points. The y-axis on the graph shows the quality scores also called the Phred quality scores.

All the samples passed the Phred quality control, reaching high scores as the maximum Phred quality for Illumina technology is 40.

3. Transcriptome assembly using Cufflinks

Cufflinks version 2.2.0 (http://cole-trapnell-lab.github.io/cufflinks/releases/v2.2.0/) was used for transcriptome assembly and differential expression analysis of RNA-Seq raw data.

All the alignment files were merged together by Cuffmerge, and the merged files were compared with the reference annotation to find novel unannotated features. Raw sequencing data was aligned to the most recent genome release: dmel_r6.03_FB2014_06 (ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/dmel_r6.03_FB2014_06/).

The new released version 2.2.0 includes two new programs, Cuffquant and Cuffnorm. Our samples were passed through Cuffquant, which quantifies the levels of expression for a single BAM file. The results were visualized using CummeRbund.

A schematic overview of the workflow is described in

Figure 19. Cufflinks version 2.2.0 (CummeRbund branch) was used in this analysis.

4. Read count normalisation and quantification

The actual RNA-seq quantification process takes in account both the molar concentration and the transcript length. The Mortazavi's[261] formula for a unit of measure of read density reflects the molar concentration of a transcript in the starting sample by normalizing for RNA length and for the total read number in the measurement. The normalization method allows comparison of transcript levels not only between the samples, but also within the same sample (Mortazavi *et al*., 2008)[261].

As a specific measure of reads density we used RPKM, as this is the most commonly used formula. RPKM indicates the Reads Per Kilobase of exon model per Million mapped reads:

RPKM = C/ N x L (**C** = number of mappable reads that belongs to exons, **N** = total resulted number of mappable reads, **L** = the sum of the exons (bp))

For example, a 1kb transcript with 2000 alignments in a sample of 10 million reads (out of which 8 million reads can be mapped) will have a RPKM=2000/ (1 x 8) = 250. A pair of reads constitutes one fragment[261].

The final results were assembled into a single merged.gtf file.

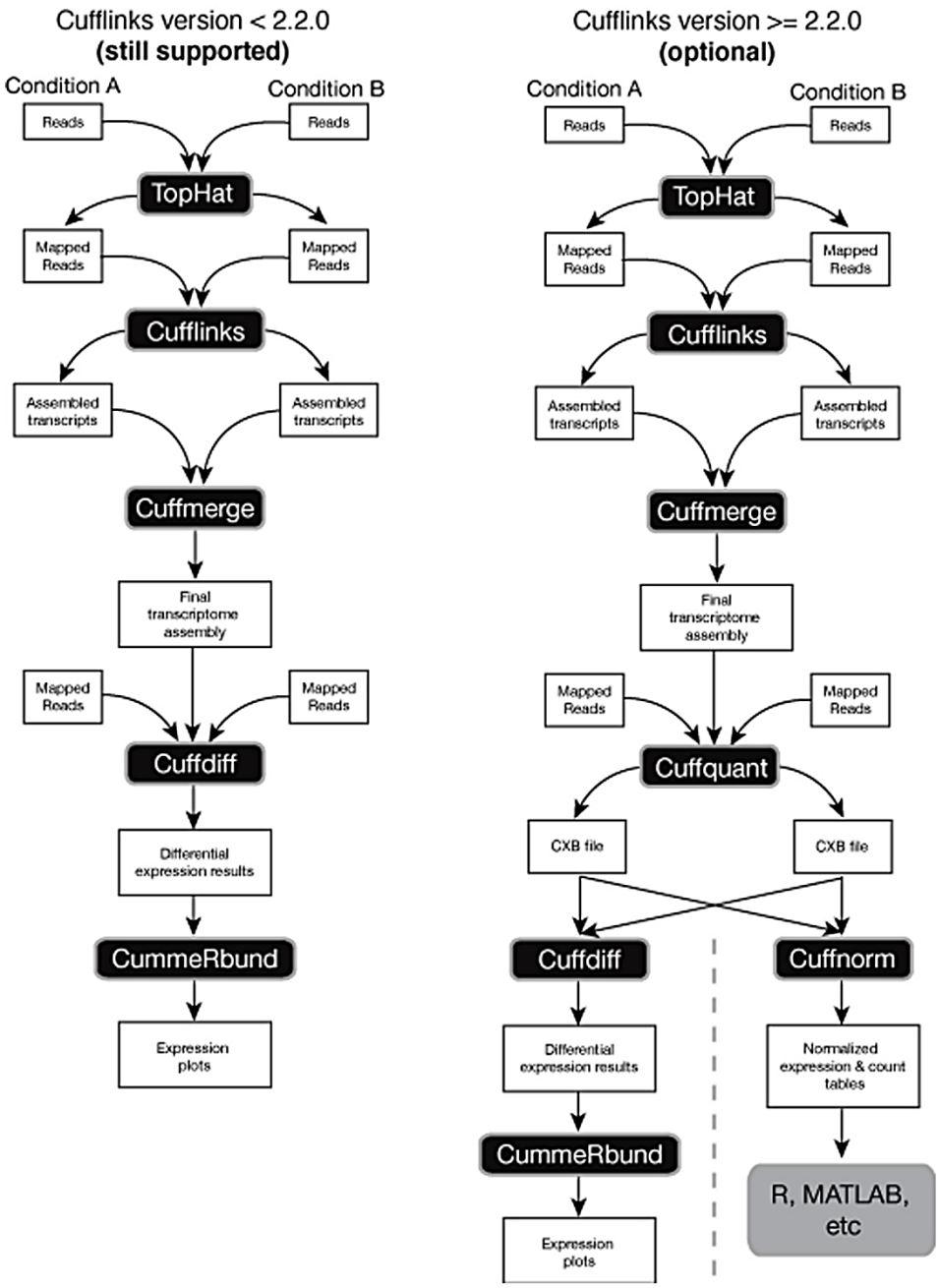

Figure 19: Overview of the bioinformatic workflow. The most recent version and the previously published workflow, described in Trapnell *et al*[262] are presented

2.7 STRING analysis

STRING (Search Tool for the Retrieval of Interacting Genes)[263] (http://string.embl.de/) was used to analyse the interactions and relationships among the significantly up- or down-regulated genes. STRING pipeline assembles the input data based on neighbourhood information, high-throughput experiments, co-occurrence, co-expression and data mining.

This database reveals the functions of proteins at the molecular level by generating the protein-protein interactions (PPI) network. Confidence scores is calculated for each interaction pair and default cutoff scores of above 0.4 were selected.

2.8 Validation by multiplex- qPCR of selected transcripts

A custom designed GeXP multiplex qPCR for gene targets was performed using 100ng of total RNA. The GeXP multiplex assay consisted of 16 pairs of chimeric primers: CG1847 transcript, two different reference genes (alpha-Tubulin at 84B, and Ribosomal protein 39) and 13 transcripts of interest. Target-specific reverse transcription and PCR amplification was performed in accordance with manufacturer's instructions (Beckman Coulter, Wycombe, UK). In brief, a master mix was prepared for reverse transcription reactions as detailed in the GeXP Start Kit (Beckman Coulter, Wycombe, UK) and performed using a G-Storm thermal cycler, using the program protocol: 48°C, 1 min, 42°C, 60 min, and 95°C, 5mins. From this, an aliquot of each reverse transcription reaction was added to PCR master mix containing GenomeLab kit PCR master mix, Thermo Scientific Thermo-Start Taq DNA polymerase.

The qPCR reaction was performed using a G-Storm thermal cycler with a 95°C activation step 10mins, followed by 35 cycles of 94°C, 30 secs, 55°C, 30 secs and 70°C, 60 secs. Products were separated by capillary gel electrophoresis using CEQTM 8000 Genetic Analysis system. Following fragment separation, peaks were analysed and matched to corresponding genes using GenomeLab Fragment Analysis software (Beckman Coulter).

ANOVA (post hoc test Fishers LSD) was applied for comparisons revealing at least a twofold change in expression levels. P values of less than 0.05 were considered significant.

Information on all primers can be found in Appendix 5.

2.9 Statistical analysis

Experimental data sets were analysed in JMP® (SAS institute). The Shapiro-Wilk and the Kolmogorov-Smirnov tests were used to check the normal distribution of the quantitative variables. Two way ANOVA or Kruskal-Wallis non-parametric if the distribution was non-normal tests were used to evaluate the significant differences. Means and standard deviations were used to report parametric data. The chi-square (I) test was used to determine whether there is a significant difference between the expected genotypes frequencies and the observed frequencies in the rescue experiments. A $P<0.05$ was considered to be statistically significant.

CHAPTER 3: ESTABLISHING A *DROSOPHILA MELANOGASTER* MODEL TO STUDY AIP FUNCTION

3.1 INTRODUCTION

At the start of my project the majority of investigations of AIP roles had focused on the developmental aspects, particularly on its involvement in limiting the normal morphogenesis processes during the embryological period.

Previous studies had examined mainly the AhR-mediated xenobiotic signalling possibly by influencing AhR nuclear translocation. However, in addition to xenobiotic metabolic processes the AIP protein is also a chaperone involved in a variety of essential molecular functions as: domain binding[142,145,172,264], signal transducer activity[140], transcription coactivator activity and transcription factor binding[265]. AIP function in adult organisms has been relatively less explored as the loss of AIP is associated with lethality during early embryological stages. It was previously reported that loss of the *Aip* gene in mice leads to lethality at e14.5 accompanied by cardiac malformations such as double outlet right ventricle, ventricular-septal defects, and pericardial edema[204].

Taken together, these data indicate an extended AIP involvement in normal mammalian biology, apart from its role as an AhR co-chaperone. Therefore, it was reasonable to hypothesise that AIP might play a role as a key regulator of developmental growth. The initial aim of my project was to formally examine the actual role of AIP in development and the mechanism behind the lethality associated with the loss of this protein. Furthermore, it is well known that *AIP* functions as a tumour suppressor gene in pituitary tumorigenesis and understanding what role this protein plays in abnormal growth could therefore be of significance to cancer initiation and progression.

Over the years, *Drosophila melanogaster* has evolved from being a useful animal model for investigating genetics and the mechanisms of inheritance to being one of the most valuable tools for understanding gene function. This insect proved to be a powerful model for the study of normal development and the mechanisms underlying disease pathogenesis, as well as an aid for the development of new drugs. New advanced tools and techniques are constantly being developed keeping *Drosophila* at the forefront of research.

Some of these approaches and methods were employed in the present study for a better understanding of AIP function.

3.2 BACKGROUND

We used some of the valuable tools that were developed in *Drosophila* model to gain further knowledge on the impact of CG1847 loss of function.

3.2.1 The Gal4/UAS system

With its powerful genetic tools, *Drosophila* is an important animal model for functionally characterising genes involved in specific biological processes. A breakthrough for *Drosophila* genetics came with the ingenious combination of budding yeast (*Saccharomyces cerevisiae*) and *Drosophila* genetics which resulted in the GAL4-UAS system[238]. This technique allows selective induction of gene expression in a temporally and spatially controlled manner. This system involves the creation of two fly lines carrying either a Gal4 or a UAS transgene, which are then crossed to achieve a desired genotype.

The Gal4-UAS system has two essential components. The first is the *Gal4* gene, which encodes a transcription activator protein originally found in *S. cerevisiae*, where it controls the expression of galactose metabolism genes. Gal4 can directly bind to the second element of the system, a DNA consensus sequence called "Upstream Activating Sequence" or UAS, which acts as an enhancer sequence for neighbouring genes. The versatility of the system arises from the fact that the Gal4 and UAS elements can be combined with specific promoter sequences (a driver). In other words, one of the flies (the female or the male) is a transgenic for the *Gal4CDS* coupled with either a ubiquitous promoter (e.g. *Act-Gal4*) or a tissue-specific promoter (e.g. *Elav-Gal4*). The other fly is a transgenic for a P-element containing the UAS site fused with a sequence of interest, as the CDS of a specific gene or a particular RNAi (e.g., *UAS-CG1847-RNAi* Figure 20).

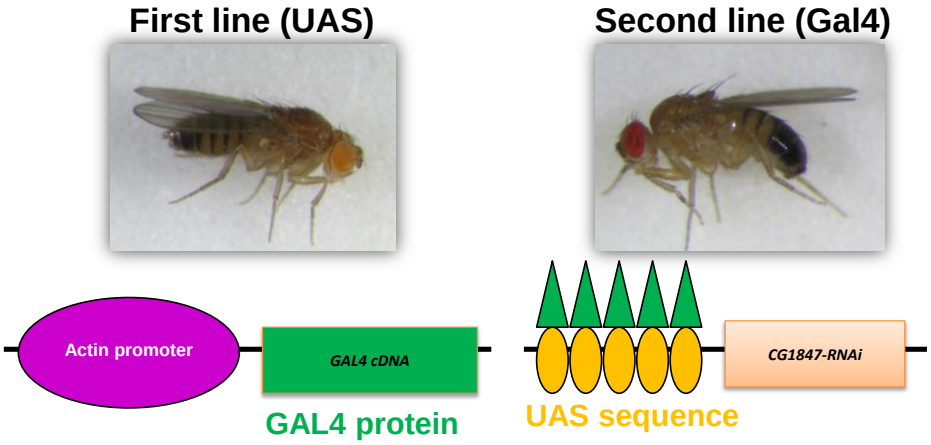

Figure 20: UAS-Gal4 system in *Drosophila*. The system consists of two parts: flies expressing the yeast GAL4 transcription factor under the control of specific promoters/enhancer named driver (red oval) and flies carrying a *CG1847-RNAi* sequence downstream of an UAS promoter region (yellow ovals). GAL4 is only expressed in cells where the driver is active. Gal4 specifically binds to UAS (green triangles) to activate transcription. Adapted from www.hoxfulmonsters.com/wp-content/uploads/2008/05/uas-gal4.jpg

The offspring of the cross between these two lines will include individuals that carry both genetic elements and, therefore, will express the desired CDS/RNAi ubiquitously (e.g., UAS-*CG1847-RNAi* expression in *actin*+ cells by *Act-Gal4/*UAS-*CG1847RNAi*) [266].

The Gal4/UAS system presents a few very important advantages. Maybe the most important is the separation of the driver (Gal4) and the effector (UAS) between in two parental lines. This prohibits the activation of the transcription in the parents, which means that it is possible to generate transgenic lines for proteins that are toxic, or lethal, which is the case for AIP.

3.2.2 Transposable elements for mutagenesis use

Another revolution in the *Drosophila* field was introduced by Rubin and Spradling in 1982[253] with the use of the P-element for transgenesis[267]. A P-element is an 8-10 kb transposon that has the ability to insert and excise itself within the genome. In order for transposition to take place a transposase enzyme is necessary. Transposase recognises inverted repeat sequences that are localised at the ends of P-elements and promotes transposon excision and reinsertion at a different genomic locus. Naturally occurring P-elements encode an internal transposase gene, whereas P-element lines developed for laboratory use have been specifically designed to lack this transposase to avoid deleterious effects of having a mobile DNA element in the genome. This allows researchers to induce transposition in a controlled manner by crossing lines that carry a P-element with a line that carries the transposase (e.g. enzyme Δ2.3).

One important molecular biological aim a *Drosophila* researcher needs to achieve is to silence a gene of interest and P-elements are valuable tools for this[268,269]. If the function of a gene cannot be disrupted by inserting a P-element, it is possible to use the P-element system to induce a deletion of the target gene. This is possible because P-elements preferably insert into the 5' UTR of genes and their excision from a genomic locus may be imprecise. Indeed, when P-elements are mobilised they often carry flanking genomic sequences. If this occurs in the vicinity of a gene, it can generate a deletion mutant for that gene. When a P-element is excised from a chromosome a double-stranded break (DSB) is created, which can be repaired either by homology-directed repair or by non-homologous repair[270]. In some cases, the DSB ends may be degraded before the repair occurs and, as a result, a deletion of the genetic material may occur, an event known as imprecise excision[271]. This event may excise also the flanking genomic DNA, which is removed with the P-element and generates deletions around the original P-element insertion point. In so-called precise excisions the size of the deleted genomic DNA occurring during such an event may vary from a few base pairs to several kilobases. The size of the remaining fragments of the P-element inverted terminal repeats ranges between 5 and 18 bp,

and are called “footprints”[272,273]. The resultant stocks are called “revertant”, and might be used as a control, if the remaining insertion is not in a coding region or if it does not introduce a frameshift mutation.

Public and commercial fly stock repositories contain transgenic fly lines with P-elements inserted in different areas of the genome. One advantage of this tool is the fact that P-elements were engineered to carry the *white* gene as a marker. Therefore, excision of the P-element and the surrounding DNA (“imprecise excision”) will also excise the *white* gene, therefore providing a means of screening those flies in which the excision has taken place.

3.2.3 PhiC31 system

In 1998 a genome integration system was developed that allows a very precise insertion of DNA elements into the genome[274]. This system is based on the site-specific phiC31 integrase. PhiC31 is a bacteriophage that encodes a serine integrase enzyme, which has the ability to mediate sequence-directed recombination between a bacterial attachment site (*attB*) and a phage attachment site (*attP*)[254] (Figure 21).

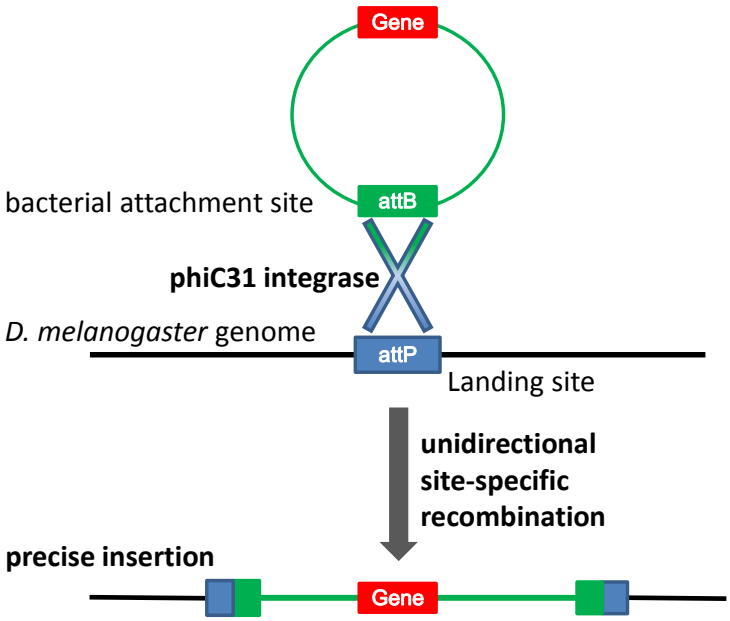

Figure 21: Schematic diagram of the phiC31 integration system into the *Drosophila* genome. A plasmid carrying a gene of interest and an attB attachment will be integrated into the Drosophila genome at an attP landing site

One of the main advantages of this system is the fact that the phiC31 integrase does not require any cofactors to mediate site-specific recombination. In 2004, this system was adapted for use

in *Drosophila*[275] and stable recombinants (transgenic flies) were generated to aid *in vivo* studies and the understanding of gene function and regulation.

Since the system's introduction, an entire library of well characterized, highly efficient landing sites that span the four chromosomes of the *Drosophila* genome have been generated. The landing sites were designed and selected to prevent them from interfering with other *Drosophila* transposon systems and to minimize: 1) effects of the insertions (e.g. interfering with the gene function at the insertion site); 2) effects on gene expression of genes on the P-element (e.g. enhancer-trap).

3.2.4 High-throughput sequencing (RNA-seq) for identifying differentially expressed genes

The transcriptome is the full range of total or messenger RNA (mRNA) and their level of expression during normal or pathologic conditions expressed by an organism. The main purpose of transcriptomic analysis is to evaluate all the transcripts in a specific sample, to quantify changes in expression levels of transcripts during development or under different physiological conditions.

These approaches were first developed at the beginning of 1990s with Sanger sequencing cDNA libraries[276]. In recent years, deep sequencing technologies based on RNA (total or fractionated, such as poly (A+)) were highly improved. The recent development of high-throughput DNA sequencing methods has provided a new tool, termed RNA sequencing or RNA-seq, which allows both transcriptome mapping and quantification. Briefly, total RNA is fragmented and converted to a double-stranded cDNA library. Adaptors are attached to the ends of the cDNA strands and, after library validation, samples are normalized and mixed in a pool that is subjected to sequencing (Figure 22). Illumina Next-Generation Sequencing allows whole-transcriptome analysis via total RNA sequencing. The TruSeq Stranded Total RNA results in a complete transcriptome, including coding and noncoding RNA, and a very accurate gene expression quantification.

The resulting raw sequencing data is used to generate *de novo* transcriptome assembly as RNA-seq has an important advantage over previous techniques: previous knowledge of the genome sequence is not required[277]. RNA-seq allows genome sequencing to a single base resolution. The quantification of transcript expression levels is based on counting the number of reads corresponding to the RNA from each known exon[261].

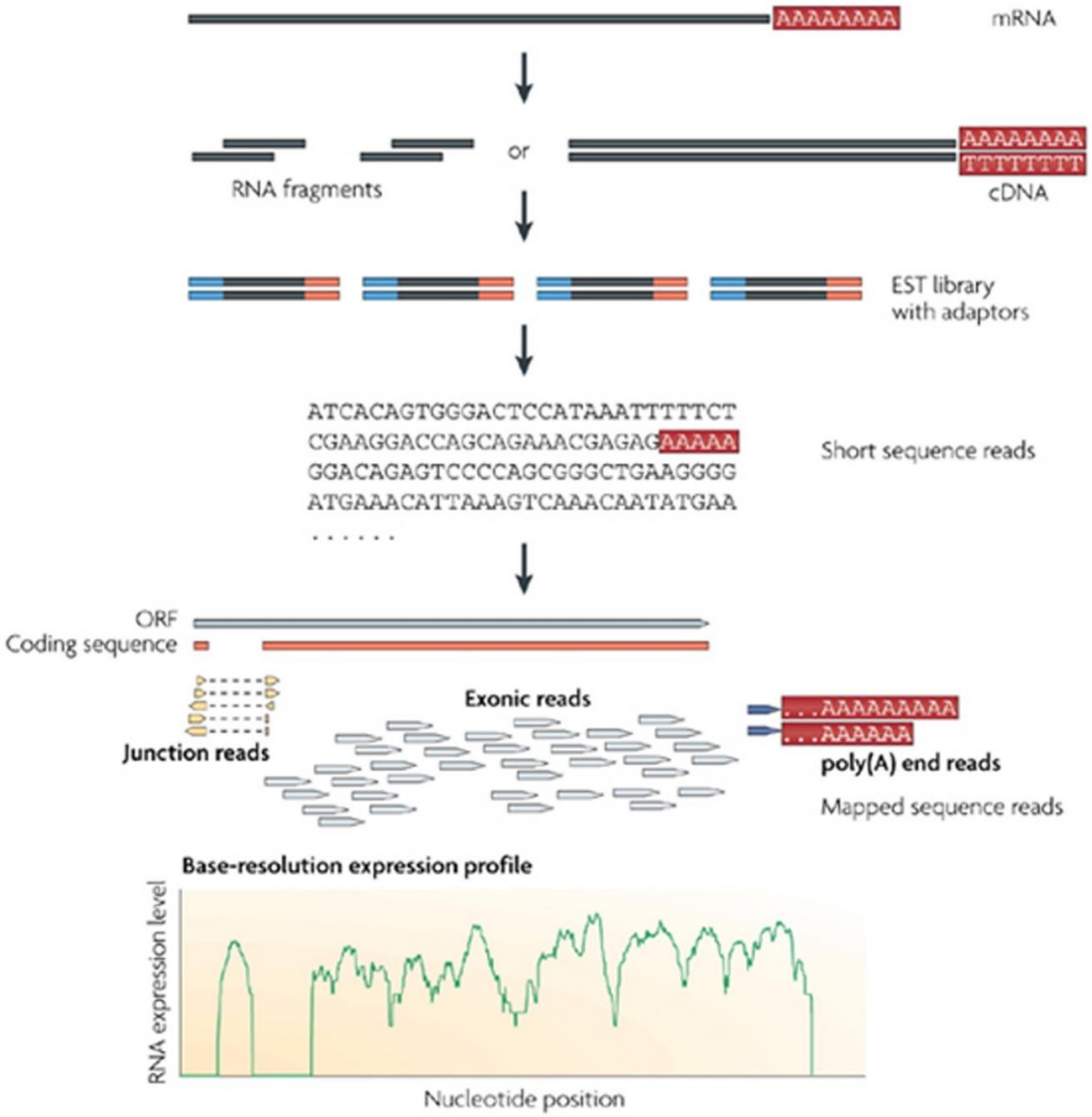

Figure 22: RNA-Seq workflow. Briefly, long RNAs are first converted into a library of cDNA fragments through either RNA fragmentation or DNA fragmentation (see main text). Sequencing adaptors (blue) are subsequently added to each cDNA fragment and a short sequence is obtained from each cDNA using high-throughput sequencing technology. The resulting sequence reads are aligned with the reference genome or transcriptome, and classified as three types: exonic reads, junction reads and poly(A) end-reads. These three types are used to generate a base-resolution expression profile for each gene, as illustrated at the bottom; a yeast ORF with one intron is shown (adapted from Wang *et al*.[277]).

In this chapter, I used RNA-seq technology to discover changes in transcript expression elicited by loss of expression of *Drosophila* AIP (*CG1847*). This approach allowed me to identify several transcripts whose expression was significantly altered in *Drosophila* AIP mutants. The results from the RNA-seq approach were subsequently confirmed via multiplex qPCR.

3.3 OBJECTIVES

AIP is involved in organism development and survival, and its deregulation leads to lethality. Genes that are essential for development are more likely to contribute to disease than non-essential genes[278], so it is important to fully understand their function. To analyse the molecular mechanisms of loss of *AIP in vivo*, I used *Drosophila* as a model system to:

a. Characterise the effect of *Drosophila AIP* (*CG1847*) knockdown, using RNAi-mediated gene expression silencing under the control of the GAL4/UAS-system.

b. Generate a *CG1847* knockout mutant by imprecise excision of a P-element (P-CG1847^{G1839}) located in the 5'UTR of *CG1847*.

c. Rescue the AIP mutant via re-expression of wt *CG1847* under the control of its endogenous promoter, to unambiguously verify that mutant phenotypes are a result of *CG1847* disruption.

d. Reveal the underlying molecular mechanisms of AIP loss by performing whole transcriptome analysis.

e. Validate via multiplex-qPCR the significantly differentially expressed genes detected by RNA-seq in mutant larvae compared to controls.

3.4 RESULTS

3.4.1 CG1847 is a *Drosophila melanogaster* AIP orthologue

Bioinformatic analysis identified *CG1847* as a putative *Drosophila* orthologue of the human *AIP* gene (identified via GenBank BLAST search). Very little is known about the role of *CG1847* as it has never been studied in the fruit fly. *CG1847* is a protein coding gene with 2 annotated transcripts (CG1847-RA and CG1847-RC), encoding 2 polypeptides. The molecular function mentioned for CG1847 in FlyBase is as aryl hydrocarbon receptor binding protein, and it is described as being involved in protein folding. No phenotypic data ara available in public databases.

Alignment of the amino acid sequences of CG1847 (FBtr0073567) and human AIP (ENST00000279146) using the Clustal Omega suite revealed that CG1847 has 37.74% overall identity with human AIP proteins (Figure 23). Clustal Omega is a multiple sequence alignment

program for proteins producing biologically meaningful multiple sequence alignments of divergent sequences (http://www.clustal.org/omega/)[279].

Figure 23: Sequence alignment of *Drosophila* CG1847 protein and human AIP. Stars indicates identity; high similarity amino acids are indicated by colons (:) while low similarity amino acids by dots (.). Source of protein sequences: Ensembl release 81 - July 2015 © WTSI / EBI http://www.ensembl.org

According to the ClustalW algorithm the human and *Drosophila* proteins share 120 identical amino acids, 80 are strongly conserved, while 34 of the amino acids are weakly conserved. The other 96 amino acids are unique between the 2 species.

Large-scale gene expression studies have been performed in *Drosophila* and they have proved to be extremely useful in providing data regarding transcription of specific genes during different developmental stages[280,281]. Overall, during the majority of development stages the CG1847 levels of expression are very low, or just moderate. The highest expression of CG1847 was found in the first 2 hours, maybe due to maternal contribution (due to the mothers supplying the eggs with their own mRNA or proteins, to support the development during the very initial stages).

Gene profiling from the earliest stages of embryonic development to adulthood revealed that *CG1847* expression varies in different tissues. The available data from the FlyAtlas[282] and modENCODE[283] platforms provide an extensive overview of *CG1847* expression during development in multiple larval and adult tissues (Figure 24).

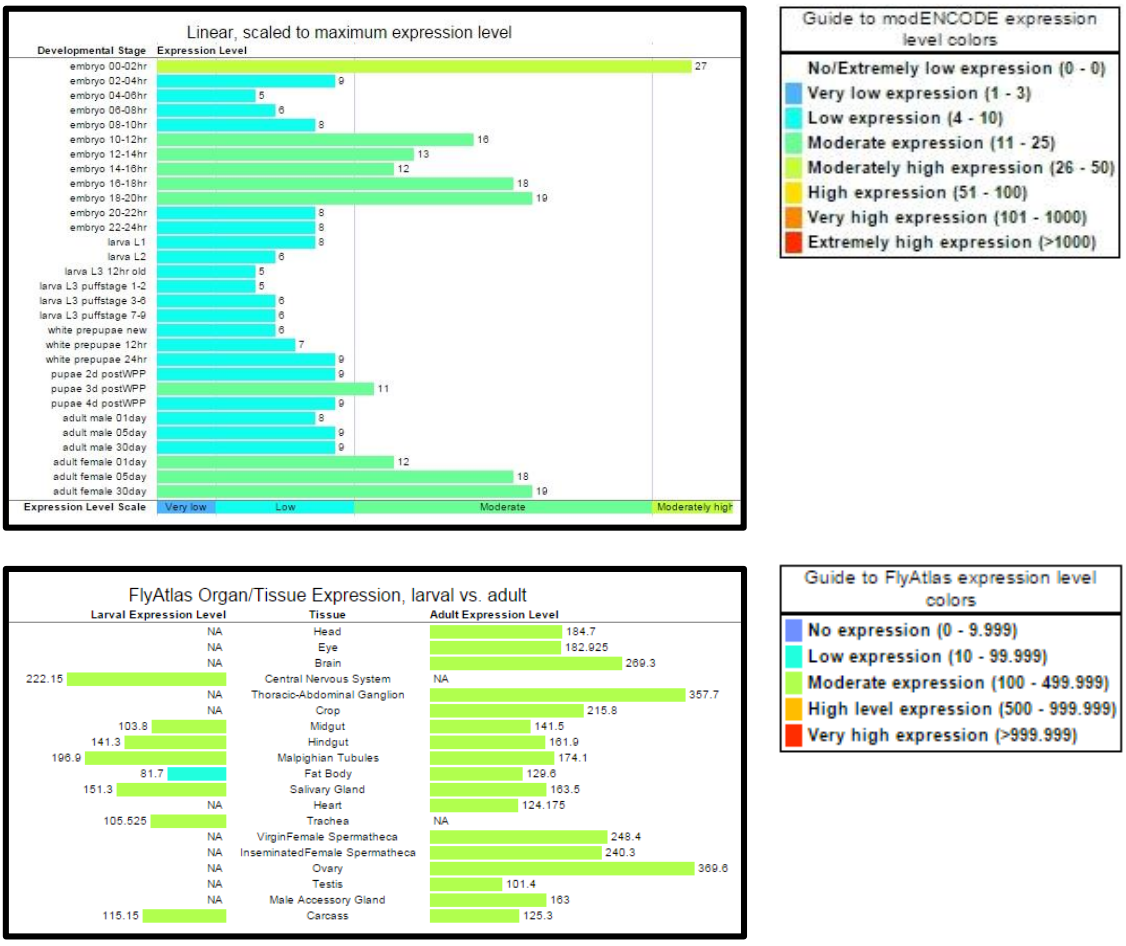

Figure 24: Expression of *CG1847* during development in different tissues and organs. Data was extracted from FlyAtlas, which reveals expression data during 30 stages of development (top panel) for 8 and 17 distinct larval and adult tissues respectively (bootom panel). The dataset comprises Affymetrix Dros2 expression arrays (representing 18770 transcripts[282]), with 4 replicates per tissue. **(**http://flybase.org/reports/FBgn0030345.html**)**

CG1847 protein has the same features as human AIP: an N-terminal FKBP-type domain and C-terminal tetratricopeptide repeats. Indeed, a three dimensional theoretical model of CG1847 generated by my collaborator Chris Prodromou revealed a protein structure that closely resembles the published AIP protein structures (Figure 25).

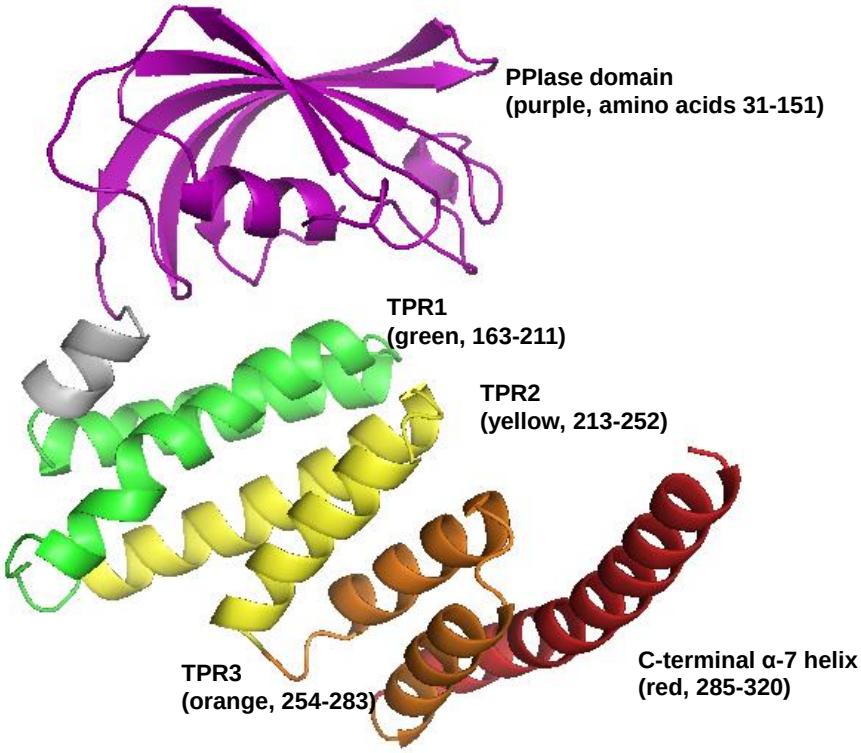

Figure 25: Three dimensional theoretical model of *CG1847*. The *Drosophila* protein has a similar predicted structure to its human orthologue: it has all three pairs of conserved anti-parallel alpha-helices (tetratricopeptide domains, TPR) and the final extended alpha-helix, alpha-7. This model was based on the crystal structure of human AIP[142,145,234] and was prepared by Chris Prodromou (Sussex University).

3.4.2 CG1847 is essential for Drosophila development

Essential genes are defined as those necessary for growth and survival under a given condition. To investigate the function of CG1847 in *Drosophila*, I examined the consequences of disrupting *CG1847* gene expression in developing flies using RNA interference via the GAL4/UAS-system[238]. To achieve this, I used 4 fly stocks that carry different UAS-RNAi constructs targeting CG1847, which were expressed during fly development under the control of a ubiquitous (*Act-Gal4*) or neuron-specific (*elav-Gal4*) promoter. Here, I present and discuss the results obtained with 3 of these stocks as they were the ones that yielded more promising phenotypes.

When UAS-*CG1847-RNAi-R2* males were crossed with *Act-GAL4* females, as actin is ubiquitously, and rather abundant, expressed, this cross resulted in strong knockdown of *CG1847* expression in all tissues. No viable adult offspring were observed, suggesting that complete AIP-knockdown is not compatible with viability. In contrast, UAS-*CG1847-RNAi*-R1 produced a few viable offspring. However, RT-PCR analysis revealed that the knockdown of *CG1847* expression with RNAi-R1 was only partial (Figure 26 A). As RNAi constructs can occasionally generate phenotypes due to off target effects, I used new RNAi stocks with a different non-overlapping sequence of

CG1847 (section 2.2.1). Using the new RNAi stocks, I obtained similar results to those obtained with UAS-*CG1847-RNAi-T2* line, that is, expression of *CG1847* RNAi under the control of *Act-Gal4* resulted in 100% lethality. As previous ubiquitous knockdown resulted in lethality, I crossed UAS-*CG1847-RNAi-T2* males with females carrying a tissue specific driver. Elav-Gal4 was used to silence *CG1847* expression only in nervous cells. We choose this driver for two different reasons. First, based on the microarray data *CG1847* has a higher level of expression in the brain (Figure 24). Second, most of the knockdowns in the fruit fly nervous tissue do not have a lethal outcome[284].

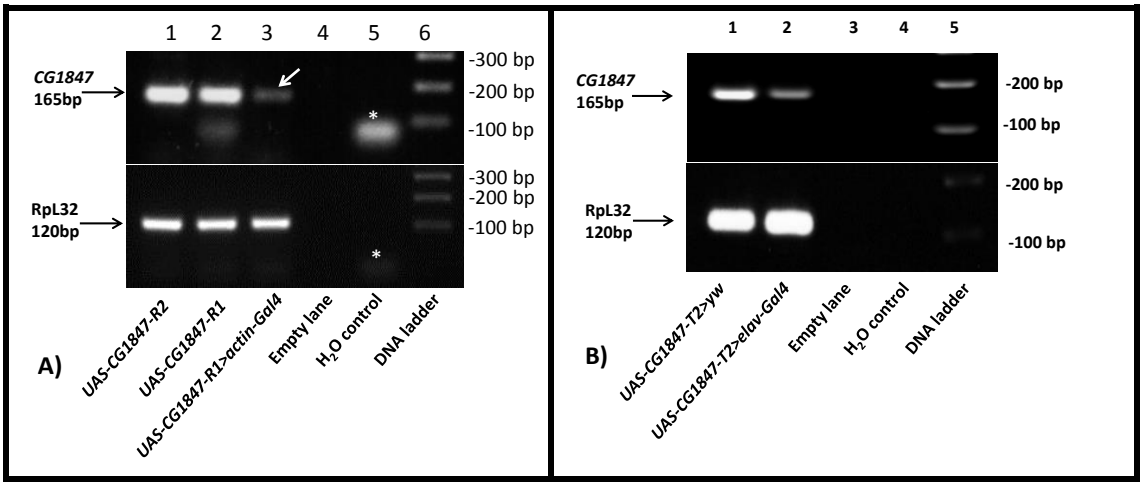

Figure 26: CG1847 RNAi knockdown. RT-PCR was performed using RNA extracted from adult fly heads. Upper panels *CG1847* amplicons, 165 bp product; Ribosomal protein 32 (RpL32) expression was used as housekeeping gene control (lower panels). A) KD with UAS-*CG1847-RNAi-R1.* Lanes 1 and 2 – UAS-RNAi stocks R1 and R2 were used as a control and tested for leak expression of the IR constructs. Lane 3 *CG1847-R1>Act-Gal4* flies showing knockdown of CG1847 expression, see arrow. Stars indicate primer dimers B) KD with UAS-*CG1847-RNAi-T2.* Lane 1 *CG1847-T2* flies were crossed with *y w* flies, in order to be used as a control and to prove that the RNAi construct is not expressed in absence of a Gal4 driver. In Lane 2 *CG1847-T2> elav-Gal4* flies show knockdown of CG1847 expression.

RNAi-mediated depletion of *CG1847* using UAS-*CG1847-R1* did not cause lethality with either *Act-Gal4* or *elav-Gal4*, with some flies reaching the adult stage. This enabled the quantification of the efficiency of *CG1847* knockdown using RT-PCR (Figure 27). For UAS-*CG1847-R1* and *R2* the average amount of CG1847 transcript in the parental lines was compared to the average amount of CG1847 transcript in the *Act-GAL4*-driven UAS-*RNAi* lines, and an expression ratio was calculated. For UAS-*CG1847-T2* the average amount of CG1847 transcript in the control crosses was compared to the average amount of CG1847 transcript in the *elav*-GAL4-driven UAS-RNAi lines. For quantification the above images were uploaded into the ImageJ software and the intensity of the bands was evaluated. The average of the target genes was normalized to the average of the control gene RpL32 and expressed in intensity units.

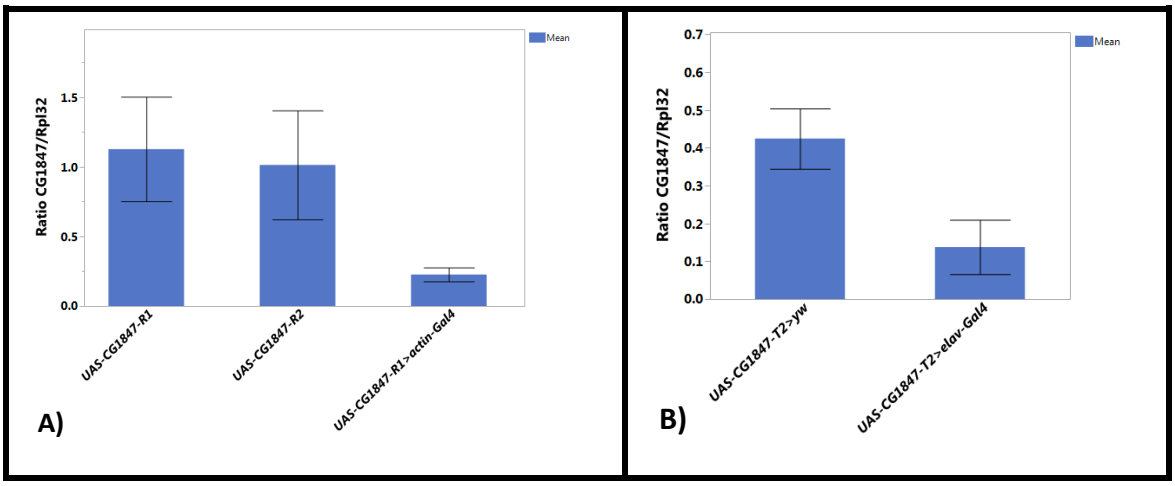

Figure 27: Quantification of *CG1847* RNAi knockdown efficiency. Relative *CG1847* expression was quantified by semiquantitative RT-PCR. RT-PCR was performed on total RNA isolated from adult fly heads. A) Quantification of *CG1847* expression in parental lines and UAS-*CG1847-R1 Act-Gal4*-driven knockdown. B) Quantification of *CG1847* expression in control flies and UAS-*CG1847-T2 elav*-Gal4 driven knockdown. Shown are the averages of the ratio of the expression of *CG1847* and *RpL32* from two independent experiments. Error bars represent SE.

In summary, a very efficient RNAi-mediated depletion of *CG1847* was obtained using all RNAi lines. The average levels of *CG1847* was less than 13.3% of control for UAS-*RNAi-R1* expressed with a strong driver as *Act-Gal4.* With e*lav-Gal4*, even though the expression of UAS-*RNAi-T2* is restricted to nervous tissue, resulted in reducing the CG1847 expression to less than 33.3% of controls.

Next, taking advantage of the temperature sensitivity of the GAL4 system[285], it was assessed whether increasing the potency of UAS-*CG1847-RNAi-R1* might result in a lethal phenotype. Higher temperature increases Gal4 activity and when animals were grown continuously at 29°C no adult UAS-*CG1847RNAi-R1>Act-Gal4* flies were recovered.

3.4.3 *CG1847*$^{exon\ 1\_3}$, a null *CG1847* allele, confirmed the essentiality of this gene

To confirm the results obtained with RNAi-mediated depletion of *CG1847*, I performed an imprecise P-element excision screen[267] to generate a "classical" mutation in *CG1847.* For this, I used an existing fly stock carrying a P-element insertion in the 5' UTR of *CG1847* (P-CG1847^{G1839}). The P-element was mobilized using standard genetic techniques (detailed protocol in section 2.2.2). In the second generation of the excision screen, 200 single crosses were set up, each with a single virgin female and 3 FM6 males. In the following generation, I screened for virgins with white eyes, that is, those who lost the P-element. I identified 49 virgin females with white eyes, which were used to generate independent putative mutant stocks, by backcrossing to FM6

males. Out of these 49 lines, 18 stocks were lethal as no non-FM6 males were recovered in the offspring. Table 15 summarises the different steps of the excision screen. Based on the PCR screening described below, the majority of the stocks carrying putative mutant alleles were discarded as the deletions were too big, affecting either the upstream or the downstream genes in addition to *CG1847*.

Stocks	Number
Total crosses	200
White eye virgin females identified	49
Lethal alleles (putative mutants)	18
Stocks discarded (very large deletions)	17

Table 15: Summary of P-element excision screen.

For a precise characterisation of these putative mutant alleles and to identify the extent of the deletion, genomic DNA from heterozygous females was analysed by PCR (primer sequences Appendix 5). Flanking primers (Dm_EP_EL_F and Dm_EP_EL_R) and PCR conditions were designed to amplify a fragment resulting from a deletion in the *CG1847* locus, but not a P-element-containing fragment (more than 8000bp in our case). These primers were also designed at the border with the upstream and downstream neighbouring genes in order to detect deletions that will be too large. Identification of the *CG1847* control stock where the P-element has undergone precise excision (revertant) was based on the detection of a larger amplicon when using the second pair of primers (Dm_EP_F and Dm_EP_R). See Figure 28 for a schematic representation of primer annealing:

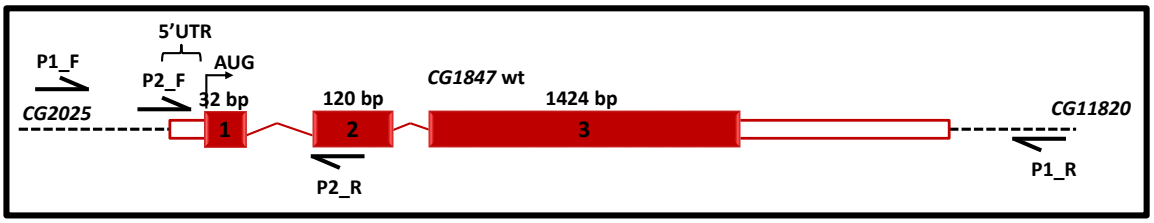

Figure 28: Schematic representation for primer annealing. For detecting the genomic deletion were used 2 pairs of primers. Pair P1 (Dm_EP_EL_F and Dm_EP_EL_R) was design to bind into the neighbouring genes and to amplify a 2487 bp amplicon. Pair P2 (Dm_EP_F and Dm_EP_R) amplifies a 298 bp fragment, inside *CG1847* gene.

As a control to verify that the PCR conditions were correct, I used *y w* flies and homozygous *FM6* females that eclosed in these stocks. After validating the efficiency of the PCR approach in the

controls, I screened the putative mutant stocks from the excision screen to identify a *CG1847* mutant stock (representative PCR images in Figure 29).

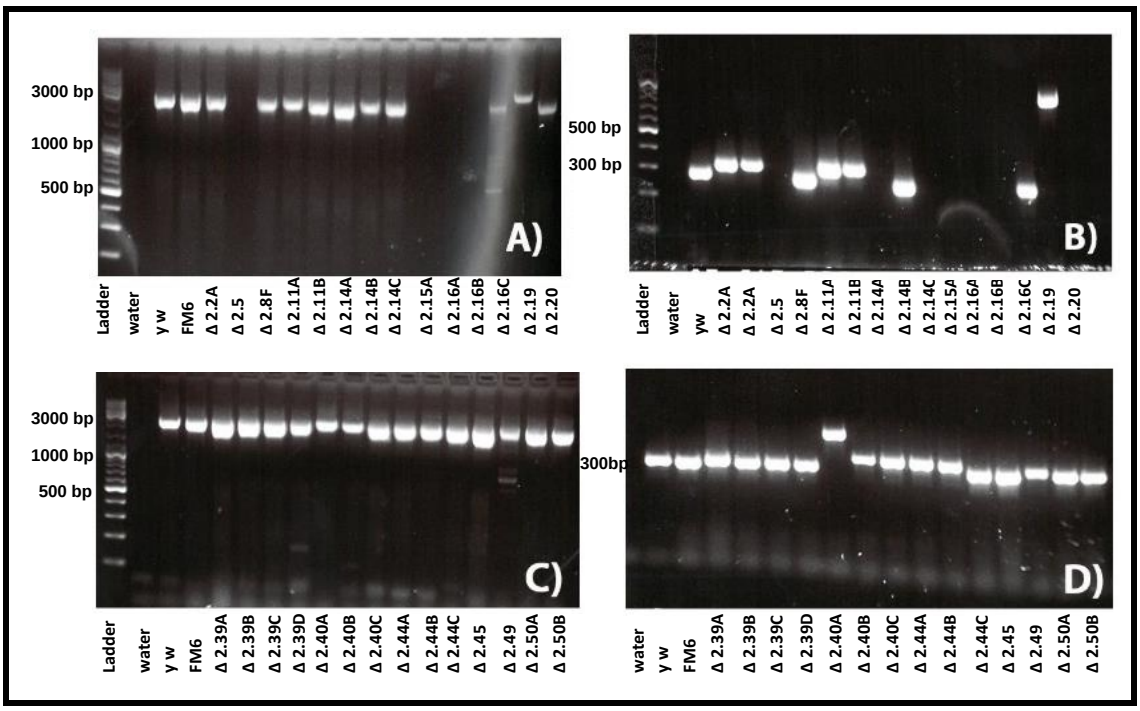

Figure 29: Agarose gel for selected number of *CG1847* mutant alleles. Identification of alleles containing a loss of genomic DNA. Each sample was amplified with 2 different pairs of primers. (a) and (c) The first pair (Dm_EP_EL_F and Dm_EP_EL_R) was designed to amplify a 2478 bp. Different alleles show different size amplicons from the control y w. (b) and (d). The second pair (Dm_EP_F and Dm_EP_R) was designed to amplify a 298 bp.

The first pair of primers (Dm_EP_EL_F and Dm_EP_EL_R) was designed to amplify a 2478 bp amplicon in wt flies (*y w*) or flies carrying the FM6 balancer (Figure 29 a) and c)). In the revertant the PCR analysis also resulted in a 2478 bp amplicon, as the P-element footprint does not produce a significant amplicon size difference. In flies carrying deletions, the PCR analysis revealed a band smaller than 2478bp or no amplicon. Failure to detect an amplicon was likely due to a large deletion that prevented primer annealing and DNA amplification. In this case, the stock was discarded as the deletion may affect not only *CG1847* but also surrounding genes (Figure 28). The second primer of pair Dm_EP_F and Dm_EP_R was design to amplify a 298 bp amplicon in *y w* flies or flies carrying the balancer chromosome FM6 (Figure 29 b and d). Therefore, I expected that "non mutant" alleles would produce a similar 298 bp DNA amplicon (or slightly larger in case -the footprint of the P-element was still present).

One excision Δ125A (named *CG1847*$^{exon1\_3}$ thereafter) exhibited promising results that led to its selection as a loss-of-function allele. This allele carries a deletion of 1512bp in the *CG1847* locus that encompasses exon 1, 2 and part of exon 3 (Figure 30).

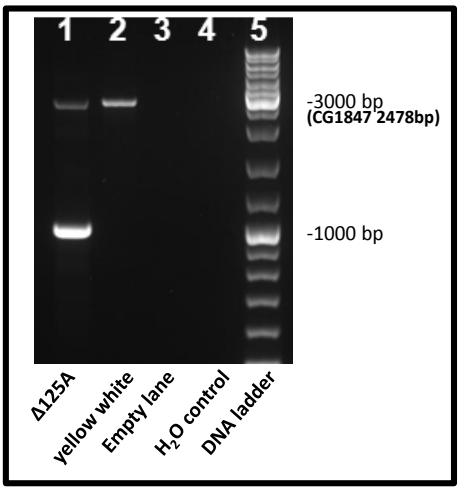

Figure 30: Creation of *CG1847* knockout fly line. *CG1847*$^{exon1\_3}$ is a loss-of-function allele, having deleted 1512bp containing exon 1, 2 and part of exon 3. PCR from adult fruitfly DNA using primers designed for a 2478bp product. Lane 1: sample 125A extracted only from females (heterozygous) as there were no viable males. In this PCR the normal size CG1847 was less amplified probably due to the polymerase primarily amplifying the smaller mutant fragment (approx. 1000p) rather than the larger (2478bp) wild-type fragment. Lane 2: yellow white fly was used as controls; Lane 3: empty lane, Lane 4: water control, Lane 5: size marker.

To confirm the deletion, we sequenced the DNA amplicon obtained in *CG1847*$^{exon1\_3}$ flies using the diagnostic primers (Figure 28). The Sanger sequencing identified the excision of 1512bp (exon 1, 2 and exon 3) creating the *CG1847*$^{exon1\_3}$ mutant. A schematic representation of the mutant is shown in Figure 31.

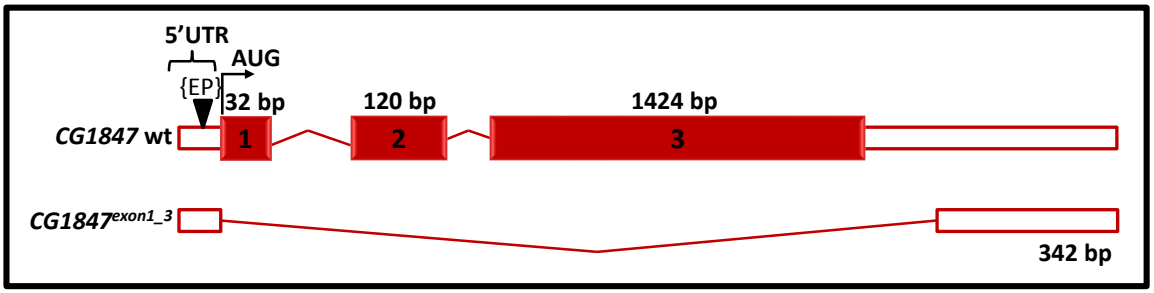

Figure 31: Schematic diagram of the CG1847 locus in wt and *CG1847*$^{exon1\_3}$ mutants. Intron and exon boundaries are based on the sequencing results of *CG1847*$^{exon1\_3}$ mutants.

This mutant allele is homozygous and hemizygous lethal, confirming the results obtained with *CG1847* RNAi and reinforcing the notion that *CG1847* is an essential gene. The *CG1847*$^{exon1\_3}$ mutant was used in subsequent experiments to analyse *CG1847* function.

In the same mutagenesis experiment, I also generated a control stock via precise excision of the P-element. As this revertant originates from the same genetic background as the mutant, it can therefore be used as a control (see below for details). For initial characterisation of stocks from the P-element excision screen, genomic DNA from heterozygous flies was analysed by PCR technique (primer sequences Appendix 5). For detecting the control stock the flanking primers were designed in such a way that they will amplify a 298 bp amplicon around the area of P-element insertion. Identification of *CG1847* revertant stocks was based on detection of a higher

amplicon than the 2 controls (y w stock and FM6/FM6 females resulting from the same cross). Figure 32 is a representative agarose image gel for detecting the revertant stock.

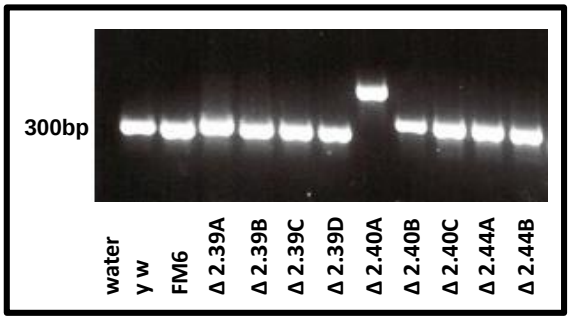

Figure 32: Agarose gel of PCR results - selected stocks resulted from the P-element excison screen. Identification of alleles which have lost the P-element via precise excision. The pair of primers (Dm_EP_F and Dm_EP_R) was designed to amplify a 298 bp. Some of the stocks show different size amplicons from the 2 different controls (y w and homozygous FM6 females). Samples 2.39A-2.39D shows an amplicon only slightly longer than the controls.

The primer pair Dm_EP_F and Dm_EP_R was designed to amplify a 298 bp amplicon in wt flies (*y w*) or flies carrying the FM6 balancer chromosome. In the PCR screening, a "revertant control" was expected to display a slightly larger amplicon than the controls, due to the fact that a few bp of the P-element are usually not excised when the transposon is mobilised and is excised correctly (the transposon footprint). As the transposon footprint indicates that the P-element was correctly excised, it also indicated that these alleles did not contain a genomic deletion of the *CG1847* gene.

Based on the PCR screening, four excision stocks 2.39A-2.39D were selected as precise excision controls for RNA-seq. As all 4 stocks showed similar footprint sizes, I confirmed the sequence of one of them (stock 2.39A) by Sanger sequencing. DNA was extracted from males which are hemizygous for P-element excision (Figure 33).

Figure 33. Comparation between the *CG1847*wt sequence and the chromatograms of *CG1847*$^{2.39A}$. The *CG1847*$^{2.39A}$ stock presents a 15 bp expected insertion (footprint).

The females lacking FM6 balancer chromosome (homozygous for P-element precise excision) are also viable. Although the precise excision control lacks the transposon, it is possible that the P-element footprint sequence may disrupt *CG1847* expression, if it is located in a conserved genomic area that has regulatory functions. To determine whether this was a likely possibility, I performed a BLAST alignment of the *CG1847* region in the vicinity of the P-element footprint, using as a reference the *Drosophila melanogaster* UCSC Genome Browser (Figure 34).

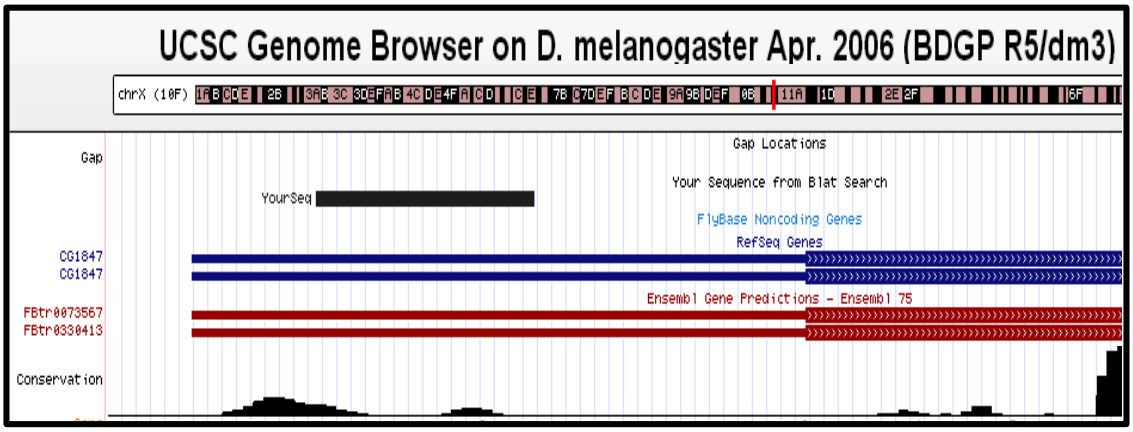

Figure 34: BLAST alignment. The DNA sequences of the region immediately surrounding the foot print of the P-element was blasted using UCSC Genome Browser to determine if the area has a high degree of conservation.

Despite the fact that the original P-element was inserted in the 5'UTR of *CG1847*, the BLAST analysis revealed that the remaining P-element footprint is not in a conserved region and, therefore, it is unlikely that it affects *CG1847* expression levels. However, for *Drosophila melanogaster* it was not possible to evaluate if in the proximity of this area are any consensus binding sites for transcription factors (the online databases are not as complete as human databases). Consequently I was not able to determine if the footprint of the P-element might disrupt such a binding site for a transcription factor, with further consequences on the expression of CG1847.

3.4.4 *CG1847*$^{exon1\_3}$ mutant display total loss of *CG1847* expression

To confirm that *CG1847*$^{exon1\_3}$ is a *CG1847* null allele, *CG1847* gene expression was analysed by RT-PCR performed using RNA extracted from 48 h hemizygous mutant male larvae. As controls RNA samples extracted from either 48 h *y w* larvae or from *y w* male adults were used (Figure 35). However, none of the controls were the appropriate ones as the RNA extracted from 48 h larvae started from a mix of males and females collection, while the adults are known to have different levels of CG1847 expression.

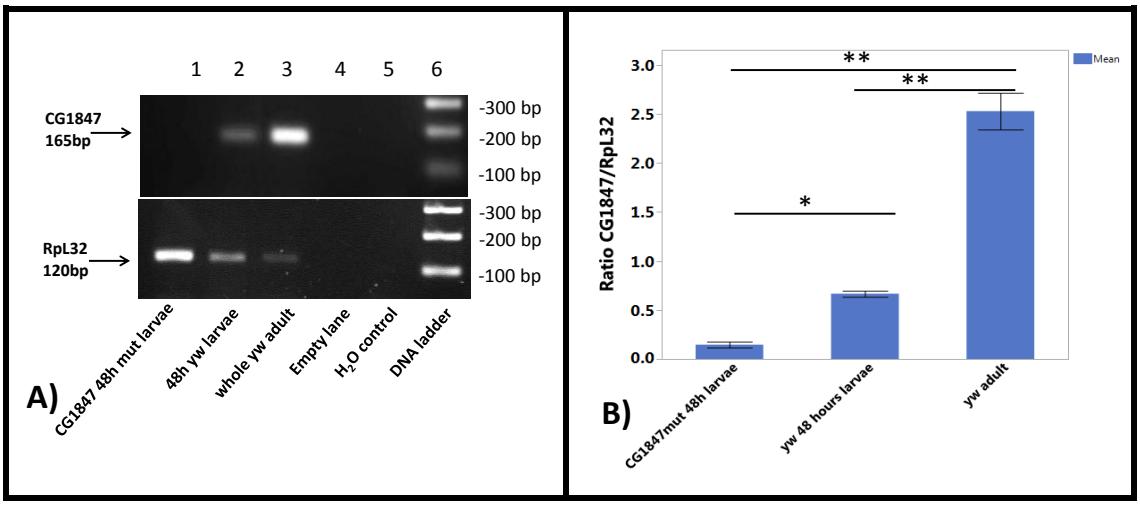

Figure 35: *CG1847* expression in mutant male larvae. A) RT-PCR amplification products from *CG1847* 48h mutant male larvae and controls (48 h *y w* larvae and *y w* adult). Upper panel: *CG1847* RT-PCR (165 bp product); Lower panel *RpL32* RT-PCR (housekeeping gene, 120 bp product). B) *CG1847* gene expression in mutant larvae was quantified by semiquantitative RT-PCR. RT-PCR was performed using specific primers on total RNA isolated from mutant male larvae and controls. Data shown as average of mean target gene normalized to the control gene RpL32 of three independent experiments. Error bars represent SE. Asterisks indicate statistical significance as determined by Student's t-test (*P<0.05, **P<0.01).

CG1847 expression was not detected in male larvae hemizygous mutants indicating that this is a null allele. This was associated with 100% lethality. Interestingly, the levels of expression of the housekeeping gene were increased in mutant larvae revealing possible increased proliferation in the mutant male larvae. For more details see Discussion section.

3.4.5 Lethality occurs in late larval stage of development

To determine at which developmental stage lethality occurs in *CG1847* mutants, I balanced the *CG1847* mutant allele over a balancer chromosome containing a fluorescent marker (*FM7c, Dfd::YFP* (Table 4)).

Embryos were collected for 4 h (details in section 2.2.10). I counted the larvae produced at various time points during development. Two separate collections were performed for each stage of development. In this stock four different genotypes are possible in each generation, due to allele segregation and combinations of X and Y chromosomes. Consequently, each genotype should count for 25%. However, homozygous *FM7c* females are not viable, so each genotype should count for around 33.3% at any time point during development. However, the homozygous *FM7c dfd::YFP* females cannot be differentiated from the heterozygous *FM7c dfd::YFP* larvae and, consequently, the moment of their lethality was not able to be precisely established.

	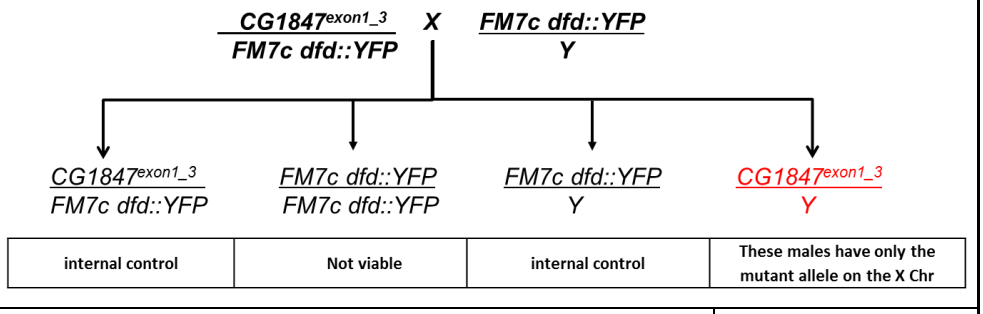
24 h	194 (66.9%) / 96 (33.1%)
48 h	65 (72.3%) / 25 (27.7%)
72 h	471 (89.1%) / **58 (10.9%)**
96 h	218 (100%) / **0 (0%)**

Table 16: Determination of lethality stage during development. Larvae were counted at 24, 48, 72 and 96 h AEL, and selected based on the presence or absence of fluorescence (dfd-YFP). Numbers represent total number of larvae counted in 2 separate experiments. Red represents significantly reduced numbers than expected.

From a total of 290 larvae counted at 24 h AEL, there were 194 controls (*FM7c, Dfd::YFP)* and 96 hemizygous mutant males (*CG1847^{exon1_3}/Y*). This was in accordance with Mendelian inheritance as the mutant genotype should count for around 33% out of the total number of larvae (Figure 36). On the other hand, from a total of 529 larvae counted at 72 h AEL there were 471 controls and just 58 hemizygous mutant male larvae (just 10% of the total viable larvae –Table 16). A chi-square analysis indicated that significant mortality of *CG1847^{exon1_3}/Y* males was associated as the ratio of the genotypes obtained (8:1) was significantly different from the expected (2:1) ratio based on Mendelian segregation. The total number of larvae counted was 1127. The observation that our genotyping at 48 h AEL indicated a control /hemizygous mutant frequency of 2.6:1, indicates that the lethality of mutant *Drosophila* males occurs after this developmental point and before reaching 72 h AEL.

Together, these data indicate that *CG1847* is an essential gene in fruit flies, as total deletion of *CG1847* leads to lethality during the 2nd instar stage of larval development. This stage begins immediately after the first larval molt and takes approximately 24 hours at 25°C. Larvae are very small, but they are active and very mobile in the food. The salivary glands extend to the first abdominal segment and the larvae are actively feeding with the food medium. Unfortunately the available information regarding the development of *Drosophila* during the second instar stage (L2) is very scarce. Detailed description of fruitfly normal development was presented by Hartenstein in 1993, but his investigations were focused mainly on embryonal, 3rd instar (L3) and

pupal stages[286]. Consequently it is very difficult to speculate which might be the tissues and organs affected by loss of CG1847.

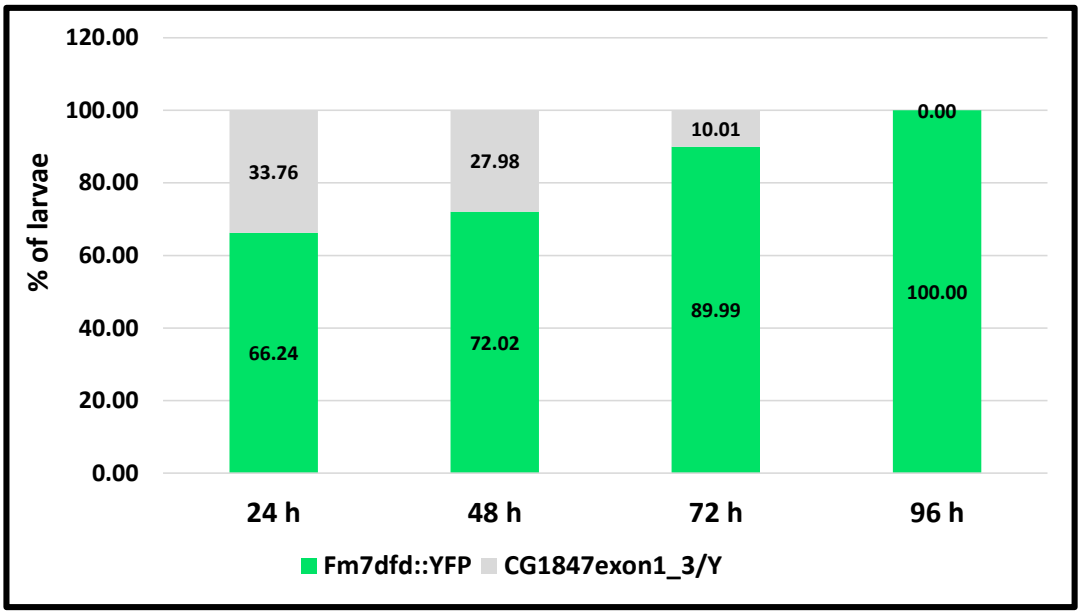

Figure 36: Survival rate between 24 and 96 h AEL. The numbers represent averages of percentages of 2 separate countings for mutant (grey) versus control fluorescent (green) larvae. At 24 h AEL the percentage of mutant larvae reaches the expected value of 33%, but is dramatically decreased by 72 h AEL. No *CG1847^{exon1_3}/Y* larvae could be found by 96 h AEL in 2 separate countings of a total of 218 larvae

3.4.6 Abnormal larval development in *CG1847^{exon1_3} mutants*

To further analyse the impact of *CG1847* loss of function, we focused on the early larval stages, before lethality occurs. I compared the phenotype of mutant male larvae with the offspring that expressed the YFP fluorescent marker present in the same 4 h AEL eggs collections (Figure 37).

Apart from the obvious stop in larval growth, I was unable to detect any obvious phenotype that could predict the cause of lethality, or might be indicative of the organs and tissues affected by loss of function of CG1847 (Figure 37A).

I first investigated these larvae ability to eat. However, evaluation of food consumption in the fruit fly is quite challenging as, in contrast to mammals, food ingestion cannot be properly quantified. To overcome this problem, I evaluated the food intake of mutant larvae by feeding them on gelled media marked with a visible dye[287,288]. Surprisingly, I was able to notice the presence of food even in the gut of the 72 h mutant larvae.

I also attempted larvae dissection and Trypan blue staining which did not reveal any necrotic tissues.

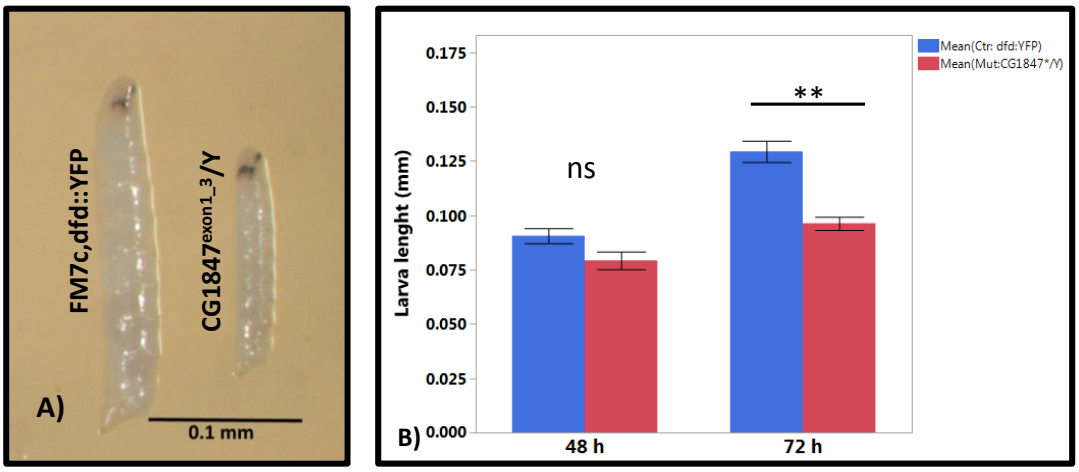

Figure 37: Delayed larval development in *CG1847*$^{exon1\_3}$ mutants. The mutant male larvae show retarded growth. (A) Third instar larvae (72 h AEL) orientated with the anterior end to the top. Hemizygous 72 h *CG1847*$^{exon1\_3}$ mutant larvae are shorter and slimmer than the control dfdYFP larvae. (B) Larval length analysis at different developmental stages. Error bars represent SE. Asterisks indicate statistical significance as determined by Student's t-test (**P<0.01, ns=not significant).

Apart from being shorter and slimmer than the control dfdYFP larvae, the 72 h *CG1847*$^{exon1\_3}$ mutant larvae did not displayed any other significant phenotypic differences. However, the 72 h mutant males appeared to be less mobile than the controls. The length difference between hemizygous *CG1847*$^{exon1\_3}$ mutant larvae and control dfd::YFP is detectable at the beginning of the second instar larva (L2), although without reaching a statistical significance. The halt in development is significantly more obvious by 72 h AEL.

Further more detailed investigations are required to determine more specific phenotypes which could be indicative of the cause of death.

3.4.7 The lethality of *CG1847* mutants can be rescued by expression of *CG1847*^{wt} under the control of its own promoter

To confirm that the lethality of *CG1847*$^{exon1\_3}$ is due to the absence of *CG1847* and not due to the P-element imprecise excision affecting additional genes, I tested whether the *CG1847*$^{exon1\_3}$ mutant could be rescued by reintroducing the normal allele on the mutant background. A genomic rescue construct containing the entire *CG1847* gene with its own promoter and regulatory elements was generated (section 2.4.3) and injected into *Drosophila* embryos to obtain transgenic lines. Transgenic flies were subsequently crossed to the *CG1847*$^{exon1\_3}$ mutant and examined for their capability to rescue the lethality. The degree of rescue was analysed by counting all the males and comparing the percentages of each viable male genotype in the second generation (Figure 38).

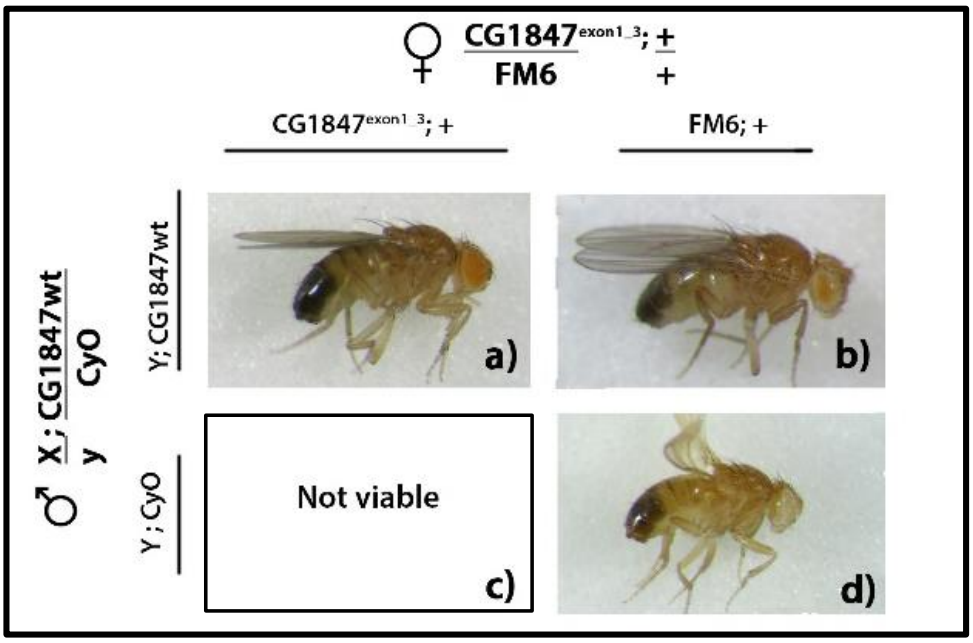

Figure 38: Results of rescue experiments. Images of male genotypes resulted from rescued cross with full-lenght wild-type CG847 gene. The segregation of the alleles and the possible combinations are shown in the lateral panels. a) Rescued males. They inherited only the mutant allele on the X chromosome (from their mothers), but are rescued by CG1847 normal allele on the second chromosome. Males c) are not viable because they inherited the other second chromosome from father (CyO), which can not compensate for the deletion inherited from mother. This genotype (C) also acts as an internal negative control.

Males *CG1847*$^{exon1\_3}$/Y; GC1847/+ (Figure 38A) are the rescued genotype. On the X chromosome they have the mutant allele, inherited from the mother. They are rescued by the wt *CG1847* gene reintroduced on the second chromosome and inherited from the father. The genotype *CG1847*$^{exon1\_3}$/Y; +/*CyO* is not viable as the balancer chromosome (CyO) is inherited from the father. Consequently, they have no wt copy of *CG1847***.**

NOTE: the genotype *CG1847*$^{exon1\_3}$/Y; +/*CyO* (c) should not be viable and should act as an internal negative control. However, I did find a reduced number of males that were phenotypically similar. In reality these males were the result of a non-disjunction phenomenon and they carried a normal *CG1847* allele on the X chromosome, inherited from paternal line. For more details on the non-disjunction phenomenon see Discussion in Chapter 5. As among the rescued genotype there were also a few males which were the result of non-disjunction, in all the crosses we checked if the differences between the rescued males and the internal negative control (*CG1847*$^{exon1\_3}$/Y; +/*CyO*) reached a significant statistical difference**.**

The *CG1847* rescue construct allowed rescue of the lethal phenotype. Mutant males expressing this construct developed normally and exhibited normal behaviour. To evaluate the rescue

capacity of the *CG1847* construct we used males from 2 transgenic stocks (CG1847-1M and CG1847-3M) and for each of these stocks the rescue crosses were performed in triplicate. The percentages of each of the above genotypes are plotted below (Figure 39).

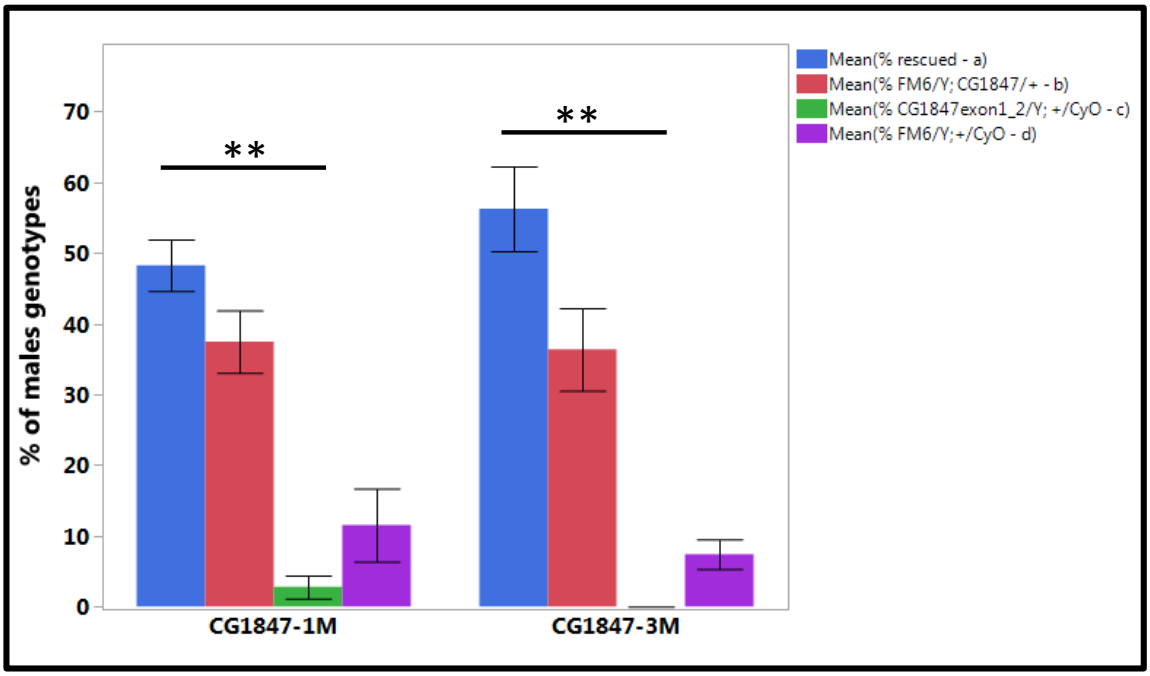

Figure 39: Rescue experiment of the CG1847 lethal phenotype with a wt CG1847 construct. The associated letters (a-d) correspond to the phenotypes depicted in Figure 38. Analysis of the statistical data reveals that the *CG1847* rescue construct is capable of rescuing significantly the lethality of the *CG1847* mutant, based upon the relative normal percentage of rescued males. CG1847-1M and CG1847-3M – 2 of the transgenic stocks carring the rescue construct. N=4 experiments. Error bars represent the standard error of the mean. Asterisks indicate statistical significance as determined by Student's t-test (**P<0.01).

Interestingly, the CG1847wt construct rescued the lethality of the mutant males very efficiently as, in both experiments, this genotype is the most represented genotype amongst all offspring. Based on the expected Mendelian distribution, the percentage of each male genotype should be around 33% of the total number of viable adult males (only 3 genotypes are viable). The percentage of rescued males varied between 33.5% and 65.9% in all 4 experiments (an average of 52.2%). When males from transgenic line CG1847-1M were used, I detected a small percentage of flies that appeared to have the phenotype of the lethal combination (*CG1847*$^{exon1\_3}$/Y; *+/CyO).* These were in fact the result of chromosomal non-disjunction and were found in a significantly lower percentage than the percentage of rescued genotype.

To further validate these experiments, I also counted the eclosed females from each experiment. Four female genotypes were expected according to Mendelian inheritance (combination of 2 X (X and FM6) chromosomes and 2 second (wt *CG1847* gene and CyO) chromosomes). The

percentage of each genotype was used to evaluate if the female genotypes are in normal distribution, as this would validate the results obtained with the males (Figure 40).

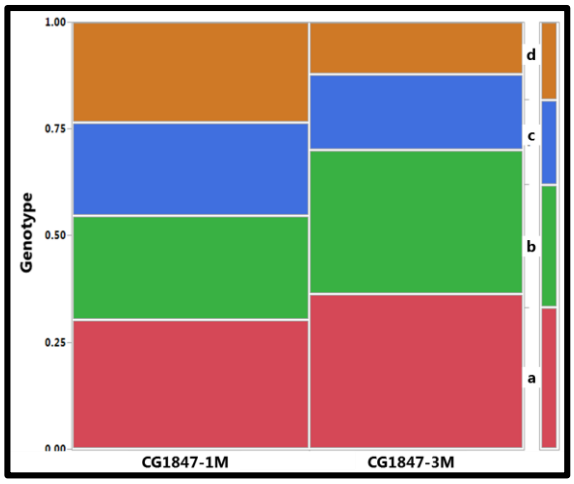

Figure 40: Chi squared contingency test for the distribution of female's genotypes. In the rescue experiment with the CG1847wt construct 4 different female genotypes are possible (as result of combinations of X and second chromosomes). For plotting purposes each female genotype received a code (a - d) and a colour: **a** ($CG1847^{exon1\_3}$/X; CG1847/+) – red; **b** (FM6/X;CG1847/ +)– green; **c** ($CG1847^{exon1\_3}$/X; CyO/+) – blue; **d** (FM6/X; CyO/ +)– brown. There is no significant difference in the distribution of female genotypes between the F1 generation of CG1847wt rescued experiment with the 2 transgenic stocks (P = 0.07).

Figure 40 depicts a chi square analysis of the 4 possible female genotypes in the F1 generation of the rescue cross. For biological replicates were used males from two different transgenic stocks, carrying the same CG1847 construct (CG1847-1M and CG1847-3M) generated by BestGene. This analysis reveals that the 4 female genotypes do not differ significantly between the 2 experiments with the 2 transgenic stocks, as expected. These results validate the rescue and the previous data regarding the rescue of male lethality providing extra proof that this experiment was successful.

3.4.8 Tissue specific hAIPwt expression cannot compensate for *CG1847* deletion

The Gal4-UAS system has evolved into a widely used and valuable tool for the temporal and spatial control of gene expression in *Drosophila*[285]. As mentioned above, I was unable to detect any overt phenotype that might account for the lethality of *CG1847* mutants, or which might indicate the organs and tissues affected by *CG1847* loss (section 3.3.5). Therefore, I decided to use the Gal4-UAS system to perform rescue experiments in a tissue-specific manner that is, expressing *CG1847* under the control of tissue-specific promoters.

Among the *GAL4* drivers available in public databases, I chose 10 drivers that allowed me to overexpress human AIP cDNA in various tissues and organs. The panel of 10 *GAL4* drivers selected for overexpression of hAIP in the *CG1847* mutant background (Table 4 and Figure 41) has an expression pattern that ranges from ubiquitous to restricted to a specific tissue or cell type such as fat body, haemocytes, insulin secreting cells, muscle, nervous cells, glial cells, gut, malpighian tubules cells, and heart specific drivers.

The rescue experiments were performed using virgin heterozygous *CG1847* mutant females expressing Gal4 under the control of tissue-specific drivers, which were crossed with males carrying a UAS-hAIP^{wt} transgene on the second chromosome (*hAIP^{wt}/CyO*; two different stocks were used: UAS-hAIP^{wt}-1M and UAS-hAIP^{wt}-2M). The tissue-specific promoters were introduced in the heterozygous mutant background either on the second chromosome (*CG1847*^{exon1_3}/FM6; Gal4/CyO) or third chromosome (*CG1847*^{exon1_3}/FM6; Gal4/TM3). All progeny (males and females) were counted and the numbers of rescued males, non-FM6 non-CyO/non-TM3, were determined.

NOTE**:** more details regarding the validation of the rescue experiments with hAIPwt constructs can be found in Chapter 5.

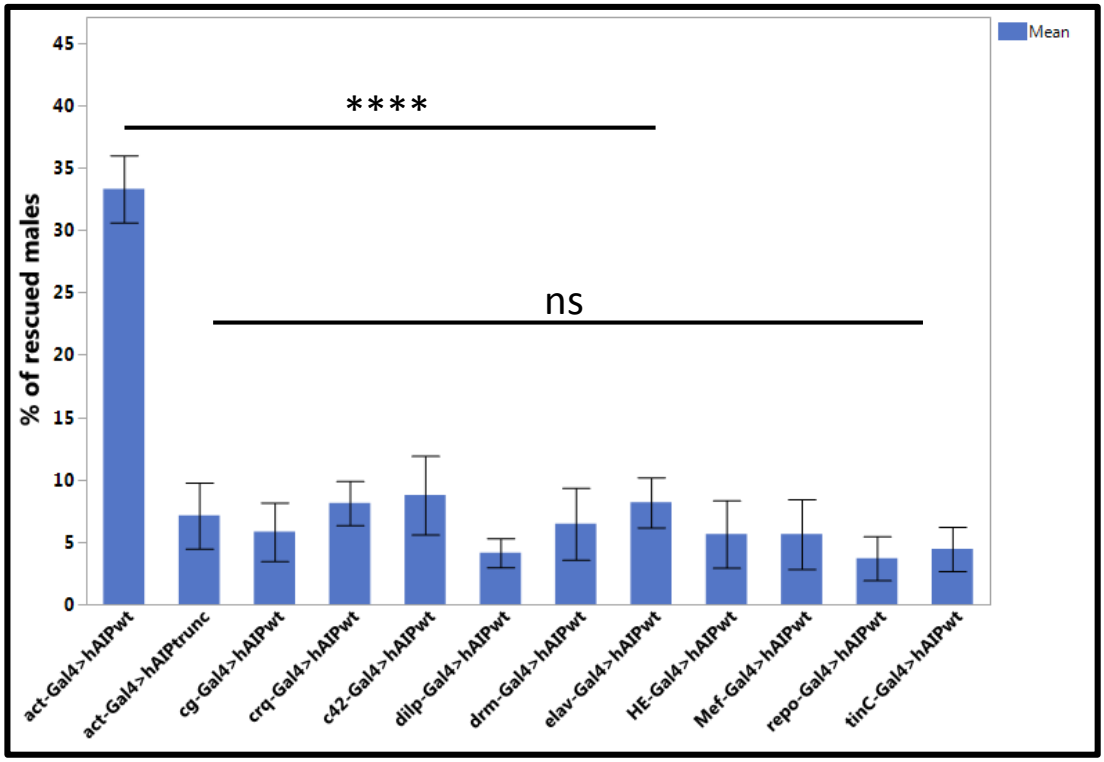

Figure 41: Rescue of the *CG1847* lethal phenotype by ubiquitous or tissue-targeted overexpression of wt hAIP. Analysis of the statistical data reveals that the pUAS-hAIPwt construct rescues the lethality of the *CG1847* mutant only when expressed under the actin promoter. None of the tissue-specific promoters was able to prevent lethality. Asterisks indicate statistical significance as determined by Student's t-test (****P<0.0001).

Expression of hAIP under the control of the ubiquitously expressed *Act-Gal4* driver rescued the decreased viability of the *CG1847*$^{exon1\_3}$ hemizygous males. I obtained *CG1847*$^{exon1\_3}$/Y; hAIPwt/*Act-Gal4*) progeny at very high percentages (33% of total viable males). These results indicate that hAIP is able to compensate for loss of *CG1847* during development, which strongly suggests that, indeed, CG1847 is the fruit fly orthologue of AIP.

However, when tissue-specific drivers were used, the lethality associated with *CG1847* deletion was not rescued. A very small number of flies with the apparent phenotype of the lethal genotype combination were recovered during these experiments but, as discussed above and in Chapter 5, these were in fact the result of the non-disjunction phenomenon).

As seen in Figure 24 during larval stages the CG1847 is already expressed in some of these cells/ tissues (fat body, muscle, nervous cells gut, and malpighian tubules cells); however, the levels of expression are very low. While it seems that at this stage of development CG1847 is not expressed in heart cells, there are no available information regarding the expression levels on glial cells, haemocytes, and insulin secreting cells.

3.4.9 Identification of differentially expressed genes by high-throughput RNA sequencing

The main aim of the RNA-seq analysis was to identify the genes whose expression is affected in *CG1847* mutants, as this might allow a deeper understanding of the AIP involvement during development and the pathogenic mechanism resulting from *AIP* mutations in humans. To investigate the changes in the *Drosophila* transcriptome elicited by loss of *CG1847*, I used the *CG1847* null mutant stock and the control stock, both generated in the same imprecise/precise excision screen. Both stocks were balanced over a balancer chromosome carrying a fluorescent marker (*FM7c*, *dfd:YFP*) and male larvae were selected based on the lack of fluorescence (section 2.2.10).

Total RNA was extracted from pooled collections of male larvae (either mutant or control) at 48 h AEL. The RNA quality of the RNA samples was assessed using an Agilent 2100 bioanalyzer, which performs a reliable RNA integrity test based on the RIN values (RNA Integrity Number) (Figure 42).

The Agilent RNA bioanalyzer software relies on the human rRNA Ratio [28S / 18S] to generate RIN values. However, since *Drosophila* possesses different rRNA values, I was unable to calculate the RIN values for the RNA samples used in the RNA-seq approach. Consequently, I evaluated the quality of the RNA samples based on their gel migration pattern. Four biological replicates of control larvae and four biological replicates of *CG1847* mutants with highest quality were submitted for comprehensive RNA-seq in order to achieve a better understanding of the mechanism(s) by which loss of AIP promotes organism lethality.

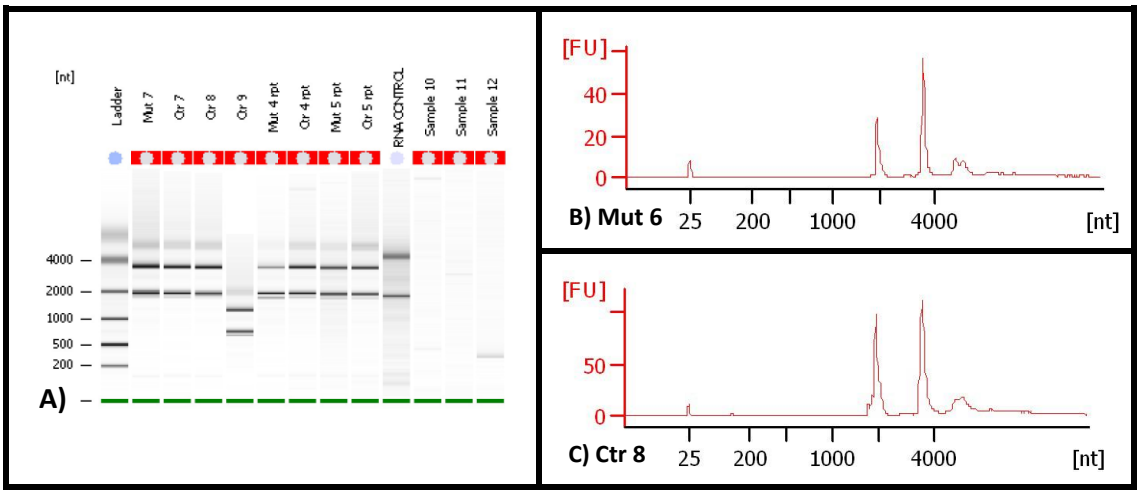

Figure 42: RNA Quality Control. A) Gel Image (from the Agilent 2100 Bioanalyzer) for *Drosophila* larvae total RNA. Lane 1: size ladder. Lanes 2-8 show very high quality RNA for 4 control and 3 mutant samples (high-quality sample appear as two distinct bands corresponding to the 18S and 28S ribosomal RNAs). B) and C) Electropherograms (from the Agilent 2100 Bioanalyzer) for *Drosophila* larvae total RNA for a control and a mutant sample (two well-defined peaks corresponding to the 18S and 28S ribosomal RNAs can be observed, similar to a denaturing agarose gel).

The raw results from the Illumina Hi-Seq were cleaned for removing the adaptors, then the FASTQ files were exported for analysis in Bowtie[257] and TopHat[258] to generate the database of transcripts. FastQC, a key quality control step, was applied to reveal possible problems regarding the quality of RNA-seq raw data (FastQC results in Material and Methods section). For the next steps Cufflinks was run for each sample separately to obtain the fold-change of differentially expressed genes. The final results were assembled into a single merged.gtf file.

An example of merged.gtf file for the top significantly changed transcripts is shown in Table 17. This file contains, for each transcript: test id (XLOC_...), gene name, genomic locus, sample values, logarithmic fold change, p value, q value, and the corresponding FlyBase number. This is only a small part from the original merge.gtf file, with some of the most significantly changed transcripts. The original file is too big to be included in the thesis, as in total 15011 transcripts where detected by the RNA-seq, of which 448 were significantly changed compared to controls (p value <0.05). The table with all these 448 transcripts can be found at Appendix 11.

From the initial list, the top 400 transcripts (cut off fold change ±1; p value <0.05) were chosen for further analysis. As expected, *CG1847* was one of the most significantly downregulated transcripts in *CG1847* KO mutants (fold change –2.6, p value = 0.00005).

test_id	gene	locus	Ctr	Mut	log2(fold_change)	p_value	q_value	oID_genes
XLOC_009430	Hsp70Bbb	3R:8328231-8330822	6.44653	0.0622877	-6.69343	0.0001	0.00842937	FBgn0051354
XLOC_000064	Lsp1beta	2L:898643-901316	21.4665	0.259455	-6.37046	5.00E-05	0.00482773	FBgn0002563
XLOC_011176	αγ-element:CR32865	3R:8295701-8304065	170.163	7.67604	-4.47041	5.00E-05	0.00482773	FBgn0052865
XLOC_005304	IM23	2R:14270208-14270737	15.65	0.990021	-3.98256	5.00E-05	0.00482773	FBgn0034328
XLOC_011881	Hsp68	3R:19880139-19883029	186.155	14.6101	-3.67146	5.00E-05	0.00482773	FBgn0001230
XLOC_003924	CG10073	2R:15268047-15272470	81.7048	6.95278	-3.55476	5.00E-05	0.00482773	FBgn0034440
XLOC_002448	Tepl	2L:15888638-15893811	0.736172	0.0691253	-3.41276	5.00E-05	0.00482773	FBgn0041183
XLOC_003925	CG10081	2R:15273427-15276823	50.3374	4.9139	-3.35669	5.00E-05	0.00482773	FBgn0034441
XLOC_005673	CG3264	2R:18096411-18098247	106.57	12.7437	-3.06395	5.00E-05	0.00482773	FBgn0034712
XLOC_014779	CG34330	X:18962305-18962925	153.186	20.898	-2.87385	5.00E-05	0.00482773	FBgn0085359
XLOC_013279	**CG1847**	**X:11763220-11765201**	**7.15024**	**1.00986**	**-2.82384**	**5.00E-05**	**0.00482773**	**FBgn0030345**
XLOC_008002	Hsp26	3L:9369517-9370475	506.525	77.7273	-2.70414	5.00E-05	0.00482773	FBgn0001225
XLOC_009431	Hsp70Bb	3R:8331514-8334105	51.4734	8.05832	-2.67528	5.00E-05	0.00482773	FBgn0013278
XLOC_009432	Hsp70Bc	3R:8334797-8337183	14.8785	2.36561	-2.65295	5.00E-05	0.00482773	FBgn0013279
XLOC_003992	CG13422	2R:16413831-16414331	8.78264	1.50562	-2.54429	0.00015	0.0112194	FBgn0034511
XLOC_001134	ninaD	2L:18081629-18083608	2.54273	0.43689	-2.54104	5.00E-05	0.00482773	FBgn0002939
XLOC_011175	Hsp70Ba	3R:8291025-8293500	9.46209	1.6779	-2.4955	5.00E-05	0.00482773	FBgn0013277
XLOC_004016	IM14	2R:16757896-16758183	28.9807	5.21566	-2.47417	0.0012	0.04902	FBgn0067905
XLOC_006565	Hsp23	3L:9374981-9375865	380.894	70.5371	-2.43294	5.00E-05	0.00482773	FBgn0001224
XLOC_006566	Hsp27	3L:9377162-9378382	191.486	35.5828	-2.42798	5.00E-05	0.00482773	FBgn0001226
XLOC_014469	mamo	X:13744880-13884528	0.754739	0.140359	-2.42686	5.00E-05	0.00482773	FBgn0030532
XLOC_001238	CG16772	2L:19962678-19963844	3.01629	0.563837	-2.41943	0.0001	0.00842937	FBgn0032835
XLOC_001566	CG15353	2L:2006762-2007193	296.896	57.2366	-2.37495	5.00E-05	0.00482773	FBgn0040718
XLOC_003841	IM2	2R:14274101-14274535	41.7845	8.13862	-2.36011	0.0001	0.00842937	FBgn0025583
XLOC_013810	CG13360	X:678823-684312	123.1	24.1893	-2.34739	5.00E-05	0.00482773	FBgn0025620
XLOC_003839	IM1	2R:14271456-14271883	58.4099	12.4957	-2.22478	0.0001	0.00842937	FBgn0034329
XLOC_007314	CG32444	3L:21630046-21632160	59.4285	12.8388	-2.21065	5.00E-05	0.00482773	FBgn0043783
XLOC_005615	CG30289	2R:17406747-17409482	2.97581	0.667241	-2.157	0.00045	0.0246363	FBgn0050289
XLOC_006292	CG11350	3L:4482808-4484370	1205.94	271.73	-2.14991	5.00E-05	0.00482773	FBgn0035552
XLOC_006562	Hsp22	3L:9365821-9368064	184.278	42.7333	-2.10845	0.0005	0.0268207	FBgn0001223
XLOC_006562	Hsp67Bb	3L:9372722-9374964	184.278	42.7333	-2.10845	0.0005	0.0268207	FBgn0001228
XLOC_007387	CG13239	3L:22853886-22854590	2.45051	10.5349	2.10403	0.00025	0.0161905	FBgn0037197
XLOC_003096	PGRP-SC1b	2R:4600948-4601587	19.1541	82.845	2.11276	5.00E-05	0.00482773	FBgn0033327
XLOC_011912	CG17780	3R:20199175-20210770	5.05373	22.4148	2.14903	5.00E-05	0.00482773	FBgn0039197
XLOC_010154	CG13616	3R:20218631-20219502	5.15558	24.2132	2.23159	5.00E-05	0.00482773	FBgn0039200
XLOC_010088	CG31148	3R:19526359-19528324	7.66618	36.3397	2.24497	5.00E-05	0.00482773	FBgn0051148
XLOC_009691	CG31268	3R:12857538-12859237	11.4359	55.3391	2.27473	5.00E-05	0.00482773	FBgn0051268
XLOC_000087	CG42329	2L:1219317-1229802	0.947499	4.65085	2.2953	5.00E-05	0.00482773	FBgn0259229
XLOC_004842	Damm	2R:7751665-7753908	4.32486	21.288	2.29932	5.00E-05	0.00482773	FBgn0033659
XLOC_009915	TotC	3R:16698709-16699310	1.90684	9.76976	2.35714	5.00E-05	0.00482773	FBgn0044812
XLOC_000287	CG33003	2L:4400952-4403142	3.33855	17.4334	2.38456	5.00E-05	0.00482773	FBgn0053003
XLOC_005578	Cht4	2R:16952884-16954592	26.7254	146.95	2.45904	5.00E-05	0.00482773	FBgn0022700
XLOC_009223	CG8147	3R:5320033-5322675	0.801251	5.04261	2.65384	5.00E-05	0.00482773	FBgn0043791
XLOC_003686	CG15919	2R:12680161-12680585	5.52152	37.6471	2.7694	5.00E-05	0.00482773	FBgn0040743
XLOC_006653	CG12522	3L:11101334-11101832	69.8909	477.68	2.77287	5.00E-05	0.00482773	FBgn0036131
XLOC_008168	CG42397	3L:11966184-11966890	3.11117	22.0361	2.82434	5.00E-05	0.00482773	FBgn0259748
XLOC_001815	CG9021	2L:5903358-5904674	0.258776	1.97669	2.93331	0.0002	0.0137935	FBgn0031747
XLOC_009842	CG17751	3R:15432975-15435044	5.39754	42.5566	2.97901	5.00E-05	0.00482773	FBgn0038717
XLOC_008095	CG32071	3L:11096520-11096973	18.3088	161.103	3.13738	5.00E-05	0.00482773	FBgn0052071
XLOC_008037	CG6749	3L:9724337-9726907	0.629835	5.62862	3.15974	0.00025	0.0161905	FBgn0036040
XLOC_010837	snRNA:7SK	3R:3300274-3300718	1.6693	37.6917	4.49693	5.00E-05	0.00482773	FBgn0065099

Table 17: Example of an excerpt from merged.gtf file The top down (<-2.0) and up-regulated (>2.0) transcripts are represented (separated by a thick line)

The first step in RNAseq data analysis was to search for the significant changes of AIP known partners, such as AhR, a major AIP partner. The *Drosophila* orthologue for human AhR is called spineless (ss) and its levels of expression were surprisingly not significantly changed in the

mutant larvae. On the other hand, among the top downregulated genes, members of the heat shock family of proteins emerged as candidates from the RNA-seq analysis. Interestingly, almost all affected heat shock gene transcripts were downregulated in the *CG1847* mutant compared to the control samples (Figure 43).

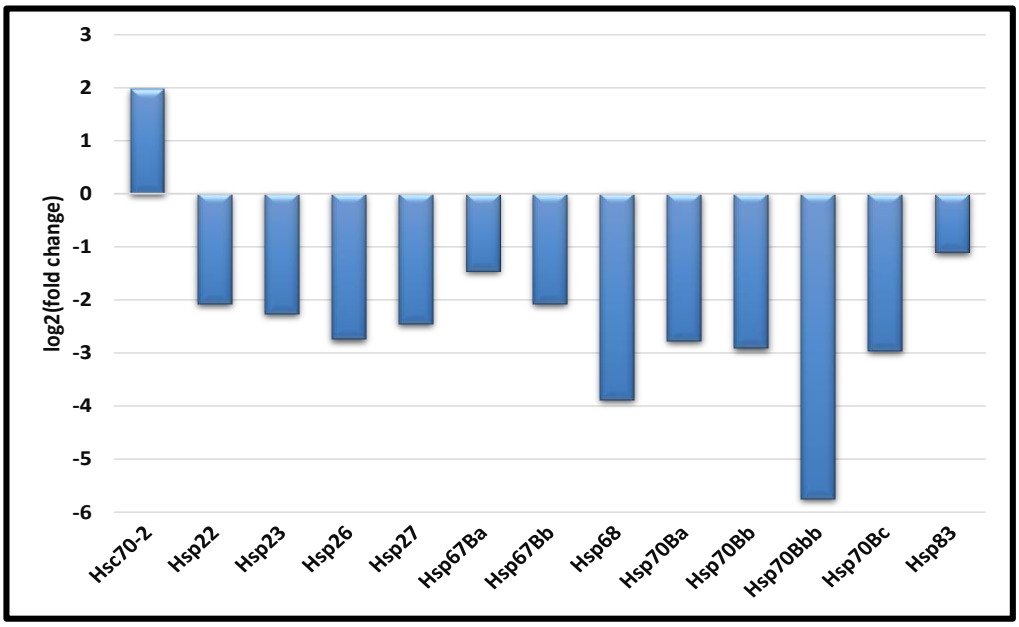

Figure 43: RNA levels for heat shock protein genes are downregulated in *CG1847* mutants. With the exception of Hsp70-2 (detected significantly upregulated), all transcripts beloging to the heat shock protein family were detected in RNA-seq as significantly dowregulated ($P<0.05$). Data is shown as the logaritmic fold change of expression levels in the mutant samples.

To find out the functional relationship among relevant genes, the top 400 differential expressed genes were mapped using STRING database, by selecting *Drosophila* melanogaster as a model organism. 393 transcripts were recognized by the database and displayed as a network of nodes (proteins) connected by coloured edges representing functional relationships of known and predicted protein interactions. The map was filtered for non-connected proteins. A medium combined score of protein pairs (confidence score) > 0.4 was considered as the cutoff value. The STRING confidence score is calculated based on the number and types of evidences that support each association (low confidence scores: <0.4; medium: 0.4 to 0.7; high: >0.7).

As shown in the Figure 44, only one interaction was found for CG1847, with one of the heat shock proteins (Hsp83). In humans AIP–Hsp90 is one of the most investigated AIP interactions. A number of different groups (Bell & Poland 2000[148], Meyer *et al*. 2000[149], Laenger *et al*. 2009[150]) have shown that specific conserved amino acids of the AIP TPR domains are important for this

interaction. Consequently, in CG1847 deficient *Drosophila* model the members belonging to heat shock proteins might be downregulated due to losing these interactions.

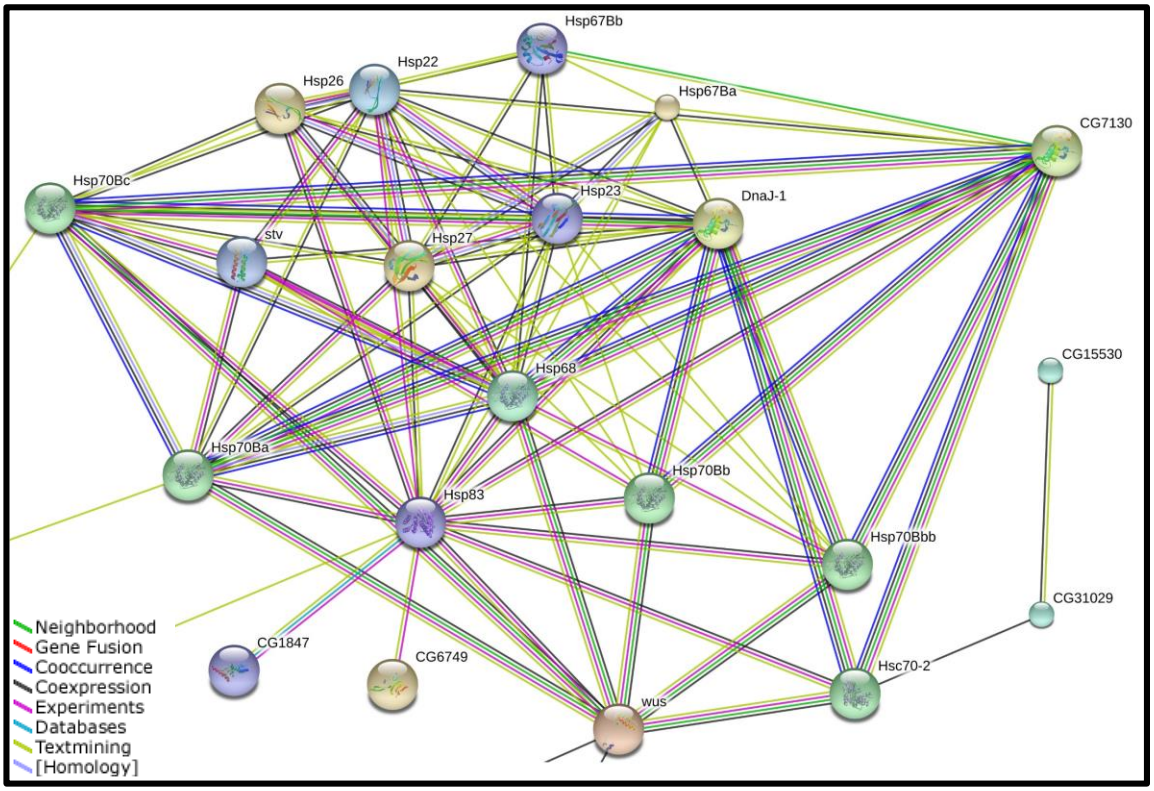

Figure 44: STRING analysis: the heat shock protein cluster. Nodes represent proteins whose transcripts were identified in the RNA-seq analysis as being differentially expressed in the absence of *CG1847*. Lines connecting individual nodes indicate predicted functional associations. Among these significant genes, different members of Hsp family were identified.

The STRING analysis also highlighted a cluster of a very interesting group of proteins called Osiris proteins. The Osiris genes belong to a large conserved family which was for the first time described in 2003 by Dorer *et al*. in *Drosophila melanogaster*[289]. These proteins have a secretion signal peptide and four domains that identify them as Osiris family members; however, there are limited publications and their function is still unknown. 12 out of 24 members of this protein family were found significantly upregulated in the RNA-seq data. Strikingly, among this cluster (Figure 45), were found a few Twdl proteins (TwdlG, TwdlS, and TwdlF). Additionally, different members of cuticular proteins (Cpr100A, Cpr47Eg, Cpr49Af, Cpr65Eb, Cpr66Cb, Cpr66D, and Cpr97Eb) were also hub nodes in this network. So far no interactions between CG1847 and these proteins have been suggested. Twdl and cuticular proteins were previously shown to be involved in body size regulation and normal development, as the exoskeleton of insects (cuticle) is mainly

formed of chitin and cuticle proteins. The Tweedle proteins are secreted by the ectodermal tissues and members of these proteins were shown to also contribute to formation of the cuticle[290].

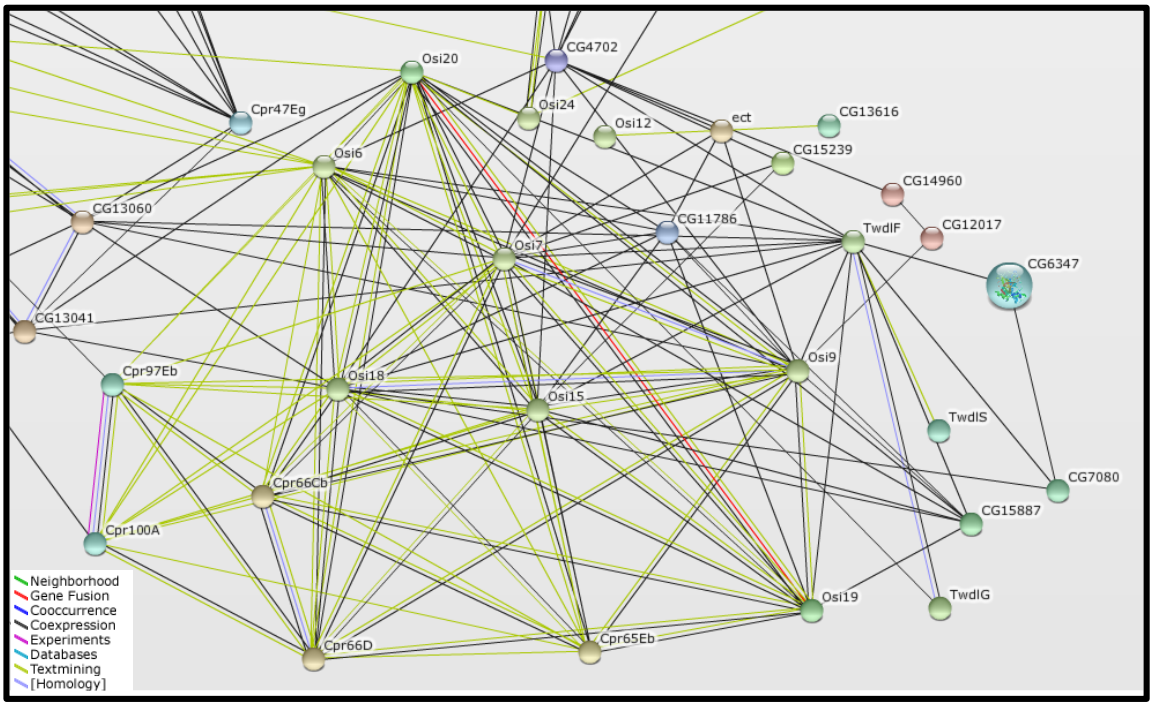

Figure 45: STRING analysis: the Osiris, Tweedl and cuticule proteins cluster

Surprisingly, all transcripts belonging to the Osiris and Tweedl group were upregulated in *CG1847* mutants (Figure 46).

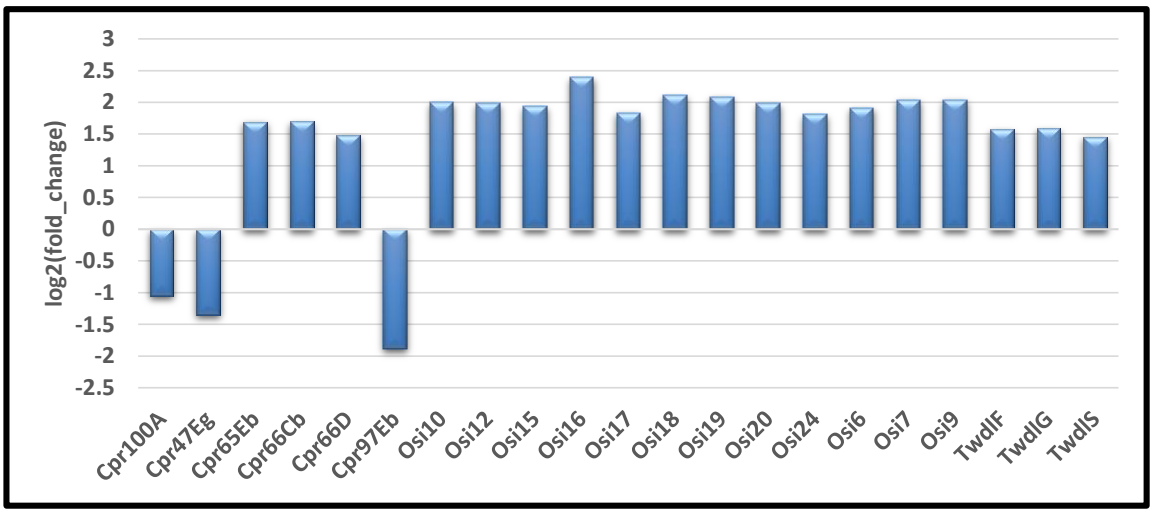

Figure 46: RNA levels for Osiris and Tweedl transcripts are significantly changed in *CG1847* mutants. With the exception of some of the cuticular proteins, all transcripts beloging to one of these 3 clusters were detected in the RNA-seq as significantly upregulated (P<0.05). Data is shown as the log fold change in the CG1847 deficient larvae.

To identify the human orthologues for Osiris protein candidates found in the RNA-seq approach, I used the BioMart free software (http://www.ensembl.org/) (Table 18).

Ensembl Transcript ID	Associated Gene Name	Homology Type	Transcript ID	Human Ensembl Gene ID
FBtr0085710	Cpr100A	ortholog_none		
FBtr0088183	Cpr47Eg	ortholog_none		
FBtr0087921	Cpr49Af	ortholog_none		
FBtr0070815	Cpr5C	ortholog_none		
FBtr0333582	Cpr65Ax2	ortholog_none		
FBtr0077022	Cpr65Ax2	ortholog_none		
FBtr0076928	Cpr65Eb	ortholog_none		
FBtr0076601	Cpr66Cb	ortholog_none		
FBtr0076567	Cpr66D	ortholog_none		
FBtr0074933	Cpr76Bb	ortholog_none		
FBtr0085135	Cpr97Eb	ortholog_none		
FBtr0082707	Hsc70-2	ortholog_none		
FBtr0100558	Hsp22	ortholog_many2many	FBpp0100010	ENSG00000109846
FBtr0309504	Hsp23	ortholog_many2many	FBpp0301282	ENSG00000109846
FBtr0076496	Hsp26	ortholog_many2many	FBpp0076224	ENSG00000109846
FBtr0076454	Hsp27	ortholog_many2many	FBpp0076182	ENSG00000109846
FBtr0076495	Hsp67Ba	ortholog_many2many	FBpp0076223	ENSG00000109846
FBtr0303473	Hsp67Bb	ortholog_one2many	FBpp0292525	ENSG00000215845
FBtr0084589	Hsp68	ortholog_none		
FBtr0082679	Hsp70Ba	ortholog_none		
FBtr0082636	Hsp70Bbb	ortholog_none		
FBtr0082638	Hsp70Bc	ortholog_none		
FBtr0332873	Hsp83	ortholog_none		
FBtr0073040	Hsp83	ortholog_none		
FBtr0078600	Osi10	ortholog_none		
FBtr0078601	Osi10	ortholog_none		
FBtr0078602	Osi12	ortholog_none		
FBtr0078607	Osi15	ortholog_none		
FBtr0078608	Osi16	ortholog_none		
FBtr0078610	Osi17	ortholog_none		
FBtr0078611	Osi18	ortholog_none		
FBtr0081707	Osi19	ortholog_none		
FBtr0301493	Osi19	ortholog_none		
FBtr0081708	Osi20	ortholog_none		
FBtr0078591	Osi24	ortholog_none		
FBtr0078596	Osi6	ortholog_none		
FBtr0078597	Osi7	ortholog_none		
FBtr0078599	Osi9	ortholog_none		
FBtr0078975	TwdlF	ortholog_none		
FBtr0078982	TwdlG	ortholog_none		
FBtr0085054	TwdlO	ortholog_none		
FBtr0085049	TwdlQ	ortholog_none		
FBtr0085034	TwdlS	ortholog_none		
FBtr0113190	TwdlU	ortholog_none		

Table 18: Human orthologues for *Drosophila* transcripts

No human orthologues were found in the human genome for any of members of the Osiris family. This is in agreement with the fact that the Osiris gene cluster is a family of genes that is present in all insects, but not present in mammals[289]. Conversely, for a few *Drosophila* heat shock proteins there are more than one single human orthologue. However, some of these proteins are *Drosophila* specific. Interestingly, no orthologue could be found for Hsp83, which is the only interacting protein with CG1847 revealed by STRING analysis.

3.4.10 Multiplex qPCR validation of RNA-seq results

RNA-seq is an extremely useful tool for whole transcriptome quantification, but the results need further validation via an alternative molecular technique. Toward this purpose, I employed a multiplex qPCR method, performed with the help of collaborators from Royal Veterinary College (Dr. Rob Fowkes and Dr Samantha M. Mirczuk). The GenomeLab™ GeXP Genetic Analysis system provides accurate simultaneous quantifications for hundreds of samples using very small amounts of total RNA. The system uses chimeric primers which are a combination of gene-specific and universal tags.

The main goal was to validate the changes in the expression levels of the transcripts possibly involved in lethality of *CG1847* deficient fruit flies. For this purpose the same larval collections as for RNA-seq were used and the RNA was extracted from a pool of male larvae. As the STRING analysis of RNA-seq data revealed clusters of interacting proteins of several families (section 3.3.7), I chose a few transcripts of each cluster/protein family for validation (Figure 49).

A key feature is the fact that the gene ratio between mutant and control in RNA samples is maintained during the PCR process. In the initial stage the reverse transcription of RNA was performed with the reverse chimeric primers (containing the gene-specific sequences and universal primer sequence). At the end of this step the cDNA library had universal tag sequences incorporated into the amplicons. The chimeric primers were designed for 2 housekeeping genes, *CG1847* and 8 various transcripts with a possible involvement in the lethality mechanism (Appendix 5 for primer sequences). In a second step transcripts amplification with a forward pair of chimeric primers was performed. In the last step fluorescent dye-labelled forward and unlabelled reverse universal primers resulted in a relative and equivalent amplification of all the gene targets (Figure 47).

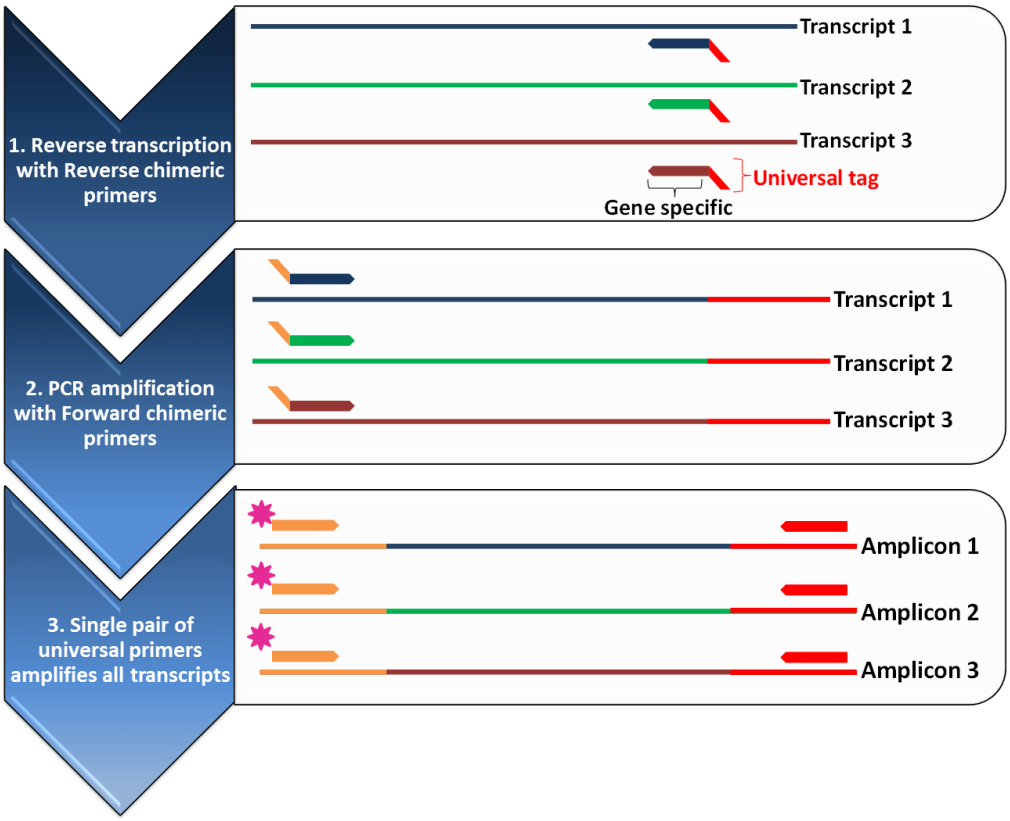

Figure 47: Schematic representation of the 3 main stepts of GeXP Genetic amplification. In the first step (top panel) chimeric primers were used containing the gene-specific sequences (blue, green and brown arrows) tagged with a universal primer sequence (red). In the second step (middle panel), the generated cDNA library, having incorporated the universal tag are further amplified via PCR using the forward chimeric primers, with a different universal tag (beige). In the third step a fluorescent (purple star) dye-labelled forward primer and a unlabelled reverse universal primer are used for amplification of all the gene targets

The amplicon mixture was then analysed with fluorescence capillary electrophoresis to identify the peak area and then converted to gene-expression values (Figure 48).

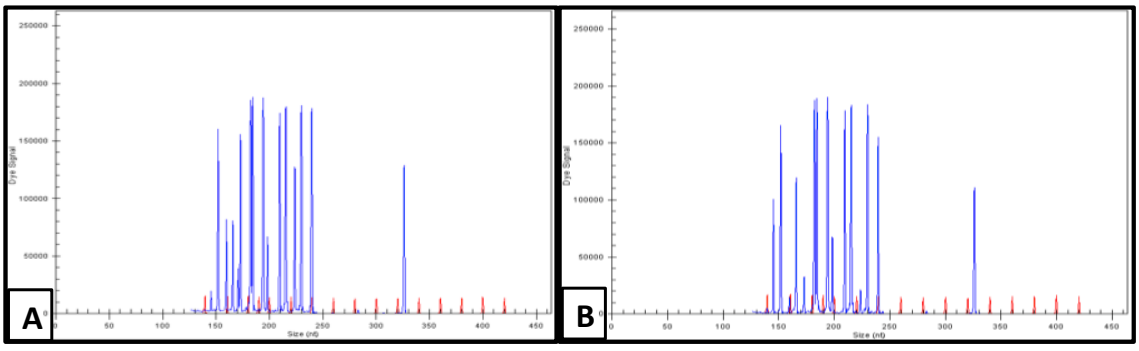

Figure 48: Representative electropherograms corresponding to gene expression profiles. Representative data generated from mutant (A) and control (B) mRNA samples are shown. The intensity versus the size is depicted within a window of 100 to 350 bp. The red peaks are the size ladder, which ranges from 140nt to 425nt. The blue peaks are the amplicons corresponding to the genes of interest, which were detected by the multiplex qPCR system

The gene-expression values were normalized to the *RpL32* housekeeping gene, included in the multiplex. The final normalized values were used to compare levels of gene expression between control and mutant samples (Figure 49).

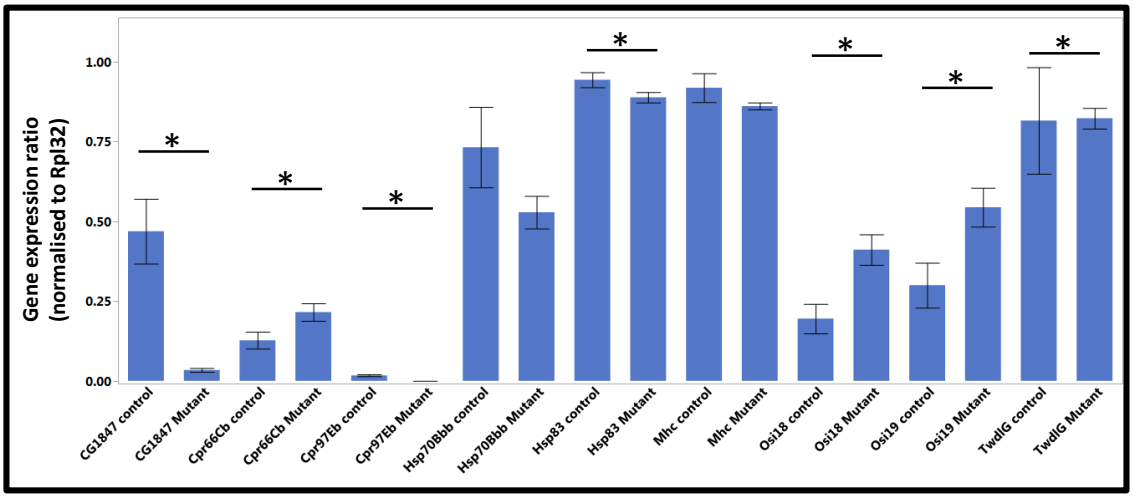

Figure 49: Multiplex qPCR results. The mRNA levels of 9 specific transcripts that were significantly changed detected by the RNA-seq and that were represented in different STRING clusters. Relative expression levels are shown as a bar graph. Expression values were normalized to control in order to present expression ratios. *RpL32* was used as housekeeping gene for normalization of genes of interest. Data is shown as mean ± SE. ANOVA was used to assess for statistical differences (*P<0.05).

In this study I sought to determine if data obtained from the RNA-seq is accurate by testing if different techniques reproduce the same results. Data analysis confirmed that the expression levels of some of the specific genes were substantially different in mutant larvae, similar to the previous RNA-seq data.

Figure 49 shows the gene expression profile obtained from comparing the level of expression between control and *CG1847* deficient *Drosophila* larvae. The results show a marked downregulation of *CG1847*, validating the statistically significant changes detected by RNA-seq in the mRNA expression levels. There was also a significant increase in the mutant larvae of the expression of both Osiris transcripts selected for validation. This recapitulates the increase in Osiris gene expression previously observed in the RNA-seq analysis. TwdlG and both cuticular mRNA transcripts found to be involved in the same biological processes as shown by STRING analysis, were also validated. *Mhc* was also tested; this is a protein found to be involved in the same biological process as the Osiris, Twedl and cuticular proteins. This transcript was significantly downregulated in the mutant samples; however, without reaching a fold change of at least -1 and this change in the expression level was not confirmed by the multiplex qPCR.

Some of the most drastically differentially expressed genes detected by the RNA-seq were clustered with CG1847 and belong to the heat shock family of proteins. The available data confirmed that AIP acts as co-chaperone to form a complex with HSP90 and AhR, in order to allow nuclear translocation. The majority of heat shock proteins were strongly downregulated in the RNA-seq results and this was confirmed by the multiplex qPCR. Of the 2 chosen transcripts I was only able to validate the downregulation of Hsp83 in *CG1847* mutants. Hsp83 was the only direct CG1847 interacting partner found in the STRING analysis (Figure 44). However, despite the fact that it was one of the genes with the highest fold change in expression levels in the RNA-seq analysis (-5.7 Table 17), Hsp70bb transcript surprisingly was not confirmed in the multiplex qPCR analysis.

This technique is simple, very efficient, and cost effective, providing an excellent solution for overcoming challenges, such as validation of a high number of transcripts, limited amount of sample and costly analysis. Further experiments are required for validation of other transcripts detected as significantly changed by the RNA-seq data in order to investigate other possible affected pathways by the loss of CG1847.

3.5 DISCUSSION

In an effort to elucidate the involvement of AIP during development and the pathogenicity of human pituitary adenomas associated with loss of function of AIP, I used an *in vivo* approach to assess the function of the fruit fly orthologue of AIP, CG1847. In this chapter, I focused on the generation of an *in vivo* genetic mutant of *CG1847* and the initial characterisation of its function. Firstly, I used a bioinformatic approach to identify *CG1847* as the fruit fly orthologue for human *AIP*. *CG1847* is located on the X chromosome. The structure and function of AIP is conserved in *Drosophila melanogaster*. Amino acid alignment of *Drosophila* CG1847 and human AIP showed a high degree of sequence similarity between the 2 proteins, with one third of the amino acids being highly conserved. According to the ClustalW algorithm there is a 37.74% overall identity between human and *Drosophila* proteins. The amino acid sequence of CG1847 is highly conserved and shares all the proteins domains, both in the N-terminal and C-terminal parts, with members of other species. This suggests that the protein retains the same molecular functions as the other orthologues. The high degree of conservation of AIP is also an indication that its physiological function would also be conserved. To determine whether CG1847 plays a role in development, I used in this study three *UAS-CG1847-RNAi* stocks, carrying 2 different constructs targeting different parts of the *CG1847* gene. As previously AIP was described as an essential

gene in mammals, studies were performed to determine the effect of CG1847 knockdown on fly survival. Lethality was classified if no GAL4-driven UAS-*RNAi* progeny emerged from a cross. When crossed to the *Act-GAL4* driver, 2 out of 3 UAS-*RNAi* stocks (UAS-*CG1847-R2* and UAS-*CG1847-T2*) showed no surviving adult flies. The third stock (UAS-*CG1847-R1*) resulted in greatly reduced numbers of the expected genotype (at 25°) and no RNAi-driven progeny being observed (at 29°C). Complete AIP-knockdown lethality is suggestive for *CG1847* being an essential gene and all these results correlate with the lethality of AIP deficient mice[204,291]. The result that flies with only around 30% residual expression of *CG1847* are viable is also a confirmation of previous studies which reported that hypomorphic AIP mice have almost normal development[181]. The available human data support the theory that homozygous AIP mutations are not compatible with life, while one normal copy is enough for normal development.

The Gal4/UAS system presents a few very important advantages. Maybe the most important is the separation of the driver (Gal4) and the effector (UAS) between two parental lines, which in our project allowed us to avoid the parental lethality. Temperature dependence[285] of GAL4 is another important feature of this system. By simply exposing the flies to lower (around 16°C) or higher (29°C) temperatures during development the potency of the system can be decreased and increased, respectively. By increasing the ambient temperature, a higher level of expression can be achieved[266], and this temperature sensitivity feature was used to enhance the expression of a weaker RNAi construct, and confirmed the lethality noticed with the other UAS-RNAi stocks at lower temperature.

However, the Gal4/UAS system also has some disadvantages that I had to be aware of. In gene expression knockdown experiments based on RNA interference (RNAi) technology, some results might be due to off-target effects. These are the result of the cross-hybridization phenomenon between the siRNAs molecules and non-target areas in endogenous RNA sequences and outside the gene of interest[292-294]. The occurrence of off-targets effects in RNAi experiments can be extensive and significant[295]. Randomly selected siRNAs can induce toxic changes by reducing cell viability in a target-independent fashion[292]. As a consequence, the resulting phenotypes might mislead the functional interpretation of gene silencing[294]. Whenever possible, good interpretation of results should be based on using more than one siRNA, or should be validated through an alternative method. The fact that using different RNAi constructs targeting different regions of *CG1847* RNA resulted in partial or full lethality indicates that it is unlikely that this result was caused by an off target silencing.

Furthermore, I applied an alternative method. To investigate the mechanisms of CG1847 driven lethality and to help understand its role during *Drosophila* development, a proper loss of function mutant was generated via P-element mobilization. Following the imprecise excision resulted a *CG1847*$^{exon1\_3}$ allele that carries a deletion of the entire CG1847 ORF. The total loss of expression (Figure 35) together with the fact that homozygous *CG1847*$^{exon1\_3}$ females and hemizygous males are not viable indicates that *CG1847*$^{exon1\_3}$ is a null allele and that *CG1847* is an essential gene.

Unfortunately, this technique also has some potential problems and pitfalls. First, the frequency of such imprecise excisions may vary between loci (from 1 in 5 to 1 in 100 events)[296,297] and the frequency of imprecise excision cannot be predicted from the outset of the experiment. In some cases hundreds of crosses need to be generate to achieve an imprecise excision. However, the P-element inserted into the 5' of *CG1847* was quite mobile (9%). Another problem that can occur is that the mobilized P-element can reinsert. The precise site in the genome where this event occurs is unpredictable. If it still carries the white gene, than these re-insertions are easy to detect. If the white gene is lost, the second hits are difficult to detect and might lead to mis-interpretation of the observed phenotypes. The transposase activity can be controlled by adding or removing the transposase source. To prevent further mobilisations of the P-element, the source of transposase is separated in the subsequent generation by appropriate crossing.

The third problem regards the issue that the size of the P-excisions are random[268], therefore there is no guarantee that the desired mutation would not extend to other genes. The size of the deletion from the imprecise excision needs to be confirmed by molecular methods, including PCR on genomic DNA from heterozygous flies. Care should be taken with designing the necessary primers, as it could be extremely challenging to properly map the deletion. As the *Drosophila* genome is very small, upstream and downstream genes might be very close to the gene of interest and the design of the primers should be done in such a way to detect those excisions that are too large and affect the surrounding genes. One of the generated mutant alleles in my screen, *CG1847* $^{exon1\_3}$, was confirmed via PCR not to affect the neighbouring genes.

Consequently, the next obvious step was to validate the fact that the lethality is solely due to deletion of *CG1847*, and that no other genes were affected during P-element excision. Transgenic lines carrying wt CG1847 on the second chromosome were able to produce rescued males in the second generation (Figure 39). Surprisingly, the percentage of the rescue males from the total number of counted viable males was much higher than expected. The percentage of the rescued males (from the total number of viable F1 males) suggest that these are very fit

for survival during larval stages and they are strong competitors for nutrients. Another possible reason might be also the influence of the genetic background. The rescued males (*CG1847*$^{exon1\_3}$*/Y; CG1847/+)* have no balancer chromosomes. All the other male genotypes in the cross have one or even 2 balancer chromosomes (FM6 or/and CyO) and this might reduce their fitness. Balancer chromosome carry multiple inverted repeats which can weaken the fly.

The last but not the least important issue regarding this technique is that some of these P-element excisions events may only be partially successful – as they may leave some very small sequences behind ("footprints"). Around 75% of the footprints are usually very short, 4 or 7 nucleotides, while the remaining 25% of the footprints might vary between 14–18 nucleotides of both inverted terminal repeats[298]. However, this issue was turned into an advantage as it helped generate the control stock, with the same genetic background.

RT-PCR was used to assess CG1847 RNA expression in mutant male larvae in comparison to the same stage control larvae and control *y w* adult flies (Figure 35). Quantification after normalisation to the housekeeping gene RpL32 showed that knocking-out the *CG1847* gene resulted in total loss of expression. This kind of data normalisation is a standard step in the quantification RT-PCR analysis[299,300].

The reliability of any quantitative RT-PCR data is based on including a stable endogenous control (reference gene) which allows for proper correction of sample variations. However, the levels of expression of any gene, including that of 'typical' housekeeping genes, might vary between tissues or between stages of development[301]. It is very important to choose the appropriate one. RpL32 was traditionally considered a very good internal control gene as it has equal transcript levels in all cells and in all stages of development[302]. This is due to the fact that ribosomal proteins are involved in protein biosynthetic pathways [303] and play critical roles in the development and growth of organisms[303]. Its levels are also stable in different experimental conditions.

Surprisingly, the levels of expression of the chosen housekeeping gene for our mutant (RpL32) were increased in mutant larvae. I was surprised by this observation since, usually, reduced growth rates and survival rates are associated with reducing or even abrogating the efficiency of protein synthesis[304,305]. Generally, an increase in ribosome biogenesis is a characteristic of an increased cellular proliferation during tissue and organ growth[306] and this might be related to the increase in proliferation associated with loss of AIP in human pituitary tumorigenesis. However, another factor to consider for quantification of gene expression levels is the genetic background. As the 2 controls (48 h *y w* larvae and *y w* adults) have similar levels of expression

for RpL32 suggests that the difference I noticed in mutant male larvae might be due to the fact that it has a different genetic background. As a consequence, in RNA-seq experiments I took great care in finding the perfect control to compensate for any possible differences in the genetic background, by generating both the mutant and the control *Drosophila* stock during the same experiment of P-element excision.

The lethality of the *CG1847* mutants was further investigated and I established that the flies are dying in the larval stage, suggesting that CG1847 is critical for survival beyond this point in development. These hemizygous mutant male larvae were much smaller in size after 48 h of development. No other obvious phenotype was noticed, therefore this did not help in understanding of the actual cause of lethality. Regarding the reduced size of the mutant larvae, it is necessary to take into consideration a few factors that might influence the larval development, in addition to the impact of the loss of CG1847.

In *Drosophila* the body size is under the influence of the genetic background, the available food (the amount and the quality of food), and/or the environmental temperature during development[307]. It is well-established that temperature influences body size as fruit flies exposed to lower temperatures during development are larger than those exposed to higher temperature[308]. It was also found that male larvae have a significantly smaller size (35.6% reduction) than females ($P = 0.008$)[309]. As a consequence, we have to be cautious before claiming that the smaller body size of the mutant male larvae is due solely to the *CG1847* deletion. As the length of male larvae was compared to the length of all the other fluorescent larvae in that collection (including females) this might introduce a bias in the analysis. On the other hand, if the loss of CG1847 has an influence in these mutant larvae fitness, then they had a more restricted access to nutrients comparing to the other genotypes, and as a result, their size is smaller due to undernutrition (which might be the case as they are not only shorter, but also thinner – Figure 37).

Another important advantage of the UAS-Gal4 system is the possibility to target the expression of any gene or RNAi construct in a variety of spatial and temporal ways by using distinct Gal4 drivers[238], which can be ordered from public *Drosophila* libraries. For example, a high number of GAL4 drivers are publicly available at the Bloomington *Drosophila* Stock Centre: http://flystocks.bio.indiana.edu/gal4.htm.

Overexpression of $hAIP^{wt}$ cDNA using the UAS/GAL4 system was employed to investigate if CG1847 is the functional orthologue of human AIP and to understand the role of this product in different cells/tissues. Numerous previous studies have reported a ubiquitous expression of AIP

in human and murine tissues, both at the mRNA and at a protein level[139,149,154,159,171]. As a consequence, it is difficult to determine solely on the literature data in which tissues AIP is more important for a normal function in a developing organism.

As the UAS-GAL4 system can be spatially defined by choosing GAL4-drivers with tissue specific promoters, I utilised it to gain a better understanding of where *CG1847* expression is more important. Surprisingly, overexpression of *CG1847* throughout the fly during development using the actin-GAL4 driver was capable of rescuing lethal effect of loss of CG1847. However, none of the tissue specific Gal4 drivers was able to rescue this phenotype, suggesting that this gene is involved in many different processes and it has important function overall in the body. Thus, in analogy with other species, CG1847 deficiency is developmentally lethal. To confirm and to understand these results I investigated the public available data from modENCODE (Table 19). modENCODE (the model organism ENCyclopedia Of DNA Elements) is a key research project of the National Human Genome Research Institute (NHGRI)[283].

My data were confirmed by the results of the RNAi screen available on modENCODE, as none of the tissue targeted knock down of *CG1847* expression resulted in lethality.

Screen	Gene	Phenotype	Score	PubMed ID
Heat nociception (1)	CG1847	Increase heat avoidance	2.46	21074052
Lipid storage	CG1847	none	-0.71	19067489
Notch pathway regulation (4)	CG1847	none	0	19363474
Adiposity regulation (1)	CG1847	none	0.48	20074523
Heart development and function (1)	CG1847	none	0.8	20371351
Dendrite pattern formation	CG1847	none	np	16547170
Muscle morphogenesis and function (1)	CG1847	none	np	20220848
Cell size and cell-cycle regulation (1)	CG1847	none	sp	16496002
Glycosylation regulation (1)	CG1847	none	sp	21203496

Table 19: modENCODE RNAi screen for CG1847 (http://intermine.modencode.org/release-33/results.do?noSelect=true&table=coll854&trail=%7Cresults.coll854)

Studies involving mouse models revealed that AIP is essential for development, as lack of this product led to lethality during embryonic development. Further investigations described AIP involvement in cardiac development as the embryos displayed a range of heart deformations, including double outlet right ventricle, ventricular-septal defects, and pericardial oedema[204]. Unfortunately the exact molecular mechanisms and the exact partner via which AIP total deficiency leads to loss of viability is still not fully understood, and further work is required to identify the key proteins and pathways involved.

In summary, ubiquitous *CG1847* expression is essential for survival to adulthood and the generated null mutant can be used for further investigations necessary for understanding its involvement.

RNA-seq is a revolutionary tool developed for deep sequencing and quantification of transcripts and their isoforms[261]. RNA-seq provides accurate identification of novel transcripts, of alternative splicing events, and allele-specific expression[277,310]. Consequently I decided to use this technology for investigating the CG1847 knockout model's transcriptome. Comparing to previous hybridization-based approaches, RNA-seq has an extremely important advantage as it is not limited to detecting transcripts that correspond to existing genomic sequence and brings the advantage of de novo assembly of the genome and of the transcriptome[311,312]. It was already undoubtedly proven that RNA-seq can detect the exon-exon borders which enable this technique to accurately detect novel isoforms, and even to discover completely new genes[310,313]. This might lead to the identification of novel transcripts in *GC1847* deficient model. Regarding the quantification of expression levels, RNA-seq has a very wide dynamic range with very limited background level and it is capable of detecting and quantifying even the transcripts with extremely low levels of expression[314,315] which might happen is this case as the RNA is extracted from fruitfly larvae. It also allows the exploration of different cellular pathways at the same time[316]. RNA-seq requires less starting material (total RNA), a valuable practical aspect as this brings a significant advantage in cases of limited sample availability.

However, this technology does have a number of limitations which we have to be aware of[277]. As this method is based on fragmentation and amplification of RNA samples to produce the cDNA libraries, this step carries the risk of introducing bias and artefacts into the system. Another downside of the library preparation method itself is that it requires careful removal of polyA or ribosome RNA in order to avoid producing different transcriptome profiles. An additional disadvantage is brought by the levels of deep sequencing and lack of complete coverage which, for more complex organisms might result in under representation of some of the genes/transcripts/isoforms. Some genes/transcripts may not be detected as they might have a restrictived pattern of expression (tissue specific or in specific stages of development). This implies careful design of the RNA-seq study[317,318]. RNA-sequencing, is an extremely powerful method which has the potential to rapidly reveal the changes in normal development or in pathological processes, but the costs involved by the whole transcriptome sequencing are prohibitive, and limit the use of this method in routine research.

In parallel with the fast developing of RNA-seq technologies, many mapping tools have also been developed. The actual mapping methodologies could have a negative influence on the RNA-seq results as they could fail in detection of splice junctions[319]. In 2011 Roberts *et al*. reanalysed previous data using a different mapping methodology and identified many more novel transcripts[320]. Other authors also confirmed that the choice of appropriate methodology is important[321].

TopHat is maybe the most popular splice junction mapper for RNA-Seq reads. It aligns the raw RNA-Seq reads to genomes using the ultra-high-throughput short read aligner Bowtie to actually identify the exon splice junctions and to reveal novel gene or novel alternative splicing. For reconstructing the whole transcriptome, the most commonly tool used program is Cufflinks[322] which, similar to TopHat, is a freely available public domain software.

TopHat, Bowtie and Cufflinks pipeline is the only available pipeline so far that includes all the required software for full analysis for RNA-seq data. However, a downside is that it requires a good annotation of the genome. The software were designed to be used with the Illumina data format, and, requires significant data storage. The large amount of data produced might be up to 1000 times bigger than the amount of data produced by microarrays, as a FASTQ files alone might be estimated to be between 20 and 40 Gb.

As RNA-seq allows quantification of the entire transcriptome. I decided to use this technique to reveal the changes in expression levels between wt CG1847 and knockout *Drosophila* larvae. I also wanted to have a better understanding of AIP orthologue role during normal development by revealing the interacting proteins affected by CG1847 loss of function. As expected, CG1847 was one of the most significantly downregulated transcripts in mutant samples. The normal level of expression of CG1847 at this stage of development (48h AEL) was very low. To confirm, understand and compare the RNA-seq results with data regarding normal fruit fly development, publicly available data from modENCODE was used. modENCODE's main purpose was to identify previously unannotated transcripts in *Drosophila melanogaster* in 27 distinct stages of development. It identified 1,938 new transcribed regions not linked to any previously annotated gene[323]. Given the fact that this project was studied on normal development, it can serve as a useful comparison of my data. On examining the *CG1847* expression patterns during *Drosophila* developmental stages in KO larvae I found significant differences with our control larvae. Within each stage of development (from embryo to adult fly), *CG1847* had very low expression. The highest intensity was in the first 2 hours possibly due to maternal contribution - Figure 24. The level of expression are slightly increased in the later embryonic stages, and then they are

reduced again during all larval stages. The same reduced gene expression levels are maintained during pupal stage and adulthood in males and is slightly increased in females, but not more than during the larval stage. Furthermore, in third instar larvae, when we collected our mutant and control males, the normal level of expression of *CG1847* are some of the lowest of all stages. The fact that at this stage the fruit flies' requirement of CG1847 is minimal might be a possible explanation for the fact that even though our mutant is a KO and has no *CG1847* expression (Table 17) it is not the transcript with the highest fold change as we would have expected.

One of the most investigated AIP functions is the involvement in the xenobiotic processes. Numerous articles published that AIP is part of the AIP-AhR-Hsp90 complex which facilitates AhR translocation into the nucleus where it binds to ARNT[148,154,324]. In the next step the AhR:ARNT complex is able to bind a dioxin-responsive element (also known as xenobiotic responsive element). As AhR is an important AIP partner, I investigated this in the RNA-seq results. The *Drosophila* orthologue for human AhR is called spineless (ss) and its levels of expression were surprisingly not significantly changed in the mutant larvae. I then noticed that the levels of expression even in the normal control larvae at this stage of development were extremely low (0.35). Further, I investigated the modENCODE database and I had the confirmation that at the stage when I performed the larvae collection the levels of expression for AhR are actually zero. These results are in accordance with previous published data which suggest that AIP has a very early developmental expression, before the appearance of AhR transcript. This idea is supported by AIP whole mount *in situ* hybridisation of mice embryos. It was revealed that at embryonic day 9.5 (e9.5) AIP expression can be detected while AhR is first expressed only around e13.5[159,325]. Taken together the available data from humans, mice and fruit fly suggest that during normal development AIP is expressed previous to AhR, and, at least in the very early embryonic stages, AIP function is independent of its interaction to AhR and of xenobiotic processes.

AIP plays important roles in mammalian development, apart from xenobiotic receptor signalling. This was confirmed on the other members of the immunophilin family which appeared to have physiological importance outside of xenobiotic signalling[326]. The available data from human and mammalian systems point out that AIP protein is localised in the cytoplasm as part of multiprotein complexes with well-known partners such as HSP90 p23 and AhR. Many other cytoplasmic partners were also described, and these partners involve AIP in a number of different pathways. Known interacting proteins, as presented in detail in Table 3 are cyclic adenosine monophosphate (cAMP), chaperone proteins (HSP90 and HSP70), G proteins (Gα13 and Gαq) and phosphodiesterases (PDE4A5, PDE2A3). Other interactions, not confirmed, have been described with a cytoskeletal protein (actin), a growth factor receptor (EGFR) and a cardiac-

specific kinase (TNNI3K). Human data identified AIP as a tumour suppressor gene. Mutations of this gene lead to pituitary adenoma formation, which are, as discussed above, benign neoplasm of the anterior pituitary.

Regarding the CG1847 interacting partners, the available information is very limited. There are only 3 described partners: Nurf-38 (Nucleosome remodeling factor - 38kD), CPB (capping protein beta) and CG9986. They were discovered in 2003 in a large project which generated a two-hybrid-based protein-interaction map of the fly proteome[327]. However, none of these 3 transcripts were found significantly changed in the RNA-seq data. Even more, their human orthologues PPA2 (pyrophosphatase 2), CAPZB (capping protein actin filament muscle Z-line beta) and C12orf4 (chromosome 12 open reading frame 4) respectively are not known interactors of hAIP.

The performed RNA-seq experiments revealed a general up-regulation of genes closely associated with Osiris genes in mutant CG1847 deficient larvae. The Osiris genes are a large group of 24 orthologue genes that were first described in *Drosophila melanogaster*[289]. Out of the 24 Osiris genes, 20 are located on the third chromosome in the cytological region 83D4-5 to E1-2. This cluster is within a 168-kb region first described in 1972[328] as being both triplo-lethal and haplo-lethal (*Triplo-lethal* (*Tpl*) locus). Previous studies have shown that flies with either one or three copies of Tpl die as late embryos or early first instar larvae. This group of proteins is a highly conserved insect-specific class. The structure of the proteins belonging to this family is characterized by five domains: (1) a hydrophobic region at the N-terminus that is likely a secretion signal peptide; (2) a two-Cys region; (3) a domain of unknown function, duf1676 (Pfam family: PF07898)[329]; (4) a hydrophobic putative transmembrane domain, and (5) a region including an AQXLAY motif.

As lack of CG1847 resulted in significant overexpression of most of Osiris genes, it is possible that this might be the mechanism that leads to lethality, by mimicking the existence of 3 copies of Osiris genes.

I further tested the hypothesis that the lethality of *Drosophila* model might be due to the upregulation of Osiris genes in mutant males. First I compared the phenotype of mutant CG1847 deficient larvae with the one described in Triplo-lethal locus. The lethality induced by the presence of 3 copies of this cluster usually develops either in later stages of embryogenesis or in the early first instar larvae[330], which is actually earlier than in our model. In these dying embryos/larvae at first the midgut turns brown due to intense apoptosis and cell death, a phenomenon that later extends to other tissues as the tracheae appear, which start to break

up[330]. None of these changes were present in CG1847 mutant. As I described in the Results section, CG1847 deficient larvae display a normal morphological phenotype, except the fact that their development is halted after the second instar larval stage. I could not detect any signs of cell death even after staining the larvae with Trypan Blue (a stain that is excluded by living cells[331].

A few studies have published that the lethality induced by the presence of three doses of Tpl might be prevented by inactivation, knockdown or mutations, of a locus called Suppression of Tpl (Su(Tpl))[332,333]. The mechanism behind suppressing the lethality of Tpl by the Su(Tpl) locus is not well understood. It is only known that this locus encodes for the transcriptional elongation factor dEll[334]. *Drosophila* ELL homologue (dELL) promotes RNA polymerase II (Pol II) elongation as it was shown that it has the ability to increase its the catalytic rate[335,336]. Eissenberg *et al*. showed that reduced dELL levels could further reduce the expression of each of the three-dose *Tpl* in larvae, resulting in *Tpl* gene product levels closer to wild-type levels (two *Tpl* doses). However, there are no proofs that dEll might act directly on Su(Tpl). The same group of authors proposed 3 possible models of dELL function: a) dELL may act on a specific subset of genes; b) dELL may act at a distinct kinetic phase of Pol II elongation; c) dELL may act additively or cooperatively with other elongation factors to achieve optimal Pol II elongation rates in vivo.

Strong Su(Tpl) mutations are able to prevent all the described abnormalities associated with cell death, resulting in viable, fertile adults. Weak Su(Tpl) mutations induce only a delay in the development of the phenotypes, as the larvae are dying during the third instar stage, but still displaying the tracheal and midgut phenotypes described above. I did knockdown the Su(Tpl) expression in our mutant larvae, via adding an actin-Gal4 promoter in the *CG1847*$^{exon1\_3}$ background, but this was not enough to prevent lethality. Still, this does not totally exclude a contribution of the Osiris genes in the lethality, as the levels of upregulation in our mutant were much higher than the presence of 3 copies. Some of the Osiris transcripts were up to 6 times more abundant in the mutant compared to control larvae. According to FlyBase (http://www.flybase.org) the Osiris genes are expressed during developmental stages in a wide variety of tissues, including epidermis, hindgut, foregut, and trachea.

Altered dosage of Tpl also alters the response to hypoxia[330]. The triple-lethality is suppressed by hyperoxia and this might suggest that there is a slight degree of hypoxia in larvae with one or 3 copies of Tpl. Interestingly, in CG1847 mutant most of the heat shock proteins are downregulated and these might have a further impact and enhance the lethality. It was already been published in 1993 that in response to stress factors, as a protective mechanism, all

organisms express heat-shock proteins (Hsps). By functioning as molecular chaperones the heat shock proteins confer stress tolerance and survival[337].

Six different members of Tweedle family genes (TwdlF, TwdlG, TwdlO, TwdlQ, TwdlS, TwdlU) were also significantly changed in *CG1847* mutants. Apart from TwdlO which was downregulated, all the other members of this family were upregulated. Interestingly, based on the STRING analysis, TwdlO is actually the only member of this protein family that does not cluster with members of the Osiris family. The Tweedle family is another novel insect-specific protein family. The Twdl family consists of 26 genes that encode for homologous proteins. 12 of these 26 genes are located within the same 74kb region, confirming the initial observation of sequence similarity among protein products of candidate genes. All together these 26 proteins form a new protein family, named the Tweedle (Twdl) family[290]. The Twdl proteins are produced by ectodermal tissues as epidermis, foregut and trachea. Each Tweedle gene has a specific temporal and spatial expression and localization pattern and they might be crucial in determining body shape, as suggested by Guan *et al.* in 2006[290]. The expression of secreted Twdl family proteins in these locations, therefore, strongly suggests that these proteins might contribute to the chitin-based matrix system. Previous findings already established a connection between body shape regulation and matrix proteins that contribute to the cuticle formation. Recent studies have revealed a remarkable diversity of cuticular proteins. By far the largest, and taxonomically most widespread, cuticular protein family is the CPR family, which is characterized by a conserved domain first identified by Rebers and Riddiford in 1988[338]. Zygotic disruption of any one of cuticular genes was shown to result in embryonic lethality[339]. As a result is not surprising that members of fruit fly cuticular proteins as Cpr65Eb, Cpr66Cb, Cpr66D, Cpr97A and Cpr100A were found strongly affected in CG1847 mutant larvae and Cpr proteins downregulation might be the cause of lethality of CG1847 mutants.

The animal model generated during this project will be further used in the next 2 chapters in order to evaluate the potential tumorigenic mechanisms initiated by loss of the AIP orthologue and to investigate the pathogenesis of human AIP missense variants. However, the RNA-seq data should be also further exploited to identify other possible candidates and validate new putative mechanisms involved in lethality. A special focus should be on those candidates that have human orthologues in order to find data relevant for humans.

3.6 CONCLUSION

In order to analyse the role of AIP during development I generated a *Drosophila melanogaster* model of CG1847 deficiency. The bioinformatics data, protein structure and rescue experiments results prove without doubt that AIP and *CG1847* are orthologue genes.

The RNA-seq sets provided a useful insight into the types of gene which are highly changed in the CG1847 deficient larvae. In this study, I pooled RNA from mutant and control *Drosophila* larvae and performed RNA deep sequencing using the Illumina platform. A very interesting cluster of genes emerged from the analysis of the RNA-seq data: Osiris, Twedl and cuticular proteins were all significantly changed *Drosophila* mutants. To date, there are no information in the literature regarding possible interactions between CG1847 and proteins involved in body development, cuticle formation or body size. Clearly this cluster is in need of further investigation for the roles of these proteins during development as this association remains poorly understood for the moment.

Based on my observations, I propose a model for the roles of CG1847 in *Drosophila* development. CG1847 is involved in body size and cuticular formation by interacting with different group of proteins as Osiris, Twedl and cuticular proteins, interaction that were never described before. Even more loss of *CG1847* expression result in downregulation of heat shock proteins which might further enhance the mechanism of lethality.

The data obtain in this discovery study support promising research directions that could lead to better understanding of these genes complex roles in development.

The role of AIP in human development remains an open question as all presented results are limited to insect specific mechanisms.

However, the availability of these new data will facilitate the isolation and characterization of other functional genes involved in the role of the AIP orthologue during development and different pathways, as well as to disclose possible new mechanisms involved in the development of tumorigenesis.

Data resulted from this model will be further used in the next chapter, for a possible deeper understanding of involved pathways.

CHAPTER 4: THE *DROSOPHILA AIP* ORTHOLOG IS ESSENTIAL FOR ACTIN CYTOSKELETON STABILISATION AND CELL ADHESION

4.1 INTRODUCTION

The actual mechanism by which *AIP* silencing disrupts normal function in the pituitary is still unknown, despite the fact that causality between *AIP* mutations and pituitary adenomas has been established. Clinical, genetic and experimental data suggest that *AIP* functions as a tumour suppressor gene (TSG) in the pituitary gland[89,115,116,153]. Review of the published clinical data suggests that patients with *AIP* mutations have some special characteristics such as age of onset around 10-20 years, the pituitary adenomas are mostly sparsely granulated, and have a higher disease penetrance. FIPA families with *AIP* mutations have more affected individuals than families where no *AIP* mutations are found[136].

Tumours with *AIP* mutations are in the majority of cases macroadenomas, often invasive and disease onset occurs during adolescence or young adulthood[119,340], suggesting a more aggressive behaviour when there is *AIP* loss of function. Some authors found that AIP expression is lower in more invasive somatotroph adenomas, even in case of patients in which no *AIP* germline mutations have been identified[201,341]. 80% of the cases of AIP positive pituitary adenomas have a certain degree of extrasellar extension, 56% of them invading local structures. These pituitary adenomas typically do not respond well to somatostatin analogue treatment, *AIP* mutation-positive patients developing partial or total resistance to this therapy reflected in a significant reduction of GH and IGF-I levels and smaller, or no tumour shrinkage. Even the long-term disease control, with similarly cumulative numbers of therapies, is lower in these patients[119].

The fact that *AIP* is a tumour suppressor gene is also supported by the fact that loss-of-heterozygosity was detected in pituitary adenomas from FIPA patients[342,343]. In accordance with the Knudson's "two-hit" hypothesis[137] this resulted in the loss of the wild type *AIP* allele in almost all cases[116]. No FIPA patients homozygous for *AIP* mutations have been identified so far. In addition, the AIP tumour suppressor role was firmly established by functional *in vitro* studies: it was previously shown that siRNA *AIP* knockdown results in increased cell proliferation of GH3 cells[344].

Still, more functional investigations are required to elucidate the role of *AIP* mutations in pituitary tumorigenesis. To achieve this goal we need to know more about the identity and functions of AIP interacting partners. Molecules that can induce remodelling changes in the epithelial tissue of the normal pituitary gland are of a particular interest.

4.2 BACKGROUND

4.2.1 The *Drosophila* wing development

The developing *Drosophila* wing is a widely used model system for examining how cells respond to changes in cell shape associated with rearrangement of epithelial sheets during organogenesis[345].

Wing development is a relatively simple process described in detail by Fristrom *et al*. in 1993[346]. This involves the transition from a single layered columnar epithelium to a flattened bilayer. In the larval stages the wing imaginal disc is an epithelial sac. During pupal stages, this tissue undergoes a series of folding, unfolding and flattening processes[345].

The first step towards the final pattern of the wing takes place in the early phases of the prepupal period, when the monolayered epithelium is folded into two columnar layers[346]. This is **Apposition 1** and it is immediately followed by **Expansion 1**, a stage where the two-layered epithelium suffers a series of alterations leading to an increase in wing surface area. The expansion process proceeds gradually from the wing margins towards the middle[347]. The epithelium flattens and the wing final pattern becomes more obvious. During the **Adhesion 1** stage, cuticulin is secreted and no major wing shape changes occur. Cuticulin is the first layer of the insect cuticle to be deposited and it is laid down as a sheet over the apical surface of cuticle-secreting cells[348]. Between 16 and 18 hours after puparium formation, the reapposition of dorsal and ventral epithelia takes place (**Apposition 2** stage). At the same time, the pattern of wing veins becomes obvious. The wing veins are not properly formed and remain as open spaces filled with hemocytes (circulating immune surveillance cells). As the reapposition is still not complete at this stage, the wing has a “spongy” texture, but with a general shape and vein pattern of the adult wing[346]. The **Adhesion 2** stage follows this phase and lasts approximately 10h. The dorsal and ventral cell layers are completely adhered, as the cells of the 2 layers connect via basal junctions, but only in the areas between the wing veins. After approximately 45 hours of pupariation, the wings starts undergoing a final lateral expansion (**Expansion 2**). In order to fit into the pupa, the wings are folded in a characteristic manner[349].

The final step in wing formation takes place after eclosion from pupae. The wings unfold and expand. The intervein cells suffer an apoptotic process, as any viable intervein cell degenerate, while only the cells surrounding the wing veins persist forming the wing veins[350]. A summary of these processes is depicted in Figure 50.

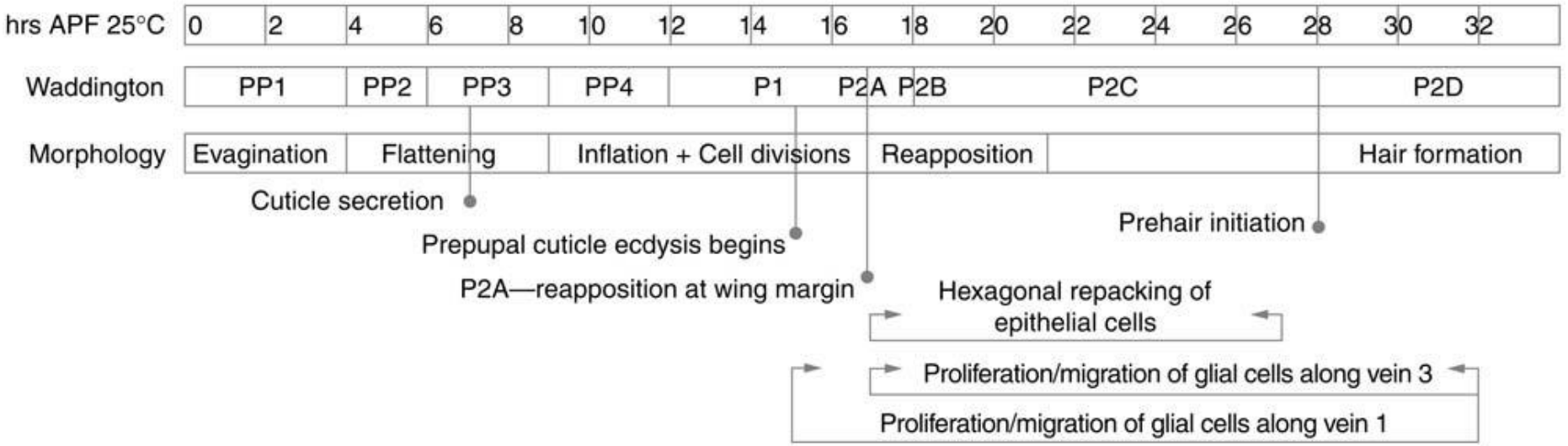

Figure 50: Early wing development. Timeline depicts morphological landmarks of wing development, the associated Waddington's stages, and the times at which they occur at 25°C temperatures. Development within the pupal case is divided into PP (pre-pupae) and P (pupae) stages that are separated by a molt. During evagination, in the first PP stage (PP1), the dorsal and ventral halves of the wing pouch begin to adhere along their basal surfaces and the wing epithelium folds along the future wing margin to assume an approximately semicircular shape. In PP2, the tissue elongates and narrows until it resembles a thick cylinder three times as long as it is wide (at about 5 h APF at 25°C). During PP3 (6–9 h APF at 25°C), the wing cells flatten and the wing dramatically increases its surface area. Cuticle deposition begins at this time, and forms a chitinous sheath during PP4, when the wing becomes inflated and the dorsal and ventral surfaces move apart. This leaves the wing looking like a balloon. During inflation, a wave of cell divisions runs through the wing. Ecdysis of the cuticle—the shedding of the chitinous cuticle from the apical side of the wing epithelium—starts late in the first P stage (P1). Waddington[351] subdivides the second P stage into substages A–D. In P2A, dorsal and ventral wing surfaces begin to reappose beginning at the wing margin, and the adhesion of the two basal surfaces spreads from the distal, anterior, and posterior ends of the wing during P2B and early P2C. P2D starts with the initiation of prehairs. Eventually hairs will cover the whole wing. Proliferation and migration of glial cells along vein 3 initiates in the beginning of P1 (by 15 h APF at 25°C). Migration, accompanied by proliferation of glial cells along vein 1 begins at approximately late P1 (by 17 h APF at 25°C). Glial cells cover the sensory nerves in vein 1 and 3 by mid P2D (32 h APF at 25°C). Abbreviations: APF, after puparium formation;hrs APF 25°C, hours APF at 25°C. (Adapted from Classen *et al*.2008)[352]

In the adhesion process the basal surfaces of the dorsal and ventral epithelia come to close contact and form basal junctions (BJs), which are mediated by the integrin family of cell surface receptors[353,354]. The extracellular domain of integrins recognises and binds to ECM proteins, whereas the intracellular domain associates with cytoskeletal elements[345] (Figure 51).

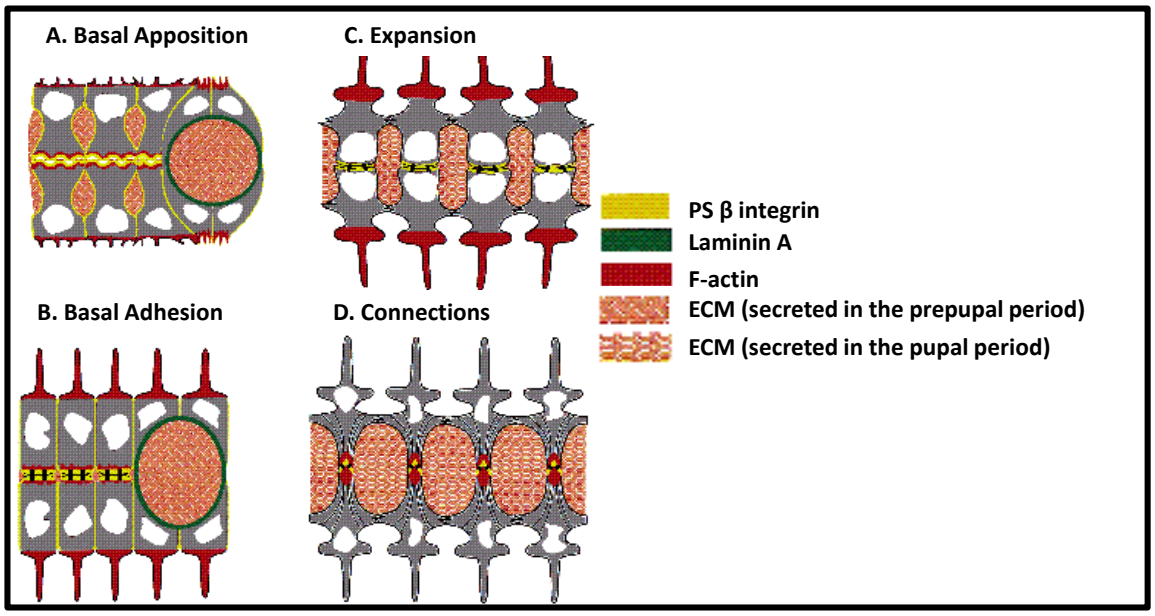

Figure 51: A schematic summary of adhesion mechanism. The localization of βPS integrin, laminin A and F-actin is depicted during the four key stages in the differentiation of pupal wings. White areas represent nuclei. Black bars (black arrows) connecting cells in B, C and D represent basal junctions. Wing veins (green) are shown only in A and B but persist unchanged for the remainder of development. Extracellular spaces containing matrix (ECM) of unknown composition in (A), disappear (B) and reappear (C, D). Adapted from Fristrom *et al*. 1993[346]

In the process of wing development integrins are essential for the cells to adhere to ECM components. Even more, these receptors are involved in the transmission of signals from the ECM to the actin cytoskeleton, hence influencing numerous cellular activities during wing development. One of the first molecular steps in the formation of the integrin–cytoskeleton link is the binding of Talin to integrins[355]. Talin is essential for further recruitment of other proteins such as ILK, PINCH and Paxillin.

4.2.2 The FRT/FLP system

As described in Chapter 3, I generated a heterozygous *CG1847* deficient *Drosophila melanogaster* animal model. One of my aims was to investigate and identify proteins that are deregulated in the absence of CG1847 *in vivo.*

A major restriction of the imprecise excision technique is the fact that, in the case of essential genes, animals can be maintained only as heterozygous stocks, which makes the study of homozygous gene deletions in adult tissue difficult. To overcome this problem, different techniques have been developed, one of which is the elegant FLP-out system which is based on site-specific DNA recombination[356]. The FLP recombinase is an enzyme native to a plasmid of *Saccharomyces cerevisiae* and rearranges DNA sequences in a very specific manner. The FLP recombinase acts on a particular 34 base pair DNA sequence, termed the FRT (FLP recombinase target) sequence. When two of these FRT sites are present on homologous strands, FLP creates double-stranded breaks in the DNA, exchanges the ends of the first FRT with those of the second target sequence, and then reattaches the exchanged strands. If this site-specific recombination occurs between the two chromosomes of the pair, in the G2 phase of the cell cycle in a proliferating mother cell, subsequent divisions will generate two clones of daughter cells that are homozygous for one or the other chromosomal arm (Figure 52). In my experiments, FLP recombinase was expressed under the control of a heat shock protein promoter[357].

Furthermore, the FLP/FRT system can also be used to produce genetic mosaics of marked loss of function, or gain of function clones in an otherwise heterozygous background[358].

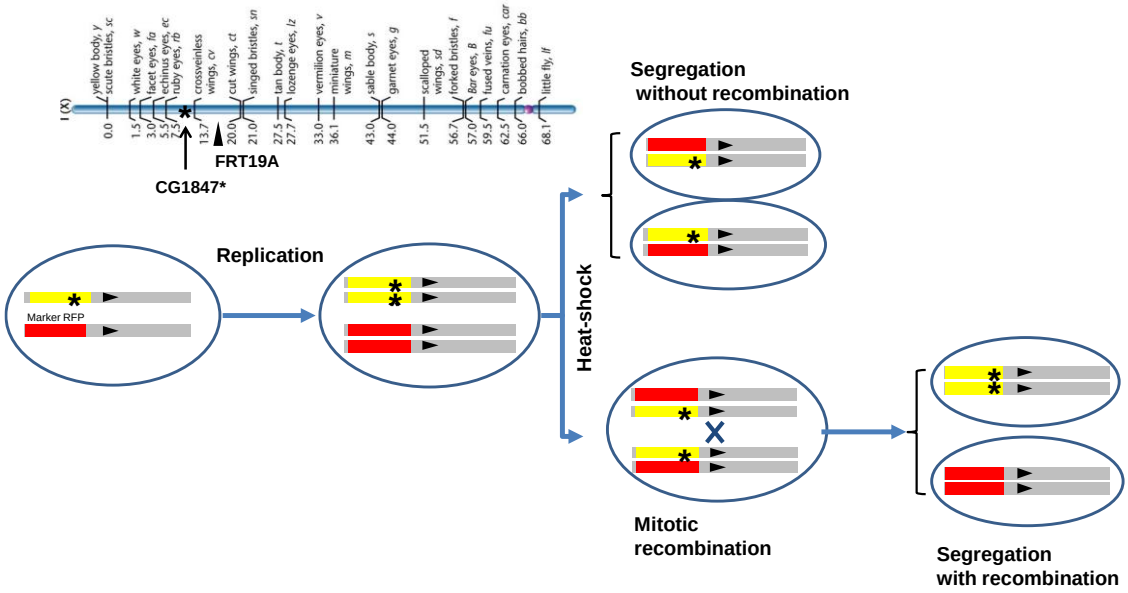

Figure 52: Schematic representation of the FLP/FRT system used to generate mosaic clones. The FLP enzyme catalyzes mitotic recombination at the FRT sites between homologous chromosomes; in this example one X chromosome carries a mutant allele (*) distal to the FRT site (triangle), the other chromosome carries a visible marker, RFP. After recombination and cell division, one daughter cell is homozygous mutant and lacking red fluorescence, the other cell is homozygous wild-type and carries both copies of the fluorescent marker (bottom right).

Pairs of daughter cells generated following a mitotic recombination event are adjacent to each other and throughout development are exposed to the same cellular environment. Thus, in the same tissue, clones of cells that represent a mutant experimental and a wild-type control are generated, whose subsequent proliferation and growth can be compared. I used the FLP/FRT system to generate mitotic clones of cells that are totally lacking *CG1847* expression in heterozygous females.

4.3 OBJECTIVES

To identify putative molecular partners of AIP with a role in tumorigenesis I investigated the molecular mechanisms of loss of *AIP in vivo* via:

a) Generation of RNAi-mediated *CG1847* knockdown under the control of the wing specific GAL4 drivers.
b) Validation of RNAi-induced phenotypes using FRT/FLP clonal analysis.
c) Live observation of β_{PS}-GFP integrin in *AIP* mutants during development.
d) Characterisation of the impact of the *CG1847* deficiency by immunostaining (where the necessary antibodies are available), in 2 stages of development:
    1) In third instar larval imaginal wing discs with homozygous *CG1847* deficient clones generated via FRT/FLP system.
    2) During pupal stages, by immunostaining *CG1847* deficient wings generated by RNAi-mediated gene silencing.
e) Validation of transcripts with possible involvement in blister formation identified in the RNA-seq screening.

4.4 RESULTS

4.4.1 *CG1847* depletion in the developing wing results in a loss-of-adhesion phenotype

To avoid the lethality associated with the use of ubiquitous drivers, I used tissue specific drivers to express RNAi constructs targeting *CG1847*. The *Drosophila* wing provides an excellent system for morphogenetic studies. The major advantage is the fact that the wing is not an essential organ and this allows important genetic manipulations without affecting viability. An additional benefit is the fact that almost any resulting phenotype is easily identifiable. I screened for possible pathways affected by CG1847 by eliminating CG1847 function in large areas of the wing

during development and examining adult wing phenotypes. I used two different Gal4 drivers to express the hairpin RNAi constructs in the wing imaginal disc: *nubbin–Gal4* (*nub-Gal4*) and *hedgehog–Gal4 (hh-Gal4)*[359]. While *nub-Gal4* leads to expression in the whole wing pouch*, hh-Gal4* promotes expression only in the posterior compartment of the wing.

Remarkably, in both cases I observed a very specific phenotype with *CG1847* RNAi flies displaying wing blisters (Figure 53).

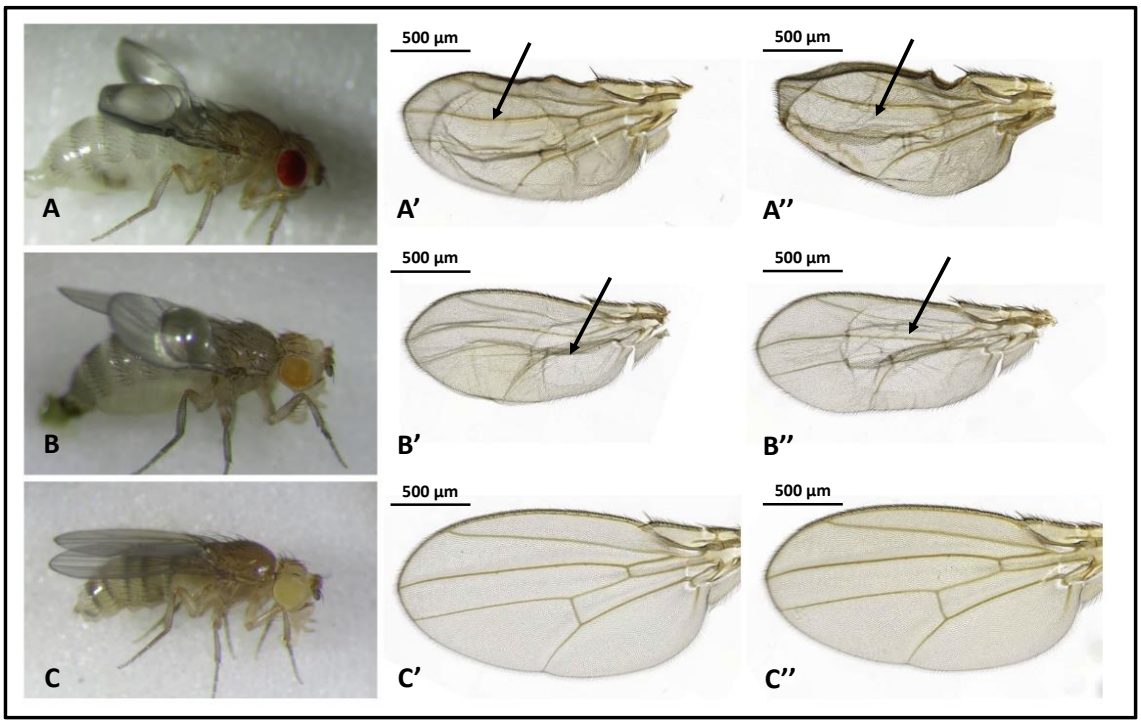

Figure 53: The CG1847 silencing in wing discs produces wing blisters. Expression of the UAS-*CG1847-RNAi-R2*, under the *hh-Gal4* driver (A) or *nubbin-Gal4* (B) drivers produces wings with blisters. C) UAS-*CG1847-RNAi-R2* crossed with *y w* (control) do not develop wing blisters. Live images and mounted wings. Wing blistering induced by *hh-Gal4* driver is much stronger. Scale bar 500µm.

Although the knockdown wings look smaller, their real size cannot be properly evaluated. The blister itself may put traction on the surrounding wing tissue and this mechanism may leads the appearance of wing shrinkage (Figure 54).

Interestingly, *nub-Gal4* led to a more moderate phenotype, as the blisters affected only the wing hinge area, but with a very high percentage, as approximately 95% of the adults presented this phenotype. In contrast, *hh-Gal4* resulted in a stronger phenotype, with blisters affecting the whole wing, but only in 77.5% of the cases (compare blister size, shape and localisation in Figure 53 A and B and Figure 54).

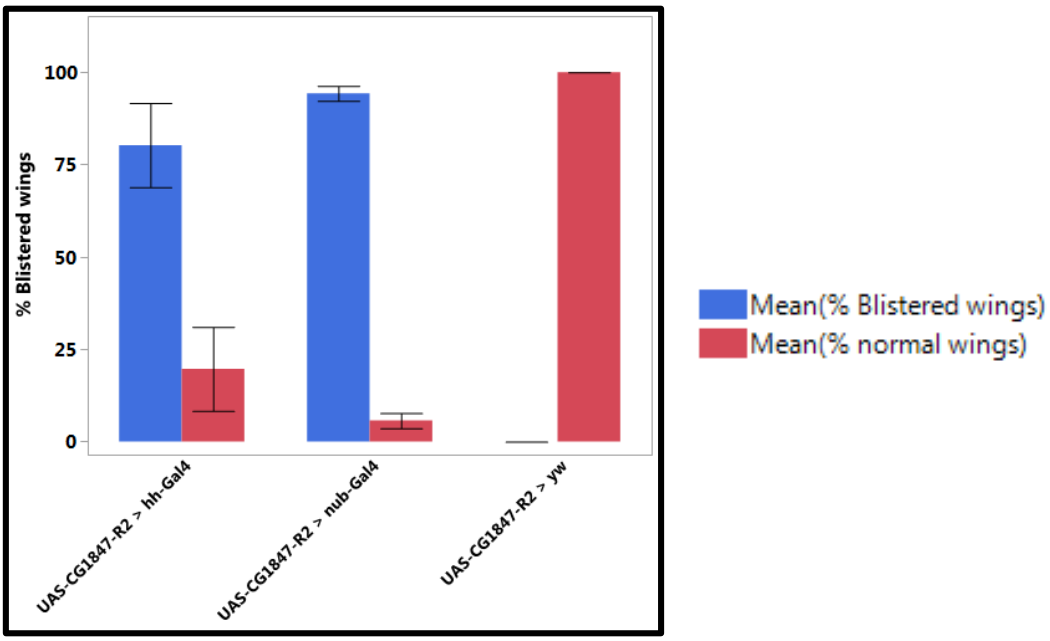

Figure 54: Quantification of blister formation. Expression of *CG1847-RNAi-R2* under the *hh-Gal4* resulted in 77.5% of blisters in *Drosophila* wings. Expression of *CG1847-RNAi-T2* with *nub-Gal4* driver resulted in 95% of blisters. *CG1847-RNAi-T2* crossed with *y w* was used as control. Error bars represent SE

To confirm that these results were not due to an off-target effect of the RNAi constructs, I repeated the experiment using a different RNAi line (UAS-*CG1947-RNAi T2*)[247]. Comparable results were obtained as flies expressing *CG1847 T2* RNAi also displayed wing blisters with both *hh-Gal4* and *nub-Gal4* (Figure 55). Similar, *hh-Gal4* resulted in a stronger phenotype and while *nub-Gal4* led to a higher penetrance of blister formation (Figure 56).

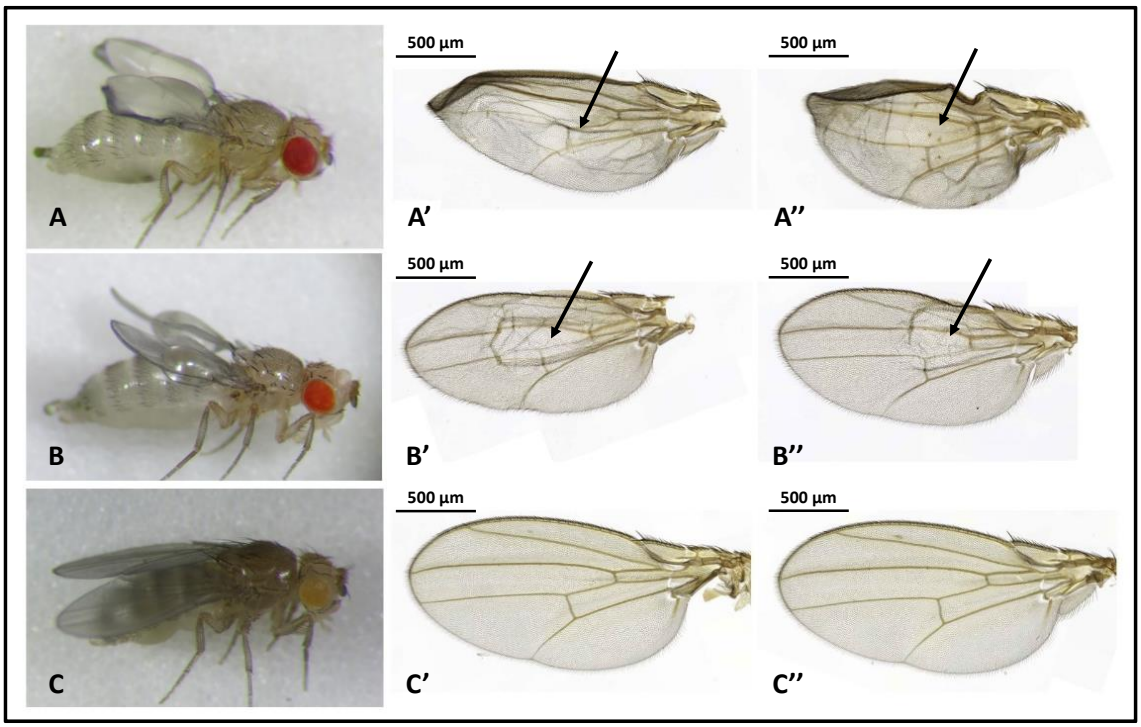

Figure 55: The CG1847 silencing with UAS-*CG1847-RNAi-T2* produces wing blisters. Expression of the UAS-*CG1847-RNAi-T2*, under the *hh-Gal4* driver (A) or *nubbin-Gal4* (B) drivers produces wings with blisters. C) UAS-*CG1847-RNAi-R2* crossed with *y w* (control) do not develop wing blisters. Live images and mounted wings. Scale bar 500µm.

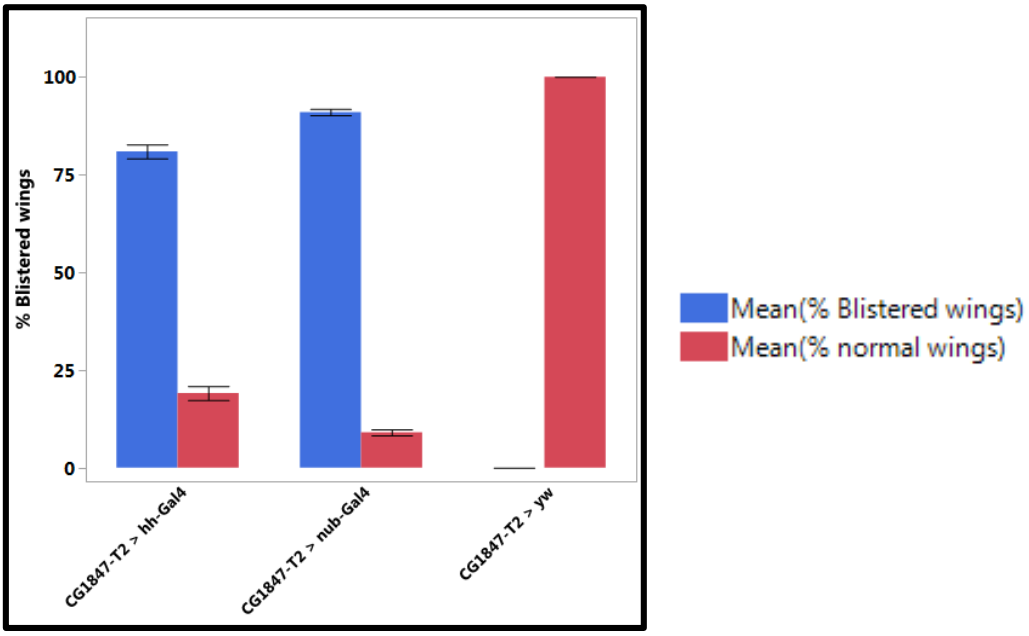

Figure 56: Quantification of blister formation. Expression of *CG1847-RNAi-T2* under the *hh-Gal4* resulted in 81% of blisters in *Drosophila* wings. Expression of *CG1847-RNAi-T2* with *nub-Gal4* driver resulted in 91% of blisters. *CG1847-RNAi-T2* crossed with *y w* was used as control. Error bars represent SE

These results suggest that the wing blister phenotype is due to *CG1847* depletion and not to an RNAi off target effect.

This phenotype was unexpected and it can be the result of a few possible causes. This phenotype strongly resembles that of loss of integrin function[345], suggesting that depleting CG1847 in the wing may have a negative effect on the integrin-actin cytoskeleton network. Wing blistering may also result from a gain of integrin function[360], which raises the possibility that CG1847 may in fact promote integrin function. However, the wing blistering phenotype is not unique to defective integrin-actin-cytoskeleton signalling and, thus, detecting the pathway that is affected by loss of CG1847 function may be challenging.

4.4.2 *CG1847* mutant clones cause wing blister formation

As the RNAi constructs carry the possibility of an off target effect, to validate the defective adhesion phenotype seen in *CG1847* RNAi flies were undertaken mitotic recombination experiments. This technique brings the significant advantage of being possible to compare mutant and control cells in the same tissue of a heterozygous animal.

To generate *CG1847* mutant clones, a recombination site was introduced in the *CG1847* mutant background (*FRT19A*^{neoR} (section 2.2.5 in Materials and methods)). Heterozygous *CG1847*$^{exon1\_3}$, *FRT19A / FM7c,dfd::YFP* recombinant females were mated with *Ubi-mRFPnls*, *hsFLP*, *FRT19A*^{neoR}*/Y* males to generate *CG1847*$^{exon1\_3}$*FRT19A*^{neoR}*/ Ubi-mRFPnls, hsFLP, FRT19A*^{neoR} females (the short version of this genotype will be *CG1847*$^{exon1\_3}$, *FRT19A /FRT19A).* Four hour egg collections were performed and submitted to heat shock treatment during early larval stages as described in section 2.2.3.

Mitotic recombination events were induced randomly, all over the body, via heat shock exposure in the very early larval stages. The resulting homozygous deficient mutant clones and twin wild type clones were examined in the wing imaginal discs of third instar larvae.

Mutant clones in wing imaginal discs are shown in Figure 57. The *CG1847*$^{exon1\_3}$ mutant regions lack the RFP marker, which readily differentiates them from the wild-type twin clones expressing two copies of RFP (and 2 normal copies of CG1847) or the heterozygous areas expressing one copy of RFP (and one copy of CG1847).

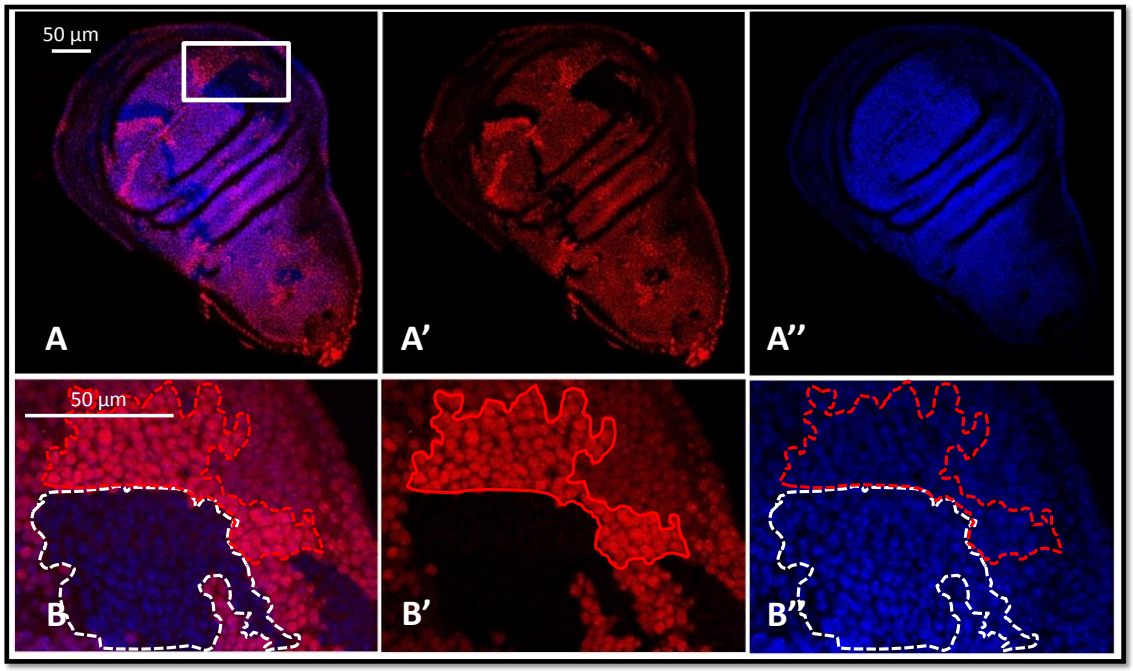

Figure 57: Homozygous *CG1847* mutant cell clones induced by the FLP/FRT system. Shown are wing imaginal discs from third instar larvae. A-A'' Low magnification (10X) view of the entire wing imaginal disc. B-B'': High magnification (63X) views of boxed area in A. Nuclei are stained with DAPI (blue). The tester stock expresses RFP (red fluorescence protein), which labels wild-type cells that have undergone mitotic recombination (bright red) and heterozygous cells (pale red). Homozygous mutant cells lack RFP expression and are therefore seen as a black area on A' and B'.B and B'': white area – the homozygous mutant clone, red area – wt clone. Scale bar 50µm

As shown in Figure 57 A', the mitotic clones are distributed throughout the wing disc, although their relative size is quite variable. As the objective was to validate the *RNAi* results (blister formation in adult wings), larvae were allowed to develop until adulthood. The wings of heterozygous *CG1847*$^{exon1\_3}$, *FRT19A /FRT19A* females were evaluated for blister formation. These females with homozygous mutant clones recapitulated the blister phenotype obtained in the *RNAi* experiments (Figure 58).

As a control, I used flies carrying the FRT construct but lacking any genetic mutation. These flies where similarly exposed to the heat shock treatment to generate neutral clones. In sharp contrast to flies displaying CG1847 mutant clones, control flies did not exhibit wing blisters. (Figure 58 and Table 21).

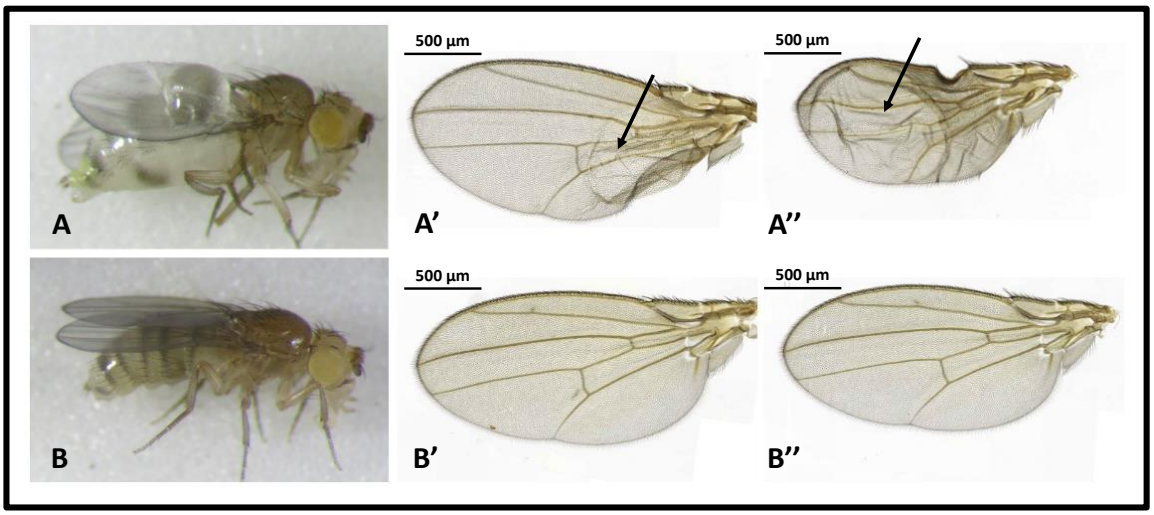

Figure 58: Blister formation in adult females with mitotic clones. Blister (arrows) formation in adult heterozygous *CG1847^{exon1_3}*, *FRT19A /FRT19A* females Top panel: heterozygous mutant females A) live image of animal with wing blisters; A' and A''- mounted wing with blister. Bottom panel: control in *w,Dm,FRT19A /FRT19A* B) live image; B' and B''- mounted wing. Scale bar 500µm.

The heat shocked heterozygous *CG1847^{exon1_3}*, *FRT19A /FRT19A* females survived to adulthood some of them developed wing blisters, exactly the same phenotype identified in the UAS-*CG1847-RNAi* knockdown animals. I analysed all eclosed adults from 3 independent experiments and calculated the percentage of heterozygous *CG1847^{exon1_3}*, *FRT19A /FRT19A* females with blisters (Table 20).

	Males			Females			
Offspring genotype	*CG1847^{exon1_3}*FRT19A Y		*FM7c dfd YFP* Y		*CG1847^{exon1_3}FRT19A FRT19A*		*FRT19A FM7c dfd YFP*
Expected phenotype	**NOT VIABLE**	**No Blisters**	**Blister**	**No Blisters**	**Blister**	**No Blisters**	**Blisters**
TOTAL	**0**	**162**	**0**	**358**	**34 (9.5%)**	**352**	0

Table 20: Quantification of blister formation in CG1847^{exon1_3}, FRT19A /FRT19A females. The numbers represent the total members of counted offspring in 3 different experiments. In parantheses the percentages of heterozygous mutant females with blisters are shown.

Only approximately 10% of the heterozygous females with mitotic clones developed blisters, while this phenotype was not identified in any of the other genotypes acting as internal negative controls. One noteworthy finding was the wide variability in the size, number and shape of the blisters in adult females. This is likely due to the fact that the clones were induced with heat shock and are therefore randomly localised within the whole animal. No blisters were detected in flies where neutral clones were induced (Table 21).

	Males		Females	
Offspring genotypes	w+,Dm,FRT19A^{neoR} / y		w+,Dm,FRT19A / RFPnls HsFLP FRT19A	
Expected phenotype	No Blisters	Blister	No Blisters	Blister
TOTAL	**121**	**0**	**152**	**0**

Table 21: Quantification of blister formation in control *w,Dm,FRT19A /FRT19A* females

The small percentage (10%) of heterozygous females with mitotic clones developing blisters may be due to the size variability of the clones induced by the FLP/FRT system as seen in Figure 57. The majority of clones may be too small to induce the detachment of the wing layers. Clone size in an adult wing is determined by the developmental stage at which the larva is heat-shocked. For example, clones induced earlier in imaginal disc development are larger, whereas those induced later are smaller and more numerous. The reason is that later in development a larger number of cells may undergo recombination; however, the number of mitotic divisions that occur after the heat shock is reduced[356]. I attempted to optimise the timing and duration of the heat shock treatment, but heat shock treatment before 48 h of development or for longer periods of times resulted in lethality of the larvae.

As *AIP* is described as a tumour suppressor gene, I also analysed the effect of *CG1847* loss on cell growth. This was evaluated by examining and comparing the size of clonally related groups of cells resulting from mitotic recombination (Figure 57).

CG1847 mutant tissue was not overtly different from the neighbouring wild-type cells. Following a more detailed analysis using ImageJ to carefully assess the area of 25 mutant clones and their twin associated wt clones, the *CG1847* deficient clones were on average larger. As shown in Figure 57A' and Figure 59 this size difference was statistically significant. It is necessary to mention that there was a large variability in the size of clones, both wt and CG1847 depleted.

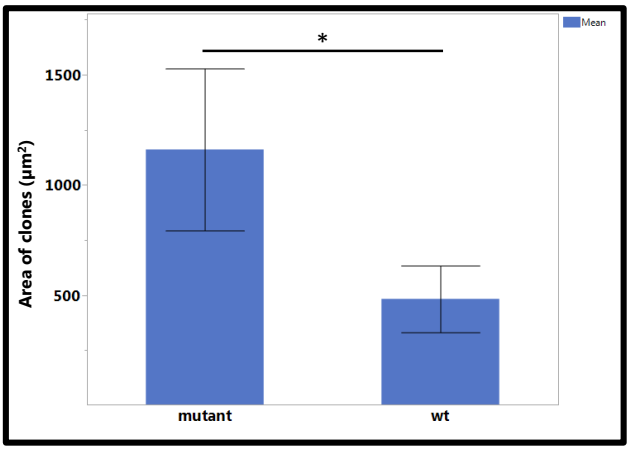

Figure 59: Quantification of the area size of of mitotic recombined clones. The area of mutant cells is signficantly increased comparing to the their neighbouring twin wild-type cells (P=0.0293) Statistical analysis was carried out using Wilcoxon/Kruskal-Wallis test. The bars represent the mean of are of the clones. N=25 clones from 10 imaginal wing discs. Error bars represent SE (* = P<0.05).

The area of *CG1847* mutant clones of was found to be significantly larger than the wild-type controls. This is in agreement with the fact that AIP has been described as a tumour suppressor gene. However, the underlying mechanism is still unknown, as AIP could be involved in cellular growth, cell survival and/or cellular proliferation.

To investigate in more detail the role of CG1847 in cell growth, I evaluated the cell density in *CG1847*$^{exon1\_3}$ mutant clones (Figure 60). Discs containing *CG1847*$^{exon1\_3}$ mutant clones were stained with the DNA binding dye DAPI for nuclei visualisation, and phalloidin staining to mark the cell boundaries.

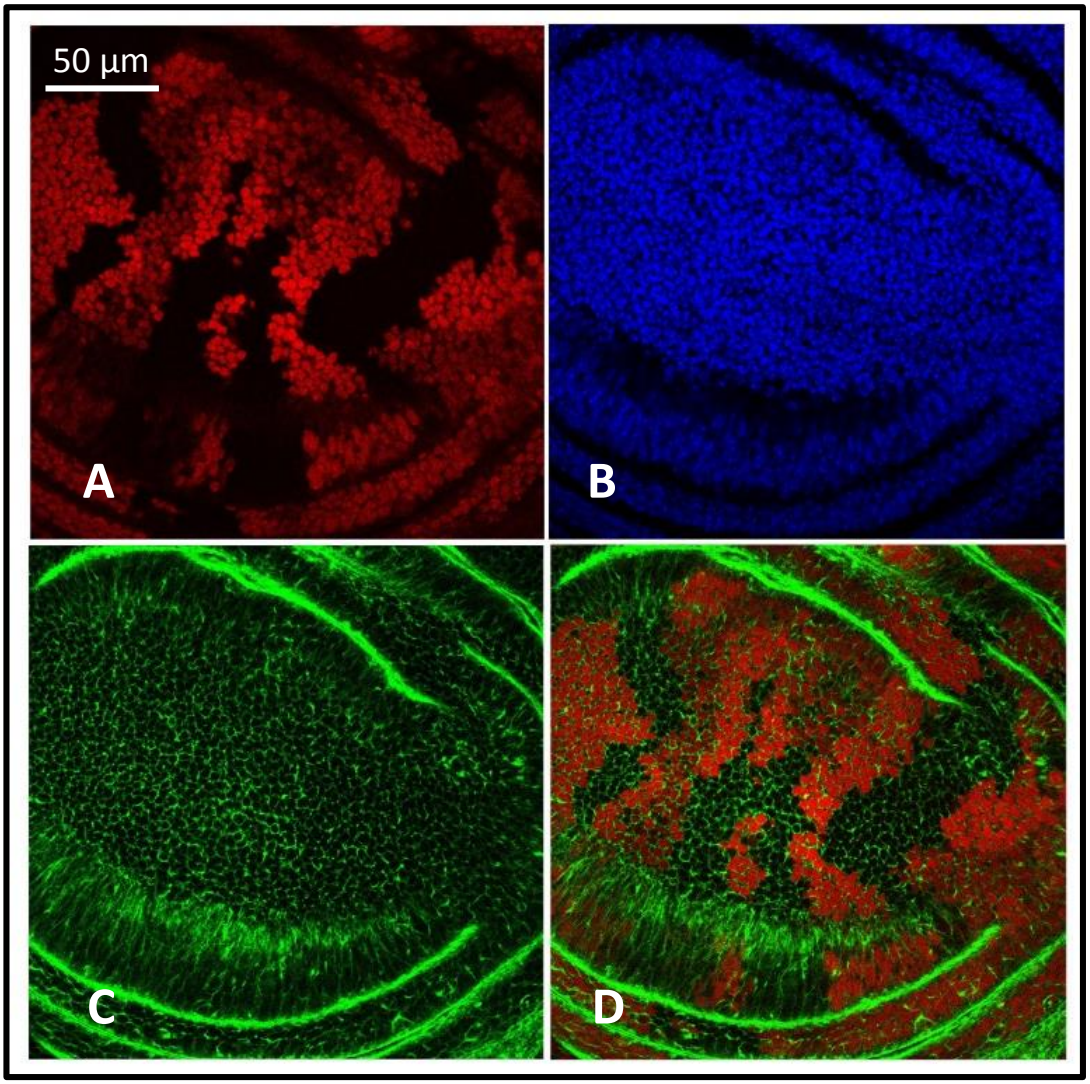

Figure 60: Wing disc clones stained with DAPI and phalloidin. 120 h AEL wing imaginal discs containing *CG1847*$^{exon1\_3}$ mutant and wild-type clones stained with DAPI (blue – B) to mark nuclei and phalloidin (green – C) to mark the cell bounderies. RFP expression (A) was used to identify the different clones (40X magnification). D: Merged chanells phalloidin and RFP. Scale bar=50 µm.

The size of cells within *CG1847*$^{exon1\_3}$ mutant clones (Figure 60) was examined, and compared with the size of the twin wt clones (for more details regarding the protocol and the antibodies see Appendix 6). I evaluated the number of cells in *CG1847* mutant clones and compared it with the number of cells in their wt twin spot counterparts by counting the number of nuclei per clone. I also determined the clone area using the RFP signal (or lack of it in case of homozygous mutant clones). Cell density was calculated by dividing the number of cells to the respective area. Homozygous mutant cells were slightly bigger than their wt neighbours, but without reaching a statistical significance (Figure 61).

The statistical analysis revealed that mutant cells are not visibly different in size from their twin spot wild-type neighbours (Figure 61) and counts of nuclei present within mutant and wild-type clones showed no difference in cell density (*CG1847*$^{exon1\_3}$ clones = 4.41×10^{-2} cells/µm^2; wild-type clones = 4.46×10^{-2} cells/µm^2).

Together, these results support the hypothesis that CG1847 exerts a tumour suppressor role, as CG1847 mutant clones are larger than controls. However, CG1847 does not seem to be involved in individual cell growth, and thus it may instead affect cell proliferation or cell survival.

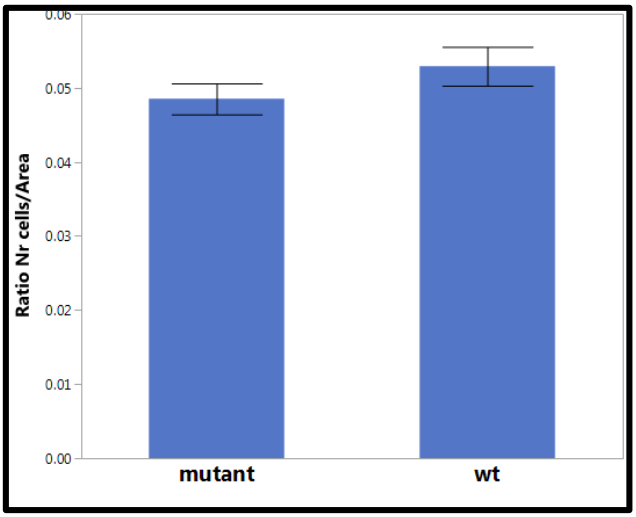

Figure 61: Quantification of cell density. Imaginal wing discs were stained with DAPI and phalloidin, and the area (µm^2) and the number of cells within each homozygous *CG1847*$^{exon1\_3}$ mutant and twin wt clone were counted. The cell density was found decreased in mutant clones, however, without reaching a statistical significance. The bars represent the mean of ratio: number of cells/area). N=25. Error bars represent SE

In summary, our results from RNAi and mitotic recombination support the involvement of *CG1847* in adhesion, proliferation and cell survival processes. As a role in adhesion has never been described for human AIP, there is virtually no information regarding the possible mechanism of action. Based on the available literature regarding blister formation in *Drosophila*, the candidate AIP-interacting partners identified belong to various classes, such as: **cell adhesion genes** *Mys* (FBgn0004657), *Mew* (FBgn0004456), *If* (FBgn0001250)[361] *blistery* (FBgn0000244)[362]; **receptors and members of signal transduction pathways**: *Gsα*[363];

transmembrane transport: *blot* (FBgn0027660)[361]; **cell migration:** ADAM metallopeptidase with thrombospondin type 1 motif A (FBgn0038341)[361]; **DNA and RNA binding proteins:** *held out wings* (FBgn0264491)[361]; *Additional sex combs* (FBgn0261823)[364]; **protein deubiquitination and regulation of proteasome assembly:** Ubiquitin specific protease *USP5* (FBgn0035402), *POMP* (FBgn0032884)[361]; **metabolic processes**: *sugarless* (FBgn0261445)[364]; **chitin based attachment**: *piopio* (FBgn0020521)[364], and **transcription factors:** *mastermind* (FBgn0002643)[364]*, blistered (FBgn0004101)*[365]. These genes were identified in three different screening studies. The first one was conducted by Prout *et al*.[364] and was trying to identify autosomal mutations which might result in blisters formation fruitfly wing. 76 independent mutations were identified in this study and the genetic interactions with mutations in the integrin gene *myospheroid* were investigated. Mutations in three new genes (*piopio, rhea and steamer duck*) that affect myo-epidermal junctions or muscle function in embryos were isolated. One year later, in 1998, Brown *et al*.[361] published the results of the second screening. The aim of their project was to use FLP-FRT system to generate clones of randomly induced mutations and to screen for those mutations that cause wing blisters.

A third study was recently published in 2014 and the authors also looked on Drosophila apposition of the dorsal and ventral wing sheets during metamorphosis. Using RNAi-silencing technique and the blister phenotype as readout, there were identify numerous novel proteins potentially involved in wing sheet adhesion: components of other cellular processes, e.g. cell cycle, RNA splicing, and vesicular trafficking[366].

Identifying the mechanisms via which loss of *CG1847* leads to loss of adhesion may provide hints on the pathology of human pituitary adenomas due AIP mutations.

4.4.3 CG1847 is not required for βPS integrin distribution at muscle attachment sites

The most studied mechanism of blister formation in *Drosophila* involves the integrin–actin cytoskeleton network. Proteins belonging to this network, such as the integrins receptors, are not only localised in the wings, but they are involved in cell adhesion in many other tissues, including the connection between muscles and between muscle and tendon cells[367,368].

To study the possible interaction between integrin-associated proteins and CG1847, the expression of different proteins involved in the integrin adhesion pathway was evaluated during development by 2 methods: live imaging in embryos and larvae, and immunostaining during larval and pupal stages.

Previously published studies have shown that *Drosophila* mutants affecting cytoskeleton components are homozygous lethal due to the detachment of somatic muscles at late embryonic stages[369]. As the *Drosophila AIP* mutant is also lethal and it may be involved in adhesion processes, I assessed whether the lethality was associated with defects in integrin distribution at muscle attachment sites (Figure 62).

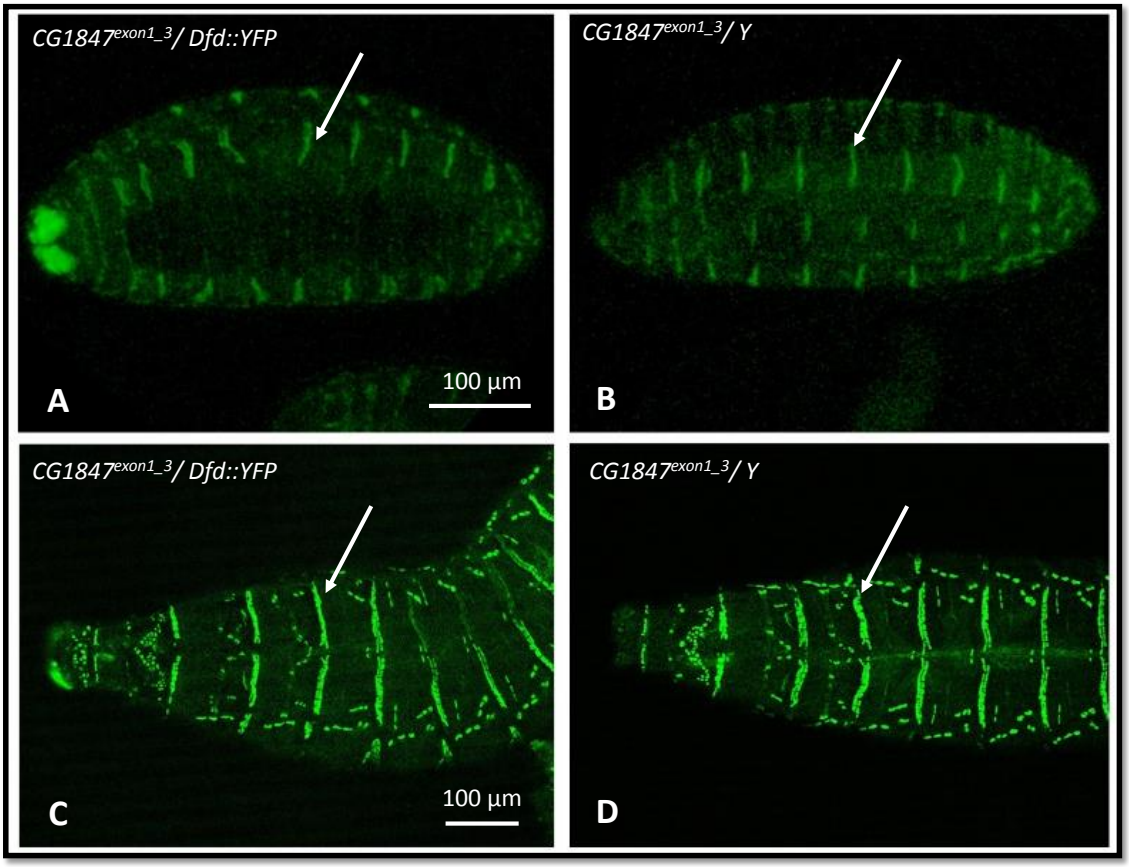

Figure 62: CG1847 deficient mutants display normal muscle attachment sites pattern. Live imaging of GFP-βPS integrin (white arrow) in females *CG1847^{exon1_3}/FM7c,dfd::YFP* (A and C) and males *CG1847^{exon1_3}* (B and D) with confocal microscopy. A) and B): 24 hours old embryos. C) and D): around 50 (± 2) hours old larvae AEL. Scale bar=100 µm.

For this, I used fly stocks where the *CG1847* deletion was recombined with *Mys*-GFP (*myospheroid* (Mys) encodes for the β_{PS} integrin subunit) and balanced over the same *FM7c, Dfd::YFP* chromosome (section 2.2.4). This allowed me to determine whether lack of CG1847 impacts on the expression or localisation of GFP- β*PS* integrin[246] in *CG1847^{exon1_3}/Y* males embryos. Females *CG1847^{exon1_3}/ FM7c, Dfd::YFP* were used as control for the normal *Mys* distribution.

Live imaging of embryos lacking *CG1847* failed to reveal any defects in adhesion or displacement of β*PS* integrin from muscle attachment sites.

4.4.4 CG1847 deficient mutants display normal expression of integrins and integrin-associated proteins in the wing imaginal discs during larval stage

As the CG1847 depleted adults display blister formation I examined the distribution of integrins and integrin associated-proteins in wing imaginal discs of *CG1847*$^{exon1\_3}$*, FRT19A /FRT19* females (Figure 63). For this purpose I performed immunostaining with the relevant available antibodies (Appendix 6).

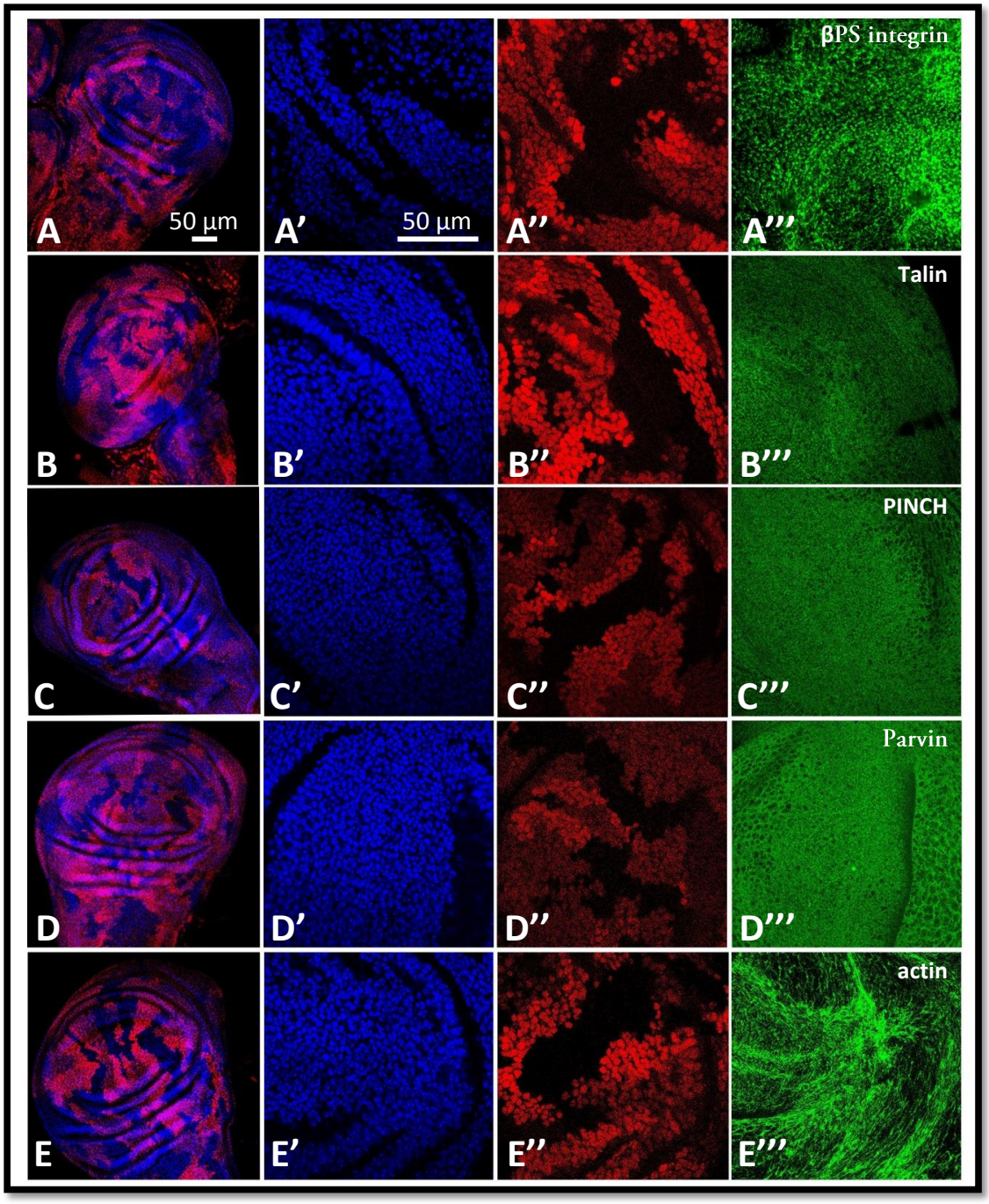

Figure 63: Normal distribution of βPS integrin, integrin-associated proteins and actin at the basal surface. (A-E) Low magnification (10X) view of mitotic recombination clones in *CG1847*$^{exon1\_3}$, *FRT19A* /*FRT19A* female wing discs. A' A'' and A''' show disc in A at 63X magnifications. Similarly for the other panels. Note in A-E the wide variability in size, shape and location of the mitotic clones. A' - E''': wing discs stained for DAPI (blue; A' and E'), expressing mRFP (red; A'' and E''), and stained for integrin associated proteins (A''' and E'''). respectively. Note that DAPI and mRFP are images of the same slice in the Z-stack at the nucleai level, while A''' and E''' are images at the basal surface, where these proteins are normally expressed and invoved in cell adhesion. Scale bar 50 μm.

In clones of cells lacking *CG1847*, βPS integrin, Talin, Parvin, PINCH and actin expression was compared with their expression in clones of homozygous wild-type cells. No obvious alterations were observed in the expression pattern of these proteins.

4.4.5 *CG1847* knockdown induces disruption in actin cytoskeleton networks

During larval stages, the *Drosophila* wing imaginal discs are formed by a single layer of cells, and as the basal junctions (BJs) are not formed in this stage, it is not surprising that no changes were observed. As the folding and adhesion processes occur exclusively during the pupal stages, I performed subsequent experiments in pupal wings to investigate the mechanism behind the loss of adhesion in *CG1847* mutants.

To address whether the members of integrin-actin cytoskeleton network are deregulated later in development in the pupal stage, I silenced *CG1847* expression in the developing wing using *nub-Gal4*. This driver was used as it generated blister formation in more than 90% of the offspring and with a more precise localisation, as all the blisters were in the hinge area of the wings (Figure 54 and Figure 56). Pupal wings were stained with the DNA binding dye DAPI for nuclei visualisation, and phalloidin to mark the actin cytoskeleton. I evaluated the pattern of actin distribution at two different developmental stages (Figure 64). To choose the time points for this evaluation, I took into consideration the normal wing development and morphogenesis. At 16 h APF, the wing already has two layers, but the dorsal and ventral layers are completely separated, and no wing veins can be distinguished. The actual apposition starts at approximately 18h APF, when wing veins become more obvious. Between 20 h APF and 32 h APF (by the definitive stage), the adhesion process progresses from the tip of the wings towards the hinge and the wing veins are clearly formed and distinct from the surrounding tissue. Therefore, I chose to evaluate the effect of *CG1847* silencing in pupal wing morphogenesis and adhesion at 24 h and 28 h APF. These particular stages were chosen because at 24 h the adhesion process has already started and differences in the attachment of the cell layers can be evaluated in *CG1847*-depleted and control. At 28 h APF, the adhesion process should be reaching completion,

thereby allowing me to detect if there is a developmental delay resulting from *CG1847* knockdown. When considering the timing of the stages of development, different laboratories have reported significant variations: for example, Urbano *et al*. in 2009[370] reported apposition and refinement at later time points (32-40h). Such differences may be influenced by genetic background, and differing laboratory conditions (e.g. if animals were raised at room temperature, rather than 25°C as in this work).

For tissue specific knockdown I used the *nubbin-Gal4* stock as this resulted in wing blister in more than 90% of the offspring (section 4.3.1). As a control pupal wings from UAS-*CG1847-RNAi-T2* stock were used, in order to have the same genetic background. Both knockdown and control stocks were submitted to the same food and temperature environment, and pupae collection and staging were performed as described in section 2.2.11.

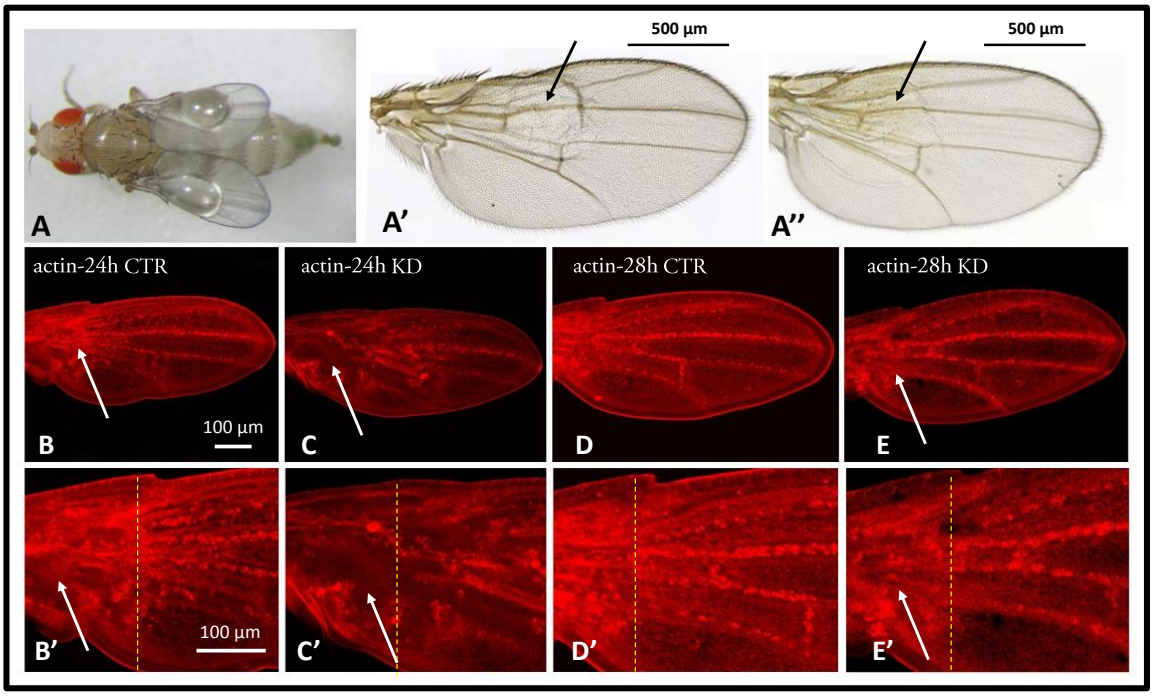

Figure 64: CG1847 silencing resulted in actin disorganisation A-A'' *CG1847-RNAi-T2* overexpressed under *nub-Gal4* driver induces blister formation (black arrow) in the hinge area of adult wings ; A) *nub-Gal4 CG1847* RNAi adult fly; A'-A'' adult wings from *nub-Gal4 CG1847* RNAi flies . B-E: low magnification (20X) visualisation of the entire wing blade in pupal stages. The genotype and developmental timing (h APF) are shown on the top. B and C: actin pattern in normal and *CG1847* knockdown pupal wings at 24 h APF. D and E: actin pattern in normal and *CG1847* knockdown pupal wings at 28 h APF.B'-E' shows higher magnification of the hinge area. By 24 h APF the adhesion process is not completed (B') as wing vein formation stops at the area marked by the yellow line At 24 h APF the adhesion process in knockdown wings (C') is severely disrupted (white arrow) compared with normal development. At 28 h APF the adhesion of dorsal and ventral wing layers is completed in control wings (D'), while in *CG1847* knockdown wings it is slightly halted (small gaps marked by white arrow).

Interestingly, loss of *CG1847* resulted in a marked disorganisation of the actin cytoskeleton. Of note, the severity of the cytoskeleton disorganisation phenotype was variable. At 24 h APF the majority of pupal wings displayed very strong phenotypes as the ones depicted in Figure 64, while others had a more mild loss of adhesion. Surprisingly, this phenotype seems to be partially recovered by 28 h APF, as the actin cytoskeleton organisation is only mildly affected in *CG1847* knockdown wings.

It has previously been demonstrated that AIP interacts directly with cellular cytoskeleton structures[156], but these results have not been convincingly validated. The results represent the first time that AIP loss has been associated with a strong cytoskeleton-related phenotype *in vivo*.

However, as aforementioned, not all wings exhibited the phenotype to the same extent, even at the earlier time point of 24 h APF. As the loss of adhesion phenotype seemed to be partially recovered by 28 h APF, I performed a more detailed analysis and specifically focused on the actin cytoskeleton at the basal surface of cell layers, where the apposition and adhesion take place.

For this analysis, I assessed the hinge area of 24 h and 28 h APF pupal wings. I evaluated the pattern and intensity of actin staining visualized at higher magnification (63X). The samples were optically cross-sectioned by generating orthogonal sections from the Z stacks (Figure 65).

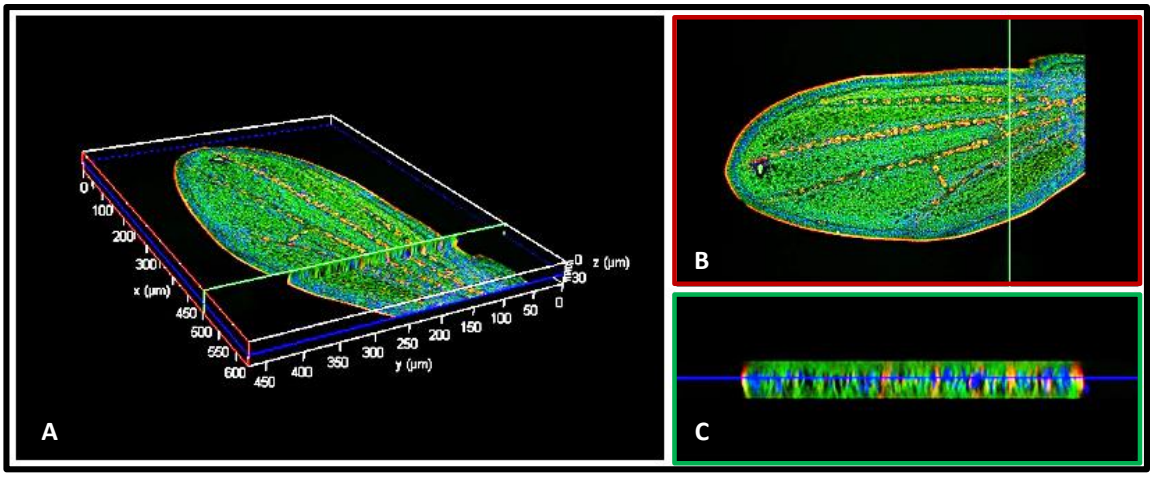

Figure 65: Schematic representation of the orthogonal sections. The orthogonal sections from the Z stacks (A) were generated via 3D reconstruction. XY view is at the level of the red line/section (B). YZ view is at the level of the green line/section (C) The colours of the borders in the B and C panels correspond to the colours of the optical sections in the A) panel.

At high magnification (Figure 66), I observed that *CG1847* depletion caused not only a failure of wing layer adhesion, but also a strong deregulation and lower intensity of actin staining.

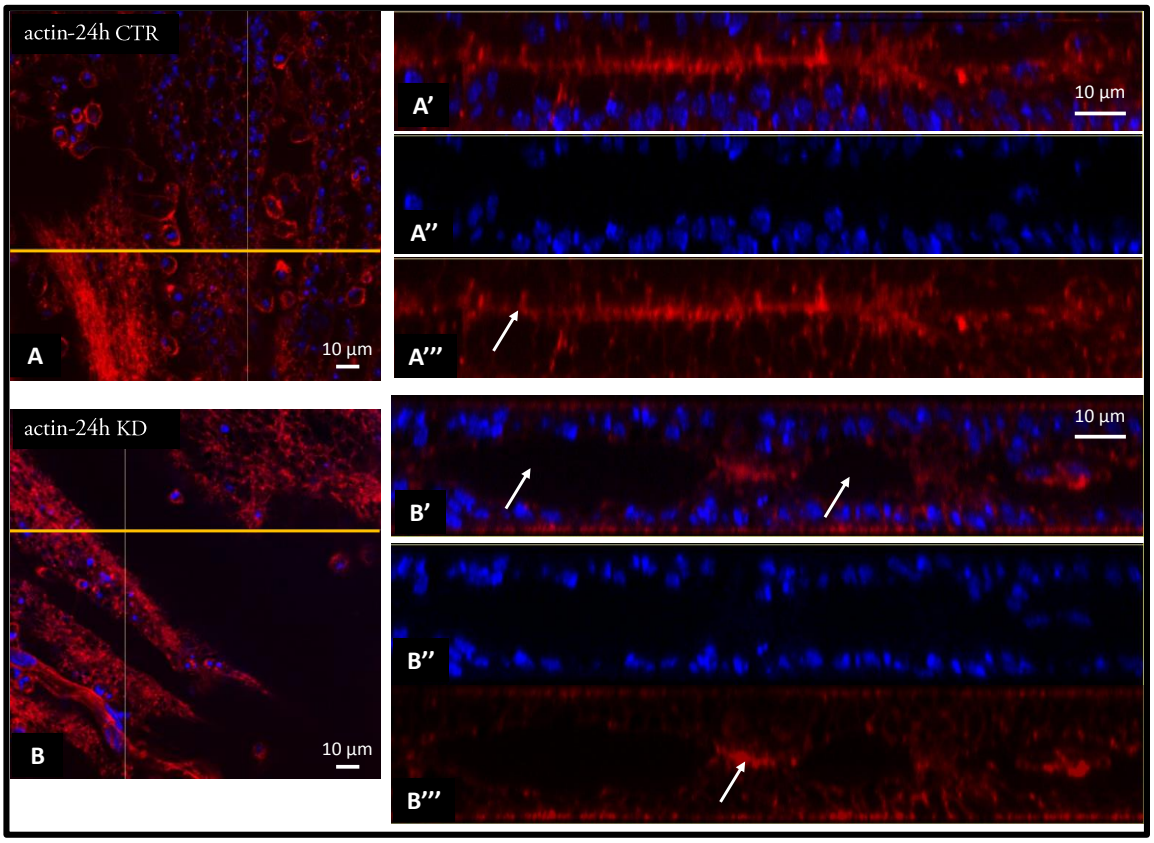

Figure 66: *CG1847* silencing specifically in the wing resulted in loss of adhesion. Gaps between the basal surfaces of wing cell layers in the hinge area are revealed in the Z sections of 24 h APF pupal wings. The *nub-Gal4>UAS-CG1847-RNAi-T2* were raised at 25°C, collected at WPP, staged and imaged at 24 h APF. A and B: high magnification (63X) single sections from Z-stacks of KD and control pupal wings.A'-A''' and B'-B''' are orthogonal views generated with ImageJ, at the level of the yellow lines. A' and B' are merged DAPI (blue) and phalloidin (red) channels. A''-A''' and B''-B''' are individual XZ sections of DAPI and phalloidin staining. Arrow in A''' points toward the actin continous layer involved in the adhesion (control wing), while arrows in B' shows the gaps resulting from loss of adhesion. The low intensity of phalloidin staining it is also obvious in KD wings, suggesting that reduced actin levels are involved in adhesion in KD wings at 24 h APF (arrow in B'''). Scale bar 10 µm.

As the loss-of-adhesion phenotype seems to be repaired overtime, at least macroscopically, by 28 h APF, I used the same approach to investigate the phenotype in more detail. However, when analysed at a higher magnification, *CG1847*-silenced wings still displayed gaps between the cell layers at 28 h APF. Moreover, *CG1847* RNAi wings exhibited a much lower intensity of phalloidin staining, indicative of problems in the actin cytoskeleton, and a dysregulation of the normal wing pattern, with nuclei not linearly arranged and at different levels (Figure 67).

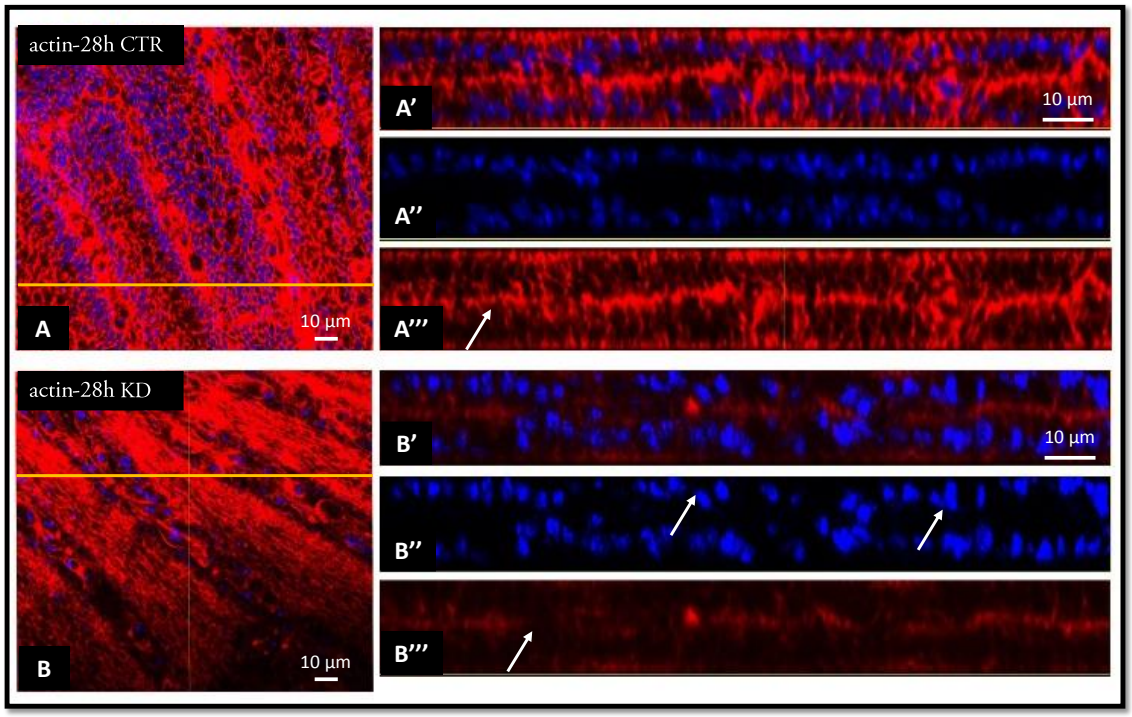

Figure 67: Overexpression of *CG1847-RNAi* in the wing causes persistence of gaps in the actin layer at 28h APF. Gaps between the basal surfaces of wing cell layers in the hinge area are revealed in the Z sections of pupal wings. The *nub-Gal4>UAS-CG1847-RNAi-T2* flies were raised at 25°C, collected at WPP, staged and imaged at 28 h APF. A and B: high magnification (63X) single sections from Z-stacks of KD and control pupal wings. A'-A''' and B'-B''' represent orthogonal views of A and B, respectively, at the level of the yellow lines. A' and B' are merged DAPI (blue) and phalloidin (red) channels. A''-A''' and B''-B''' are individual XZ sections of DAPI and phalloidin staining, respectively. In A' the arrow indicates the continuous actin layer involved in adhesion (control wing). In B'' the arrows indicate the nuclei which are not linearly arranged. In B''', arrows depict gaps in the actin cytoskeleton structure resulting from loss of adhesion. Scale bar 10 µm.

The loss-of-adhesion phenotype was partially rescued in the 28h APF wings as the gaps in the actin basal layer were more subtle than at 24 h APF. Still, at 28 h APF *CG1847* RNAi pupal wings showed a very weak phalloidin staining compared to control pupal wings. The disorganisation of the normal wing pattern, with nuclei at the different levels, not linearly arranged in KD wings at 28 h APF (arrow in B'') can also be noticed. This suggests two possible mechanisms. First, loss of *CG1847* directly affects the actin cytoskeleton, which will result in the loss-of-adhesion phenotype. Second, CG1847 controls the adhesion itself, which would indirectly be responsible for the actin cytoskeleton phenotype. Further experiments looking at adhesion molecules and at more members of integrin actin cytoskeleton network are required to determine which one of these two hypothesis is true.

4.4.6 *CG1847* silencing induces wing vein widening in early pupal development

When analysing closely the *CG1847* RNAi pupal wings at 28 h APF, I noticed a slight increase in the size of the wing veins in the *nub-Gal4>UAS-CG1847-RNAi-T2* wings (Figure 67 A and B). To determine whether this was a significant phenotype, I measured the diameter of the wing veins in 10 pupal wings of control and 10 pupal wings of *CG1847* RNAi animals. For each wing, the diameter of the L2, L3 and L4 longitudinal wing veins was measured at 3 different points (yellow marks in Figure 68).

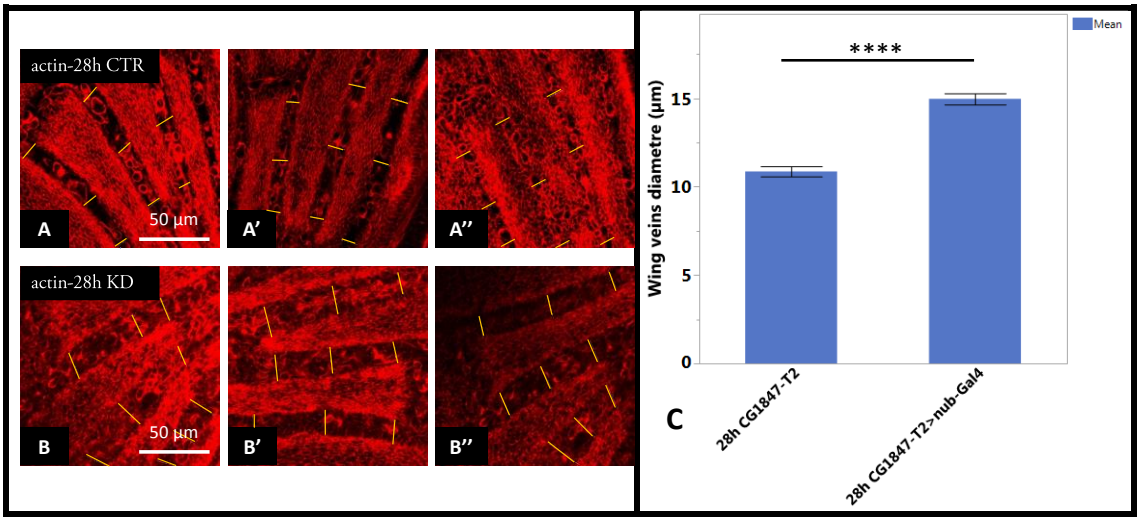

Figure 68: Overexpression of *CG1847-RNAi* in the wings resulted in wider veins. A-A'' Wing vein sizes in control samples (*UAS-CG1847-RNAi-T2)*. B-B''Vein sizes are markly increased in animals of the genotype *UAS-CG1847-RNAi-T2 nub-Gal4* (KD) raised at 25°C until the WPP stage, dissected at 28 h APF and imaged with confocal microscopy. The phalloidin staining was done following the protocol described in Appendix 6. C: There was a significant statistical difference in wing vein diameter as there is a delay in the adhesion of the intervein spaces which usualy closes the wing veins channels. Asterisks indicate statistical significance as determined by Student's t-test (n=20 wings analysed). Error bars represent SE, ****P<0.0001).

Quantification of wing vein size revealed that depleting *CG1847* in the developing wing causes vein widening at 28 h APF. At this stage of development, the actin cytoskeleton disorganisation is not as obvious as at 24 h APF, but is reflected in the significantly larger wing veins. However, no "blister-like" phenotype can be observed at this developmental stage.

4.4.7 CG1847 is required for normal expression levels of integrin associated protein PINCH

To distinguish whether the blister phenotype in wings lacking *CG1847* expression is due to the cytoskeletal defect itself or if it is the result of deregulation of integrin receptors or integrin associated proteins, I evaluated the expression levels of βPS integrin, Talin, Parvin and PINCH in pupal wings.

UAS-*CG1847-RNAi-T2* was overexpressed to repress CG1847 in whole wing during development. Pupae were collected, staged and dissected at 28 h APF. The immunostaining was performed according to the protocol in Appendix 6. Half of the pupal wings were immunostained for Talin and PINCH, while for the others we used anti-βPS and anti-Parvin antibodies. All samples were also stained with phalloidin to visualise the actin cytoskeleton and DAPI to stain the nuclei. Post immunostaining all the samples were scanned using identical laser confocal microscope parameters. The imaging of KD (nub-Gal4> *CG1847-RNAi-T2*) sample and corresponding control were taken on the same day to avoid any possible confounding factors as technical issues regarding the confocal microscope itself. As expected, I observed a dysregulation of the actin cytoskeleton in *CG1847* RNAi wings (Figure 64). Talin levels and localisation in 28 h APF pupal wings were not noticeably affected by *CG1847* depletion. However, PINCH expression levels were dramatically reduced compared to control samples. In addition, the pattern of PINCH expression was also altered. While in control samples PINCH is strongly increased at the apical border, in *CG1847* RNAi wings this pattern is not obvious (Figure 69).

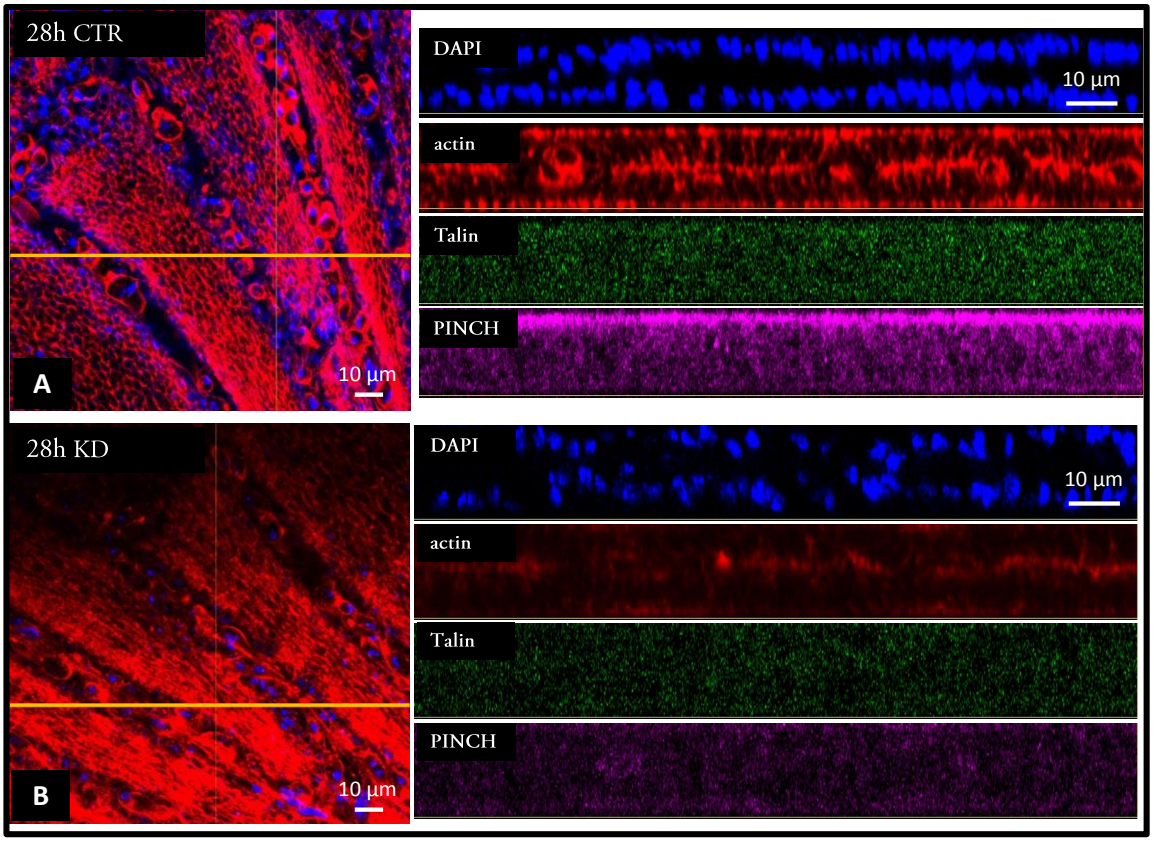

Figure 69: CG1847 controls actin and PINCH stability in pupal wings. The *nub-Gal4>UAS-CG1847-RNAi-T2* were raised at 25°C, collected at WPP, staged and imaged at 28 h APF. A and B: high magnification (63X) single sections from Z-stacks of control and KD pupal wings. Orthogonal XZ views of individual channels, (DAPI – blue; phalloidin – red; Talin – green; PINCH – magenta) were generated with ImageJ at the level marked with a yellow line in A and B, respectively.

Next, I assessed whether *CG1847* depletion induces changes in expression or localisation of βPS integrin or parvin. βPS integrin strongly co-localised with actin, especially at the basal level of the wing layers, as both proteins are involved in the establishment of adhesion. Parvin was localised preferentially at the apical level. My analysis of 28 h APF pupal wings revealed that neither βPS integrin nor Parvin were significantly affected by depletion of CG1847 (Figure 70).

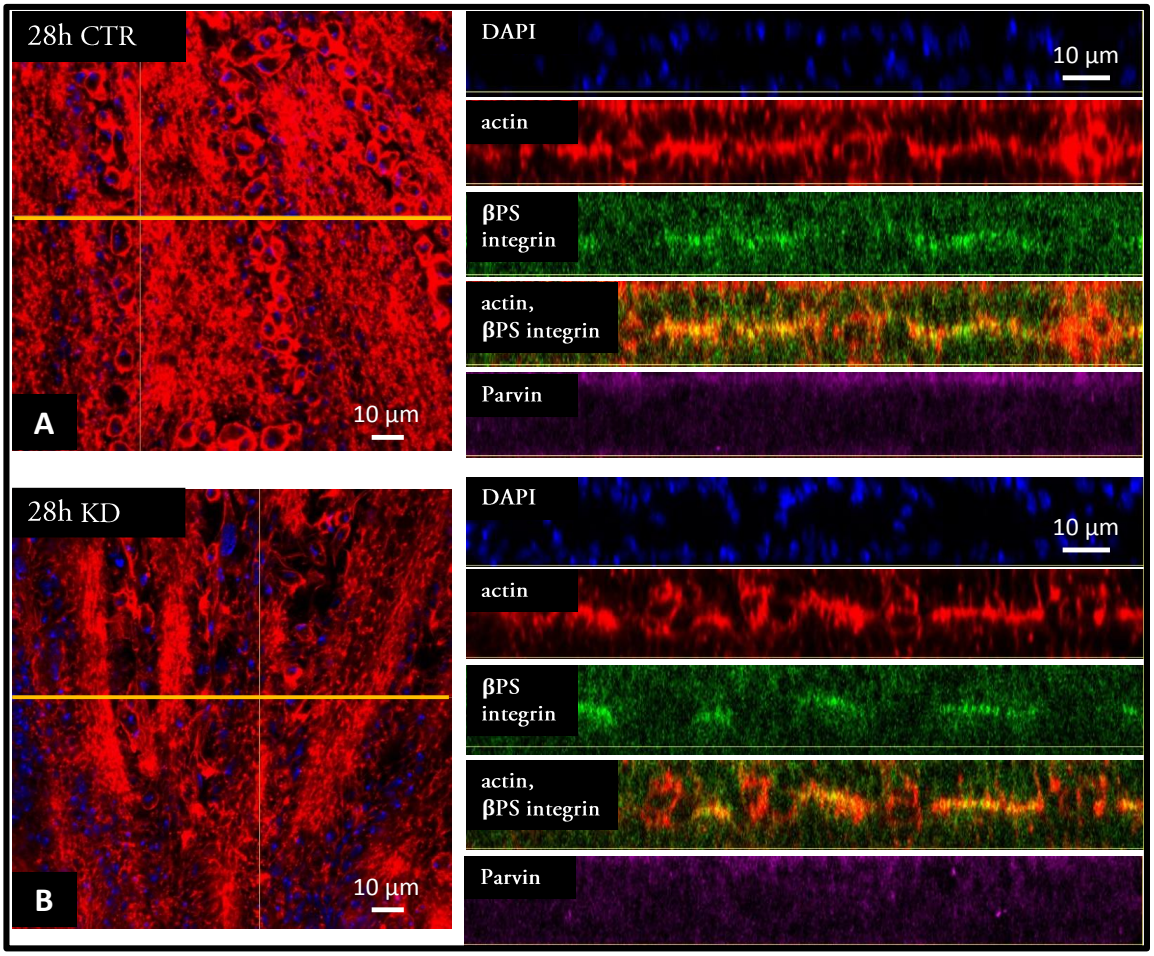

Figure 70: βPS and Parvin are normal in pupal wings lacking *CG1847*. A and B: high magnification (63X) single sections from Z-stacks of control and KD pupal wings. Orthogonal XZ views of individual chanels, (DAPI – blue; phalloidin – red; βPS – green; Parvin – magenta) and merged (phalloidin – red; βPS – green) were generated with ImageJ at the level of the yellow line in A and B, respectively. Actin, as identified by phalloidin staining, is severely disregulated in *CG1847* KD tissue. The expression of βPS follows the same pattern as actin as these proteins strongly co-localise at the basal surface of cell layers. The expression of Parvin is unchanged in the KD wings.

NOTE: As previously was described that Talin, PINCH and Parvin are localized at the site of integrin adhesion, further immunostainings are required to confirm that these results are not due to nonspecific background staining.

To confirm these results, I quantified the staining intensity of the proteins of the integrin complex in the XZ projections. For each protein, expression levels were quantified in 5 control

or RNAi samples. Phalloidin intensity was quantified in all 10 control or RNAi samples. For the quantification of βPS integrin staining, I defined a standardised area that was used to quantify the intensity of the signal exclusively in the basal level of the wing layers. This approach was used to avoid a potential effect of the βPS integrin fluorescent signal in the areas involved in lateral cell-cell adhesion. Importantly, the quantification of signal intensities (Figure 71) reflects the results shown in Figure 69 and Figure 70. Actin and PINCH were strongly downregulated in *CG1847*-depleted pupal wings, while there was no significant difference in βPS integrin, Talin or Parvin.

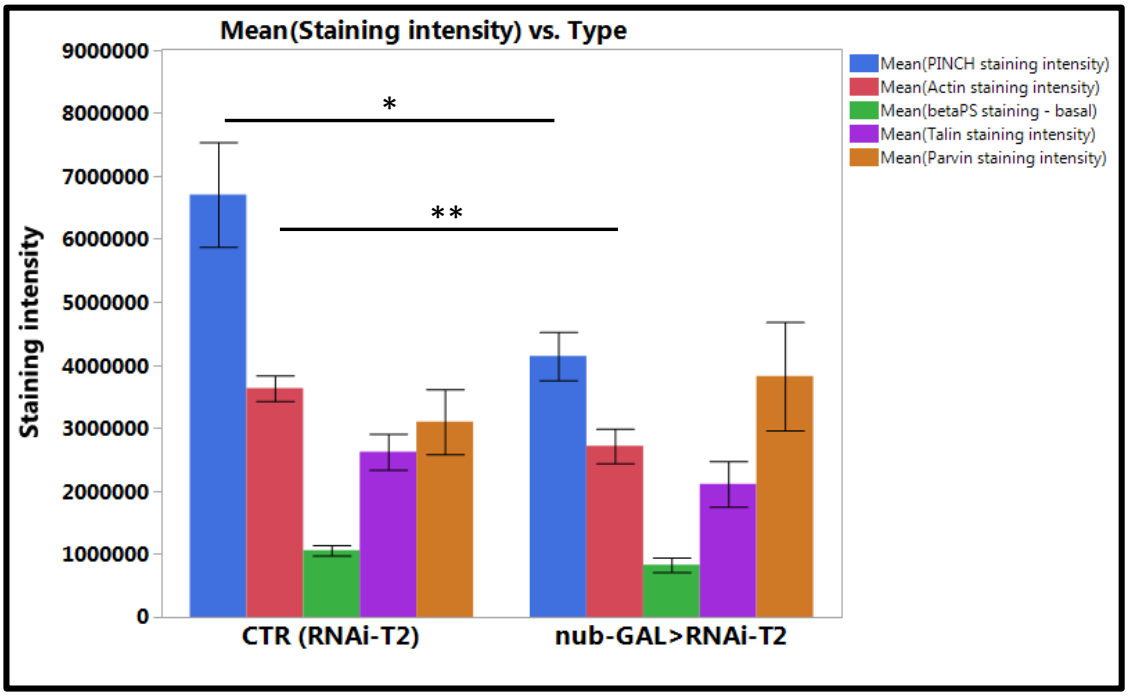

Figure 71: Significant difference in actin and PINCH staining intensity between CG1847 knockdown samples and control. Post immunostaining, samples were scanned using identical laser confocal microscope parameters. The corresponding fluorescence measurements of different proteins were evaluated using ImageJ. The staining intensity unit was in pixels. Assuming that fluorescent intensity levels accurately reflect the quantity of protein present, 28 h APF KD wings have approximately 50% less PINCH protein expressed and significanltly lower levels of actin. No statistical significant changes were detected for Talin, βPS and Parvin. Error bars represent SE. Asterisks indicate statistical significance as determined by Student's t-test (*P<0.05, **P<0.01)

My results reveal that CG1847 signalling is strongly required for the maintenance of a normal tissue actin cytoskeleton. Silencing CG1847 leads to marked loss of wing cell layer adhesion and widening of wing veins in pupal stages, which are further translated into blister formation in adult flies. Even more, in *Drosophila* pupal wings, lack of CG1847 leads to marked downregulation of integrin-associated protein PINCH.

4.4.8 Overexpressed hAIP has a cytoplasmic localisation in *Drosophila* tissues

A lack of availability of antibodies that recognise *Drosophila* CG1847 means that there is currently no experimental evidence regarding exactly where in the wing is CG1847 expressed during pupal development. Human and mouse studies revealed that AIP protein is ubiquitously expressed in both developmental and adult stages. AIP expression is abundant and has been confirmed in multiple tissues including the human heart, brain, skeletal muscle, kidney, testis, ovary and pituitary, among others. At the cellular level, AIP expression is predominantly cytoplasmic[139,154]; however, nuclear expression was also reported[149]. As *CG1847* remains a non-characterised gene in flies, precise information regarding its levels or expression pattern in *Drosophila* is lacking.

To overcome this, I used a different approach. As described in Chapter 3, human AIP protein is able to compensate for CG1847 deficiency and to rescue the lethality of mutant hemizygous males. Consequently, it is formally possible that hAIP is also expressed in the same cellular compartments in order to be able to substitute for CG1847 in fulfilling its normal functions.

Several antibodies that recognise human AIP are available (both commercial and homemade) and their specificity has been tested by western blot analysis to determine if some of the antibodies could detect the exogenous hAIP expressed in rescued *CG1847* mutant males (Chapter 5). To determine the *in vivo* subcellular localisation of hAIP (and, by proxy, the localisation of CG1847), I used the *UAS-Gal4* system to overexpress hAIP ubiquitously. *Act-Gal4*, was introduced (as previously described) in the heterozygous mutant background on the second chromosome (*CG1847^{exon1_3}/FM6;Actin-Gal4/CyO*). Heterozygous mutant females were mated with homozygous transgenic males carrying *hAIP^{wt}* cDNA on the second chromosome under the UAS promoter (UAS-*hAIP^{wt}*/ UAS-*hAIP^{wt}).*

In the next generation half of the males were *CG1847^{exon1_3}/Y;Actin-Gal4/UAS-hAIP^{wt}* and expressed hAIP protein in all tissues (Figure 72 A''). The other half of the males were *CG1847*^{exon1_3}*/Y;CyO/*UAS-*hAIP^{wt}* and acted as an internal negative control for specificity of antibody raised against hAIP protein (Figure 72 H'')).

I performed 4 h egg collections and the larvae were allowed to develop at 25°C until they reached the third instar stage. In this larval stage it was possible to visually select under the microscope the male larvae with the genotype of interest, which were then allowed to further develop. WPP were collected, staged and dissected at approximately 26 h APF. The images were taken with the confocal microscope; for each wing we took images with low (10X) and high (63X) magnification.

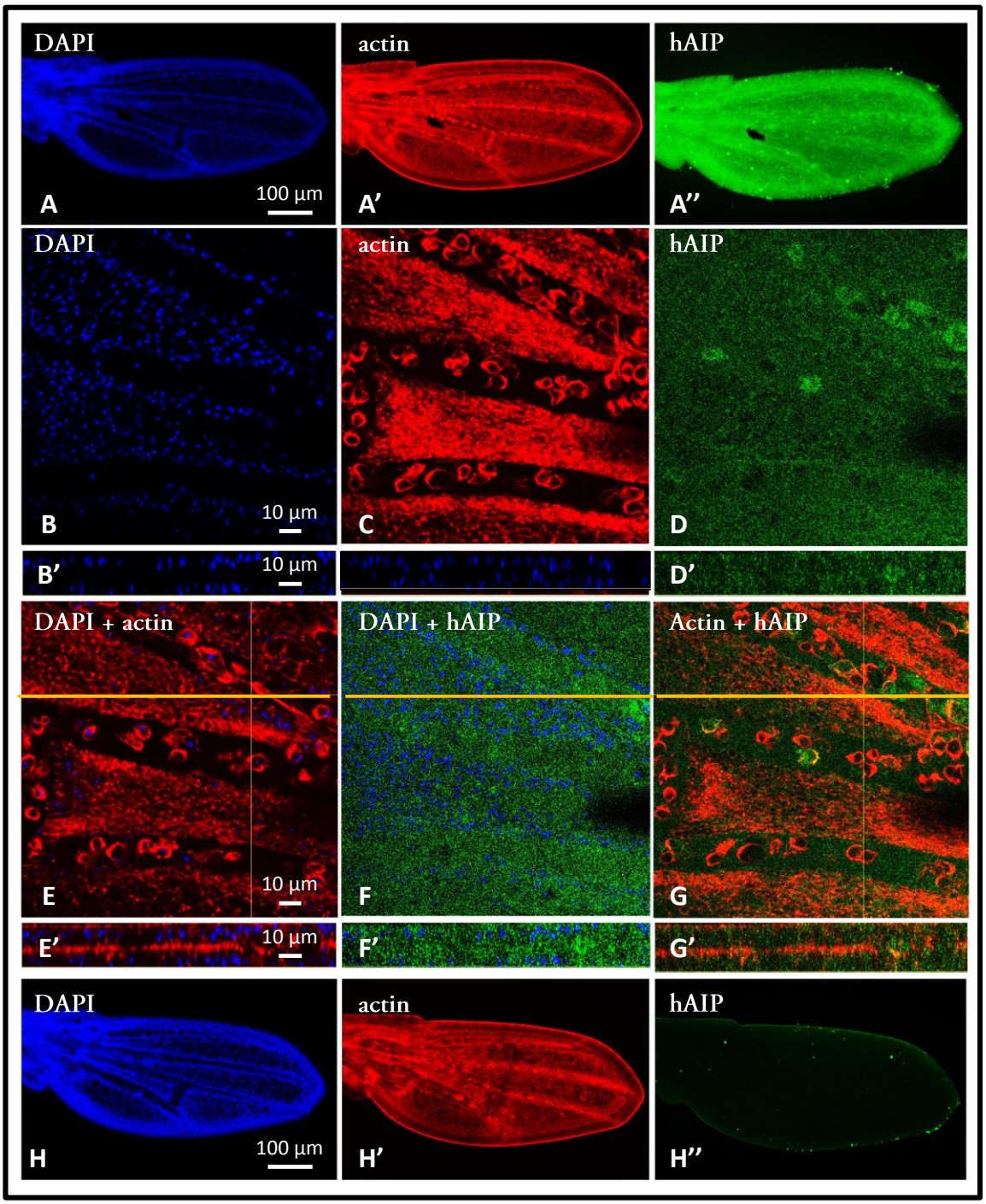

Figure 72: Overexpressed human AIP is localised in the cytoplasm of *Drosophila* wing cells. A-A'' - low magnification (10X) visualisation of the entire wing blade at 26 h APF of wings overexpressing human AIP protein. B-D: high magnification (63X) single sections from Z-stacks of pupal wings: DAPI – blue; phalloidin – red; hAIP – green. C'-D': Orthogonal XZ views of individual chanels. E-G: merged images 2 of the individual channels. E'-G': orthogonal XZ views of merged chanels. H - H'': low magnification (10X) visualisation of the entire wing blade from males *CG1847*$^{exon1\_3}$*/Y;CyO/*UAS-*hAIP*^{wt} used as internal negative control for specificity of antibody. All the optical sections were generated with ImageJ at the same level in all images (see yellow lines in E, F and G).

Human AIP protein overexpressed under actin-Gal4 is ubiquitously localised in the cytoplasm compartment of the cell. There is no co-localisation of hAIP with actin or DAPI. The results shown in Figure 72 confirm previous data, which showed that AIP is expressed throughout the cytoplasm without a particular pattern or a precise localisation.

Note: as based on the Western Blot results presented in Section 5.4.6 we are now aware that this antibody against human AIP protein cannot detect the endogenous CG1847, a further control could be performed in the future: staining of a wt wing, to check for specificity of the staining and to exclude any possible leaky effects of the pUAS-hAIP construct.

4.4.9 cAMP/PKA pathway is not significantly changed in *Drosophila* AIP mutants

As described in Chapter 3, I performed an RNA-seq study for whole transcriptome quantification of the *CG1847* mutant. The main purpose of using RNA-seq analysis was to identify the affected transcripts in *Drosophila CG1847* deficient males, as this could allow a deeper understanding of the affected pathogenic processes. As the *in vivo* data supported an involvement of CG1847 in the adhesion process, I examined the RNA-seq data for possible candidates.

The available literature regarding blister formation in *Drosophila* involves a wide range of possible candidates, apart from integrins and integrin associated proteins. Some of the alternative candidates are: Gsα, via cAMP/PKA pathway, proteins involved in the chitin based attachment and other classes as presented above. A very interesting candidate is Gsα, via cAMP/PKA pathway as the most frequently observed genetic change in pituitary adenomas is the somatic heterozygous activating mutation of the *GNAS,* the gene coding for the G protein α-subunit. This can be present in up to 40% of GH-secreting pituitary adenomas[69,70]. Even more interesting, in *Drosophila*, expression by multiple Gal4 drivers of a constitutively active form of Gsα resulted in wing blistering[363]. The mechanism behind this phenotype is that expression of the constitutively activated form of Gsα in *Drosophila* wing induces premature cell death which interferes with the normal adhesion process during wing development[371]. Consequently the Gsα/cAMP/PKA pathway appeared to be a strong candidate to be examined for possible understanding of the mechanism of blister formation due to loss of CG1847, and I investigated the RNAseq data to see if these candidates were significantly changed in this mutant.

I analysed the RNA-seq data and searched for changes in expression of transcripts known to be involved in cell-cell adhesion. Surprisingly, the majority of the genes previously described as being involved in blister formation were not significantly changed at the RNA level. The only putative candidate that was detected as significantly changed by the RNA-seq study was Act57B.

Gene	Full name	Log fold change	p_value	significant	oID_genes
Act42A	Actin 42A	0.406432	0.2532	no	FBgn0000043
Act57B	Actin 57B	-0.633148	0.0001	yes	FBgn0000044
Act5C	Actin 5C	0.276335	0.1869	no	FBgn0000042
Act79B	Actin 79B	-0.460301	0.3667	no	FBgn0000045
Act87E	Actin 87E	-0.493142	0.0848	no	FBgn0000046
Act88F	Actin 88F	-0.916564	1	no	FBgn0000047
Actbeta	Actin beta	0.515229	0.0679	no	FBgn0024913
Actn3	α actinin 3	-1.54302	1	no	FBgn0015008
Asx	Additional sex combs	-0.0169036	0.9874	no	FBgn0261823
Bj1	Highwire	-0.539097	0.2875	no	FBgn0002638
blot	bloated tubules	-0.0593775	0.8388	no	FBgn0027660
CG1136	CG1136 (insect-specific)	-0.151218	0.5831	no	FBgn0035490
Pka-C2	Protein kinase, cAMP-dependent, catalytic subunit 2	-0.138539	1	no	FBgn0000274
CG14967	CG14967	0.0663947	0.8943	no	FBgn0035420
Clc	Chloride channel-c	0.153304	0.6614	no	FBgn0024814
dl	Delta	-0.0269314	0.9375	no	FBgn0260632
Gs alpha	Gs alpha	-0.195469	0.4815	no	FBgn0001123
how	held out wings	-0.045544	0.881	no	FBgn0264491
if	inflated (αPS2 - ventral region of the wing)	-0.391494	0.2154	no	FBgn0001250
mam	mastermind	0.174134	0.5295	no	FBgn0002643
mew	multiple edematous wings (αPS1)	-0.0201201	0.9415	no	FBgn0004456
mys	myospheroid βPS	-0.0154065	0.9626	no	FBgn0004657
Pax	Paxillin	-0.453091	0.1577	no	FBgn0041789
pio	piopio	-0.350258	0.2197	no	FBgn0020521
Pomp	pomp	-0.155682	0.7813	no	FBgn0032884
rhea	talin	0.0650331	0.8494	no	FBgn0260442
Sac1	Sac1	-0.167354	0.5496	no	FBgn0035195
sfl	sulfateless	-0.029291	0.9199	no	FBgn0020251
sgl	sugarless	0.0745148	0.8112	no	FBgn0261445
shot	short stop	-0.398396	0.1734	no	FBgn0013733
stck (PINCH)	steamer duck	-0.286582	0.3821	no	FBgn0020249
Tsf2	gigas	-0.00332658	0.9898	no	FBgn0036299
Vinc	Vinculin	0.024666	0.9356	no	FBgn0004397

Table 22: List of genes with a putative involvement in blister formation

RNA-seq is a very powerful tool, but the results need further validation via an alternative molecular technique. For this, I used a multiplex qPCR method. Apart from Act57B, I chose to validate other putative candidates, which were selected on the basis of literature searches (both from human and *Drosophila* studies). Gsα and PKA (all 3 subunits: PKA-C1, -C2 and -C3 present

in fruit fly) were similarly expressed in the mutant and control samples via RNA-seq (Figure 73). However, I chose to include these genes for further quantification as they have been shown to be highly involved in the pathogenesis of pituitary adenomas, and also in blister formation.

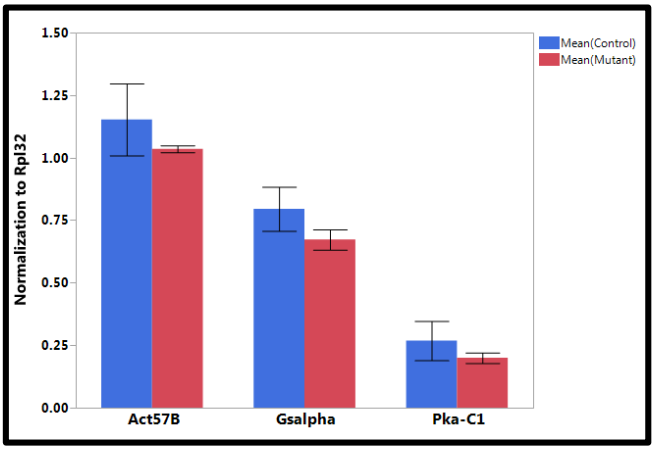

Figure 73: Quantitative PCR analysis of transcripts with a possible involvement in blister formation in *Drosophila* adult wings. None of the quantified transcripts were significantly downregulated in mutant larvae when compared to controls Target gene expression was normalised to the reference gene Ribosomal protein Rpl32 (Chapter 3). N=4 biological replicates with 2 technical replicates for each. Error bars represent SE.

The comparison of the expression profiles for the other possible candidates during *Drosophila* development revealed that these specific transcripts were not substantially perturbed in mutant larvae. The multiplex qPCR analysis failed to detect significant changes in the mRNA expression levels of Act57B, Gsα or Pka-C1.

Members of other protein families, such as the heat shock proteins, or proteins involved in cuticle formation and secretion, or chitin attachment were also validated via this method (Chapter 3). These also might have an influence on wing blistering.

A very recently published article[366] used the *Drosophila* wing tissue to investigate the components regulating the adhesion process. By using a library of *Drosophila* RNAi lines targeting 1573 protein coding genes and screening for the blister phenotype the authors identified 190 novel genes involved in apposition and adhesion of the two wing epithelial layers. Remarkably, Bilousov *et al*. revealed components of cellular processes as cell cycle, RNA splicing, and vesicular trafficking. With bioinformatics tools, they assembled the resulted data into a blisterome network.

However, not even the huge data resulted from this study did not identify CG1847 as being involved in blister formation. Further validation of RNA-seq data and more specific experiments are required to investigate in more detail the mechanism behind loss of adhesion in the CG1847 deficient mutant.

4.5 DISCUSSION

The human adenohypophysis is a glandular epithelial tissue of ectodermal origin, derived from Rathke's pouch. The majority of this gland is composed of sinuous cords of epithelial cells surrounded by vascular tissue. Epithelial cells form many different tissues, but their characteristics are the same regardless the type of tissue or the organism they belong to. Epithelia are layers of simple or stratified interconnected cells. They have an apical and a basal surface, the latter being in contact with a basement membrane[372,373]. Based on their shape, epithelial cells are further classified as squamous (flat, thin cells), cuboidal or columnar epithelia (mainly composed of tall, thin cells). Another important characteristic of epithelial tissues is the abundance of cell junctions, which are formed of protein complexes involved in either connecting neighbouring cells, or connecting the cells to the extracellular matrix (ECM) via different types of adhesion. Connections between neighbouring cells is established by apical adherent junctions (AJs). The main component of AJs is E-Cadherin (E-Cad). E-Cad is a transmembrane protein, which can homodimerise with E-Cad molecules from surrounding cells via its extracellular domain. The cytoplasmic domain of E-Cad interacts with the actin cytoskeleton via a complex network of proteins[374]. Thus, the AJs are a link between the lateral surfaces of the cells, and they are essential structures for apical-basal polarity, a main feature of epithelial cells[375]. Another type of cell junctions is involved in connecting the epithelial cells to the ECM: the basal junctions (BJs)[375]. The ECM is a very complex mix of tissue-specific proteins, involved in tissue integrity and binding of extracellular signalling molecules[376,377]. The ECM initiates crucial biochemical cues required for tissue morphogenesis and differentiation during development[378]. The BJs involved in cell-ECM connections are mediated by integrin receptors, which are adhesion receptors for extracellular matrix ligands, that can also serve as transmembrane mechanical links to the cytoskeleton inside cells[379]. These bidirectional linkages integrate cells with their microenvironment. All integrins are heterodimeric molecules containing an α and a β subunit. The extracellular domains of the α subunit bind to specific ECM components such as laminin and collagen. The intracellular domains of the β subunit mediate intracellular connections to the actin cytoskeleton[380,381]. Thus, while AJs have an important role in connecting neighbouring cells, BJs connect the cell cytoskeleton to the ECM. The ECM is a key player in cancer development[382-384]. In tumorigenesis processes, the ECM is commonly deregulated and becomes disorganised. Abnormal ECM has a strong influence in cancer progression by directly promoting cellular transformation and motility[385]. An abnormal ECM can have increased collagen deposition, which in turn up-regulates integrin signalling, thereby promoting cell survival and increasing proliferation[386,383].

To understand the role of AIP in the pathogenic mechanisms of pituitary adenomas, it is essential to investigate which are the interacting partners affected by the loss of this protein and what is their influence on the structure of the epithelial tissue, or connections with the ECM. The developing *Drosophila* wing is a very simple system which enables scientists to investigate the molecular mechanisms involved in cell shape changes associated with rearrangement of epithelial sheets during normal development or tumorigenesis. As previously presented in the Introduction section, *Drosophila* wing development is a relatively simple process, which involves the transition from a single layered columnar epithelium to a flattened bilayer tissue. The basal surfaces of the dorsal and ventral epithelia are in close contact to the basal membrane as during pupal stages they adhere by forming the BJs mediated by integrin receptors[353,354].

I used the *Drosophila* wing to show via two independent methods that silencing or deletion of *CG1847* during development causes wing blister formation in adults. This phenotype strongly resembles that of a decrease in integrin function; homozygous mutant clones for integrins induced in the wing disc result in blisters. This is due to the fact that the dorsal and ventral wing epithelia fail to adhere in and around the integrin mutant clone[387-389]. The adhesion process is mainly regulated by the different proteins belonging to the integrin-actin-cytoskeleton network, and eliminating almost any of the components of this complex network results in failure of adhesion between wing surfaces[390]. Previous studies have found that the same phenotype, a fluid-filled blister, might be the result of silencing βPS integrin signalling in the wings[381,389,391,392], or of a loss of PINCH[393]. Mutations of *blistery*, the gene which encodes for tensin, also result in loss-of-adhesion in the adult wing[362], as well as *integrin-linked kinase* (ILK) mutations[394]. As a result, it is likely that a failure in any of the members of this integrin-actin cytoskeleton complex may lead to a loss-of-adhesion phenotype. However, there are no descriptions of an *actin* subunit mutation resulting in blister phenotype. As actin is essential for cell viability[395], null mutations of this gene will result in lethality[396].

Besides the blister phenotype which strongly suggests defects in wing cell adhesion, our results reveal that CG1847 signalling is required for the maintenance of a normal actin cytoskeleton as actin is significantly downregulated. In addition, we show that apart from actin, in *Drosophila* pupal wings, loss of CG1847 leads to marked downregulation of the integrin-associated protein PINCH. Finally, another surprising result was the observation that CG1847 mutants display widened wing veins during the pupal stages. As a consequence, my results strongly suggest that CG1847 may be an important factor of cell-to-ECM adhesion via an intricate relationship of mechanisms involving the integrin-actin-cytoskeleton and the wing vein formation.

Integrins are a family of heterodimer receptors formed by the association of an α- and a β-integrin subunit. The *Drosophila* genome encodes 5 α (αPS1-5) - and 2 β-subunits (βPS and βv). While βPS forms tissue specific heterodimers with all 5 α subunits, for the second β-subunit, βv, only one α partner has been identified so far[223,224,397]. Integrins have an extracellular, a transmembrane and a cytoplasmic domain. These receptors are classified based on the ECM molecules they bind to via their extracellular domain: some recognise fibronectin, while others bind specifically to laminins[398]. An important part of the adhesive function of integrins is their ability to connect to the actin cytoskeleton, therefore integrating cells with their microenvironment[379,399,400]. The integrins do not possess intrinsic enzymatic activity; however, numerous proteins and kinases (such as FAK, Ras, Raf, Mek, PI3K and PKA) have been involved in mediating the integrin involvement in signalling pathways that control cell migration, proliferation, differentiation and survival[398,401]. After binding to ECM ligands, integrins undergo an activating conformational change. This leads to focal adhesion kinase (FAK) and tyrosine auto-phosphorylation of their cytoplasmic associated proteins. Integrin-ECM interaction is followed by the recruitment of integrin-interacting proteins that form complexes known as focal adhesions[402]. Some of the most studied integrins functions regards their involvement with the ECM, especially the role in mediating the dynamics of cellular migration. Integrin null mutations are lethal at early stages of development, and the mechanism is most often due to failure of integrin-mediated adhesion[369,390].

Actin is an abundant, essential protein found in all eukaryotic cells. It exists in two major conformations: G-actin (globular) or F-actin (filamentous). Numerous direct or indirect interacting partners have been described for actin, which are involved in a tight regulation of actin dynamics. Different ECM stimuli activate integrins, which bind to their ligands, become activated and send signals to the actin cytoskeleton[403,404]. In response, additional cytoskeleton proteins are recruited and these will influence ATP-bound actin monomers to assemble into filaments. The final result is that the newly assembled actin filaments will determine, also via integrins, the necessary changes in cell motility. It is well known that actin is involved in pathogenic states, and it plays a particularly important role in tumour invasion and metastasis[405,406]. Malignant cells migrate by invading adjacent tissues and the vasculature. These are multi-step processes, usually initiated by polymerization of actin filaments which leads to formation of membrane protrusions. Actin cytoskeleton signals were shown to be upregulated in invasive cancer cells[407].

Talin is a major integrin-binding protein as it has an atypical FERM (band 4.1, ezrin, radixin, and moesin) domain. The four subdomains (F0-F3) do not adopt the classic clover conformation of

FERM domains; however, it was shown that talin directly interacts with the cytoplasmic domain of β-integrin subunits[246]. The integrin-talin FERM domain complex plays a critical role in actin polymerization[408]. In *Drosophila*, mutations in *rhea*, which encodes talin, are also lethal during early embryonic development, while loss of this protein specifically in the wings causes failure in adhesion and blister formation[381]. As shown in Table 22, the *rhea* levels of expression are not significantly changed in *CG1847* mutants.

The ILK-PINCH-Parvin (IPP) complex has been shown to be downstream of talin and to connect talin to the actin cytoskeleton. As such, the IPP complex has a central role in the integrin adhesome network[409,410]. ILK is the first protein of this complex that is activated upon integrin signalling and is recruited to focal adhesion sites[411]. ILK has two different domains involved in binding to the other members of the complex. While the ILK ANKRs (five tandem ankyrin) repeats connect to PINCH, the kinase-like domain binds to parvin. Upon binding, both PINCH and parvin connect to actin[412]. Disruption of any member of the IPP complex significantly impairs the other two, and leads to changes in cell shape and motility[413]. In *Drosophila*, the IPP complex has been thoroughly investigated in two well-established models of integrin-mediated adhesion: muscle attachment sites and the wing epithelium. Surprisingly, at muscle attachment sites, ILK and PINCH do not require parvin, suggesting that parvin functions to strengthen the integrin–actin link in the muscles[414]. On the other hand, studies in the wing epithelium revealed that the stability and localization of ILK/PINCH/parvin are interdependent in wing epithelium. Staining in wing imaginal discs revealed that knockdown of either PINCH, ILK or parvin resulted in reduced expression of the other two, whereas talin levels remained unchanged[414].

In summary, there is interdependence of all three IPP-complex components for their stability and subcellular localization at the basal side of wing epithelia, unlike the central role of ILK in IPP-complex assembly at the embryonic muscle attachment sites (MASs), as depicted in Figure 74.

Interesting, these results were confirmed by the findings of another group who proposed that mutations PINCH function is required for the stable adhesion between epithelial layers in the wing as mutations which inactivate PINCH in *Drosophila* wings leads to wing blisters. However, investigating the distribution of this network in the muscles of *Drosophila* embryos, PINCH inactivation do not result in ILK mislocalization, as ILK appears normal in PINCH mutant embryos[393].

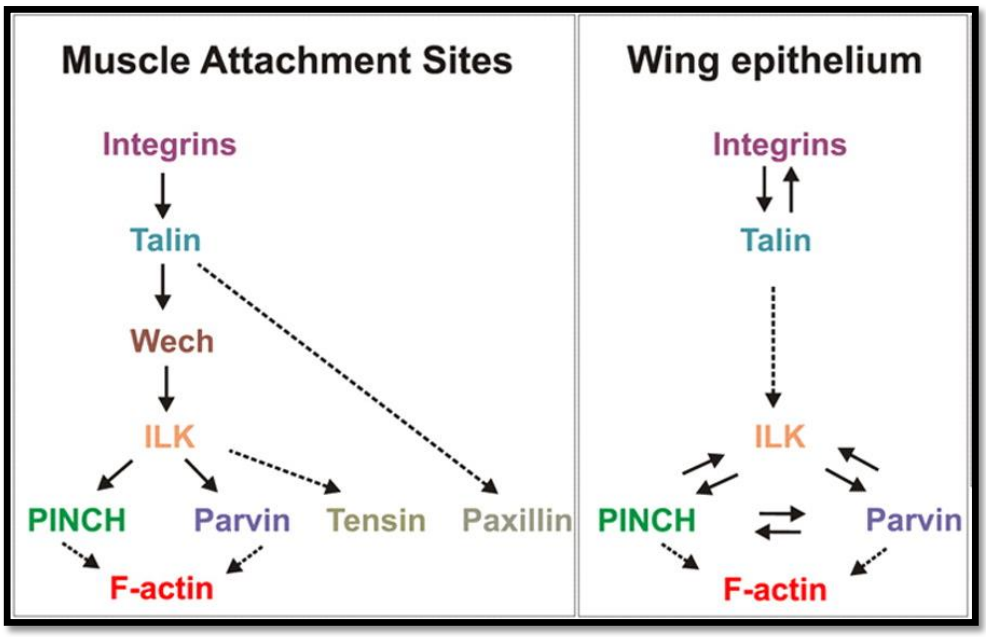

Figure 74: Genetic hierarchy of integrin–actin linker complex assembly at muscle attachment sites and in wing epithelium. (A) At muscle attacment sites, assembly of the integrin–actin linker complex follows a linear pathway of genetic interactions. IPP complex assembly depends on ILK and there is no interdependency between parvin and PINCH. (B) In the wing epithelium, the genetic hierarchy can be classified at two levels: Integrins and talin are mutually dependent and both are required for PIP-complex stability. All proteins of the IPP complex are mutually dependent. This network is highly conserved in humans, mice but also in organisms such as fruit fly (*Drosophila melanogaster*), nematode (*Caenorhabditis elegans*), slime mold (*Dictyostelium discoideum*), and yeast (*Saccharomyces cerevisiae*)[415].

Given the blister phenotype upon CG1847 loss of function, and the identical phenotype of βPS, mutants, I assessed whether CG1847 is required for the subcellular localization of integrin β-subunit at muscle attachment sites. We performed this investigation in male mutants, when they were just a little older than 48 hours of development (the moment when they start dying). We found identical distribution of βPS integrin in *CG1847* male mutants and control larvae. However, βPS distribution was found normal also in pupal wings lacking CG1847. Hence, my results suggest that *CG1847* loss of function does not affect βPS distribution in muscle attachment sites or in the wing epithelium. It remains to be determined whether loss of *CG1847* affects the signalling activity of integrins.

Whereas actin and PINCH were significantly downregulated in pupal wings, talin expression levels were unchanged, this being in accordance with previously published results[414]. Likewise, the intensity of parvin staining was similar between knockdown and control samples which although is in disagreement with the results of Vakaloglou *et al*.[414], it confirms the data published by Clarks *et al*. in 2003[393].

In summary, my results show a novel connection between CG1847 and the integrin-actin cytoskeleton in the *Drosophila* wing epithelium. As I could not detect a significant change of actin at the RNA level, neither in the RNA-seq data nor by Multiplex qPCR, I propose that the loss of CG1847 leads to lower actin levels via post-transcription mechanisms. However, the actual mechanism of this interaction remains unclear and the direct interaction between AIP and actin remains controversial. The ECM–integrin–actin is an intricate complex which signals bidirectional and, therefore, it is not entirely clear which factors initiate a process involving this network.

Some early studies involving *Drosophila* and *C. elegans* suggested that integrins are the origin of the stimulus for the linkage between ECM and the cell cytoskeleton as integrin mutants are not viable[416,417]. However, subsequent studies focused in understanding ECM structure have proposed that changes in the ECM components dictate the intracellular assembly of actin filaments. This is achieved via integrins and integrin-associated proteins; however, integrins are obviously not essential to generate the ECM, as ECM components still accumulate in integrin null mutants[418,419]. To further complicate matters, other groups have shown that polarisation of the cytoskeleton precedes changes in either the ECM components or integrin activity. Their hypothesis suggests that actin polymerisation leads to accumulation of ECM ligands, which in turn determine integrin activation[420]. Other groups supported a very similar hypothesis, proposing that integrin activation comes from inside the cell, without involving ECM components. It was shown that, at the muscle attachment sites, the cytoskeleton would first undergo polarisation which will lead to integrin activation and connection to the ECM. This mechanism has been supported by the phenotypic analysis of *Drosophila* PS integrins[368,421,422].

In summary, increasing evidence supports the sequential model of an initial cytoskeleton polarisation, resulting in the tissue-specific accumulation of different ligands in the ECM. Only then the integrins are able to attach to their ligands. Following integrin attachment to the ECM components, these receptors are able to recruit additional proteins to focal adhesions, such as talin, parvin, vinculin, which further strengthen the interaction. Nevertheless, this model raises a critical question: what controls intracellular actin polarisation? Interestingly, in our *in vivo* *Drosophila* model of *CG1847* deficiency actin is severely downregulated and disorganised and this might be the causative factor for the various CG1847 phenotypes. Furthermore, PINCH downregulation brings additional support to the hypothesis that the loss-of-adhesion phenotype generated by lack of the *Drosophila* AIP orthologue involves the actin-integrin network.

In addition to the blister phenotype, inhibiting CG1847 signalling resulted in wing vein widening in pupal stage. When induced ubiquitously, *CG1847* RNAi led to a strong defect in wing vein formation. Broader wing veins have been commonly associated with wing blistering[365,423-425]. The cascade of morphogenetic events during normal wing development is crucial to understand the mechanism of this phenotype. Following eversion, dorsal and ventral surfaces of the wing reach complete apposition at their basal surfaces, forming a bilayer epithelial tissue. At approximately 18 h APF, longitudinal bands of cells send extensions of basal processes to connect with their partners on the opposite layer[346]. At this stage, the presumptive veins are not visible. The 'intervein bands' first appear around the periphery of the wing and increase progressively from anterior to posterior. At 20 h APF, the veins emerge as areas which do not form basal connective processes, do not adhere to the opposite layer and appear as a network of vein lumens, with a pattern similar to the adult wing. Between 21 and 30 h AFP, all the cells in the intervein areas form BJs with the cells from the other layer, progressively 'zipping' the intervein epithelia. As a consequence, vein development is based on the persistence of dorsal-ventral separation. While all intervein cells become apposed basally via BJs mediated by integrin receptors, the vein channels remain unopposed and the vein network is fully formed. The vein cells remain relatively unspecialized, do not express integrins, do not form BJs and do not connect with the cells on the opposite layer. At approximately 30 h APF, the veins lumens become lined with laminin[346].

In *Drosophila* integrin mutants (e.g. *inflated, myospheroid*) the mechanism of cell adhesion in intervein areas is defective and vein channels are wider resulting in wing blisters[387,391,392]. However, in integrin mutants no evidence of blisters in pupal wings could be found; the intervein regions were basally apposed and were not laminin-lined. Even mutations that in adults led to a completely ballooned wing had an appearance of normal apposition between the cell layers during pupal stage. Consequently, it seems that blisters form later in development, during the expansion stage, 50-60 h AFP or during eclosion[414].

Interestingly, the wing vein widening phenotype has also been described in *Drosophila* when there are alterations in cadherins[426]. Cadherins are adhesion molecules involved in mediating the cell-cell adhesion via AJs, and in the control of the cell shape and the cell polarity[427]. The *Drosophila* E-cadherin (DE-cad) is encoded by the *shotgun* (*shg*) gene. Vein cells lacking *shg*, and consequently the AJs, failed to form a lumen. The vein cell fate specification in the developing *Drosophila* wing is mediated by Egfr, as there is evidence that Egfr controls *shg* both at transcriptional and posttranslational levels[428]. In summary, Egfr signalling is essential for normal DE-cad localization[426] and alterations in DE-cad-mediated adhesion result in vein broadening. This mechanism might be relevant for this study, as E-Cad associates with the actin cytoskeleton

and negatively regulates cell migration in *Drosophila* by actin cytoskeleton remodelling[429]. Furthermore, E-Cad involvement in tumorigenesis is well known and it is strongly involved in epithelial-mesenchymal transition. Beside this, EGFR was reported to interact with AIP[166], and EGFR inhibition has a repressive effect on rat pituitary tumour growth[430]. However, a report published by D'Souza-Schorey *et al.* in 2005[427] did not describe the blister phenotype in wings of adult fruit flies with E-cad mutations. Moreover, there are no data supporting the hypothesis that mutations in E-Cad result in wing blistering in *Drosophila*.

Briefly, my data supports the hypothesis that the *Drosophila* AIP orthologue is required for cell adhesion. The mechanism of CG1847 involvement in cell adhesion is via actin deregulation and integrin signalling, supported also by the wing vein widening. This novel possible interaction with cytoskeletal proteins and a role in cell adhesion indicate a putative involvement of AIP in cell motility and tumour behaviour. Cytoskeletal disorganisation is an important feature of epithelial-mesenchymal transition (EMT), the process by which cells develop increased migratory capacity, invasiveness, and resistance to apoptosis, resulting in the development of a more malignant cellular phenotype[431,432]. During epithelial-mesenchymal transition, epithelial cells lose their polarity and their ability to adhere to surrounding cells and become mesenchymal cells.

Unpublished mass spectrometry and affimetrix data from other members of the laboratory also suggest the possibility of a functional interaction between AIP and the actin cytoskeleton as well as an involvement of this AIP-actin partnership in the pathogenesis of pituitary adenomas.

In mammals, the combination of 18 α- and 8 β-integrin-subunits regulate cellular migration, invasion, proliferation and survival. Some oncogenes may require integrin signalling in order to initiate tumour growth and invasion[385]. In several tumour types the expression of specific integrins correlates with increased disease progression and decreased survival[433]. For other tumours, drugs have been already designed to target this pathway as for examples the integrin αv inhibitor cilengitide, which is now in Phase-III clinical trials as combined therapy for newly diagnosed glioblastoma patients[434]. Regarding pituitary adenomas, the specific pattern of integrin expression is altered compared to the normal pituitary[435]. While normal pituitary cells express $\alpha_3\beta_1$ and $\alpha_6\beta_4$ and stromal cells express $\alpha_1\beta_1$, adenomas express $\alpha_v\beta_3$ and lose $\alpha_3\beta_1$ and $\alpha_6\beta_4$. These alterations are similar to those that occur in malignant tumours but, since pituitary tumours very rarely metastasise, these changes are probably important for their invasive behaviour.

Recent studies have suggested that the AIP-partner AHR activates FAK, thereby promoting integrin activation and increased cell migration[22]. On the other hand, another AIP-partner, Hsp90, is known to interact with ILK in tumorigenic processes[436].

With regards to the possible involvement of AIP in tumour growth, our data support the hypothesis that CG1847 exerts a tumour suppressor role, as the area of *CG1847* homozygous mutant clones was significantly larger than controls. Still, CG1847 doesn't seem to be involved in cell growth, as we failed to identify any differences in cell density. Hence, CG1847 deficiency may have an important role in promoting cell proliferation or cell survival. Further investigation is needed to distinguish between the two alternatives as previous data exists to support both mechanisms.

Cell proliferation is tightly controlled by cyclins and their respective cyclin dependent kinases[437,438]. Although there are numerous studies regarding the control mechanisms of the cell cycle, the actual link between cellular proliferation and AIP-driven adenoma development remains incompletely understood. Cyclin-dependent kinase 9 (cdk9) is a cdc2-related serine/threonine kinase that has been shown to interact with AIP[164]. The mechanisms required for regulation of CDK9 and its involvement in processes regulating cellular growth and proliferation are poorly understood. It has been shown that Cdk9, together with cyclin T1, forms a heterodimer called P-TEFb, which is involved in cell cycle progression by stimulating transcriptional elongation via RNA polymerase II phosphorylation[439,440]. The P-TEFb complex is recruited to chromosomes at mid- to late anaphase. Inhibition of this process was shown to reduce the binding of P-TEFb and expression of key G1 and growth-associated genes, leading to G1 cell cycle arrest and apoptosis[441]. Consequently, CDK9 is a promising target for an antiproliferative drug in oncologic pathologies[442], CDK9 inhibitors being currently under clinical investigation[443].

AIP was also shown to suppress apoptosis as Kang *et al.* published in 2006[444] that AIP directly interacts with survivin and is required for its stability. Survivin is a member of the Inhibitor of Apoptosis family, which include evolutionary conserved members with key properties in regulation of mitosis and apoptosis[445]. Survivin protects cells against apoptosis and can reduce the sensitivity of tumour cells to apoptosis stimulation by increasing the survival capability of tumour cells. It was also shown that survivin changes can affect individual susceptibility to tumour formation[446,447]. All these characteristics make this protein a novel target for cancer therapeutics[448].

By using the fly model I also fulfil the strategic priority of replacement in research using animals. So far, most studies used rodent models in the fields of pituitary tumour, therefore this project is entirely novel. In UK alone there are around 10 different labs using mouse models for pituitary tumorigenesis. We are aware of at least 5 more labs worldwide using mouse models for studying the mechanisms of pituitary tumorigenesis. It is difficult to estimate the total number of animals used in pituitary tumorigenesis studies in these labs, because for a single ongoing project in our lab a large number of mice are sacrificed which do not have the desired genotypes. From an average of 8-9 pups born in every litter only 2-3 pups are kept. For financial constraints all the unused pups are killed before weaning (day 21 postnatal). For a single project a relatively stable colony size is 120 mice in a year. Colony size control, dictated by practical and financial reasons leads to some of the adult mice being sacrificed, with an estimate of around 10 adult mice killed per month. Obtaining embryos (average of 7-8 per pregnancy) for the developmental part of this study unfortunately also requires sacrificing the mother. For most studies at least 15 pregnancies are being examined, meaning that on average an additional 15 adult females and around 100 embryos will be sacrificed. In summary, each year around 120-150 mice and 100 embryos are being used for a single study, with around 200 adult and 300 young (before weaning) more mice being sacrificed in a single project on pituitary tumorigenesis. Regarding costs, the basic housing costs for this study are approximately 1100 GBP/month (around 1300 GBP/ month). The colonies need to be maintained for long periods of time, as pituitary tumour penetrance is age-dependent (Raitila *et al*. demonstrated in 2010 that full penetrance of pituitary tumours in AIP-KO mice was reached at the age of 15 months).

This *Drosophila* model will be further developed to study the pituitary tumorigenesis, with estimated savings of hundreds of mice/year. This number may increase even further with the adoption of this model by other laboratories and future projects in our laboratory.

4.5 CONCLUSIONS

In conclusion, the study of *CG1847,* the *Drosophila* orthologue of human AIP revealed an exciting novel role for this protein in cell adhesion. I have shown that CG1847 is required for actin cytoskeleton stabilisation, with a possible involvement in the regulation of integrins. The possible interactions with actin indicate a possible strong influence of AIP in cell motility and migration, cellular functions which are essential for tumour growth, invasion and metastasis.

In addition, I have shown for the first time a possible unique contribution of CG1847/AIP to integrin–actin cytoskeleton signalling.

Finally, my data supports the role of *AIP* as a tumour suppressor gene. However, further studies are required to uncover the exact molecular mechanisms involved.

Although my results still need thorough investigation for translation to human data, this study provides an important insight into how the loss of AIP may promote adenoma formation and local invasion. Furthermore, the *in vivo* model developed in this study supports the use of *Drosophila melanogaster* as a system to elucidate the molecular mechanisms of human tumorigenesis.

Replacing mammalian animal models with *Drosophila* ones enables research projects that would otherwise not be practically feasible, except for considerably large animal and financial expense.

CHAPTER 5: AN *IN VIVO* SYSTEM TO TEST THE PATHOGENICITY OF AIP MISSENSE MUTATIONS

5.1 INTRODUCTION

Despite significant progress on the understanding of AIP-associated pituitary adenomas, the etiology and pathogenesis have remained elusive. Many questions remain unanswered, some of them involving the pathogenicity of AIP variants. AIP-related pituitary adenomas are autosomal dominant with incomplete penetrance[116]. While different groups have reported a low prevalence of *AIP* mutations in unselected patients with pituitary adenomas (<4%)[122,138] this is significantly increased in patients with young onset GH-excess (24%) and FIPA (17.1%)[117]. In the subgroup of the patients diagnosed with gigantism, the prevalence of *AIP* mutations is close to 50%[208,209,449].

A few hotspot mutations in AIP, as well as a few cases of founder mutations have been described, and these have been shown to mostly affect CpG islands. CpG sites are regions of DNA where a cytosine lies next to guanine. "CpG" stands for a cytosine that is separated by a phosphate from a guanine (—C—phosphate—G—) on the same allele. In CpG sites, cytosines can be methylated by DNA methyltransferases, which attach a methyl group to carbon 5 of cytosine, thereby altering the DNA structure. In normal conditions, a spontaneous deamination of cytosine forms uracil, which is subsequently recognised and removed by DNA repair systems. A methylated cytosine which deaminated will form thymine instead of uracil, and this might not be recognised by the repair enzymes. This conversion will ultimately result in a transition mutation[450].

Examples of *AIP* founder mutations were found in Finland, Italy[451] and Ireland, which helped the identification of the *AIP* gene. In Ireland, a DNA sample from an 18th century patient matched the haplotype of numerous current families carrying the p.R304* nonsense mutation[452]. Using coalescent theory, it was estimated that the common ancestor lived about 57 to 66 generations earlier. Interestingly, the 304 residue is at a CpG site, which is a mutational hotspot. Numerous cases, in at least 20 kindreds, of truncating (c.910 C>T, p.R304*) and missense (c.911 G>A, p.R304Q) mutations have been identified across the globe[117,120].

While pathogenicity is beyond doubt for some mutations, such as truncated or stop variants that alter important AIP domains, for other mutations this is not immediately obvious. An interesting synonymous variant, c.249G>T, was shown to be pathogenic as it results in changes in AIP

splicing[136]. Conversely, some non-synonymous changes have been shown not to affect AIP function (e.g. p.Q228K)[344]. In the case of missense mutations, the change in amino acid sequence can affect protein 3D structure, with consequences for its folding and stability. A missense mutation can also alter the binding to interacting proteins. While some of these genetic variants have been classified as benign, with no functional consequence on protein function, for others there it is still a debate on their effect.

One of the challenges in the management of patients carrying a missense AIP variant is to determine whether the variant is a disease-causing mutation or not[102]. Several groups have exploited *in silico* predictions to test the pathogenicity of mutations. Even though *in silico* prediction pipelines could correctly predict the pathogenicity of some splice-site mutations[453], in cases where clear deleterious effects were not present, caution should be used before labelling them as deleterious mutations. Different *in silico* prediction platforms have been used (reviewed in Thusberg *et al*.[453]*),* but no single method can evaluate all parameters involved. Therefore, other methods should be employed for evaluating whether missense variants are indeed pathogenic or not.

To date, more than 70 different *AIP* variants have been identified, the majority (75%) resulting in a truncated AIP protein[120], missing the C-terminal tetratricopeptide-repeat (TPR) domains. As the frequently used term "a pathogenic variant" might be confusing a 5 categories classification was proposed: a) pathogenic, b) likely to be pathogenic, c) unlikely to be pathogenic, d) not pathogenic and e) variants of unknown significance. In an article published in 2012, Korbonits *et al.*[454] analysed the available data and grouped *AIP* variants into these five categories. Out of the fifty-two AIP variants identified at the time, 12 were previously reported as SNP or rare variants with apparently no pathological significance. Some of the AIP changes were classified as “not available” because the data regarding their pathogenicity was missing. Since then, progress in this area has been limited due to the lack of a reliable assay capable of differentiating between disease causing mutations and benign ones.

New *AIP* variants identified in FIPA patients keep on being identified. The most recent article was published by a Turkish group which detected 2 homozygous missense single-nucleotide polymorphisms (rs641081 [Q228K] and rs4930195 [Q307R]) and their frequencies were significantly higher in FIPA patients compared to controls[455]. rs641081 was previously published[344,456,457] and it is not considered to be a disease causing mutation. However, for other *AIP* variants the available data is more contradictory and there is an increased need for an *in vivo* test.

5.2 BACKGROUND

Up to 60% of fly genes are evolutionarily conserved in human. Numerous technologies have been developed in the last decades to manipulate the fly genome allowing significant insight into the role of genes, especially for those ones whose mutated human orthologues were shown to cause genetic disorders. Many genes that are essential and conserved during evolution were associated with human diseases. Multiple studies have attempted to bring a better understanding into the mechanism by which these specific mutations cause genetic diseases. In 2014 Yamamoto *et al*.[458] conducted a genetic screen of lethal mutations. They identified 165 genes on the *Drosophila* X chromosome, with a possible involvement in the development and function of the nervous system. An investigation of rare variant alleles in 1,929 human exomes from families with unsolved Mendelian disease allowed this group to identify disease-associated mutations in six families and to provide insights into microcephaly associated with brain dysgenesis.

The recent progress in sequencing technology and bioinformatics allows not only a direct identification of mutations, but also the impact of different types of mutations on the levels of expression of the genes. However, these technologies bring the real challenge in the interpretation of such genomic data and functional experiments are required. These data can be obtained using human cell cultures, however *Drosophila* is an extremely useful model organism to obtain *in vivo* data. This degree of identity between fly and human genomics helped the functional annotation of evolutionarily conserved genes.

Once a putative *Drosophila* orthologue is identified, a null allele or a strong hypomorph mutation should be generated as this will generate a specific phenotype and will provide a reference. In the next step, the human cDNA orthologue of a fly gene can be tested for rescuing ability.

A description of the strategy that was used in this thesis is shown in Figure 75.

Imprecise excision of a P-element from a gene of interest generated a null allele. Upon identification of the phenotype in the fly (lethality), rescue experiments by the UAS-hAIP cDNA transgene ubiquitously overexpressed under the actin-Gal4 driver allowed the testing of the functional conservation of *AIP* gene function between fly and human.

Using this principle, the function of variants found in human patients can be further assessed. In the next step, comparing the rescue efficiency of wt hAIP cDNAs versus mutant versions is a rapid method of assessing whether a particular variant found in a human patient might be affecting the normal function of this gene.

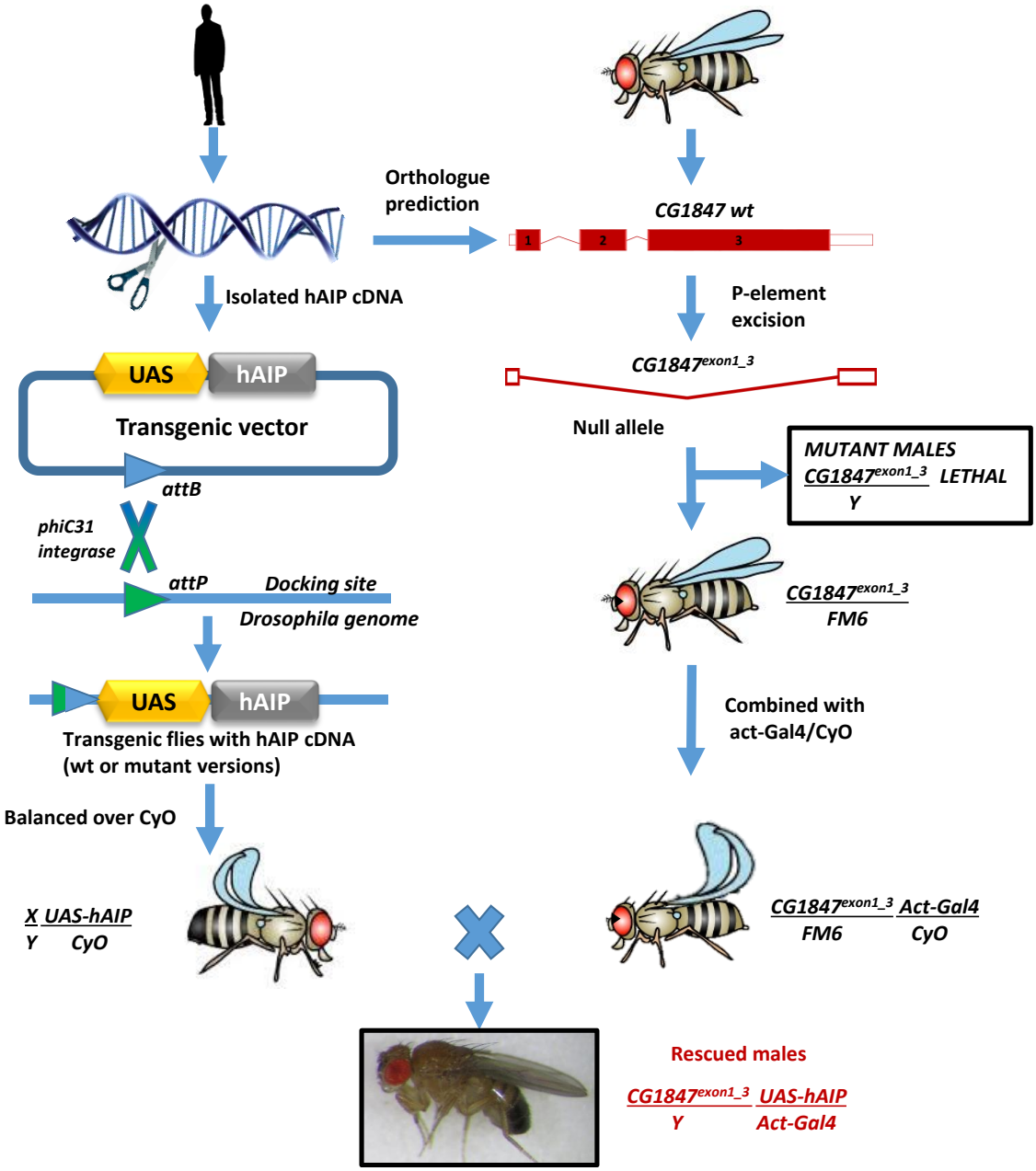

Figure 75: Schematic representation of the rescue experiments. First, the potential fly orthologue of human *AIP* was identified. A mutant fruit fly was generated via imprecise excision of a P-element. Only heterozygous females are viable, as hemizygous males are lethal. The *Act-Gal4* driver was introduced in the mutant background and the stock was balanced over *CyO*. The hAIP cDNA was cloned into a transgenic UAS vector. This allows generation of transgenic stocks for hAIP, which were then balanced over *CyO*. By crossing males that carry the transgene of the hAIP cDNA under the control of UAS with heterozygous mutant females (with *Act-Gal4*) it can be determined if hAIP is able to rescue the fly lethality.

5.3 OBJECTIVES

The fruit fly model could be utilised to identify which missense variants affect human AIP function and are likely to be a disease-causing mutation (or not) via:

1. Testing the functional conservation between human and fly protein
2. Determining the functional importance of the C-terminal domain
3. Testing which hAIP variants do or do not have a significant impact on protein function
4. Evaluating the expression of human wild-type and missense mutations in *Drosophila melanogaster* via immunoblotting

This structure-function analysis will determine which variant impairs the conserved function of AIP required for viability. This approach will ultimately help to determine the pathogenicity of human AIP missense mutations, with a significant impact on the genetic counselling of FIPA patients.

5.4 RESULTS

As seen in Chapter 3, I have shown that *CG1847*$^{exon1\_3}$ is a null mutation for *CG1847*, the *Drosophila* orthologue of mammalian *AIP*. Since *CG1847* is located in the X chromosome, males lacking *CG1847* function are lethal, while heterozygous females are fully viable. Therefore, *CG1847* mutant flies are a very useful model to test the functional conservation between human and *Drosophila* proteins. Expression of mammalian AIP can functionally complement CG1847 and rescue the lethality of *CG1847* mutant males (Figure 38). This functional complementation can be extended to AIP mutant variants, in order to determine whether they are likely to be pathogenic or not, in an *in vivo* setting. Therefore, if a specific AIP variant rescues the lethality phenotype of *CG1847* mutant flies, it would strongly suggest that the variant does not cause a major functional disruption of AIP function. Conversely, a failure to rescue the lethality phenotype would indicate that the variant is likely to be non-functional and can possibly account for the disease. In summary, this *in vivo* approach has the potential to identify the functional relevance of the hAIP protein domains and the pathogenicity of human AIP missense variants.

The AIP protein contains several identifiable domains, which are similar between human and *Drosophila*, (Chapter 1, Figure 6 and Figure 11). Structurally, in its N-terminal region AIP shares a significant degree of homology with immunophilins, as it has a peptidyl-prolyl *cistrans*

isomerases (PPIase)-like domain. However, this protein does not function as an immunophilin. The PPI-like domain has been described as a protein-protein binding domain[264]. It was observed that it is required for the stabilisation and nuclear translocation of the dioxin receptor-hsp90-XAP2 complex[194,459]. The C-terminal part of AIP contains three antiparallel helices that constitute TRP domains and a final α-helix. The α-7 helix is essential for AIP function, and its role as a protein-protein interaction domain has been well established[142]. To investigate the functional conservation between the human and the fly protein, the importance of the α-7 helix and to test the pathogenicity of 4 human missense variants in our *in vivo* model, 6 different *UAS-hAIP* constructs were generated and injected into *Drosophila* embryos to obtain transgenic lines.

5.4.1 Selection of AIP variants

For this study constructs with wild-type AIP, truncated and missense AIP variants were generated. The 4 missense mutations were selected as detailed below.

AIP p.R16H (rs145047094, c.47G>A, chr11:67483205)

p.R16H is a very controversial missense AIP variant, which to date has not been conclusively classified as a true mutation or as a rare benign polymorphism. However, a considerable amount of data supports the idea that this is a benign variant. Based on the work of Guaraldi *et al.*[460], the AIP p.R16H variant does not segregate with the disease. Moreover, the R16H mutation does not affect the AIP-RET interaction[162]. Consequently, this variant was selected as a positive control.

AIP p.C238Y (rs267606569, c.713G>A, chr11:67490383)

In contrast to p.R16H, the AIP p.C238Y missense variant was shown to affect both cell proliferation[89] and the PDE4A5 binding assays[136]. Based on the crystal structure of AIP, this amino acid is predicted to be involved in protein folding. In 2012 Morgan *et al*. described that their attempts to purify C238Y resulted in protein aggregating suggesting that the protein was at least partly unfolded. The C238Y mutation causes destabilisation of the packaging of α and β helices of the second TPR motif[142]. All the available data indicate that C238Y is a true disease-associated mutation, with a strong pathogenic role and, therefore, it is not expected to compensate for *CG1847* loss of function.

AIP p.A299V (rs148986773, c.896C>T, chr11:67258367)

Functional studies for the A299V AIP mutant variant have shown that the mutation does not affect the AIP-RET interaction[162]. In addition, this mutation causes only a slight reduction in PDE4A5 binding *in vitro*[136]. However, similar to C238Y, the attempts to purify A299V resulted in protein aggregation[142]. The A299V mutant variant was first described in a sporadic acromegaly patient[461]. This mutation was identified in 5 subjects of a family where the p.R304* mutation was also detected. Two unaffected patients carried both changes, one unaffected patient carried only A299V, while one young female carried only A299V and was diagnosed with a microprolactinoma at the age of 30 years. As the 2 subjects with the compound heterozygote genotype (double mutants for p.A299V and for the certainly pathogenic p.R304*) were unaffected, A299V variant is unlikely to have a functional impact as this will lead to possibly totally silencing of the second AIP alleles. Based on the available clinical data, the AIP A299V mutant variant is not considered pathogenic, but there further functional studies are required to support this conclusion.

AIP p.R304Q (rs104894190, c.911G>A, chr11:67258382)

One of the most frequently reported changes is p.R304Q. R304Q is categorised as pathogenic mainly based on clinical data. Multiple groups have identified this change in patients with pituitary adenomas[89,122,136,162,208,209,456,461-463]. However, none of the functional data support the pathogenic role of the R304Q mutation. Taking into account the increased discrepancy between the numerous clinical cases and functional studies, I decided to test the p.R304Q variant in the *in vivo* model.

Assembly of hAIP protein constructs

The pUAS-k10-hAIPwt plasmid was used for the assembly of hAIP truncation or missense constructs. Deletion of the last α helix and the missense mutations were performed via PCR-mediated mutagenesis as described in section 2.3.8. in Material and methods and Figure 76 below, with primer combinations corresponding to the specific change (Appendix 5).

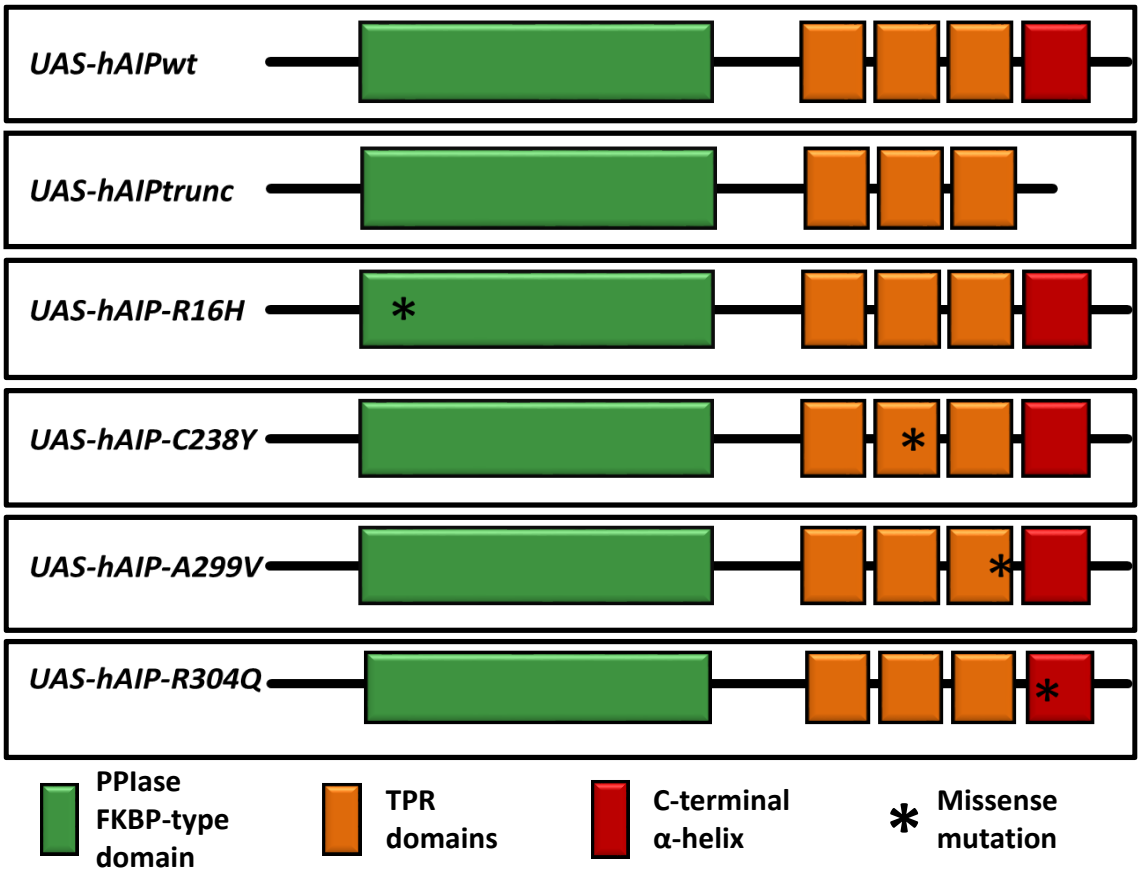

Figure 76: Schematic diagram of the *UAS-hAIP* constructs. The full-length construct *UAS-hAIPwt* is shown at the top. The deletion of the 7th α helix is represented in the second image. The position of amino acid changes introduced to match the missense mutations into *hAIPwt* domains are indicated with a star. Protein domains are indicated by the colour code shown below the deletion construct assembly.

The positive clones for the mutations were identified via sequencing, as shown in Figure 77. Constructs were microinjected into fruit fly embryos harbouring *attP40* landing sites, which enabled the generation of transgenic stocks with human AIP constructs on the second chromosome. Given that all transgenes are inserted in the same genomic locus, they are predicted to be expressed at similar levels and to not have differences due to positional effects. These transgenic fruit flies stocks were balanced over the balancer chromosome *CyO.* Once available, transgenic males were crossed into the *CG1847* deficient mutant background and their ability to rescue the lethality of *CG1847* mutants was examined. The ubiquitous *actin-Gal4* driver was used to drive the expression of the *UAS-hAIP* constructs during fly development. The degree of rescue was analysed based on the percentage of non-FM6, non-CyO males within the total number of viable adult males in the next generation.

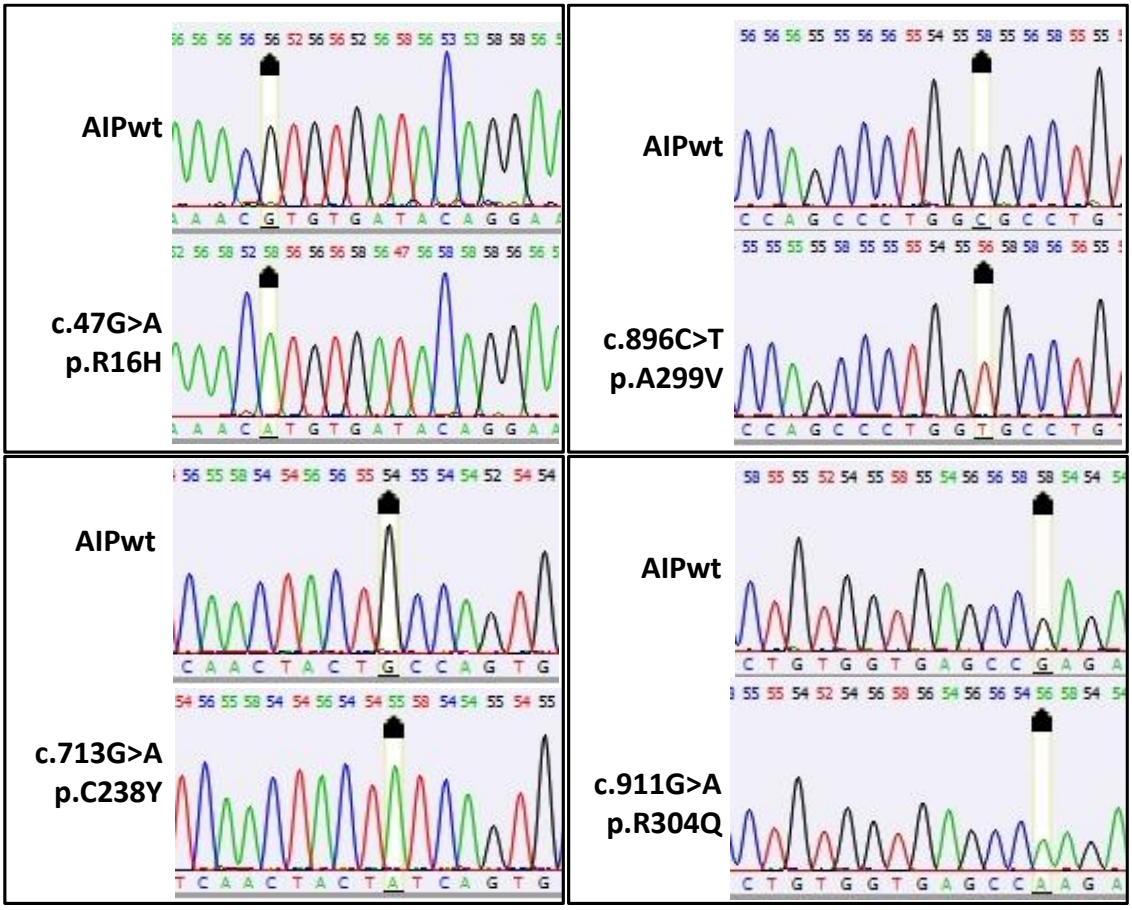

Figure 77: hAIP chromatograms. Partial nucleotide sequences of pUASK10 hAIPwt compared to the sequences of UAS-hAIP missense mutant constructs.

5.4.2 Human *AIP* in *Drosophila* is able to functionally compensate for *CG1847* loss of function

All rescue crosses were performed at 25°C using females heterozygous for the loss-of-function mutation *CG1847*$^{exon1\_3}$ and carrying the *actin-Gal4* ubiquitous driver on the second chromosome (section 2.2.13). Heterozygous *CG1847*$^{exon1\_3}$*/FM6; actin-Gal4/CyO* females were then crossed to transgenic males carrying different human AIP variants (*X/Y; UAS-hAIP/ CyO*). To analyse the data from two independent biological replicates, males from two transgenic lines of each hAIP genotype were individually crossed with *CG1847* heterozygous mutant females and each cross was performed in triplicate. Overall, six different crosses were performed for each hAIP variant. Successful rescue of lethality was scored as the presence of non-FM6 F1 males (with the genotype *CG1847*$^{exon1\_3}$*/Y; actin-Gal4/UAS-hAIP*), which lack endogenous *CG1847* as they inherited the deleted allele from their mothers. Figure 78 depicts the result of the rescue

experiment for the stock carrying hAIPwt on the second chromosome. Only the male genotypes are shown, as only mutant CG1847^{exon1_3} hemizygous males will suffer lethality in the absence of Actin-Gal4 driven transgene expression.

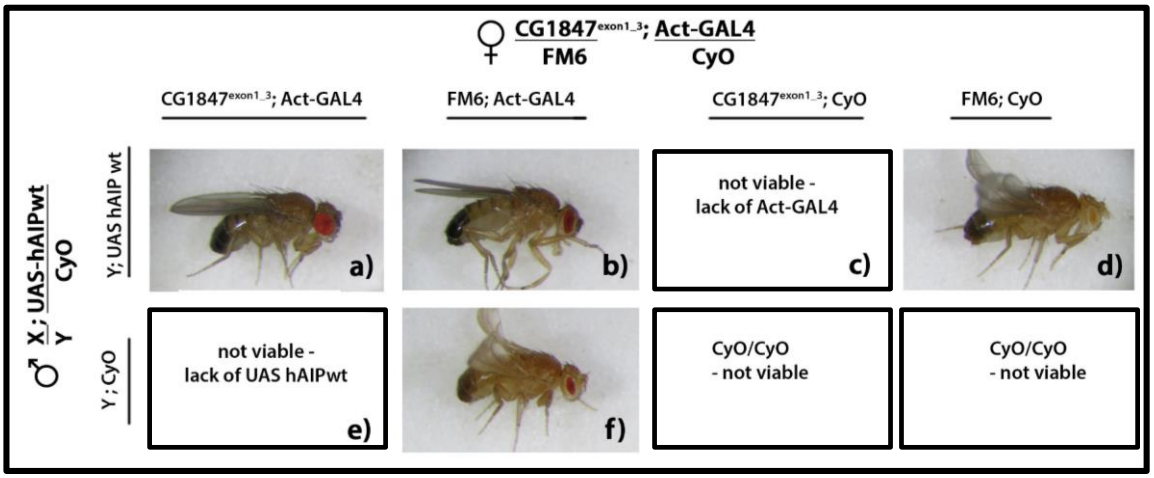

Figure 78: Results of rescue experiment with wild-type AIP. Images of F1 viable males showing: a) Rescued males with full-length wild-type human AIP gene. These males inherited from their mothers the mutant *CG1847* allele on the X chromosome, but are rescued by expression of the human AIP transgene on the second chromosome; c) and e) Males inheriting the mutant *CG1847* allele on maternal chromosome X are not viable in the absence of hAIP expression due to either lack of driver – c) or rescue transgene – e) of the UAS-Gal4 system. These two genotypes also serve as internal negative controls; b) and d) Males expressing a wild-type CG1847 allele on the FM6 chromosome X balancer are viable. The parents genotypes and allele segregation are shown above and to the left of images

Expression of wild-type *hAIP* resulted in a high percentage of rescued males, as this genotype accounted for almost 33% of the total number of viable F1 males. Therefore, ubiquitous expression of UAS-hAIPwt was able to rescue the lethality of *Drosophila CG1847^{exon1_3}* mutants, confirming that *AIP* gene is functionally conserved between flies and humans.

Based on these results, I demonstrated that *CG1847* is the functional *Drosophila* orthologue for human *AIP*.

5.4.3. The AIP C terminal α-7 helix is essential for the conserved function of AIP

To evaluate if the C-terminal portion of AIP, or more precisely, the last α helix is required for AIP function *in vivo*, I generated a rescue construct lacking this helix (section 2.4.5 and Figure 76). To obtain the genomic rescue construct pUASk10attB-AIPtrunc (Figure 79), 907 bp of hAIP together with the upstream Myc tag were amplified from a pcDNA3-Myc-AIPwt vector. The last 86 bp of AIP, which encode for the 7th alpha helix, were deleted to generate a truncated hAIP variant (hAIPtrunc).

Foward frame 1

994 nucleotides, 331 amino acids, structure: sequence C

```
  1 GRGNSMAGTM EQKLISEEDL MADIIARLRE DGIQKRVIQE GRGELPDFQD GTKATFHYRT
 61 LHSDDEGTVL DDSRARGKPM ELIIGKKFKL PVWETIVCTM REGEIAQFLC DIKHVVLYPL
121 VAKSLRNIAV GKDPLEGQRH CCGVAQMREH SSLGHADLDA LQQNPQPLIF HMEMLKVESP
181 GTYQQDPWAM TDEEKAKAVP LIHQEGNRLY REGHVKEAAA KYYDAIACLK NLQMKEQPGS
241 PEWIQLDKQI TPLLLNYCQC KLVVEEYYEV LDHCSSILNK YDDNVKAYFK RGKAHAAVWN
301 AQEAQADFAK VLELDPALAP VVI*KKSLVN S
```

Figure 79: *pUAS-hAIPtrunc* DNA to protein sequence translation. Schematic representation of the protein encoded by the construct containing a truncated AIP version. Myc-Tag (blue) the Start (green) / Premature Stop (pink) positions are coloured in the protein sequences. Translator, online free software (http://www.fr33.net/translator.php)

Transgenic males for hAIPtrunc were crossed with females heterozygous for CG1847 deficiency. A small number of non-FM6 non-CyO males were found in F1 generation. Between zero and 5 adult males were recovered in most of the vials. DNA was extracted from individual males to determine whether these males inherited the normal or mutated *CG1847*$^{exon1\text{-}2}$ allele via a diagnostic PCR approach (below Figure 85). No rescue males were found, and altogether, the rescue experiments with the hAIP truncation construct confirmed previous published data indicating that the last α helix is necessary and essential for AIP function.

As a proof of principle, this construct was used as a negative control for subsequent *in vivo* rescue experiments. The ubiquitous expression of the truncated hAIP transgene is not sufficient to compensate for the mutant CG1847^{exon1_3} on the X chromosome and this validates our model.

5.4.4. Human AIP missense variants differ in their ability to rescue *CG1847* loss of function mutants

Having demonstrated that human AIP expression in *CG1847* mutant flies is able to rescue the male lethality phenotype, while a truncated construct does not, I next tested the rescue proprieties of hAIP variants found in FIPA families.

To explore the pathogenicity of several hAIP missense mutations, four additional constructs were tested. Using a similar approach to the one described above, I used the *UAS-GAL4* system to express human *AIP* transgenes of the c.47G>A (p.R16H), c.713G>A (p.C238Y), c.896C>T (p.A299V), and c.911G>A (p.R304Q) missense variants and assessed their capacity to rescue male lethality.

To investigate if the p.R16H, p.C238Y, p.A299V and p.R304Q hAIP variants indeed compensate for loss of *CG1847*, I performed new crosses, similar to the one described above. Successful rescue was scored by the presence of non-FM6 F1 males (of the genotype *CG1847*$^{exon1\_3}$*/Y; actin-*

Gal4/UAS-hAIP variant). For statistical analysis, for each missense variant, the percentages of the rescued male genotype (from the total of viable males) were compared in with the percentages of non-FM6, non-CyO males resulting from the rescue crosses with wt and truncated hAIP constructs.

When the percentages of rescued males within each of the hAIP constructs rescue crosses were evaluated and compared to the results obtained with the wt and truncated constructs, hAIP variants separated into 2 classes (Figure 80).

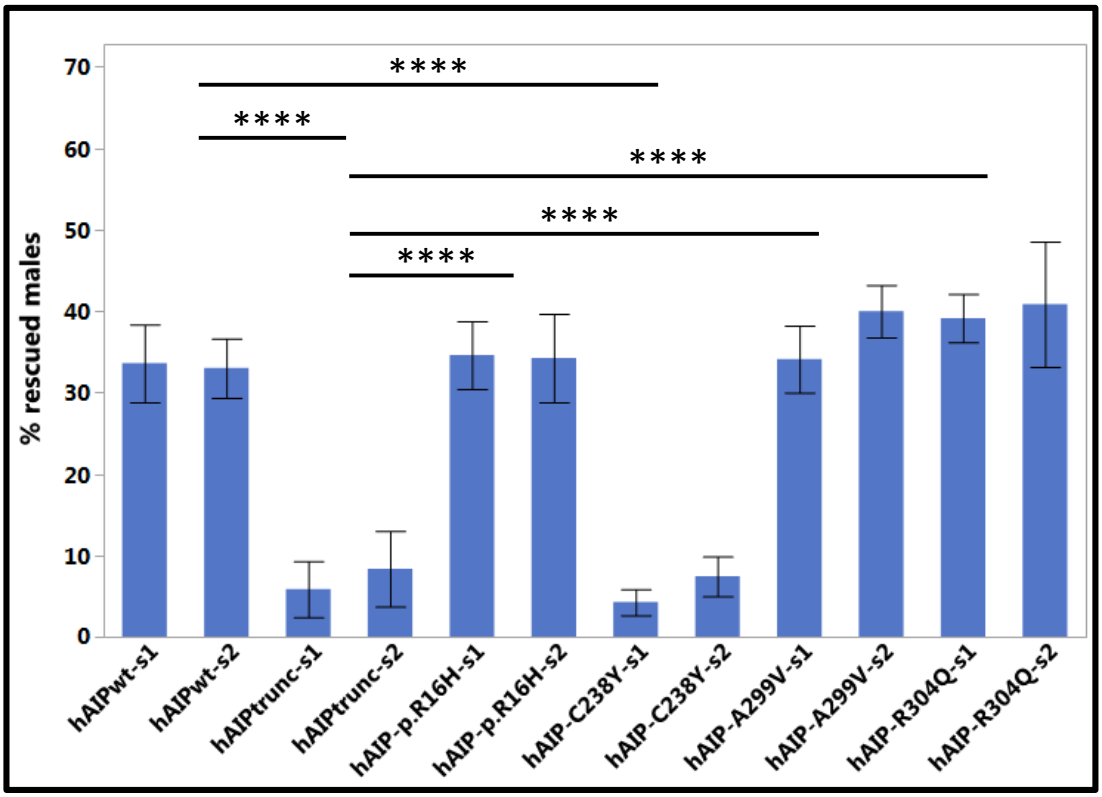

Figure 80: Quantitative analysis of *in vivo* rescue experiments using hAIP missense variants. The rescue of the lethality phenotype by expression of UAS-hAIPwt, UAS-hAIPtrunc and the various missense mutant constructs is represented as the percentage of F1 males of the desired genotype relative to the entire F1 male population. Bar graph shows the quantification of the rescue results, with 2 different stocks (s1 and s2) for each construct (each stock in triplicates). Error bars represent SE. Significant differences between samples are indicated by asterisks (****P<0.0001; Oneway Anova with post-hoc Tukey-Kramer analysis test).

hAIP variants are separated into 2 classes by the proportion of rescued males. Ubiquitous expression of p.R16H, p.A299V and p.R304Q variants (at 25°C) rescued very efficiently the lethal *CG1847*$^{exon1\_3}$ mutation, at a similar rate to the ubiquitous expression of wild-type hAIP. There was no significant difference between these constructs regarding the percentage of rescued

males. In contrast, like the truncated version of hAIP, ubiquitous expression of the pathogenic variant p.C238Y was unable to rescue the male lethality of *CG1847* mutants (p=0.0001).

In fact, it should be noted that all the males scored as rescued males in the progeny of crosses with the truncated hAIP or p.C238Y mutant were actually flies with aberrant genotypes, as shown via PCR genotyping (Figure 84 and Figure 85).

The p.R16H, p.A299V and p.R304Q missense variants have similar rescue capacities as the wt hAIP

A statistical analysis comparing the distribution of all the male genotypes resulting from the rescue experiments show that there is no statistical difference between the capacities of the 3 hAIP missense variants (p.R16H, p.A299V, and p.304Q) vs wt hAIP in rescuing the male's lethality.

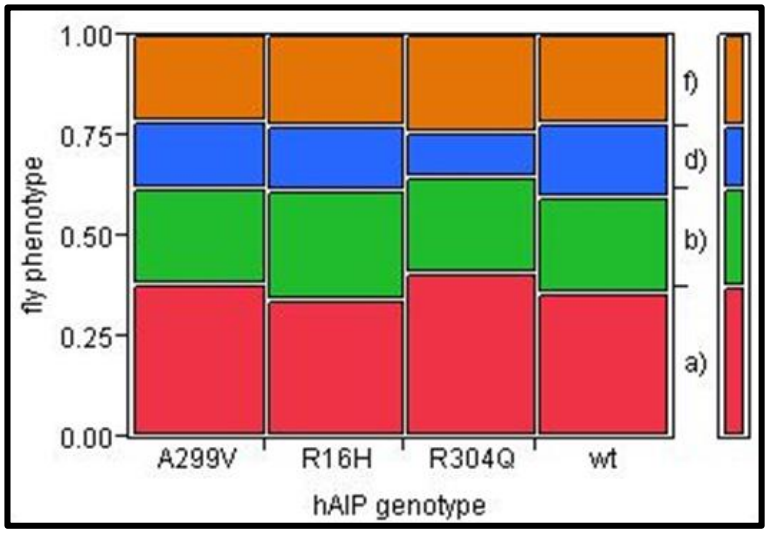

Figure 81: Similar distribution of the males phenotypes To test the distribution of phenotypes in the male progeny of crosses with the four hAIP variants that rescued the lethality of CG1847^{exon1_3} mutants, a Chi square contingency test was performed. Each viable male phenotype received a colour code: a) rescued males (genotype *CG1847^{exon1_3} /Y; actin-Gal4/UAS-hAIP variant)* are indicated in red; b (*FM6/Y; actin-Gal4/UAS-hAIP)* - green, d) (*FM6/Y; UAS-hAIP/CyO)* - blue, and f) (*FM6/Y; actin-Gal4/CyO)* - brown. There was no significant difference in the distribution of male genotypes between the F1 generation of ubiquitously expressed p.R16H, p.299V and p.304Q hAIP variant and the F1 generation of the wt hAIP (positive control) (p=0.9790).

As seen in Figure 81, Chi square contingency test for the three hAIP missense variants and wt hAIP shows that there is no significant difference in the distribution of male genotypes between

the F1 generation of ubiquitously expressed p.R16H, p.299V and p.304Q hAIP variants and the F1 generation of the wt hAIP (positive control) (p=0.9790).

I found that wt hAIP, p.R16H, p.299V and p.304Q hAIP constructs had the same rescue capacity as there was no significant difference in the distribution of the four possible male phenotypes (p=0.97, Chi-square statistical analysis).

5.4.5 Validation of rescue experiments by PCR genotyping

To confirm the genotypes of the F1 offspring, genomic DNA was extracted from a proportion of the rescued males from each of the rescue crosses. For each male, three different PCR reactions were performed to determine the genotype of the males: one for the detection of the mutant *CG1847*$^{exon1\_3}$ allele, a second for detection of the *hAIP* transgene, and a third one for the amplification of *PPr-Y. PPr-Y* is a gene located in the Y chromosome, which was used to detect the presence of the Y chromosome in the animals that were phenotypically males. This is necessary because a wild type male phenotype can also be produced by offspring with a X0 genotype, which can occur by non-disjunction of the X chromosome pair in the mother (details in the Discussion section).

Figure 82 shows the PCR products of genomic DNA separated by agarose gel electrophoresis for 12 males from the rescue cross with hAIP wt. The expected PCR products in the rescued males are of a 1000bp for CG1847^{exon1_3} (instead of 2500bp the amplicon corresponding to the normal copy).

Eight out of the 12 males carried the deleted *CG1847* allele and were rescued by hAIP present on the second chromosome. Three males (m5, m6, and m12) were viable due to the presence of the wild-type copy of *CG1847* (Figure 82 top panel) inherited from the male parents (Discussion section). The m6 and m12 males do not carry the Y chromosome as they result from a nondisjunction event in females. Occurrence of chromosome nondisjunction in females leads to the production of eggs containing either both maternal X chromosomes (XX) or none (0). The eggs with both X maternal chromosomes (XX) can combine with the Y chromosome from males resulting in individuals that are XXY and are phenotypically female. Eggs without maternal contribution (0) will inherit the X chromosome from males and the resulting X0 combination will have a normal *CG1847* allele and will appear phenotypically male.

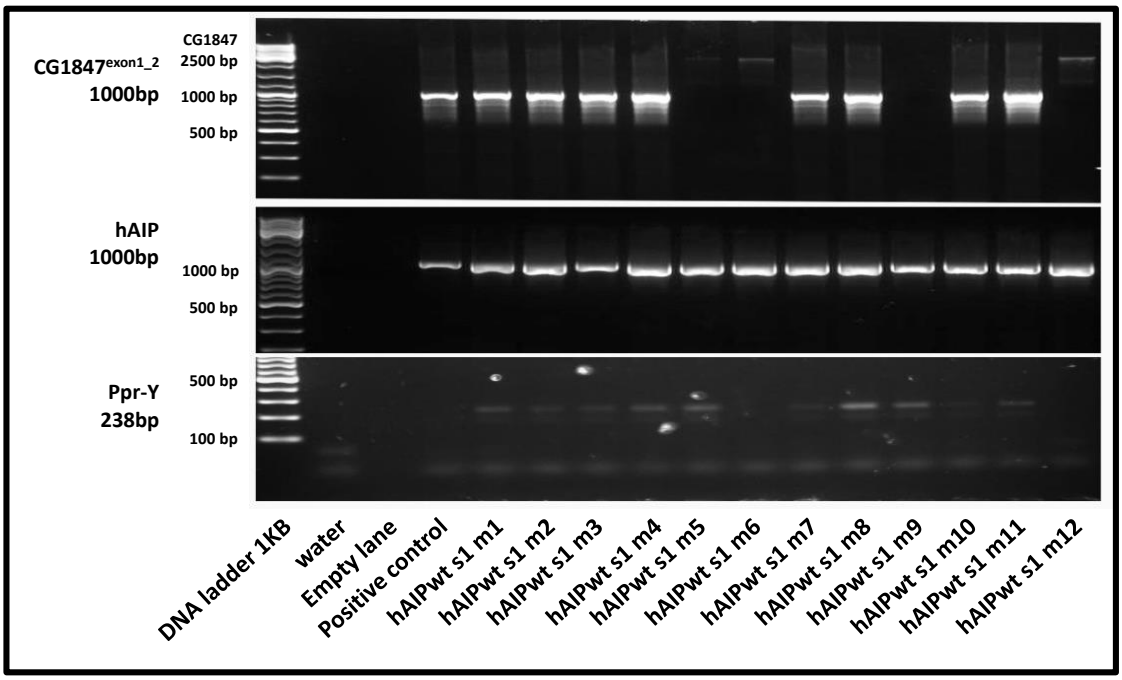

Figure 82: Genotyping of rescued males with hAIPwt. The resulting rescued males (Figure 78A) were genotyped by PCR. The *CG1847* gene was amplified using a pair of primers that produce a 2500 bp amplicon for the wt allele and a 1000 bp for the mutant allele (top panel). The hAIP transgene was detected using primers against human AIP cDNA (middle panel). In addition, the presence of Y chromosome (bottom panel) was detected using a set of primers for the *Ppr-Y* gene. Abbreviation on the figure: hAIPwt s1 – the males used in the rescued cross were from first of the 5 individual lines generated by BestGene; m1-m12 – individual labelling of each of the 12 males collected for genotyping. Eight out of 12 males had the expected rescue genotype as they carried the $CG1847^{exon1\_3}$ allele and hAIP cDNA on the second chromosome. Males m5, m6, and m12 do not carry the mutant *CG1847* allele, and therefore are the result of maternal chromosome X non-disjunction. Although phenotypically male, males m6 and m12 do not carry the Y chromosome. For sample m9 I could not amplify the CG1847 amplicon possibly due to a technical issue.

Additionally, I observed 2 types of adult males in the F1 progeny that were phenotypically similar to genotypes $CG1847^{exon1\_3}$*/Y; UAS-hAIP/CyO* (Figure 78C) and $CG1847^{exon1\_3}$*/Y; actin-Gal4/CyO* (Figure 78E). These males are not expected to be viable, since they do not have the proper combination of both GAL4 driver and UAS-hAIP transgene on the second chromosome (Figure 83). However, although these males are missing either the GAL4 driver or the UAS-hAIP they are viable as they are the result of the same nondisjunction phenomenon mentioned above. Consequently they have aberrant genotypes.

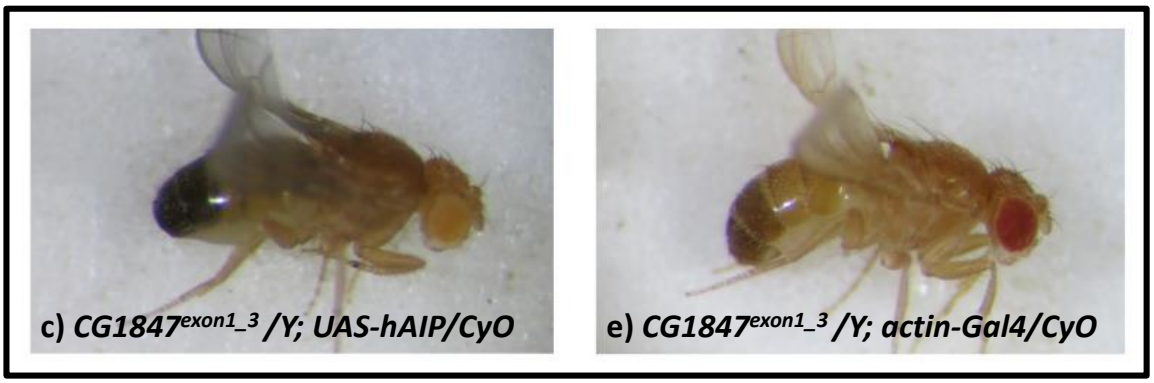

Figure 83: Two types of F1 males due to nondisjunction. These males are phenotypically similar to genotypes *CG1847^{exon1_3} /Y; UAS-hAIP/CyO* (genotype "c" in Figure 78) and *CG1847^{exon1_3} /Y; actin-Gal4/CyO* (genotype "e" in Figure 78). These males are non-FM6 balancer chromosome (see the normal round shape of the eyes). As they have the *CyO* balancer chromosome inherited on the second chromosome; consequently they are missing one component of the UAS-GAL4 system: either *UAS-hAIP* (c) or the *actin-GAL4* (e), therefore could not express hAIP.

These phenotypes were found in extremely low numbers (in average 4% of viable male offspring). I genotyped the majority of the males with these 2 phenotypes in at least one cross for all the 6 constructs. Figure 84 depicts the result of the diagnostic PCRs for individuals with the aforementioned aberrant phenotypes from the progeny of crosses with hAIP wt. The same genotyping system described above was used for these experiments.

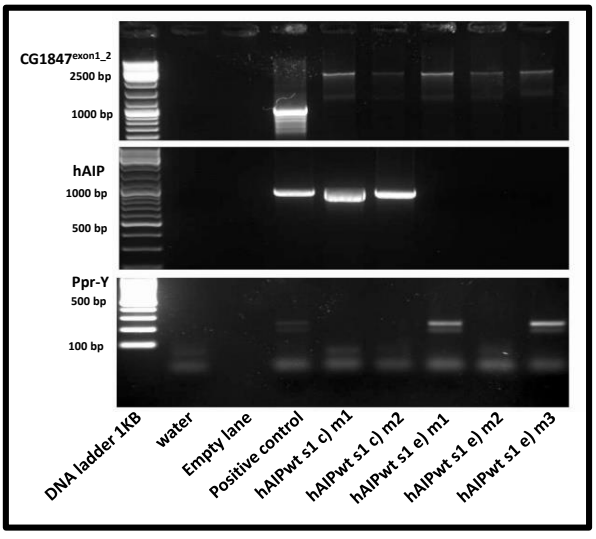

Figure 84: Genotyping for the two aberrant male types. The two types of males described above were genotyped using PCR. All 5 males have the wt *CG1847* copy (top panel, amplicons at 2500 bp). In the first two males (expected genotype *CG1847^{exon1_3} /Y; UAS-hAIP/CyO*) a copy of hAIP cDNA was detected (inherited from the father). As expected, the hAIP copy was not detected in the other three males for which the expected genotype, based on their phenotype, is *CG1847^{exon1_3} /Y; actin-Gal4/CyO*. In addition, in some of these phenotypical males the Y chromosome was not detected.

All 5 males have the normal *CG1847* allele, which is the reason for their viability in the absence of the FM7 balancer chromosome or without the proper combination of UAS-GAL4 system components. The hAIPwts1c) m1 and m2 males have the hAIP construct on the second chromosome, which demonstrates that they inherited this chromosome from transgenic male

parents (carrying the *pUASk10-hAIPwt* construct). On the other hand, in the 3 *CG1847*$^{exon1\_3}$*/Y; actin-Gal4/CyO* males the hAIP cDNA could not be detected as they contain the *actin-GAL4* construct inherited from female parents. Some of these flies were also negative for the Y chromosome, which brings extra proof that they are result of nondisjunction.

Together, these results suggest that the non-FM6 males are due to chromosome nondisjunction in F0 females (more details in the Discussion section). These particular phenotypes, observed in low numbers in subsequent experiments, were considered aberrant genotypes and, therefore, were excluded from statistical analysis.

Regarding the individuals that are phenotypically identical to the rescued males found in the F1 generation of the rescue crosses for hAIPtrunc construct: these males were also the result of the nondisjunction phenomenon. To confirm the genotypes of the F1 offspring, DNA was extracted from all the viable males from one of these crosses and diagnostic PCRs were performed to detect either the wt or the mutant *CG1847*$^{exon1\_3}$ allele.

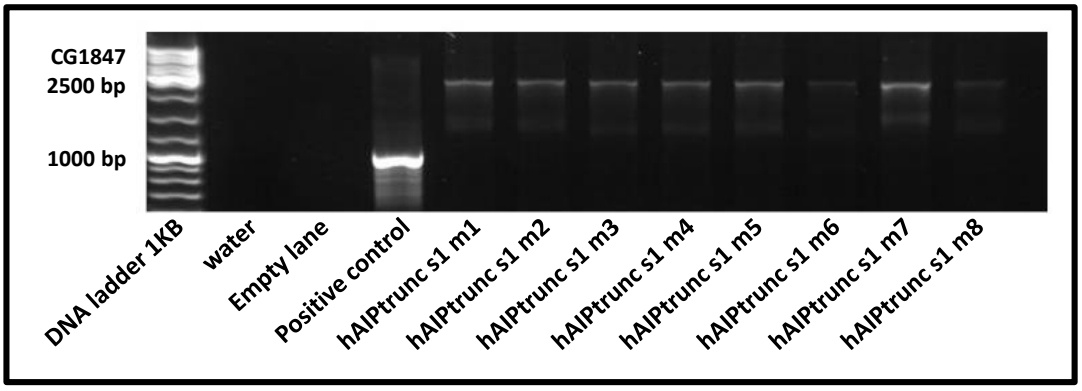

Figure 85: Genotyping of non-Fm6 non-CyO males resulting from crosses with the hAIPtrunc construct. The males were genotyped by using primers for *CG1847.* This gene was amplified using the same set of primers that produce a 2500 bp amplicon for the wt allele or a 1000 bp amplicon for the mutant allele. All 8 males have the normal copy of *CG1847*.

Figure 85 shows the results of the PCR genotyping. All non-Fm6 non-CyO males have the normal copy of CG1847 (2500 bp instead of expected 1000bp size of the amplicon in the CG1847^{exon1_3}). All eight males actually inherited the normal *CG1847* allele, and this is the explanation for their viability. These particular male phenotypes were found in very low numbers in subsequent experiments. The same result was seen in case of hAIP C238Y construct.

The aberrant males in the rescue experiments with hAIPtrunc and hAIPC238Y constructs were used in the statistical analysis as “non-rescue” to evaluate the “rescued or not rescued result” based on a significant difference.

5.4.6. hAIP rescue constructs have equivalent expression levels

To determine the *in vivo* expression levels of AIP in the fruit flies, I used the UAS-GAL4 system to transgenically express human AIPwt, AIPtrunc and AIP with missense mutations (R16H, C238Y, A299V, and R304Q). In the Results section of this chapter it has been shown that ubiquitous expression of AIPwt, R16H, A299V and R304Q during development resulted in rescue of lethality. Total proteins were extracted from the rescued males and Western blots were performed using an antibody against the hAIP. The human protein was detected as a 37 kDa band, similar to the band detected in human HEK293T cells used as positive control (Figure 86). Wild-type flies (w^{iso}), without any of the UAS-GAL4 components, were used as a negative control for the specificity of the antibody to see if the endogenous CG1847 expression is also detected. In addition, transgenic males from the stock carrying the UAS-hAIP constructs (without Gal4 driver) were tested for leaky potential expression of the construct.

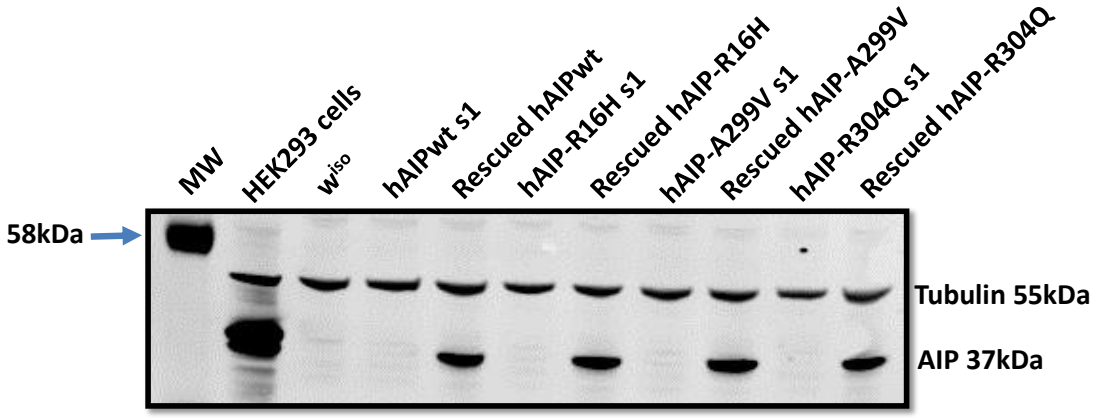

Figure 86: Overexpression of pUAS-hAIP constructs driven by actin-Gal4. The proteins were extracted from adult male heads. Expression level appears equivalent for the different hAIP constructs ubiquitously expressed using the actin-Gal4 driver. In addition, the hAIP expression was not detected neither in wild-type flies (w^{iso}) nor in the transgenic F0 males (as these only carry the UAS-hAIP construct and not the *actin-Gal4* driver). Loading control is anti β-tubulin. The primary antibodies were Novus monoclonal AIP/ARA9 Antibody (35-2) and DSHB monoclonal β tubulin (E7).

These results were consistently observed in four independent Western blots, with independent biological samples representing different protein extractions. All hAIP missense constructs which rescued the lethality resulted in a robust expression, similar to the hAIPwt construct. The 37 kDa band of hAIP missense mutations was detected as having equal intensity when compared to wild-type hAIP.

Based on the Western Blot results it is possible to conclude that the R16H, A299V and R304Q missense variants encode for proteins which not only have the same capacity of rescuing the lethality as the wt hAIP, but they also have the same stability.

Regarding the pathogenic hAIP variants (truncated and C238Y) further experiments are required in order to investigate their levels of expression and protein stability.

5.5 DISCUSSION

In this project, I developed and optimised a new *in vivo* system to test the pathogenicity of AIP mutations using Gal4/UAS transgenes generated with the phiC31 integrase system[275]. The Gal4/UAS system brings significant advantages as it offers the possibility of expressing an exogenous gene in a transgenic animal while controlling the pattern of expression. The site-specific phiC31 phage allows the integration of different constructs into the same landing site and, as the genetic background remains unchanged, the results of the different constructs are directly comparable[464]. This system has been previously successfully used to generate *Drosophila* transgenics for human genes, in order to understand their functional role or the importance of specific protein domains. In 2011, to unravel the underlying genetics, Grossman *et al*. selectively overexpressed mammalian candidate genes in the fly heart. The selected genes were thought to be involved in cardiac heart defects associated with Down syndrome. They investigated the effects of over-expressing the candidate genes by evaluating the basal heart rate and the frequency of heart failure following exposure to stress[465]. The evolutionary conservation of specific genes can also be investigated using a similar approach. In 2014, Ikmi *et al.* examine the Yap/Yorkie (yes-associated protein), which controls the organ size in both *Drosophila* and mammals by generating *Drosophila* transgenic stocks for Yap/Yorkie orthologues from mammalian lineages, including human, and their unicellular relatives. The results brought significant insights into the evolutionary history of Yap protein structure and function[466]. In yet another example, a group studied the implication of different pathological L1CAM mutations in a broad spectrum of neurological and non-neurological phenotypes. By overexpressing different missense mutations of the human gene L1CAM in *Drosophila* it was possible to significantly rescue the neuron guidance defects, showing that some of those variants were not pathogenic[467].

I exploited the power of *Drosophila* genetics approaches to evaluate the degree of functional conservation between fly and human AIP, by testing whether *CG1847*$^{exon1\_3}$ mutant flies can be rescued by human *AIP.* The last AIP α helix had previously been described as essential for AIP function and, as proof-of-principle, I also tested a truncated mutant AIP lacking this domain. Additionally, I tested whether the expression of 4 different missense variants identified in FIPA patients can compensate for the loss of *CG1847*. My data showed that the lethality of

CG1847$^{exon1\_3}$ mutants could be rescued by expression of human *AIP*, demonstrating that *CG1847* is a true functional homologue of AIP and that human AIP is functionally conserved throughout evolution. Conversely, a C-terminus truncated human AIP mutant transgene failed to rescue the defects that cause lethality in *CG1847* mutant flies. These results are the first to provide insights into the functional importance of the AIP α-7 helix in an *in vivo* system. The lethality of *CG1847* mutants was also rescued by 3 of the 4 tested missense variants: p.R16H, p.A299V, and p.304Q. The p.C238Y missense variant, which is considered pathogenic, was unable to rescue the lethality of the *CG1847* knockout. When compared with the truncated or C238Y variant, all the other hAIP missense constructs showed a strong ability in rescuing the lethality, reaching the same statistical significance as the wt hAIP.

Moreover, the Western Blot results showed that the R16H, A299V and R304Q missense variants encode for proteins which have the same stability as the wt hAIP. As the hAIP truncated C238Y variants did not result in rescued males I was not able to extract proteins for Western Blot to evaluate their level of expression. However, this could be achieved via overexpressing these pathogenic variants in the wt background, as now the results in Figure 86 showed that the antibody used against hAIP protein does not detect the endogenous CG1847 protein.

Some of the genes that encode for proteins involved in different cellular processes were proved to be essential (fundamental for organism viability). Many recent studies were interested in determining the gene essentiality and several comparative genomic analyses already confirmed that essential genes are conserved during evolution[468-471]. Even more, most of these genes remain essential in different organisms, as the orthologues of one gene tend to be essential in other organisms. However, some genes have become non-essential maybe due to environmental conditions, or have been lost from the genomes of other organisms[469,471,472].

My data support the evolutionary conservation of the *AIP* gene. As evidence for evolutionary conservation of the AIP gene, I have found that deletion of the endogenous *Drosophila* gene resulted in lethality, similarly to previously data from mice[181,204]. Furthermore, the human gene was able to functionally compensate for the deletion of the *Drosophila* orthologue, *CG1847*.

Regarding the experiments with the truncated protein, no viable rescued progeny resulted, supporting the previous findings that the last α helix is essential for protein-protein interactions. It has been previously reported by that the TPR domain of AIP is required for the interaction with HSP90[459]. In 2012 Morgan *et al.* reported that the highly conserved C-terminus of AIP and the 7th α helix in particular, are involved in AIP interactions with different partners as AhR and PDE4A5. Consequently, loss of these domains leads to loss of interactions with these partners[142].

AhR is best known for being involved in the defence mechanisms against halogenated dioxins and carcinogenic environmental chemicals[473]. In the view of the fact that AIP interacts with the chaperones at the level of the TPR motifs, it is possible that the loss of the last α helix could interfere with the normal folding of the TPR motifs. Under our experimental conditions, the results obtained with truncated AIP also predict that this final structure is essential for the normal function of the protein and binding to its partners. Consequently, the loss of the last α helix impaired the binding of AIP to different partners and results in lethality. Although this interferes with the translocation of the AhR into the nucleus, resulting in loss of activation of dioxin response elements, this might not be the mechanism leading to lethality as *Ahr*-null mice are viable[205]. Further investigations are required to determine the actual cause of lethality.

In accordance with presented results my literature review regarding the p.R16H AIP variant revealed that most arguments favour a non-pathogenic role.

The R16H variant was first reported in 2007 by Daly *et al*. in two cousins with acromegaly[151]. The initial hypothesis was that this variant is a disease causing one, but it was refuted only one year later by Georgitsi *et al.*[460,461]. Other groups also support the fact that it represents only a rare SNP[179]. A recent article reported a 3 generation Italian FIPA family with the R16H AIP change identified in 8 different individuals. However, only 2 individuals were diagnosed with pituitary adenomas and their conclusion was that this variant should be regarded as a rare polymorphism[474]. This variant does not segregate with the pituitary adenomas; however, there are no sufficient data to exclude a possible involvement of the p.R16H variant in these families, as part of an oligogenic model. The LOH status was investigated in pituitary tumorous tissues only in two patients and the wild-type allele was present[456,461].

This amino acid residue is moderately evolutionarily conserved[151]. The basic R amino acid is substituted by another basic amino acid. Consequently, the overall change in the protein structure is expected to be with no or minor deleterious effects. Furthermore, this variant seems to be relatively stable, it degrades at similar speed as the wild-type protein as shown in Figure 86 and in an independent project (Hernandez-Ramirez *et al*., unpublished data).

The impact of this amino acid change on AIP splicing has never been tested systematically. Two studies which performed a series of *in silico* prediction tests (as Berkeley *Drosophila* Genome Project, ESEfinder 2.0, Splice Scan, Alternative Splice Site Predictor, and NetGene2 programs) concluded that p.R16H is likely pathogenic[208,475]. At the same time, PolyPhen2[453] reports this variant as possibly damaging. However, our investigation using 2 prediction softwares, Alamut and Pathogenic-or-Not Pipeline (PONP) concluded that this AIP variant is not pathogenic. The

frequency of rs145047094 (p.R16H) is low. The dbSNP[476,477] reports an uncertain clinical significance and a frequency of 0.1% positive individuals (http://www.ncbi.nlm.nih.gov/snp/?term=rs145047094), while the EVS[478] (Exome Variant Server) estimates frequencies up to 0.7% for the White American sample (n= 4295 individuals) and up to 0.14% for the African American sample (n=2200 individuals) (http://evs.gs.washington.edu/EVS). In the ExAC database there is one case reported homozygous for this *AIP* variant. Therefore, there are no convincing data in favour of a pathogenic role for AIP p.R16H variant which is supported by our *in vivo* model.

The p.C238Y variant was first described by Leontiou *et al*. in 2008[89]. This change was identified in a Mexican family with 3 members being diagnosed with acromegaly. All the arguments are in favour of a pathogenic role for this variant. The functional investigations have revealed that it has a reduced ability to block cellular proliferation[89] and no interaction with PDE4A5 was detected[89]. The LOH status was investigated in pituitary tumour of the three Mexican patients and the wild-type allele could not be detected in any of the samples[124]. This is in agreement with the second hit hypothesis.

This amino acid residue is moderately evolutionarily conserved. The amino acid substitution is conservative as a polar amino acid, cysteine, is exchanged with tyrosine, which is also a polar amino acid. However, there is a possible loss of a disulphide bound. The functional impact is significant, with serious disruption on the folding of the TPR domains, the packing of the hydrophobic core and severe steric clashes of the 3D protein structure[142]. Unpublished data from our lab reports that this protein is very unstable, with a very short half-life comparing with the wild-type AIP, and is rapidly degraded[479].

PolyPhen2 reports this variant as damaging with a very high score of 0.994. Prediction based on the Alamut software came to the same conclusion. This variant is not reported in dbSNP or EVS. The ExAC database[480] reports 1 allele out of 118770 was found.

In summary, all the available data, including our *in vivo* model, support the pathogenic role of p.C238Y.

The p.A299V variant (rs148986773) was first identified in 2007 by Georgitsi *et al*.[461] in a Dutch patient with a GH secreting pituitary adenoma. All the clinical and in vitro data are in favour of a non-pathogenic role for this variant and that A299V change may be just a rare polymorphism[136,481]. In a very interesting FIPA family five family members were identified with the A299V variant, two of them harbouring a truncating AIP mutation, (p.R304*), on different

alleles and without pituitary adenoma. The A299V missense mutation does not segregate with the disease and the LOH status of these patients was never investigated.

A299V variant is localised at the border between the second and the third TPR domains, in a pocket essential for the interaction of AIP with HSP90. This AIP region is among the most conserved ones[142] and it is very important for proper folding of the protein. Misfolded structures could be unstable and result in rapid degradation of the protein[482]. The amino acid substitution is conservative as a nonpolar amino acid (alanine) is exchanged with valine, also a nonpolar amino acid.

Although PolyPhen2 reports this variant as possible damaging, PONP and Alamut consider it as not pathogenic. The frequency of rs148986773 is low. The dbSNP reports an uncertain clinical significance and global MAF of 0.001. In the EVS the frequencies of this variant are 0.14% for the White American samples (n=4293 individuals) and 0.023% for the African American samples (n=2195 individuals).

In accordance with our results, a literature review regarding the p.A299V AIP variant revealed that most arguments favour a non-pathogenic role.

The most controversial missense variant is the relatively frequent exon 6 genomic change c.911G>A, p.R304Q. The mutation is located in the C-terminal α-7 helix of the AIP protein, which is a CpG island hotspot.

It`s pathogenicity is based mainly on clinical data. Multiple groups identified this change in patients with pituitary adenomas. A recently published article performed a literature review of these cases and revealed that p.R304Q seems to be much more common than the previous discussed missense mutations. So far 23 patients have been reported. The majority of these, 20 patients, are familial cases, while the other are just sporadic cases[117]. However, none of the functional available data do not support the pathogenicity: neither the disruption of PDE4A5[136] or RET interaction[162]. Moreover, no LOH was detected in a patient's tumour tissue (unpublished data).

On the other hand, the predicted changes in the protein structure bring arguments in favour of a pathogenic involvement. This amino acid residue is moderately evolutionarily conserved, especially in mammals. The arginine, a long side chain positively charged amino acid, is exchanged with glutamine, a polar and slightly shorter, uncharged, hydrophilic amino acid[463]. This mutation does not affect the HSP90 binding site directly; however, an abnormal C-terminal tail may disrupt the folding of the third TPR domain, and this may result in disruption of the

normal binding to the HSP90 MEEVD motif[142]. The half-life of the mutated protein was found to be slightly shorter than that of the wt protein (Hernandez *et al*., unpublished data).

Additional information based on studies which used web prediction tools were contradictory. Occhi *et al*[483] published that according to PolyPhen R304Q is a deleterious mutation. On the other hand, using a newer version of the same prediction pipeline, PolyPhen2, Tichomirowa *et al*[208] reported that R304Q is a benign polymorphism. Our own investigation using 2 other *in silico* web tools (Alamut and PONP) showed that this variant is not suggested to be pathogenic. This is a rare variant as EVS reports a MAF of 0.0693 (the frequency for European American population is 0.1048, while for African American is 0). Interestingly, the ExAC database reports the presence of this variant in homozygosity in two out of 37894 Europeans and none of 20413 non-Europeans, further increasing the uncertainty about its pathogenic role.

The results of our project also not favour a pathogenic role for the R304Q mutation, even though the clinical suspicion is very high. As this AIP variant is able to rescue the *Drosophila* lethality and 2 individuals were reported homozygous in the ExAC database, our conclusion is that p.R304Q alone is not the cause of pituitary adenomas. One possible hypothesis is that p.R304Q may change the AIP activity in addition to a different gene (oligogenic model), which would explain the retention of the wild-type allele and no LOH in the pituitary adenoma tissue. One other possibility is that p.R304Q might be just a surrogate marker, in reality another mutation with which it is associated by linkage of disequilibrium, is the cause of these pituitary adenomas. One last hypothesis is that maybe the AIP function is very complex, different amino acids being more or less important in different processes. As a consequence, a residue which is important for tumorigenesis may not interfere with normal fly development.

Nondisjunction

phiC31 recombination brings a series of significant advantages to *Drosophila* genetics. Maybe the most important feature of this system is the non-random integration. As the catalysed phiC31 integrase recombination process is site-specific, it allows a precise pre-selection of the insertion location. By eliminating the randomness of transgenesis while keeping the same genetic background, this system allows the integration of different constructs into the same landing site and the advantage of a direct comparison between results[464]. In this project I designed a vector-based system to facilitate *in vivo* AIP missense mutation analysis using phiC31 recombination.

Our system offers a number of advantages as well as some limitations. First, by design, using the same backbone to insert the different constructs at the same *attP* specific site in the *Drosophila*

genome allows equivalent transgene expression. We proved this by evaluating the level of expression of different variants capable of rescuing the fly lethality by immunoblotting (Figure 80 and Figure 86). Homogenous expression for different constructs is essential for the comparison of wild-type and mutant transgenes[484]. Another important advantage of our constructs, a key feature for transgenic animal identification, is the use the mini-*white* gene, a common and easy to use phenotypic marker.

Second, the pUAS-K10 vector used for generating *Drosophila* transgenics for different AIP constructs was selected based on results of previous studies which had shown a low basal and highly inducible expression of integrated transgenes[485]. This backbone has the significant advantage of avoiding non-specific 'leaky' transgene expression, as shown in Figure 86.

Third, the majority of constructs used in this project were generated via site directed mutagenesis starting from the same vector. This system allows multiple parallel cloning of different missense mutations into the same vector, and simplifies the generation of transgenes. This strategy overcomes the disadvantages of other techniques such as the Gateway technology which may also be more expensive. Therefore this system has several advantages and we hope that the research community interested in the pathogenicity of AIP mutations will take advantage of it.

Regardless of the presented results and the numerous advantages, our *in vivo* model faces some limitations among which is the issue of non-disjunction phenomenon which lead to a degree of false positive results. The first data regarding the discovery of non-disjunction phenomenon and its effect upon sex-linked inheritance was published in 1913 by Bridges C.B.[486], and was explained in detail a few years later[487,488]. This first study of spontaneous X chromosome nondisjunction in *Drosophila*, although completed more than 100 years ago provided the insights regarding the mechanisms of non-disjunction of the chromosomes during meiosis. Over the last few years, studies on *Drosophila* genetics have produced significant advances regarding the molecular mechanisms of how meiotic chromosome pairing, synapsis and segregation take place[489-491].

Despite several controversies regarding which is the critical stage of the cell cycle when the normal separation of homologous chromosomes is impaired resulting in non-disjunction, there are a few aspects on which there is agreement. First of all, spontaneous recombination in *Drosophila melanogaster* males is an extremely rare event and the nondisjunction phenomenon takes place mainly in females. Second, the X chromosome nondisjunction is much more frequent in XXY females than it is in normal XX females. These XX-Y segregation events were called

''secondary nondisjunction"[487]. Third, it has been revealed that abnormalities in genetic recombination will consequently perturb normal meiotic chromosome segregation. The different types of mutations which lower the possibility of recombination, including the ones in balancer chromosomes, will certainly increase the frequency of nondisjunction. It has been shown that the association of the two X and a Y chromosomes occurs and is maintained at a much higher rate in cases where the crossing over is suppressed. A study published in 2006 has proved that frequency of secondary nondisjunction in *FM7/X/Y* females is much higher as the inter-chromosomal genetic material exchange (recombination) is absent. A twenty-fold higher frequencies of X nondisjunction in *FM7/X/Y* females (70.8%) was reported when compared with the frequencies of X nondisjunction in XXY females (around 3%)[492].

In this project, as the deletion of *CG1847* in not compatible with life, the mutant stock was balanced over the *FM6* balancer chromosomes. Consequently, the rate of nondisjunction was increased.

During gamete formation, the alleles for each gene segregate during meiosis and will later recombine following the Mendelian laws of inheritance.

		CG1847^{exon1_3}/ FM6	
		CG1847^{exon1_3}	***FM6***
X/Y	**Y**	*CG1847^{exon1_3}/Y*	*FM6*
	X	*CG1847^{exon1_3}/X*	*FM6*

Figure 87 Overview of normal meiosis in the *CG1847^{exon1_3}* mutant stock. In bold: The segregation of maternal and paternal alleles during normal meiosis results in four haploid gametes, each containing one set of chromosomes. Normal segregation of a nonrecombinant chromosome pair results in normal disjunction.

Out of the four possible chromosome combinations depicted in Figure 87, the nondisjunction phenotype occurs at highest rate in *FM6 /X* females.

In the primary type of nondisjunction normal *CG1847^{exon1_3}* and *FM6* female chromosomes fail to segregate during meiosis and consequently both will be found in the egg.

The combination of such *CG1847^{exon1_3}/ FM6* and zero eggs with the X and the Y male chromosomes will result in four new types of zygotes, as shown in Figure 88.

		CG1847^{exon1_3} / FM6	
		CG1847^{exon1_3} / FM6	**0**
X/Y	**Y**	*CG1847^{exon1_3} / FM6/ Y*	*0 /Y* *Not viable*
	X	*CG1847^{exon1_3} / FM6/ X*	*0 /X* Male phenotype

Figure 88: Primary nondisjunction in *CG1847^{exon1_3}* mutant stock. A Punnett square for segregation and recombination of *CG1847^{exon1_3}* and *Fm6* chromosomes through primary non-disjunction, and the possible resulting progeny.

The *CG1847^{exon1_3}/FM6* females are the result of the primary nondisjunction. These females remain in the stock as they are not phenotypically different than *CG1847^{exon1_3}/FM6.*

In addition, Bridges noticed that almost all cases of X nondisjunction in XXY females is due to XX-Y segregation. This particular type of nondisjunction was very well described in *Drosophil*a; the result of such non-disjunctional events might result in females who inherit both X chromosomes from their mothers, while no sex chromosome comes from their fathers. Bridges observed that X chromosome nondisjunction is much more frequent in these types of females, an event called ‘‘secondary nondisjunction.’’ The frequency of secondary nondisjunction is significantly increased in females with reduced X chromosomal exchange due to the presence balancer chromosomes[493].

		CG1847^{exon1_3} /FM6	
		CG1847^{exon1_3} / FM6	***Y***
X/Y	**Y**	*CG1847^{exon1_3} / FM6 / Y*	*Y/Y* *Not viable*
	X	*CG1847^{exon1_3} / FM6 / X*	*X/Y* *Male rescued-look like phenotype (but the X chr inherited from male parents)*

Figure 89: Secondary non-disjunction in the female. A Punnett square for segregation and recombination of *CG1847^{exon1_3}* and *Fm6* chromosomes through secondary non-disjunction, and the possible resulting progeny.

In our *in vivo* assay the problem of non-disjunction was demonstrated by developing a PCR-based genotyping system and by using statistical analysis to determine the significance of the results.

Consequently, this system can be used to test human *AIP* missense variants, where pathogenicity cannot be easily determined based on clinical data. This is crucial for genetic counselling and management of the proband's family[102,136]. The benefit of cascade genetic screening for these families would be immediate[481,494], as clinical screening can identify the disease early and at a more manageable stage as it was shown already[481,494] (screening and treatment are already available via NHS).

5.6 CONCLUSIONS

This chapter supports the evolutionary conservation of the *AIP* gene and helps to understand the pathogenesis of *AIP* mutations. As evidence for conservational evolution of the AIP gene we report that deletions of the endogenous *Drosophila* gene resulted in lethality of the flies while the human gene is able to functionally compensate for the deletion of the *Drosophila* orthologue, *CG1847*.

Transgenically expressed AIP proteins with p.R16H, p.A299V, and p.R304Q variants had similar rescue capacities as the wild-type human protein and this allows us to conclude that these genetic changes do not have a significant impact on the AIP function at least in our model. Our data bring further support towards the hypothesis that these are just non-pathogenic SNPs.

In view of the presented data, which conclusively demonstrated a benign role of the AIP p.R16H and p.A299V missense variants, it is not justifiable to offer AIP genetic screening to family members. Contrarily, for p.C238Y carriers, after appropriate counselling, all family members should be genetically and clinically tested, in order to avoid missing cases which developed or might develop pituitary adenomas.

Regarding the pathogenic role of p.R304Q, it was demonstrated to have a benign role during organismal development, but the involvement in tumour development cannot conclusively be discarded. Until more detailed investigation will be available, it would be justifiable to offer AIP genetic screening to all family members of p.R304Q probands.

In conclusion, p.R304Q remains an AIP variant with significant clinical suspicion and due to its very high frequency in pituitary adenoma patients, clarification of its effect is mandatory.

CHAPTER 6: GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

Three projects are presented in this thesis. They hopefully represent a step forward for a better understanding of AIP functions during development and tumorigenesis. The main results and conclusions identified during my studies are summarised in this chapter, specially highlighting the novel findings. Future directions for further studies are also presented here.

Our main goals were:

a. To characterise the effect of *Drosophila AIP* orthologue *CG1847* silencing via RNAi-mediated gene knockdown and imprecise excision of a P-element.

b. To test he possible mechanisms explaining the knockdown/knockout phenotype.

c. To determine gene expression profiles using an established analysis pipeline and to identify key pathways that are significantly altered in the mutant, and that are related to embryonic development or survival.

d. RNA sequencing and confirming its results.

e. To functionally test the homology between *hAIP* and *CG1847* and to test whether wild-type hAIP, a truncated hAIP and four missense mutations can rescue the *CG1847* knockout phenotype.

6.1 *CG1847* is a *Drosophila melanogaster AIP* orthologue and is essential for normal development

In order to analyse the role of AIP during development we generated *Drosophila melanogaster* model of CG1847 deficiency. The *CG1847* loss of function results in lethality during larval stages, showing that in *Drosophila*, similar to human and murine data, this gene is an essential one. In addition, this is supported by the bioinformatics data, protein structure and rescue experiments with human protein. Furthermore, the lack of CG1847 is not compatible with normal development, as the larvae present an obvious delay in development by 72 h AEL.

Nevertheless, the actual mechanism of lethality remains unknown. The answer to this question might be provided by future research involving GC1847 immunostaining during the very early stages of development. This might be possible by designing an antibody, since CG1847 is not a well investigated and no antibody is available at the moment. Designing an antibody for CG1847

will further help to identify the location of the protein in the cells during normal or abnormal development.

6.2 RNA sequencing reveals possible new underlying CG1847 molecular mechanisms

In order to gain further insight into the molecular mechanisms disturbed by loss of CG1847 function a RNA sequencing was performed using the Illumina platform. The top differentially expressed genes were mapped to STRING database to construct the protein-protein interactions network and to reveal the underlying molecular mechanism of CG1847 deficiency. A very interesting cluster of proteins involved in body size, body regulation or cuticle formation was revealed by this study, implying that CG1847 may play an important role in body size and cuticular formation by interacting with different group of proteins such as Osiris, Twedl and cuticular proteins, interactions which were never described before. Meanwhile, heat shock proteins family was significantly downregulated transcripts in the mutant *Drosophila* larvae, and may be important for the underlying mechanisms of the lethality process. Results from this study might provide the groundwork for the understanding the role of AIP in organ development and tumorigenesis.

Further experiments are needed in order to prove the role of the AIP in human or murine development, as the results from this thesis only describe the process in insects. The answer might be provided by a thorough search and validation of other expressed genes. The enormous advantage of our *Drosophila* model is that we can use the mutant *CG1847*-deficient flies to test whether transgenic flies that express different cDNAs of the validated downregulated transcripts (e.g. members of Hsp family) can rescue their lethality. In case of the upregulated transcripts an RNAi screen can be perform to identify components of affected signalling pathways. An *in vivo* validation of the RNAi screening will be possible by generating transgenic flies carrying mutated versions of the identified candidates and use them in our rescue experiments.

The finding that numerous members of the heat shock protein family are significantly downregulated, suggest a possible further list of a new repertoire of interacting partners of AIP. Future studies could focus on determining the exact nature of Hsp-AIP interactions, bringing a possible new light in AIP roles.

6.3 The cytoskeletal disorganisation might be related to CG1847 loss of function, this being the mechanism for the tumour suppressor function of AIP

Cytoskeletal network disorganisation and loss of normal cellular adhesion are signature for invasive tumours. The study of CG1847, the *Drosophila* orthologue for human AIP revealed an exciting novel involvement of this protein in the cell-to-ECM adhesion process. Furthermore, loss of CG1847 resulted in significant actin cytoskeleton deregulation and a possible involvement in regulation of integrins receptors. The interactions with actin indicates a possible strong influence of AIP in cell motility and migration, cellular functions which are essential for tumour growth, invasion and metastasis, hence providing encouraging insights into how the loss of AIP might promote adenoma formation and local invasion. This *in vivo* model supports the use of *Drosophila melanogaster* as a system to elucidate the molecular mechanisms of human tumorigenesis.

The actual mechanisms by which AIP acts as a tumour suppressor gene remains still unknown. Cellular proliferation (pH3 staining, BrdU incorporation) or apoptosis (caspase staining, tunnel) assays are necessary to reveal if in *Drosophila in vivo* model AIP promotes tumorigenesis via increased proliferation or decreased cellular death. To further investigate the interactions between CG1847 and integrins-actin-cytoskeleton, if the design of a specific antibody is successful, the protein interactions could be tested by co-immunoprecipitation with or without cross-linking agents in *Drosophila* larvae.

Does the human AIP interact with the same partners as CG1847 *in vivo*? To confirm that the human AIP does interacts with the same candidate genes as CG1847 *in vivo* (data obtained from the RNA sequencing) one possibility will be to use the human AIP rescued males and to perform Co-IP studies. The results could be further validated in the rat pituitary cell line (GH3) or human HEK293 cells. Other possible studies for data translation might involve investigation of integrin-actin cytoskeleton expression in human pituitary adenomas from patients with AIP mutations and comparison to AIP-mutation negative adenomas. These tissues are available in the supervisor's laboratory as part of her large cohort of FIPA families. RT-qPCR for actin and the most relevant integrin associated proteins might be performed in normal pituitary, sporadic adenomas and AIP-mutation positive adenomas (available as frozen tissues from transsphenoidal operations).

6.4 Human AIP variants have different capacities for compensating for CG1847 loss of function in an *in vivo* model

To functionally test the homology between h*AIP* and *CG1847*, we used the Gal4/UAS system to perform rescue experiments using different UAS::hAIP constructs. We subsequently tested whether hAIP could rescue the lethality of $CG1847^{exon1\_3}$ mutants by expressing UAS::hAIP under the control of a ubiquitous promoter (i.e. *actin*) during fly development. Strikingly, hAIP expression is sufficient to rescue the lethality of *CG1847* mutants, demonstrating that *CG1847* is the functional homologue of *AIP*. As a proof-of-principle, we have shown that, in contrast to wild-type hAIP, a truncated mutant AIP failed to rescue the lethality of the $CG1847^{exon1\_3}$ mutant. Additionally, *CG1847* mutants failed to be rescued by a h*AIP* transgene carrying the p.C238Y variant, a pathogenic missense mutation identified in FIPA patients, known to disrupt a conserved, structurally important amino acid in the 2nd TPR domain. The rescue results for the other 3 missense variants (p.R16H, p.A299V, p.R304Q) support a rather non-pathogenic role than a disease-causing association. These results were in accordance with the literature data, except for p.R304Q missense variant for which the available information are highly contradictory.

Given the fact that human AIP can functionally substitute for CG1847 *in vivo*, this *Drosophila*-based *in vivo* assay might be used to discriminate between pathogenic and non-pathogenic AIP mutations on the basis of their ability to rescue phenotypes associated with loss-of-function of *CG1847*. The immediate benefit to patients and their families can be direct and immediate as to date 19 *AIP* missense variants have been described and the pathogenicity of 15 of these mutations remains questionable. Further rescue experiments could be performed in the future to test other *hAIP* missense variants. Whenever the lethality is rescued, it might be evaluated the development and lifespan of the flies for determining whether they acquire tumours. In addition, it can also be tested whether the rescued flies display changes in the RNA and/or protein level of known AIP interaction partners. For the pathogenic variants, which will not result in the rescue of mutant males, as the hAIP truncated of C238Y variants, their levels of expression could be further evaluated by overexpressing these pathogenic variants in the wt background (the antibody used against hAIP protein does not detect the endogenous CG1847 protein). Previous experiments in our lab showed that used antibody can detect both the truncated and C238Y proteins.

These results demonstrate that *Drosophila* is a useful system in the study of human AIP missense variants pathogenicity.

APPENDICES

Appendix 1 AIP sequence

AIP cDNA sequence

Transcript ID: ENST00000279146

Length 1221 bp, 6 exons

CCCTCAACCAAAATGGCGCTAGCTCGGAAGCTGCCGAGGTGCTAGGAGTTGCCGAAGCAAGTCCGGAAGC
TACCGAGCGAGTCCGGAAGTTGCCGAAAGGGAGCAGCGGGGAAGGAGGATGGCGGATATCATCGCAAGAC
TCCGGGAGGACGGGATCCAAAAACGTGTGATACAGGAAGGCCGAGGAGAGCTCCCGGACTTTCAAGATGG
GACCAAGGCCACGTTCCACTACCGGACGCTGCACAGTGACGACGAGGGCACCGTGCTGGACGACAGCCGG
GCTCGTGGCAAGCCCATGGAGCTCATCATTGGCAAGAAGTTCAAGCTGCCTGTGTGGGAGACCATCGTGT
GCACCATGCGAGAAGGGGAGATTGCCCAGTTCCTCTGTGACATCAAGCATGTGGTCCTGTACCCGCTGGT
GGCCAAGAGTCTCCGCAACATCGCGGTGGGCAAGGACCCCCTGGAGGGCCAGCGGCACTGCTGCGGTGTT
GCACAGATGCGTGAACACAGCTCCCTGGGCCATGCTGACCTGGACGCCCTGCAGCAGAACCCCCAGCCCC
TCATCTTCCACATGGAGATGCTGAAGGTGGAGAGCCCTGGCACGTACCAGCAGGACCCATGGGCCATGAC
AGACGAAGAGAAGGCAAAGGCAGTGCCACTTATCCACCAGGAGGGCAACCGGTTGTACCGCGAGGGGCAT
GTGAAGGAGGCTGCTGCCAAGTACTACGATGCCATTGCCTGCCTCAAGAACCTGCAGATGAAGGAACAGC
CTGGGTCCCCTGAATGGATCCAGCTGGACCAGCAGATCACGCCGCTGCTGCTCAACTACTGCCAGTGCAA
GCTGGTGGTCGAGGAGTACTACGAGGTGCTGGACCACTGCTCTTCCATCCTCAACAAGTACGACGACAAC
GTCAAGGCCTACTTCAAGCGGGGCAAGGCCCACGCGGCCGTGTGGAATGCCCAGGAGGCCCAGGCTGACT
TTGCCAAAGTGCTGGAGCTGGACCCAGCCCTGGCGCCTGTGGTGAGCCGAGAGCTGCAGGCCCTGGAGGC
ACGGATCCGGCAGAAGGACGAAGAGGACAAAGCCCGGTTCCGGGGGATCTTCTCCCATTGACAGGAGCAC
TTGGCCCTGCCTTACCTGCCAAGCCCACTGCTGCAGCTGCCAGCCCCCCTGCCCGTGCTGCGTCATGCTT
CTGTGTATATAAAGGCCTTTATTTATCTCTC

AIP protein sequence

Protein ID: ENSP00000279146

Length: 330 aa

MADIIARLREDGIQKRVIQEGRGELPDFQDGTKATFHYRTLHSDDEGTVLDDSRARGKPMELII
GKKFKLPVWETIVCTMREGEIAQFLCDIKHVVLYPLVAKSLRNIAVGKDPLEGQRHCCGVAQMR
EHSSLGHADLDALQQNPQPLIFHMEMLKVESPGTYQQDPWAMTDEEKAKAVPLIHQEGNRLYRE
GHVKEAAAKYYDAIACLKNLQMKEQPGSPEWIQLDQQITPLLLNYCQCKLVVEEYYEVLDHCSS
ILNKYDDNVKAYFKRGKAHAAVWNAQEAQADFAKVLELDPALAPVVSRELQALEARIRQKDEED
KARFRGIFSH

Appendix 2 CG1847 sequence

CG1847 cDNA sequence

Transcript ID: FBtr0073567

Length: 1758 bp

TAACGTCTGGTATCGAAAGGAAATTGTGTTGATTCCAAATAAATCCGATTCCAGCGGAGATGCAGTCGCG
CAGCAAGTCCGATATGAAGCCCATACGAAAGGAGATCCTCAATCCGGGAAACGCCTACATCGAGCTAACC
CCGGGCACCAGGGTGAAGTTCCACTTTCAAACGCGGAGGGCCGGCGACAGTCGCATCATCGATGATAGCC
GCAAGATGGAGAAGCCCATGGAGCTGGTCCTGGGGAAGAAGTTTAAGCTAGAGGTCTGGGAGCTGATTGT
GCAGCAGATGTCCCTTAACGAAGTGGCCAAGTTCACGGTACATAAGTCGCTCTGCGCTCAATATCCTTTT
ATATCCAAGACCCTGCGGGACATTGGCAAGAAACCGGAGGAGCGACGTCACTGCTGCGGAATGACATTGC
AGAACGAGGGCATTGGGTACACCGACCTGGATGAGCTGCTGCAAAATCCTTCCGATCTGGAGTTCATCAT
TGAACTGTTCTCCATTGAGCTGCCCGAGCAGTACGAAAAAGAGCGCTGGCAGATGTCCGACGACGAAAAG
ATGCTGGCCACCAGTACGCTGCGCGAACGGGGCAACAACTTCTATAAGGCCAGTCGGTTCACAGAGGCGG
AGACCTGCTACCGCGAGGCTGTCGGAATTGTGGAGCAGCTGATGCTAAAGGAGAAGCCGCACGACGAGGA
GTGGCAGGAGCTGGCGGCCATCAAGACACCGCTGTTGTTGAACTACGCGCAATGTCGGTTGATCGCCGGC
GACTTCTACGCTGTGATCGAGCACTGCAACGAGGTGCTCACCCTGGATCCGCGCAATGTCAAGGCACTTT
TTCGTCGGGCCAAGGCCCATGCGGGTGCCTGGAATCCAGCACAGGCACGTCGCGACTTCCTCGACGCCTT
GGCCTTGGACGCCAGCCTCAAGTCGACCGTGTCTAAGGAGCTCAAGTCCATCGAGGATCAGCAGCAGGCA
CGTAACGTCCAGGATCGCATTCACATGCAGAAGCTCTTCTAGAACATAAGTTGCGTCAACGTGCTGCTCA
TGCTGCTTGTCTATTGGAGCAGCTACCTGCAGCGCTAGCAGTACTTGTCACTACCTTCTTTTCTTTCGCT
GGCCGTCAGCTTCGCTCATGGTCTCGTGCTGCTCTGCTCGATGGCGAACCTATTCCTCTGCTGCTGGACC
CTCAAGAAGCTGCTGCGAGCAATGCAGGGGCTCAGCTGGTGATTTTGCGACAACATAACACAATCAACCA
ACTCGGTACCATACCACCTCATTTTGTGAGAGCTGCATTTTGGGGCACTATATGCCCATACTCATCCTCC
GCCTCGATTCACATTCAACGGTGTAGGCTAAAGGGTCTAAATATAATGTAAATGTACCATCCAGATGCTT
GTGTGGAATTGTAATCGTGTGTATATGGAATGAAAAATGTTGTTTGCGTAGCGTTACCAAAAATAGTAAT
CAAATGTTTCACATTTGTTGTCCATAGTCGTATATGTATGTTTGTATTGTATTGTATATCCCTATATGCG
ATATTTACTCGTAGCTAGAATCTACTCTAAATCTAGACAAATTGTGTAAGAAGTAATAAATGTGCTTTTG
ACCGCTGTCTAATTGTTAATTGTAGTTGAAGACTTTATTGTTTGTATACATATATGCGGCAAGCATACGT
AATAATATGCATTTGCATAAGAGAGAATACAGTGTATAAAATAAATATCCAAACCTTTAAATGGACAAAT
AAGGTCGT

CG1847 protein sequence

Protein ID: FBpp0073411

Length: 320

aaMQSRSKSDMKPIRKEILNPGNAYIELTPGTRVKFHFQTRRAGDSRIIDDSRKMEKPMELVLG
KKFKLEVWELIVQQMSLNEVAKFTVHKSLCAQYPFISKTLRDIGKKPEERRHCCGMTLQNEGIG
YTDLDELLQNPSDLEFIIELFSIELPEQYEKERWQMSDDEKMLATSTLRERGNNFYKASRFTEA

ETCYREAVGIVEQLMLKEKPHDEEWQELAAIKTPLLLNYAQCRLIAGDFYAVIEHCNEVLTLDP
RNVKALFRRAKAHAGAWNPAQARRDFLDALALDASLKSTVSKELKSIEDQQQARNVQDRIHMQK
LF

Appendix 3: Fly food recipe

The fly food was prepared in Professor Ralf Stanewsky fly facility as follow:

• Water 1 litre • Agar 10g • Sucrose 15g • Glucose 33g • Yeast 35g • Maize meal 15g	• Wheat germ 10g • Treacle 30g • Soya our 1 table spoon • Nipagin 10 ml • Propionic Acid 5 ml

Appendix 4: Structure of Inverted Repeats (IR) for RNAi constructs

Using UAS-GAL4 system, the RNAi flies produce double stranded RNA *in vivo*, inducing post-transcriptional gene silencing. This system is based on transcription of inverted repeats (IR), which are commonly used as they have high efficiency in making dsRNA. The IR are cloned in vectors backbones under a UAS sequence. Usually the final constructs are injected into 30-100 eggs and the transformants adults are selected to establish IR fly lines. A phenotypical marker (usually white+) is used to select the transformants. Genotyping is usually performed by using traditional genetic methods (PCR).

Although this technique is very efficient and IR sequences are choose to target very specific areas of gene of interest, there are possible some cross-reactions between IR fragments and off-target genes. Consequently further investigation are necessary to proof that the effects are not due to an off-target effect.

SUPPORTING TABLE 1.The structure of IR constructs

Stock ID	IR fragment full Sequences
1847R-1	TCCGATATGAAGCCCATACGAAAGGGAGATCCTCAATCCGGGAAACGCCTACATCGAGCTAACCCCGGGCACCAGGGTGAAGTTCCACTTTCAAACGCG GAGGGCCGGCGACAGTCGCATCATCGATGATAGCCGCAAGATGGAGAAGCCCATGGAGCTGGTCCTGGGGAAGAAGTTTAAGCTAGAGGTCTGGGAGC TGATTGTGCAGCAGATGTCCCTTAACGAAGTGGCCAAGTTCACGGTACATAAGTCGCTCTGCGCTCAATATCCTTTTATATCCANNGAccCTGCGGGACAT TGGCAAGAAACCGGAGGAGCGACGTCACTGCTGCGGAATGACATTGCAGAACGAGGGCATTGGGTACACCGACCTGGATGAGCNGCTGCAAAATCCTTC CGATCTGGAGTTCATCATTGAACTGTTCTCCATTGAGCtGCCCGAGCAGTACGAAAAAGAGCGCTGGCAGATGTCCGACGACGAAAAGATGCTGGCCACC AGTAC
1847R-2	TCCGATATGAAGCCCATACGAAAGGGAGATCCTCAATCCGGGAAACGCCTACATCGAGCTAACCCCGGGCACCAGGGTGAAGTTCCACTTTCAAACGCG GAGGGCCGGCGACAGTCGCATCATCGATGATAGCCGCAAGATGGAGAAGCCCATGGAGCTGGTCCTGGGGAAGAAGTTTAAGCTAGAGGTCTGGGAGC TGATTGTGCAGCAGATGTCCCTTAACGAAGTGGCCAAGTTCACGGTACATAAGTCGCTCTGCGCTCAATATCCTTTTATATCCANNGAccCTGCGGGACAT TGGCAAGAAACCGGAGGAGCGACGTCACTGCTGCGGAATGACATTGCAGAACGAGGGCATTGGGTACACCGACCTGGATGAGCNGCTGCAAAATCCTTC CGATCTGGAGTTCATCATTGAACTGTTCTCCATTGAGCtGCCCGAGCAGTACGAAAAAGAGCGCTGGCAGATGTCCGACGACGAAAAGATGCTGGCCACC AGTAC
43701	AGTCGGTTCACAGAGGCGGAGACCTGCTACCGCGAGGCTGTCGGAATTGTGGAGCAGCTGATGCTAAAGGAGAAGCCGCACGACGAGGAGTGGCAGGA GCTGGCGGCCATCAAGACACCGCTGTTGTTGAACTACGCGCAATGTCGGTTGATCGCCGGCGACTTCTACGCTGTGATCGAGCACTGCAACGAGGTGCTC ACCCTGGATCCGCGCAATGTCAAGGCACTTTTTCGTCGGGCCAAGGCCCATGCGGGTGCCTGGAATCCAGCACAGGCACGTCGCGACTTCCTCGACGCCT TGGCCTT
43702	AGTCGGTTCACAGAGGCGGAGACCTGCTACCGCGAGGCTGTCGGAATTGTGGAGCAGCTGATGCTAAAGGAGAAGCCGCACGACGAGGAGTGGCAGGA GCTGGCGGCCATCAAGACACCGCTGTTGTTGAACTACGCGCAATGTCGGTTGATCGCCGGCGACTTCTACGCTGTGATCGAGCACTGCAACGAGGTGCTC ACCCTGGATCCGCGCAATGTCAAGGCACTTTTTCGTCGGGCCAAGGCCCATGCGGGTGCCTGGAATCCAGCACAGGCACGTCGCGACTTCCTCGACGCCT TGGCCTT

Appendix 5: Primers used in this study

SUPPORTING TABLE 2: Primers used in this study.

Primer name	5'-3' Sequence - Forward	5'-3' Sequence-Reverse	Annealing °C
Dm_CG1847_RT-PCR	ataagtcgctctgcgctcaa	tgaactccagatcggaagga	57.7
Dm_RpL32_RT-PCR	cgatatgctaagctgtcgcaca	cgcttgttcgatccgtaacc	60
Dm_EP	ggacaaataacgtctggtatcg	gaaaaggccaatatcatgagga	62.4
Dm_EP_EL	atccgtatcaactgggatcg	tcagccaaccaaccacaata	60.9
Ppr-Y	ccaagctttgccttaattgc	tcaattaaattatttccaaggctga	58
Dm_NeoRgene	atcaagagacaggatgaggatcgtttcg	gcggcggtggaatcgaaatctcgtgatg	62
FOR CLONING			
CG1847wt	ggatccgcaaacgcaaaggcaactat	gcggccgcccaaagattttcctagctca	60
hAIPwt	acgagcggccgcatggcggatatcatcgcacgcctcc	aatagcggccgctcaatgggagaagatcccccggaac	60
MycAIP_trunc	ggcaggtaccatggaacaaaagttga	atattctagatcaccacaggcgccag	71.1
FOR DIRECT SITE MUTAGENESIS			
ins_7nt_sense	cagccctggcgcctgtggtgaaatcgaattcccgc	gcgggaattcgatttcaccacaggcgccagggctg	60
c.47G>A (p.R16H)	ggacgggatccaaaaacatgtgatacaggaaggcc	ggccttcctgtatcacatgtttttggatcccgtcc	60
c.713G>A (p.C238Y)	tgctgctcaactactgaccagtgcaagctggt	accagcttgcactggtcagtagttgagcagca	60
c.896C>T (p.A299V)	acccagccctggtgcctgtggtgag	ctcaccacaggcaccagggctgggt	60
c.911G>A (p.R304Q)	ctgtggtgagccaagagctgcgggc	ggcccgcagctctcagctcaccacaggc	60
FOR SEQUENCING CONSTRUNCTS			
M13 primers	gttttcccagtcacgac	caggaaacagctatgac	
AIP _Ex4B_F	gacccatgggccatgacagacgaaga		
AIP_Ex4A_R		gcatgtgaaggaggctgctgccaag	

SUPPORTING TABLE 3. Genes included in validation set by the Multiplex-qPCR

Gene symbol	Full name	RefSeq	Function
RpL32	Ribosomal protein L32	NM_170461.	Housekeeping gene
AlphaTub84b	alpha-Tubulin at 84B	NM_057424.4	Housekeeping gene
CG1847	CG1847	NM_132530.4	Drosophila orthologue for human *AIP*
Hsp70Bbb	heat shock protein 70Bbb	NM_176486.2	Mediating response to heat;response to hypoxia
Hsp83	heat shock protein 83	NM_079175.4	ATP binding;ATPase activity, unfolded protein binding
Osi18	Osiris 18	NM_141382.3	Protein of unknown function
Osi19	Osiris 19	NM_001170058.2	Protein of unknown function
TwdlG	TweedleG	NM_001275316.2	Chitin-based cuticle development; body morphogenesis
cpr66Cb	cuticular protein 66Cb	NM_139952.3	Chitin-based cuticle development
cpr97Eb	cuticular protein 97Eb	NM_143273.3	Chitin-based cuticle development; neurogenesis.
Mhc	myosin heavy chain	NM_001259121.2	Epithelium migration; adult somatic muscle development; protein stabilization;
Act57B	actin 57B	NM_079076.4	Cytoskeleton organization; heart development, skeletal muscle fiber.
Pka-Cl	Protein kinase cAMP-dependent, catalytic subunit 1	NM_057629.4	Regulation of embryonic development; regulation of apoptotic process
Gsalpha	G protein alpha s subunit	NM_001299869.1	regulation of cAMP biosynthetic process; tissue development; cell-cell signalling; regulation of cAMP metabolic process
Octβ2R	octopamine beta2 receptor	NM_001170125.3	G-protein coupled receptor signalling pathway; positive regulation of adenylate cyclase activity involved in G-protein coupled receptor signalling pathway.

SUPPORTING TABLE 4. Primers used for Multiplex-qPCR

Gene symbol	Left Sequence w/o Universal Tags	Right Sequence w/o Universal Tags
RpL32	AGGTGACACTATAGAATAATGCTAAGCTGTCGCACAAA	GTACGACTCACTATAGGGAGAACTTCTTGAATCCGGTGG
AlphaTub84b	AGGTGACACTATAGAATAAACCTGAACCGTCTGATTGG	GTACGACTCACTATAGGGACGTAGGTCACCAGAGGGAAG
CG1847	AGGTGACACTATAGAATAAGGTCTGGGAGCTGATTGTG	GTACGACTCACTATAGGGACCGCAGGGTCTTGGATATAA
Hsp70Bbb	AGGTGACACTATAGAATACAAAATCGCAGAGGACATGA	GTACGACTCACTATAGGGAATCTCCTCGGGAGCAAATCT
Hsp83	AGGTGACACTATAGAATAGTCTACATGACCGAGCCCAT	GTACGACTCACTATAGGGAGGACTTCATCAGCTTGCACA
Osi18	AGGTGACACTATAGAATACAGTTCTGCCCTTCCTTCTG	GTACGACTCACTATAGGGATGCACCACCTCGTAGTTGAC
Osi19	AGGTGACACTATAGAATACACCTTTAGCTCCGTTCCTG	GTACGACTCACTATAGGGAACTAGCTGGCTCCAAACTGC
Twd1G	AGGTGACACTATAGAATAGCAACAACGGAATTTCATCC	GTACGACTCACTATAGGGAAATATGGCTGCAGAGTCGCT
cpr66Cb	AGGTGACACTATAGAATAGAGCTGCACGAACACCACTA	GTACGACTCACTATAGGGAGTGGATATGCTTGTCCCTCC
cpr97Eb	AGGTGACACTATAGAATATCAACCTCTACACCGGTTCC	GTACGACTCACTATAGGGAACTCTCTGGCCCAACTCAGA
Mhc	AGGTGACACTATAGAATAGAGGAGTCTCGCACTCTGCT	GTACGACTCACTATAGGGATCTTGGCTTCGTTCAGGAGT
Act57B	AGGTGACACTATAGAATAAGGACCTGTACGCCAACATC	GTACGACTCACTATAGGGACACCGATCCAGACGGAGTAT
Pka-Cl	AGGTGACACTATAGAATACGGCTATGCGGGTATTTTTA	GTACGACTCACTATAGGGATTTGCTTTTCCATTTTCGCT
Gsalpha	AGGTGACACTATAGAATAGATATTCTTCGGTGCCGTGT	GTACGACTCACTATAGGGACTTGAGCACGCAGTTACGAA
Octβ2R	AGGTGACACTATAGAATAACACACGAACTGAATGCCAC	GTACGACTCACTATAGGGAATGTTGTCCAGCCAATCCTC
Kan(r)	AGGTGACACTATAGAATAATCATCAGCATTGCATTCGATTCCTGTTTG	GTACGACTCACTATAGGGAATTCCGACTCGTCCAACATC

Appendix 6: Immunostaining protocols

Imaginal wing discs immunostaining

Buffers and solutions

1X PBS: Phosphate buffered saline (Oxoid - Product Code: 10209252). 1 tablet makes 100 ml of solution. Store at RT.

Fixation buffer4%: 16% Formaldehyde vials (10 ml) and adjust with 1X PBS to a final volume of 40 ml. Store at -20°C.

Permeabilisation buffer (PBT 0.2%): 500 ml PBS+1000 µl Triton X-100. Store at RT.

Blocking buffer (0.5%): 100 ml PBT+500mg BSA (Bovine Serum Albumin, SIGMA-ALDRICH, A2058). Store at 4°C.

Method

1. While in PBS, rip the larvae in half and turn the anterior half inside-out. Get rid of fat tissues and gut remaining. Keep for maximum 15-20min in PBS, in 1500 µl tubes.
2. Remove the PBS with a Pasteur pipette, add 500 µl of fixation buffer and incubate for 25 min at RT on rocker, protected from light (make sure that the samples are floating around).
   Can rinse out the fix with 2 PBS rinses and keep samples 2-3 days at 4˚C
3. Remove the fixation solution with a Pasteur pipette carefully not to absorb the samples. Add 500 µl PBT 0.2%, 1 quick rinse to dilute the fix.
4. Incubate the samples with 500 µl of permeabilisation buffer (PBT 0.2%), twice, for 15 min each on the rocker.
5. Remove the permeabilisation buffer with a Pasteur pipette and add 500 µl of blocking buffer for 1 h at RT.
6. Remove the blocking buffer and add of the appropriate primary antibodies (adjust dilution in each case in 250 µl PAT). Rotate/rock overnight at 4˚C.
7. Wash the samples twice with 500 µl of PAT for 15 min at RT, on the rocker. Remove PAT and do 2 more washes in 500 µl of PAT for 30 min each at RT, on the rocker.
8. Add the appropriate secondary antibodies diluted in 250 µl of PAT and incubate for 2 h at RT or overnight at 4˚C, protected from light.
9. Wash the samples with 400 µl of PBT for 5 min three times, at RT.
10. Incubate in DAPI solution and wash the samples 3 times, for 15 min each, in PBT.

11. Dissect out the wing discs and add them on a glass slide in a drop mounting medium Vectashield (Vector Laboratories, ref H-1000).
12. Add a few more drops of mounting medium, cover with a cover slip, ensuring no bubbles are left, seal the edges with nail polish and store overninght at 4°C, protected from light, until analysis. For longer term, keep at -20°C, protected from light.

Pupal wing immunostaining

Buffers and solutions

1X PBS: Phosphate buffered saline (Oxoid - Product Code: 10209252). 1 tablet makes 100 ml of solution. Store at RT

Fixation buffer4%: 16% Formaldehyde vials (10 ml) and adjust with 1X PBS to a final volume of 40 ml. Store at -20°C

Permeabilisation buffer (PBT 0.3%): For 500 ml use 498.5 ml PBS + 1.5 ml TritonX/Tween. Store at RT

Blocking buffer (0.5%): For 500 ml use 483.5 ml PBS + 1.5 ml Triton X-100 + 15 ml BSA (Bovine Serum Albumin, SIGMA-ALDRICH, A2058). Store at 4°C

Method

1. Collect white pupae on double sticky tape. Keep them at 25°C for 24-28 h until they reach the appropriate stage. Dissect them while attached to the sticky tape. Grip tail and then remove operculum. In order to avoid fat going into the wings punch a hole in the pupae' head while they are still in the pupae shield. Gently remove the hard cuticle. Put them very fast on PBS (no detergent). To protect the pupae each of them was put in one well of a 96wells plate.
2. In each well add 300 µl of 4% formaldehyde in PBT and incubate overnight at 4°C protected from light (pre-fixation).
3. NEXT DAY: Remove the fixing solution with a pipette, and wash the pupae in PBS.
   In this step pupae can be kept in PBS at 4°C for a few days, before moving to next step.
4. Move the pupae into a dissecting dish. Grip the wing from the shoulder and remove the cuticle (starting from the shoulder – where the separation between wing and the cuticle is more obvious). Move each wing into a well of a Terasaki 60 Microwell Plate.
5. Add 10 µl of fixation buffer and incubate for 10 min at RT, protected from light (fixation).
   From this step forward all the washes have to be done under the microscope.

6. Remove the fixation solution with a pipette and wash the cells thrice with 15 µl of PBS
7. Wash in twice in PBT, for 10 min, at RT.
8. Incubate the pupal wings in 15 µl of permeabilisation buffer for 20 min at RT.
9. Remove the permeabilisation buffer and wash twice in PBT for 10min at RT.
10. Add 15 µl of blocking buffer. Incubate 1 h at RT.
    If just phalloidin staining go to step 22.
11. Remove the blocking buffer and add the appropriate primary antibodies (adjust dilution in each case) diluted in 15 µl of blocking buffer. Incubate overnight at 4°C, on a rocker.
    Pre - absorption of the primary antibody was done against third instar larvae to reduce the non - specific binding of the antibodies in the tissue of interest.
    Two mixes were prepared in order to perform double staining in the same time:

MIX 1 primary antibodies	**MIX 2 primary antibodies**
anti talin – Mouse monoclonal	anti βPS2 – Mouse monoclonal
anti-PINCH - Rabbit polyclonal	anti-parvin - Rabbit polyclonal

12. NEXT DAY: Remove the primary antibodies and rinse three times in 15 µl PAT.
13. Wash twice for 20min in 15 µl PAT.
14. Block as above, in 15 µl of blocking buffer for 1 h at RT.
15. Add 15 µl the appropriate secondary antibodies dilutions in PAT and incubate for 1 h at RT protected from light. One mix was prepared for all the samples:

MIX
Fluorescein (FITC) Donkey Anti-Mouse
Cy™5 Pure Donkey Anti-Rabbit
phalloidin

16. Rinse 3 times with 15 µl of PBT and then wash 10 min in PBT.
17. Move the pupal wings with the tip of a syringe needle on a glass slide in drop of 24 µl mounting medium Vectashield (Vector Laboratories, ref H-1000).
18. Cover with a cover slip, ensuring no bubbles are left, seal the edges with nail polish and store overnight at 4°C, protected from light, until analysis. For longer term keep at -20°C, protected from light.

Appendix 7: Antibodies used in this study

SUPPORTING TABLE 5: Antibodies used in this study

Primary antibodies	Dilution	Usage		Cat number	Reference
anti βPS2 – Mouse monoclonal (HB CF.6G11)	1: 15	Immunostaining	Gift from N.Brown – Gurdon Institute		Brower et al., 1984
anti talin – Mouse monoclonal	1:50	Immunostaining	Gift from N.Brown – Gurdon Institute		Brown et al., 2002
anti-parvin - Rabbit polyclonal	1:500	Immunostaining	Gift from C.G. Zervas – Academy of Athens		Vakaloglou et al.,2012
anti-PINCH - Rabbit polyclonal	1:500	Immunostaining	Gift from M. Beckerle– University of Utah		Clark et al., 2003
anti-AIP/ARA9 Mouse Monoclonal	1:20	Immunostaining	Novus Biologicals	NB100-127	Kasuki et al 2011
anti-AIP/ARA9 Mouse Monoclonal	1:1000	Western Blot	Novus Biologicals	NB100-127	Kasuki et al 2011
anti Tubulin, beta Mouse monoclonal	1:15000	Western Blot	Developmental Studies Hybridoma Bank	E7	Chu et al 1989
Secondary antibodies	**Dilution**	**Usage**		**Cat number**	**Reference**
Alexa Fluor® 488 Goat anti-Mouse	1:250	Immunostaining	Invitrogen	A11029	
Alexa Fluor® 488 Goat anti-Rabbit	1:250	Immunostaining	Invitrogen	A-11008	
Alexa Fluor 488® phalloidin	1:1000	Immunostaining	Invitrogen	A12379	
Fluorescein (FITC) Donkey anti-Mouse IgG H+L)	1:250	Immunostaining	Jackson ImmunoResearch Laboratories	715-096-151	
Fluorescein (FITC) anti-Rabbit IgG (H+L)	1:250	Immunostaining	Jackson ImmunoResearch Laboratories	711-096-152	
Cy™5 Pure Donkey anti-Rabbit IgG (H+L)	1:200	Immunostaining	Jackson ImmunoResearch Laboratories	711-175-152	
Alexa Fluor® 647 Donkey anti-Mouse IgG (H+L),	1:200	Immunostaining	Invitrogen	A-31571	
IRDye® 680 RD Goat anti-Mouse IgM	1:1000	Western Blot	LI-COR Biotechnology	926-68180	
Phalloidin-TRITC	1:250	Immunostaining	Sigma-Aldrich	P1951-1MG	

Appendix 8: Preparing competent cells

In order to prepare competent cells for plasmid DNA transformation we used the chemical method (Nicholas Renzette, Current Protocols in Molecular Biology, 2011) that has the advantages of being simple to complete, requires no special equipment and gives good transformation efficiencies. Disadvantages are that the efficiency is somewhat lower (vs. electroporation).

Materials:

- Single colony of JM109 or BL21 cells to be transformed
- LB medium
- LB amp plates (2 plates without any ampicillin + 4 with ampicillin)
- M $CaCl_2$, ice cold – autoclaved
- M $CaCl_2$+15% glycerol – autoclaved
- 42°C water bath for transformation

Recipe:

- For 1L of 0.1M $CaCl_2$ solution, add 11.10g of $CaCl_2$
- For the $CaCl_2$ +15% glycerol solution, make up 500 ml solution. 5.55 $CaCl_2$ were added to 300 ml dH_2O. Then 75 ml glycerol and the rest of the dH_2O to a final volume of 500 ml

Procedure:

Two ml LB medium were inoculated with one single colony from an LB agar plate (without antibiotics) of JM109 or BL21 bacteria and incubated overnight at 37°C and 225 RPM. 1-ml of the starter culture was used inoculate to a flask with 100 ml LB medium (in a 500 ml flask) and incubated with shaking at 37°C to OD_{600} ~ 0.25-0.3 (usually it took about 1.5-2 hours). The culture was then chilled on ice for 15 min. The 0.1M $CaCl_2$ solution and 0.1M $CaCl_2$ plus 15% glycerol were also placed on ice because all steps after harvesting the cell should be done on ice (or at 4°C). The cells were harvested by centrifuging for 10 min at 4,000 rpm at 4°C. The medium was removed and the cells were resuspended in 40 ml of ice-cold 0.1M $CaCl_2$ using the stripette. The cells were kept on ice for 30 min. After the centrifugation, the transformation medium was removed and the cell pellet was resuspended in 6 ml 0.1 M $CaCl_2$ solution plus 15% glycerol. For storage 250 µl or 500 µl aliquots of competent cells were made in already labelled and chilled cryotubes. The aliquots were quickly frozen and kept at -80°C.

In order to test the competent cells, the following tests were performed:

1) Viability of cells: 0.5 µl of untransformed cells were diluted in 30 µl water and were plated on LB plate without ampicillin.

2) Antibiotic resistance of cells: 50 µl of untransformed culture were plated out onto an LB plate with ampicillin.

3) Transformation efficiency of cells using vector of known concentration. In order to determine the transformation efficiency of the cells, 1 µg of vector DNA was used to transform the cells. During transformation, we used 500 LB media during the recovery period. From this, we plate out 5 µl, 10 µl, and 20 µl of the culture onto LB plates containing ampicillin. On the plates where the number of colonies could be determined, the transformation efficiency was calculated for the batch of competent cells using the formula:

$$\text{transformation efficiency} = \frac{\text{total \#of colonies on LB:AMP plate}}{\text{amount of DNA plated(in } \mu\text{g/mL)}}$$

Appendix 9: TruSeq Stranded mRNA Sample Preparation – Low Sample (LS) Protocol

Introduction – Overview of the protocol (ILLUMINA PROPRIETARY, Catalog # RS-122-9004DOC, Part # 15031047 Rev. E, October 2013).

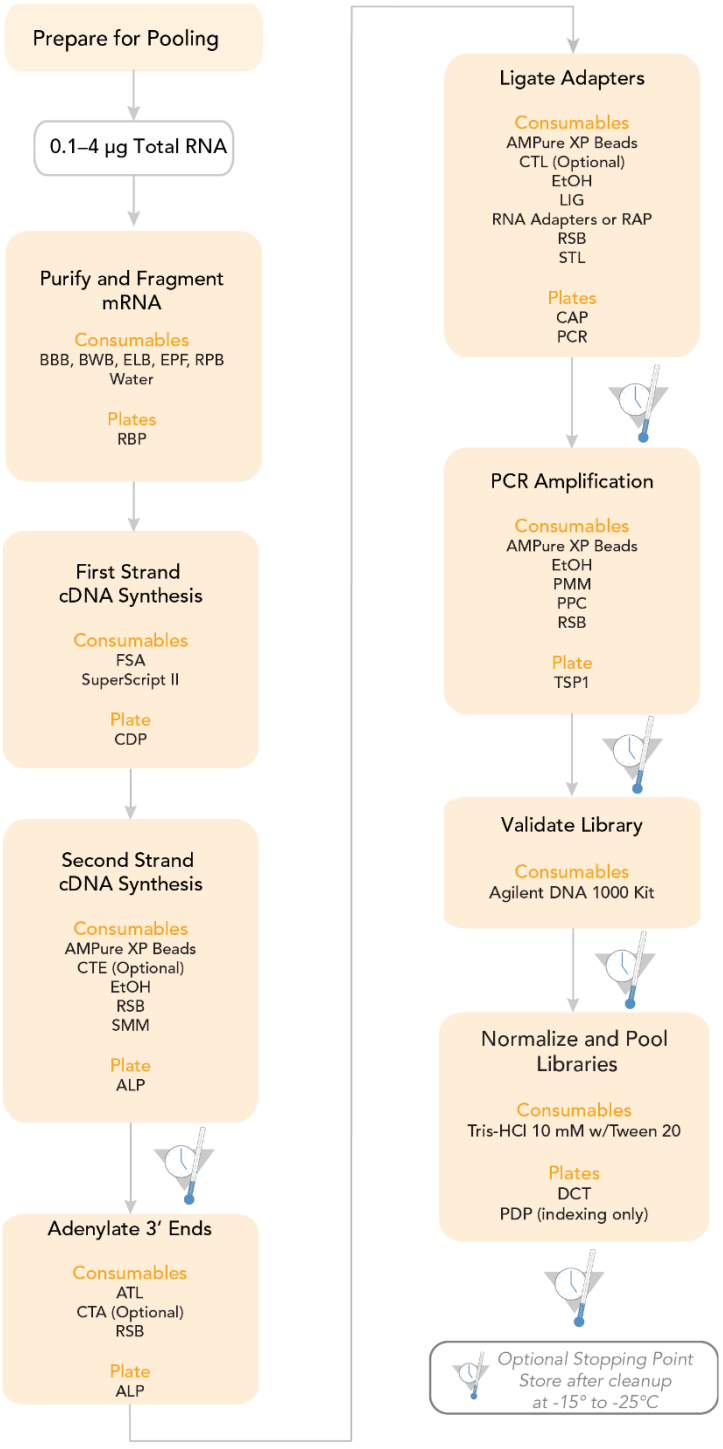

Figure 90: TruSeq Stranded mRNA Sample Preparation LS Workflow

Purify and Fragment mRNA

This process purifies the polyA containing mRNA molecules using poly-T oligo attached magnetic beads using two rounds of purification. During the second elution of the polyA RNA, the RNA is also fragmented and primed for cDNA synthesis. It is important to follow this procedure exactly to be sure of reproducibility.
NOTES: Allow the beads to fully pellet against the magnetic stand 5 minutes. Remove the supernatant from the beads immediately while the beads are still pelleted against the magnetic stand.
Pre-program the thermal cycler with the following programs: Choose the pre-heat lid option and set to 100°C

- 65°C for 5 minutes, 4°C hold—save as mRNA Denaturation
- 80°C for 2 minutes, 25°C hold—save as mRNA Elution 1
- 94°C for 8 minutes, 4°C hold—save as Elution 2 - Frag - Prime

Make RBP

1 Dilute the total RNA with nuclease-free ultrapure water to a final volume of 50 µl in the new 96-well 0.3 ml PCR plate.
2 Vortex the room temperature RNA Purification Beads tube vigorously to resuspend the oligo-dT beads.
3 Add 50 µl of RNA Purification Beads to each well of the RBP plate to bind the polyA RNA to the oligo dT magnetic beads. Gently pipette the entire volume up and down 6 times to mix thoroughly.
4 Seal the RBP plate with a Microseal 'B' adhesive seal.

Incubate 1 RBP

1 Place the sealed RBP plate on the pre-programmed thermal cycler. Close the lid and select mRNA Denaturation (65°C for 5 minutes, 4°C hold) to denature the RNA and facilitate binding of the polyA RNA to the beads.
2 Remove the RBP plate from the thermal cycler when it reaches 4°C.
3 Place the RBP plate on the bench and incubate at room temperature for 5 minutes to allow the RNA to bind to the beads.

Wash RBP

1 Remove the adhesive seal from the RBP plate.
2 Place the RBP plate on the magnetic stand at room temperature for 5 minutes to separate the polyA RNA bound beads from the solution.
3 Remove and discard all of the supernatant from each well of the RBP plate.
4 Remove the RBP plate from the magnetic stand.
5 Wash the beads by adding 200 µl of Bead Washing Buffer in each well of the RBP plate to remove unbound RNA. Gently pipette the entire volume up and down 6 times to mix thoroughly.
6 Place the RBP plate on the magnetic stand at room temperature for 5 minutes.
7 Centrifuge the thawed Elution Buffer to 600 x g for 5 seconds.
8 Remove and discard all of the supernatant from each well of the RBP plate. The supernatant contains most of the ribosomal and other non-messenger RNA.
9 Remove the RBP plate from the magnetic stand.
10 Add 50 µl of Elution Buffer in each well of the RBP plate. Gently pipette the entire volume up and down 6 times to mix thoroughly.
11 Seal the RBP plate with a Microseal 'B' adhesive seal.

Incubate 2 RBP

1 Place the sealed RBP plate on the pre-programmed thermal cycler. Close the lid and select mRNA Elution 1 (80°C for 2 minutes, 25°C hold) to elute the mRNA from the beads.

2 Remove the RBP plate from the thermal cycler when it reaches 25°C.
3 Place the RBP plate on the bench at room temperature. Remove the adhesive seal from the RBP plate.

Make RFP

1 Add 50 µl of Bead Binding Buffer to each well of the RBP plate. This allows mRNA to specifically rebind the beads, while reducing the amount of rRNA that non-specifically binds. Gently pipette the entire volume up and down 6 times to mix thoroughly.
2 Incubate the RBP plate at room temperature for 5 minutes.
3 Place the RBP plate on the magnetic stand at RT for 5 minutes. Discard all of the supernatant.
4 Remove the RBP plate from the magnetic stand.
5 Wash the beads by adding 200 µl of Bead Washing Buffer in each well of the RBP plate. Gently pipette the entire volume up and down 6 times to mix thoroughly.
6 Place the RBP plate on the magnetic stand at room temperature for 5 minutes.
7 Remove and discard all of the supernatant from each well of the RBP plate. The supernatant contains residual rRNA and other contaminants that were released in the first elution and did not rebind the beads.
8 Remove the RBP plate from the magnetic stand.
9 Add 19.5 µl of Fragment, Prime, Finish Mix to each well of the RBP plate. Gently pipette the entire volume up and down 6 times to mix thoroughly. Seal the RBP plate with a Microseal 'B' adhesive seal.

Incubate RFP

1 Place the sealed RBP plate on the pre-programmed thermal cycler. Close the lid and select Elution 2 - Frag - Prime (94°C for 8 minutes, 4°C hold) to elute, fragment, and prime the RNA.
2 Remove the RBP plate from the thermal cycler when it reaches 4°C and centrifuge briefly.
3 Proceed immediately to Synthesize First Strand cDNA on page 21.

Synthesize First Strand cDNA

This process reverse transcribes the cleaved RNA fragments that were primed with random hexamers into first strand cDNA using reverse transcriptase and random primers. The addition of Actinomycin D to the First Stand Synthesis Act D mix (FSA) prevents spurious DNA-dependent synthesis, while allowing RNA-dependent synthesis, improving strand specificity.
Pre-program the thermal cycler with the following program and save as Synthesize 1st Strand: pre-heat lid option and set to 100°C, 25°C for 10 minutes, 42°C for 15 minutes, 70°C for 15 minutes, Hold at 4°C

Make CDP

1 Place the RBP plate on the magnetic stand at RT for 5 minutes.
2 Transfer 17 µl supernatant from each well of the RBP plate to the corresponding well of the new 0.3 ml PCR plate labelled with the CDP barcode.

3 Centrifuge the thawed First Strand Synthesis Act D Mix tube to 600g for 5 seconds.
4 Add 50 µl SuperScript II to the First Strand Synthesis Act D Mix tube.
5 Add 8 µl of First Strand Synthesis Act D Mix and SuperScript II mix to each well of the CDP plate. Gently pipette the entire volume up and down 6 times to mix thoroughly.
6 Seal the CDP plate with a Microseal 'B' adhesive seal and centrifuge briefly.

Incubate 1 CDP

1 Place the sealed CDP plate on the pre-programmed thermal cycler. Close the lid, and then select and run the Synthesize 1st Strand program. a Choose the pre-heat lid option and set to 100°C, 25°C for 10 minutes, 42°C for 15 minutes, 70°C for 15 minutes, Hold at 4°C
2 When the thermal cycler reaches 4°C, remove the CDP plate from the thermal cycler and proceed immediately to Synthesize Second Strand cDNA.

Synthesize Second Strand cDNA

This process removes the RNA template and synthesizes a replacement strand, incorporating dUTP in place of dTTP to generate dscDNA. The incorporation of dUTP quenches the second strand during amplification, because the polymerase does not incorporate past this nucleotide. AMPure XP beads are used to separate the dscDNA from the second strand reaction mix.

Add SMM

1 Remove the adhesive seal from the CDP plate. 2 Do one of the following:
2 Add 5 µl of Resuspension Buffer to each well of the CDP plate.
3 Centrifuge the thawed Second Strand Marking Master Mix to 600g for 5 seconds.
4 Add 20 µl of thawed Second Strand Marking Master Mix to each well of the CDP plate. Gently pipette the entire volume up and down 6 times to mix thoroughly. Seal the CDP plate.

Incubate 2 CDP

1 Place the sealed CDP plate on the pre-heated thermal cycler. Close the lid and incubate at 16°C for 1 hour.
2 Remove the CDP plate from the thermal cycler, remove the adhesive seal and place it on the bench.
3 Let the CDP plate stand to bring it to room temperature.

Purify CDP

1 Vortex the AMPure XP beads until they are well dispersed.
2 Add 90 µl of well-mixed AMPure XP beads to each well of the CDP plate containing 50 µl of ds cDNA. Gently pipette the entire volume up and down 10 times to mix thoroughly.
3 Incubate the CDP plate at room temperature for 15 minutes.
4 Place the CDP plate on the magnetic stand at room temperature, for 5 minutes to make sure that all of the beads are bound to the side of the wells.
5 Remove and discard 135 µl supernatant from each well of the CDP plate. Leave the CDP plate on the magnetic stand while performing the following 80% EtOH wash steps (6–8).
6 With the CDP plate on the magnetic stand, add 200 µl freshly prepared 80%EtOH to each well without disturbing the beads.
7 Incubate the CDP plate at room temperature for 30 seconds, and then remove and discard all of the supernatant from each well.
8 Repeat steps 6 and 7 one time for a total of two 80% EtOH washes.

9 Let the CDP plate stand at room temperature for 15 minutes to dry, and then remove the plate from the magnetic stand.
10 Centrifuge the thawed, room temperature Resuspension Buffer to 600g for 5 seconds.
11 Add 17.5 µl Resuspension Buffer to each well of the CDP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly.
12 Incubate the CDP plate at room temperature for 2 minutes.
13 Place the CDP plate on the magnetic stand at room temperature for 5 minutes.
14 Transfer 15 µl supernatant (ds cDNA) from the CDP plate to the new 96-well 0.3 ml PCR plate.

Adenylate 3' Ends

A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. A corresponding single T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera (concatenated template) formation.

Add ATL

1 Add 2.5 µl of Resuspension Buffer to each well of the ALP plate.
2 Add 12.5 µl of thawed A-Tailing Mix to each well of the ALP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly. Seal the ALP plate with a Microseal 'B' adhesive seal.

Incubate 1 ALP

1 Place the sealed ALP plate on the pre-programmed thermal cycler. Close the lid, then select and run the ATAIL70 program: pre-heat lid option and set to 100°C, 37°C for 30 minutes, 70°C for 5 minutes, Hold at 4°C
2 When the thermal cycler temperature is 4°C, remove the ALP plate from the thermal cycler, then proceed immediately to Ligate Adapters.

Ligate Adapters

Add LIG

1 Add 2.5 µl of Resuspension Buffer to each well of the ALP plate.
2 Add 2.5 µl of Ligation Mix to each well of the ALP plate.
3 Add 2.5 µl of the thawed RNA Adapter Index to each well of the ALP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly.
4 Use the bottom of a clean eight-tube strip, with caps attached, to pierce holes in the seals of the wells that will be used for ligation. Repeat with a new, clean eight-tube strip, with caps attached, for each row.
5 Using an eight-tip multichannel pipette, transfer 2.5 µl of the thawed RNA Adapter to each well of the ALP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly.
6 Seal the ALP plate with a Microseal 'B' adhesive seal. Centrifuge the ALP plate to 280g for 1 minute.

Incubate 2 ALP

1 Place the sealed ALP plate on the pre-heated thermal cycler. Incubate at 30°C for 10 minutes.
2 Remove the ALP plate from the thermal cycler.

Add STL

1 Remove the adhesive seal from the ALP plate.
2 Add 5 µl of Stop Ligation Buffer to each well of the ALP plate to inactivate the ligation. Gently pipette the entire volume up and down 10 times to mix thoroughly.

Clean Up ALP

1 Vortex the AMPure XP Beads for at least 1 minute or until they are well dispersed.
2 Add 42 µl of mixed AMPure XP Beads to each well of the ALP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly and incubate the ALP plate at room temperature for 15 minutes.
3 Place the ALP plate on the magnetic stand at room temperature for 5 minutes or until the liquid is clear.
4 Remove and discard 79.5 µl supernatant from each well of the ALP plate. Take care not to disturb the beads. Leave the ALP plate on the magnetic stand while performing the following 80% EtOH wash steps.
5 With the ALP plate on the magnetic stand, add 200 µl freshly prepared 80% EtOH to each well without disturbing the beads. Incubate the ALP plate at RT for 30 seconds, and then remove and discard all of the supernatant from each well. Take care not to disturb the beads.
6 Repeat step5 one time for a total of two 80% EtOH washes.
7 With the ALP plate on the magnetic stand, let the samples air-dry at room temperature for 15 minutes.
8 Remove the ALP plate from the magnetic stand and add 52.5 µl Resuspension Buffer to each well of the ALP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly or until the beads are fully resuspended.
9 Incubate the ALP plate at room temperature for 2 minutes.
10 Place the ALP plate on the magnetic stand at RT for 5 minutes or until the liquid is clear.
11 Transfer 50 µl supernatant from each well of the ALP plate to the corresponding well of the new 0.3 ml PCR plate labelled with the CAP barcode. Take care not to disturb the beads.
12 Vortex the AMPure XP Beads until they are well dispersed.
13 Add 50 µl of mixed AMPure XP Beads to each well of the CAP plate for a second cleanup. Gently pipette the entire volume up and down 10 times to mix thoroughly.
14 Incubate the CAP plate at RT for 15 minutes and place the CAP plate on the magnetic stand at room temperature for 5 minutes or until the liquid is clear.
15 Remove and discard 95 µl supernatant from each well of the CAP plate. With the CAP plate on the magnetic stand, add 200 µl freshly prepared 80% EtOH to each well. Take care not to disturb the beads.
16 Incubate the CAP plate at RT for 30 seconds, and then discard all of the supernatant from each well.
17 Repeat the 80% EtOH washes. With the CAP plate on the magnetic stand, let the samples air-dry at room temperature for 15 minutes, and then remove the plate from the magnetic stand.
18 Add 22.5 µl Resuspension Buffer to each well of the CAP plate. Gently pipette the entire volume up and down 10 times to mix thoroughly or until the beads are fully resuspended.
19 Incubate the CAP plate at room temperature for 2 minutes, then place the CAP plate on the magnetic stand at room temperature for 5 minutes or until the liquid is clear.
20 Transfer 20 µl supernatant from each well of the CAP plate to the corresponding well of the new 0.3 ml PCR plate labelled with the PCR barcode. Take care not to disturb the beads.

Enrich DNA Fragments

This process uses PCR to selectively enrich those DNA fragments that have adapter molecules on both ends and to amplify the amount of DNA in the library. The PCR is performed with a PCR

Primer Cocktail that anneals to the ends of the adapters. Minimize the number of PCR cycles to avoid skewing the representation of the library. PCR enriches for fragments that have adapters ligated on both ends. Fragments with only one or no adapters on their ends are by-products of inefficiencies in the ligation reaction. Neither species can be used to make clusters. Fragments without any adapters cannot hybridize to surface-bound primers in the flow cell. Fragments with an adapter on only one end can hybridize to surface bound primers, but cannot form clusters.

Make PCR

1 Add 5 µl of thawed PCR Primer Cocktail to each well of the PCR plate.
2 Add 25 µl of thawed PCR Master Mix to each well of the PCR plate. Gently pipette the entire volume up and down 10 times to mix thoroughly.
3 Seal the PCR plate with a Microseal 'B' adhesive seal.

Amp PCR

1 Place the sealed PCR plate on the pre-programmed thermal cycler. Close the lid, then select and run PCR to amplify the plate. Choose the pre-heat lid option and set to 100°C, 98°C for 30 seconds, 15 cycles of: 98°C for 10 seconds/ 60°C for 30 seconds/ 72°C for 30 seconds. Final step 72°C for 5 min. Hold at 4°C

Clean Up PCR

1 Remove the adhesive seal from the PCR plate.
2 Vortex the AMPure XP Beads until they are well dispersed.
3 Add 50 µl of the mixed AMPure XP Beads to each well of the PCR plate containing 50 µl of the PCR amplified library. Gently pipette the entire volume up and down 10times to mix thoroughly.
4 Incubate the PCR plate at room temperature for 15 minutes.
5 Place the PCR plate on the magnetic stand at room temperature for 5 minutes or until the liquid is clear.
6 Remove and discard 95 µl supernatant from each well of the PCR plate.
7 With the PCR plate on the magnetic stand, add 200 µl freshly prepared 80% EtOH to each well without disturbing the beads.
8 Incubate the PCR plate at RT for 30 seconds, then discard all of the supernatant. Repeat 80% EtOH wash.
9 With the PCR plate on the magnetic stand, let the samples air-dry at room temperature for 15 minutes, and then remove the plate from the magnetic stand.
10 Add 32.5 µl Resuspension Buffer to each well of the PCR plate. Gently pipette the entire volume up and down 10 times to mix thoroughly.
11 Incubate the PCR plate at room temperature for 2 minutes.
12 Place the PCR plate on the magnetic stand at RT for 5 minutes or until the liquid is clear.
13 Transfer 30 µl supernatant from each well of the PCR plate to the corresponding well of the new 0.3 ml PCR plate labelled with the TSP1 barcode.

Validate Library

Illumina recommends performing the following procedures for quality control analysis on your sample library and quantification of the DNA library templates.

Quantify Libraries

To achieve the highest quality data on Illumina sequencing platforms, it is important to create optimum cluster densities across every lane of the flow cell. Optimizing cluster densities requires accurate quantitation of DNA library templates. Quantify your libraries using qPCR according to the Illumina Sequencing Library qPCR Quantification Guide (part # 11322363).

Quality Control

1 Load 1 µl of the resuspended construct on an Agilent Technologies 2100 Bioanalyzer using a DNA-specific chip such as the Agilent DNA 1000.
2 Check the size and purity of the sample. The final product should be a band at approximately 260 bp.

Normalize and Pool Libraries

This process describes how to prepare DNA templates for cluster generation. Indexed DNA libraries are normalized to 10 nM in the DCT plate and then pooled in equal volumes in the PDP plate. DNA libraries not intended for pooling are normalized to 10 nM in the DCT plate.

Make DCT

1 Transfer 10 µl of sample library from each well of the TSP1 plate to the corresponding well of the new MIDI plate labelled with the DCT barcode.
2 Normalize the concentration of sample library in each well of the DCT plate to 10 nM using a mix of Tris-HCl 10 mM, pH 8.5 with 0.1% Tween 20.
3 Gently pipette the entire normalized sample library volume up and down 10 times to mix thoroughly.

Make PDP (for pooling only)

1 Determine the number of samples to be combined together for each pool.
2 Transfer 10 µl of each normalized sample library to be pooled from the DCT plate to one well of the new 0.3 ml PCR plate labelled with the PDP barcode.
The total volume in each well of the PDP plate is 10 X the number of combined sample libraries and 20–240 µl (2–24 libraries). For example, in our case, the volume for 8 samples was 80 µl.
3 Gently pipette the entire volume up and down 10 times to mix thoroughly.
4 Proceed to cluster generation.

Appendix 10: RNA-seq – Table

SUPPORTING TABLE 5: Differentially expressed genes (the most downregulated transcripts are on top of the table)

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
1	XLOC_000064	Lsp1beta	2L:898643-901316	Ctr	Mut	21.732	0.20409	-6.734	5.00E-05	0.00245	FBgn0002563
2	XLOC_009430	Hsp70Bbb	3R:8328231-8330822	Ctr	Mut	6.53143	0.121707	-5.746	5.00E-05	0.00245	FBgn0051354
3	XLOC_011176	CR32865	3R:8295701-8304065	Ctr	Mut	172.408	7.62157	-4.5	5.00E-05	0.00245	FBgn0052865
4	XLOC_011881	Hsp68	3R:19880139-19883029	Ctr	Mut	188.604	12.7823	-3.883	5.00E-05	0.00245	FBgn0001230
5	XLOC_003924	CG10073	2R:15268047-15272470	Ctr	Mut	82.6779	7.96064	-3.377	5.00E-05	0.00245	FBgn0034440
6	XLOC_005673	CG3264	2R:18096411-18098247	Ctr	Mut	107.861	12.5462	-3.104	5.00E-05	0.00245	FBgn0034712
7	XLOC_014779	CG34330	X:18962305-18962925	Ctr	Mut	155.109	18.4714	-3.07	5.00E-05	0.00245	FBgn0085359
8	XLOC_005304	IM23	2R:14270208-14270737	Ctr	Mut	15.8438	1.92163	-3.044	5.00E-05	0.00245	FBgn0034328
9	XLOC_003925	CG10081	2R:15273427-15276823	Ctr	Mut	50.9421	6.51389	-2.967	5.00E-05	0.00245	FBgn0034441
10	XLOC_009432	Hsp70Bc	3R:8334797-8337183	Ctr	Mut	15.074	1.95085	-2.95	5.00E-05	0.00245	FBgn0013279
11	XLOC_009431	Hsp70Bb	3R:8331514-8334105	Ctr	Mut	52.1496	6.92209	-2.913	5.00E-05	0.00245	FBgn0013278
12	XLOC_011175	Hsp70Ba	3R:8291025-8293500	Ctr	Mut	9.58651	1.40245	-2.773	5.00E-05	0.00245	FBgn0013277
13	XLOC_001566	CG15353	2L:2006762-2007193	Ctr	Mut	300.451	44.8723	-2.743	5.00E-05	0.00245	FBgn0040718
14	XLOC_001134	ninaD	2L:18081629-18083608	Ctr	Mut	2.57373	0.385647	-2.739	5.00E-05	0.00245	FBgn0002939
15	XLOC_008002	Hsp26	3L:9369517-9370475	Ctr	Mut	513.165	77.0631	-2.735	5.00E-05	0.00245	FBgn0001225
16	XLOC_014469	CG11071	X:13744880-13884528	Ctr	Mut	0.763823	0.118598	-2.687	5.00E-05	0.00245	FBgn0263115
17	XLOC_014469	mamo	X:13744880-13884528	Ctr	Mut	0.763823	0.118598	-2.687	5.00E-05	0.00245	FBgn0030532
18	**XLOC_013279**	**CG1847**	**X:11763220-11765201**	**Ctr**	**Mut**	**7.23844**	**1.18461**	**-2.611**	**5.00E-05**	**0.00245**	**FBgn0030345**
19	XLOC_001238	CG16772	2L:19962678-19963844	Ctr	Mut	3.05377	0.511087	-2.579	5.00E-05	0.00245	FBgn0032835
20	XLOC_006566	Hsp27	3L:9377162-9378382	Ctr	Mut	193.978	35.6058	-2.446	5.00E-05	0.00245	FBgn0001226
21	XLOC_007314	CG32444	3L:21630046-21632160	Ctr	Mut	60.1528	11.0785	-2.441	5.00E-05	0.00245	FBgn0043783
22	XLOC_003070	Lcp3	2R:4322814-4323600	Ctr	Mut	54.0499	10.0496	-2.427	5.00E-05	0.00245	FBgn0002534
23	XLOC_006292	CG11350	3L:4482808-4484370	Ctr	Mut	1220.35	226.958	-2.427	5.00E-05	0.00245	FBgn0035552
24	XLOC_013810	CG13360	X:678823-684312	Ctr	Mut	124.579	23.3674	-2.414	5.00E-05	0.00245	FBgn0025620
25	XLOC_002297	CG9928	2L:13142493-13142902	Ctr	Mut	11.9928	2.46779	-2.281	5.00E-05	0.00245	FBgn0032472
26	XLOC_006565	Hsp23	3L:9374981-9375865	Ctr	Mut	385.882	80.4126	-2.263	5.00E-05	0.00245	FBgn0001224

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
27	XLOC_000883	CG5867	2L:13236492-13239296	Ctr	Mut	5.45954	1.16017	-2.234	5.00E-05	0.00245	FBgn0027586
28	XLOC_008415	CG42718	3L:16313929-16314372	Ctr	Mut	112.352	25.0911	-2.163	5.00E-05	0.00245	FBgn0261635
29	XLOC_007885	CG32376	3L:7534681-7535557	Ctr	Mut	1.78108	0.401468	-2.149	0.0006	0.01944	FBgn0052376
30	XLOC_005555	CG43710	2R:16664231-16665254	Ctr	Mut	3.18182	0.737946	-2.108	0.0001	0.00448	FBgn0263849
31	XLOC_000339	Cyp28d2	2L:5207266-5209345	Ctr	Mut	58.9185	13.8134	-2.093	5.00E-05	0.00245	FBgn0031688
32	XLOC_006562	Hsp22	3L:9365821-9368064	Ctr	Mut	186.688	44.4412	-2.071	0.0001	0.00448	FBgn0001223
33	XLOC_006562	Hsp67Bb	3L:9365821-9368064	Ctr	Mut	186.688	44.4412	-2.071	0.0001	0.00448	FBgn0001228
34	XLOC_002448	Tepl	2L:15888638-15893811	Ctr	Mut	0.745169	0.178166	-2.064	0.0011	0.03093	FBgn0041183
35	XLOC_004016	IM14	2R:16757896-16758183	Ctr	Mut	29.3402	7.11787	-2.043	0.0007	0.02195	FBgn0067905
36	XLOC_001953	CG7214	2L:7743676-7744841	Ctr	Mut	21.0819	5.15237	-2.033	5.00E-05	0.00245	FBgn0031940
37	XLOC_003728	snoRNA:U3:54Ab	2R:13033134-13034026	Ctr	Mut	5.7294	1.41325	-2.019	0.00125	0.03415	FBgn0065047
38	XLOC_005615	CG30288	2R:17406747-17409482	Ctr	Mut	3.01157	0.753787	-1.998	0.00025	0.00942	FBgn0050288
39	XLOC_005615	CG30289	2R:17406747-17409482	Ctr	Mut	3.01157	0.753787	-1.998	0.00025	0.00942	FBgn0050289
40	XLOC_011090	CR43283	3R:7073349-7081916	Ctr	Mut	43.613	10.9697	-1.991	5.00E-05	0.00245	FBgn0262972
41	XLOC_008403	CG34248	3L:16263468-16264002	Ctr	Mut	35.2717	8.87464	-1.991	5.00E-05	0.00245	FBgn0085277
42	XLOC_008782	CG14565	3L:21736347-21737369	Ctr	Mut	122.528	30.9103	-1.987	5.00E-05	0.00245	FBgn0037129
43	XLOC_013276	CR43908	X:11731575-11732059	Ctr	Mut	494.77	126.457	-1.968	5.00E-05	0.00245	FBgn0264509
44	XLOC_001438	CG17018	2L:22311930-22368796	Ctr	Mut	0.589344	0.152225	-1.953	0.0001	0.00448	FBgn0039972
45	XLOC_000517	CG7224	2L:7998933-8004313	Ctr	Mut	205.125	53.7754	-1.931	0.0002	0.00793	FBgn0031971
46	XLOC_014033	CG34434	X:5508586-5510403	Ctr	Mut	2.27512	0.607058	-1.906	0.0001	0.00448	FBgn0250904
47	XLOC_007806	Cpr65Ax2	3L:6143157-6143957	Ctr	Mut	69.2301	18.5246	-1.902	5.00E-05	0.00245	FBgn0042118
48	XLOC_012090	Cpr97Eb	3R:22915425-22916964	Ctr	Mut	13.7466	3.71012	-1.89	5.00E-05	0.00245	FBgn0039481
49	XLOC_008781	CG14566	3L:21732578-21733409	Ctr	Mut	208.541	58.0778	-1.844	5.00E-05	0.00245	FBgn0037127
50	XLOC_007326	CG14572	3L:21734633-21735541	Ctr	Mut	125.315	35.5982	-1.816	5.00E-05	0.00245	FBgn0037128
51	XLOC_003741	CG10764	2R:13340923-13343234	Ctr	Mut	0.589858	0.168019	-1.812	0.00165	0.04288	FBgn0034221
52	XLOC_010668	CG14661	3R:779227-780975	Ctr	Mut	7.74517	2.21541	-1.806	5.00E-05	0.00245	FBgn0037288
53	XLOC_014489	CG13403	X:14085247-14086019	Ctr	Mut	13.0043	3.78246	-1.782	5.00E-05	0.00245	FBgn0030544
54	XLOC_013400	CG11585	X:14083548-14084788	Ctr	Mut	77.8925	23.0131	-1.759	5.00E-05	0.00245	FBgn0030543
55	XLOC_007752	blanks	3L:5490622-5492058	Ctr	Mut	5.76446	1.7253	-1.74	5.00E-05	0.00245	FBgn0035608

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
56	XLOC_007797	l(3)mbn	3L:6117033-6121907	Ctr	Mut	42.175	12.7285	-1.728	5.00E-05	0.00245	FBgn0002440
57	XLOC_010531	CG15530	3R:26003598-26005903	Ctr	Mut	1.22738	0.3786	-1.697	0.00015	0.00635	FBgn0039752
58	XLOC_010488	Obp99d	3R:25540260-25540674	Ctr	Mut	19.4452	6.02601	-1.69	0.0003	0.01084	FBgn0039684
59	XLOC_008952	Hph	3R:1082762-1095297	Ctr	Mut	85.8225	26.9312	-1.672	0.0002	0.00793	FBgn0264652
60	XLOC_012147	CG43124	3R:24088456-24090504	Ctr	Mut	10.8337	3.40539	-1.67	0.0004	0.01385	FBgn0262587
61	XLOC_012147	CG43125	3R:24088456-24090504	Ctr	Mut	10.8337	3.40539	-1.67	0.0004	0.01385	FBgn0262588
62	XLOC_001224	CG13077	2L:19563903-19567359	Ctr	Mut	40.0333	12.6016	-1.668	5.00E-05	0.00245	FBgn0032810
63	XLOC_001224	CG13078	2L:19563903-19567359	Ctr	Mut	40.0333	12.6016	-1.668	5.00E-05	0.00245	FBgn0032809
64	XLOC_007393	CG11131	3L:22891843-22893362	Ctr	Mut	471.945	149.262	-1.661	5.00E-05	0.00245	FBgn0037204
65	XLOC_014285	CG15308	X:10112038-10113172	Ctr	Mut	586.355	186.877	-1.65	5.00E-05	0.00245	FBgn0040941
66	XLOC_007766	DnaJ-1	3L:5743128-5745289	Ctr	Mut	182.369	59.2321	-1.622	5.00E-05	0.00245	FBgn0263106
67	XLOC_012056	TwdlO	3R:22451729-22452419	Ctr	Mut	334.705	110.297	-1.601	5.00E-05	0.00245	FBgn0039438
68	XLOC_012952	Cpr5C	X:5698709-5699377	Ctr	Mut	11.4301	3.77434	-1.599	0.00025	0.00942	FBgn0029811
69	XLOC_001490	CG11835	2L:560567-563347	Ctr	Mut	3.37205	1.1194	-1.591	5.00E-05	0.00245	FBgn0031264
70	XLOC_001954	CG7203	2L:7752167-7753156	Ctr	Mut	324.836	107.943	-1.589	5.00E-05	0.00245	FBgn0031942
71	XLOC_006368	Lcp65Aa	3L:6144403-6144903	Ctr	Mut	261.484	87.5947	-1.578	5.00E-05	0.00245	FBgn0020645
72	XLOC_007802	Lcp65Ae	3L:6130682-6131218	Ctr	Mut	155.104	51.9618	-1.578	5.00E-05	0.00245	FBgn0020640
73	XLOC_000968	Adh	2L:14599768-14689326	Ctr	Mut	5266.79	1768.03	-1.575	5.00E-05	0.00245	FBgn0000055
74	XLOC_000968	Adhr	2L:14599768-14689326	Ctr	Mut	5266.79	1768.03	-1.575	5.00E-05	0.00245	FBgn0000056
75	XLOC_006973	CG13067	3L:16264381-16265018	Ctr	Mut	1927.84	654.105	-1.559	5.00E-05	0.00245	FBgn0036589
76	XLOC_007072	CG7497	3L:17626548-17630975	Ctr	Mut	8.54687	2.90929	-1.555	5.00E-05	0.00245	FBgn0036742
77	XLOC_000123	Nplp4	2L:2008459-2008966	Ctr	Mut	1339.23	456.409	-1.553	5.00E-05	0.00245	FBgn0040717
78	XLOC_003116	CG13748	2R:4827919-4828641	Ctr	Mut	126.142	42.9972	-1.553	0.00095	0.02805	FBgn0033355
79	XLOC_009157	Atg13	3R:4175523-4178539	Ctr	Mut	42.4394	14.5206	-1.547	0.0009	0.02717	FBgn0261108
80	XLOC_009326	Ugt86Dc	3R:6980376-6982235	Ctr	Mut	7.19726	2.48006	-1.537	5.00E-05	0.00245	FBgn0040257
81	XLOC_005277	proPO-A1	2R:13761451-13814148	Ctr	Mut	34.6074	12.0992	-1.516	5.00E-05	0.00245	FBgn0261362
82	XLOC_008797	CG7130	3L:22068155-22068856	Ctr	Mut	10.4691	3.68428	-1.507	0.0001	0.00448	FBgn0037151
83	XLOC_000678	CG33301	2L:10049481-10050956	Ctr	Mut	7.0418	2.48883	-1.5	5.00E-05	0.00245	FBgn0053301
84	XLOC_011873	CG12268	3R:19769349-19774141	Ctr	Mut	38.7036	13.9172	-1.476	5.00E-05	0.00245	FBgn0039131

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
85	XLOC_008658	CG7365	3L:20117397-20120481	Ctr	Mut	19.9037	7.168	-1.473	5.00E-05	0.00245	FBgn0036939
86	XLOC_008411	CG13044	3L:16295589-16296278	Ctr	Mut	846.606	305.344	-1.471	5.00E-05	0.00245	FBgn0036599
87	XLOC_010227	CG31103	3R:21045007-21047163	Ctr	Mut	10.4994	3.80111	-1.466	5.00E-05	0.00245	FBgn0051103
88	XLOC_005773	chrw	2R:19407361-19411096	Ctr	Mut	31.4098	11.3926	-1.463	0.0008	0.02472	FBgn0015372
89	XLOC_011984	CG10560	3R:21144271-21145801	Ctr	Mut	37.3235	13.6398	-1.452	5.00E-05	0.00245	FBgn0039325
90	XLOC_007003	CG4229	3L:16549922-16551058	Ctr	Mut	289.674	105.89	-1.452	0.00015	0.00635	FBgn0036639
91	XLOC_008003	Hsp67Ba	3L:9370901-9372634	Ctr	Mut	4.58809	1.68227	-1.447	0.0002	0.00793	FBgn0001227
92	XLOC_004604	Lcp2	2R:4321301-4322080	Ctr	Mut	29.8738	11.0082	-1.44	0.0003	0.01084	FBgn0002533
93	XLOC_008579	CG34256	3L:18981835-18982915	Ctr	Mut	5.97674	2.21397	-1.433	0.00125	0.03415	FBgn0085285
94	XLOC_006820	stv	3L:13470640-13476615	Ctr	Mut	98.0141	36.408	-1.429	0.00055	0.0181	FBgn0086708
95	XLOC_014566	CG15599	X:15578717-15582995	Ctr	Mut	5.82233	2.16531	-1.427	5.00E-05	0.00245	FBgn0030667
96	XLOC_013446	CG6324	X:15370021-15375755	Ctr	Mut	3.28488	1.22934	-1.418	5.00E-05	0.00245	FBgn0030647
97	XLOC_006983	CG13060	3L:16312955-16313554	Ctr	Mut	176.651	66.1886	-1.416	0.00015	0.00635	FBgn0036606
98	XLOC_004781	CG13218	2R:7120768-7121209	Ctr	Mut	179.559	67.4227	-1.413	0.0001	0.00448	FBgn0033587
99	XLOC_008412	CG13043	3L:16298324-16298970	Ctr	Mut	597.248	224.922	-1.409	0.0002	0.00793	FBgn0036600
100	XLOC_008778	CG14569	3L:21723824-21724650	Ctr	Mut	358.927	135.194	-1.409	0.0002	0.00793	FBgn0037123
101	XLOC_010766	CG15597	3R:2110180-2110922	Ctr	Mut	154.353	58.4937	-1.4	5.00E-05	0.00245	FBgn0037420
102	XLOC_012255	CG31029	3R:25955010-25958117	Ctr	Mut	0.845408	0.326161	-1.374	0.00135	0.03642	FBgn0051029
103	XLOC_005156	CR43730	2R:12176829-12178927	Ctr	Mut	67.0347	26.0723	-1.362	5.00E-05	0.00245	FBgn0263981
104	XLOC_008779	CG14568	3L:21725905-21726587	Ctr	Mut	229.553	89.5976	-1.357	0.0003	0.01084	FBgn0037124
105	XLOC_007831	ple	3L:6707137-6712625	Ctr	Mut	109.25	42.9276	-1.348	5.00E-05	0.00245	FBgn0005626
106	XLOC_006972	CG13068	3L:16262491-16263011	Ctr	Mut	536.478	210.899	-1.347	0.00045	0.01528	FBgn0036588
107	XLOC_004792	Cpr47Eg	2R:7165460-7165946	Ctr	Mut	5681.34	2237.57	-1.344	5.00E-05	0.00245	FBgn0086519
108	XLOC_005740	CG13545	2R:19041400-19049553	Ctr	Mut	455.401	180.392	-1.336	0.0001	0.00448	FBgn0034828
109	XLOC_010005	CG7069	3R:18198556-18201301	Ctr	Mut	0.838872	0.334754	-1.325	0.00115	0.03208	FBgn0038952
110	XLOC_003261	CG13228	2R:7121507-7121915	Ctr	Mut	119.517	47.8564	-1.32	5.00E-05	0.00245	FBgn0033588
111	XLOC_009280	CG31477	3R:5947494-5948042	Ctr	Mut	12.8956	5.16838	-1.319	0.00105	0.02992	FBgn0051477
112	XLOC_008563	CG12477	3L:18714992-18716372	Ctr	Mut	8.68613	3.50669	-1.309	5.00E-05	0.00245	FBgn0036809
113	XLOC_007664	CG12766	3L:3938623-3940125	Ctr	Mut	6.95344	2.80943	-1.307	0.0006	0.01944	FBgn0035476

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
114	XLOC_010229	CG11852	3R:21063346-21064617	Ctr	Mut	176.98	71.8569	-1.3	5.00E-05	0.00245	FBgn0039297
115	XLOC_014288	CG12643	X:10158265-10159296	Ctr	Mut	39.5161	16.2466	-1.282	5.00E-05	0.00245	FBgn0040942
116	XLOC_011190	CG5999	3R:8567679-8569402	Ctr	Mut	11.5723	4.76525	-1.28	5.00E-05	0.00245	FBgn0038083
117	XLOC_013568	CG5162	X:17109334-17111409	Ctr	Mut	10.0183	4.15535	-1.27	0.00025	0.00942	FBgn0030828
118	XLOC_006289	CG32248	3L:4471748-4472438	Ctr	Mut	192.03	79.92	-1.265	5.00E-05	0.00245	FBgn0052248
119	XLOC_011481	Edg91	3R:13436577-13437233	Ctr	Mut	95.3078	39.9	-1.256	0.0002	0.00793	FBgn0004554
120	XLOC_008416	CG13039	3L:16318182-16319862	Ctr	Mut	722.418	303.544	-1.251	0.00035	0.01228	FBgn0036609
121	XLOC_008416	CG13040	3L:16318182-16319862	Ctr	Mut	722.418	303.544	-1.251	0.00035	0.01228	FBgn0036608
122	XLOC_004644	PO45	2R:4929765-4932213	Ctr	Mut	60.5837	25.6193	-1.242	5.00E-05	0.00245	FBgn0033367
123	XLOC_008414	CG13041	3L:16312053-16312626	Ctr	Mut	333.969	141.834	-1.236	0.00025	0.00942	FBgn0036605
124	XLOC_006984	CG13059	3L:16316771-16317408	Ctr	Mut	696.059	296.916	-1.229	0.00035	0.01228	FBgn0036607
125	XLOC_004251	CG13560	2R:19635095-19635741	Ctr	Mut	876.361	374.119	-1.228	0.00015	0.00635	FBgn0034899
126	XLOC_010128	CG13606	3R:19997934-20009752	Ctr	Mut	71.9911	30.7475	-1.227	0.0011	0.03093	FBgn0039161
127	XLOC_002677	CG17570	2L:20263379-20264703	Ctr	Mut	10.007	4.27965	-1.225	0.0002	0.00793	FBgn0260000
128	XLOC_002475	CG5953	2L:16508075-16532877	Ctr	Mut	31.7608	13.6132	-1.222	0.0015	0.03966	FBgn0263555
129	XLOC_005193	Ugt37c1	2R:12730800-12732543	Ctr	Mut	23.5817	10.1546	-1.216	0.00015	0.00635	FBgn0026754
130	XLOC_009674	CG17560	3R:12443671-12445600	Ctr	Mut	10.9523	4.73	-1.211	0.0002	0.00793	FBgn0038450
131	XLOC_003262	CG13227	2R:7122758-7123260	Ctr	Mut	105.254	45.4569	-1.211	0.00025	0.00942	FBgn0033589
132	XLOC_011983	CG10553	3R:21142182-21143833	Ctr	Mut	3.58872	1.56124	-1.201	0.0017	0.04354	FBgn0039324
133	XLOC_011483	CG14324	3R:13444002-13444513	Ctr	Mut	138.158	60.753	-1.185	0.002	0.0492	FBgn0038527
134	XLOC_007730	CG32237	3L:4847182-4852708	Ctr	Mut	318.934	140.658	-1.181	0.0018	0.04544	FBgn0052237
135	XLOC_004080	CG33225	2R:17477890-17479089	Ctr	Mut	7.64226	3.38216	-1.176	0.00185	0.04626	FBgn0053225
136	XLOC_007800	Lcp65Ag1	3L:6127796-6128374	Ctr	Mut	2092.22	930.964	-1.168	5.00E-05	0.00245	FBgn0020638
137	XLOC_002404	mol	2L:14975746-14997559	Ctr	Mut	20.8796	9.30152	-1.167	0.00015	0.00635	FBgn0086711
138	XLOC_009718	CG42821	3R:13238733-13239263	Ctr	Mut	70.1204	31.4534	-1.157	0.00025	0.00942	FBgn0262003
139	XLOC_008690	CG11796	3L:20429349-20432972	Ctr	Mut	183.973	82.7544	-1.153	0.0003	0.01084	FBgn0036992
140	XLOC_006860	shd	3L:14607557-14615305	Ctr	Mut	16.1022	7.2448	-1.152	5.00E-05	0.00245	FBgn0003388
141	XLOC_010568	CG15544	3R:26653498-26667756	Ctr	Mut	3.87806	1.76018	-1.14	0.0007	0.02195	FBgn0039804
142	XLOC_011430	Scp2	3R:12400264-12409382	Ctr	Mut	33.28	15.138	-1.136	0.00015	0.00635	FBgn0020907

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
143	XLOC_003374	Cpr49Af	2R:8293509-8293948	Ctr	Mut	106.149	48.4462	-1.132	0.0002	0.00793	FBgn0033729
144	XLOC_011008	CG42857	3R:5646235-5646913	Ctr	Mut	9.68486	4.42239	-1.131	0.00105	0.02992	FBgn0262104
145	XLOC_007756	CG10625	3L:5531507-5544280	Ctr	Mut	408.193	189.606	-1.106	0.0003	0.01084	FBgn0035612
146	XLOC_005521	CG13868	2R:16196281-16204422	Ctr	Mut	280.598	130.78	-1.101	0.00105	0.02992	FBgn0034501
147	XLOC_006473	CG13678	3L:8215240-8216053	Ctr	Mut	954.63	446.702	-1.096	0.00115	0.03208	FBgn0035859
148	XLOC_006192	Hsp83	3L:3192968-3197631	Ctr	Mut	749.786	387.733	-1.095	0.0002	0.00793	FBgn0001233
149	XLOC_001646	CG17224	2L:3016590-3018308	Ctr	Mut	12.7966	6.03424	-1.085	0.00095	0.02805	FBgn0031489
150	XLOC_013564	CG5070	X:17097146-17097965	Ctr	Mut	123.881	58.6552	-1.079	0.0002	0.00793	FBgn0030824
151	XLOC_004271	Tal	2R:19827935-19829637	Ctr	Mut	501.871	238.365	-1.074	0.00035	0.01228	FBgn0023477
152	XLOC_007157	Cpr76Bb	3L:19512748-19513519	Ctr	Mut	51.5536	24.5953	-1.068	0.0008	0.02472	FBgn0036879
153	XLOC_011978	CG31097	3R:21119179-21128039	Ctr	Mut	73.7198	35.2114	-1.066	0.00175	0.0445	FBgn0051288
154	XLOC_011978	CG31102	3R:21119179-21128039	Ctr	Mut	73.7198	35.2114	-1.066	0.00175	0.0445	FBgn0051097
155	XLOC_011978	CG31288	3R:21119179-21128039	Ctr	Mut	73.7198	35.2114	-1.066	0.00175	0.0445	FBgn0051102
156	XLOC_011697	CG3301	3R:17099964-17101973	Ctr	Mut	40.183	19.3845	-1.052	0.00025	0.00942	FBgn0038878
157	XLOC_010570	Cpr100A	3R:26693244-26694908	Ctr	Mut	194.102	93.7259	-1.05	0.00175	0.0445	FBgn0039805
158	XLOC_011678	Mvl	3R:16877104-16886523	Ctr	Mut	76.7857	37.1095	-1.049	0.0005	0.01666	FBgn0011672
159	XLOC_003750	CG4847	2R:13399027-13401364	Ctr	Mut	112.724	54.5119	-1.048	0.00095	0.02805	FBgn0034229
160	XLOC_009412	GstD7	3R:8204115-8204977	Ctr	Mut	149.782	72.5543	-1.046	0.00065	0.02068	FBgn0010043
161	XLOC_013222	CG15201	X:11038152-11038842	Ctr	Mut	24.7431	12.01	-1.043	0.00155	0.04068	FBgn0030272
162	XLOC_000625	Aldh	2L:9370308-9415156	Ctr	Mut	670.262	325.788	-1.041	0.0003	0.01084	FBgn0012036
163	XLOC_007798	Lcp65Ag2	3L:6123485-6126693	Ctr	Mut	5409.42	2637.64	-1.036	5.00E-05	0.00245	FBgn002063
164	XLOC_007798	Lcp65Ag3	3L:6123485-6126693	Ctr	Mut	5409.42	2637.64	-1.036	5.00E-05	0.00245	FBgn0086611
165	XLOC_013750	CG34120	X:20994753-21012062	Ctr	Mut	6.2822	3.08376	-1.027	0.0015	0.03966	FBgn0083956
166	XLOC_007708	CG32241	3L:4465103-4466932	Ctr	Mut	492.191	242.305	-1.022	0.00105	0.02992	FBgn0052241
167	XLOC_002166	CG7296	2L:10686390-10686988	Ctr	Mut	1079.97	534.024	-1.016	0.0017	0.04354	FBgn0032283
168	XLOC_002581	CG42502	2L:18810035-18810986	Ctr	Mut	106.086	52.5898	-1.012	0.00145	0.03872	FBgn0040992
169	XLOC_004536	CG12826	2R:3557831-3558614	Ctr	Mut	42.4752	21.2215	-1.001	0.00045	0.01528	FBgn0033207
170	XLOC_009843	CG17752	3R:15435765-15438166	Ctr	Mut	93.3581	187.682	1.0074	0.00065	0.02068	FBgn0038718
171	XLOC_009803	CG14302	3R:14558422-14558756	Ctr	Mut	1528.84	3103.22	1.0213	0.0006	0.01944	FBgn0038647

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
172	XLOC_011562	CG14291	3R:14736772-14738641	Ctr	Mut	6.73496	13.6718	1.0215	0.00035	0.01228	FBgn0038660
173	XLOC_000050	CG13947	2L:779172-779655	Ctr	Mut	80.96	164.841	1.0258	0.00035	0.01228	FBgn0031277
174	XLOC_006714	CG7252	3L:11940774-11942348	Ctr	Mut	60.8624	123.922	1.0258	0.00025	0.00942	FBgn0036226
175	XLOC_005078	CG8093	2R:11101785-11103191	Ctr	Mut	7.84335	16.1261	1.0399	0.0012	0.03321	FBgn0033999
176	XLOC_000098	CG33128	2L:1494089-1495456	Ctr	Mut	343.173	707.19	1.0432	0.00015	0.00635	FBgn0053128
177	XLOC_006553	PGRP-LA	3L:9327431-9331436	Ctr	Mut	10.738	22.1476	1.0444	0.0018	0.04544	FBgn0035975
178	XLOC_011976	CG10514	3R:21104145-21105541	Ctr	Mut	54.4335	112.348	1.0454	0.0007	0.02195	FBgn0039312
179	XLOC_010214	CG11836	3R:20896852-20898434	Ctr	Mut	9.96505	20.6502	1.0512	0.00085	0.02603	FBgn0039272
180	XLOC_007561	CG5687	3L:1933632-1938740	Ctr	Mut	9.50584	19.7321	1.0537	0.0001	0.00448	FBgn0035293
181	XLOC_004931	CG13323	2R:8934237-8937132	Ctr	Mut	981.74	2042.16	1.0567	0.00115	0.03208	FBgn0033788
182	XLOC_008666	CG7017	3L:20203391-20205262	Ctr	Mut	290.07	604.283	1.0588	0.00035	0.01228	FBgn0036951
183	XLOC_003180	CG1698	2R:5653866-5694370	Ctr	Mut	14.1789	29.6999	1.0667	0.00055	0.0181	FBgn0033443
184	XLOC_006353	Jon65Aii	3L:6045420-6046315	Ctr	Mut	188.37	396.193	1.0726	0.0004	0.01385	FBgn0035666
185	XLOC_006202	CG12009	3L:3291458-3299000	Ctr	Mut	16.8805	35.5859	1.0759	0.00015	0.00635	FBgn0035430
186	XLOC_004187	CG9826	2R:18728357-18729988	Ctr	Mut	11.8752	25.0608	1.0775	0.0002	0.00793	FBgn0034784
187	XLOC_013817	CG3706	X:824963-826704	Ctr	Mut	2.0898	4.43946	1.087	0.00105	0.02992	FBgn0040342
188	XLOC_014720	CG32557	X:17746990-17750156	Ctr	Mut	8.90336	18.9303	1.0883	0.00025	0.00942	FBgn0052557
189	XLOC_006330	Blimp-1	3L:5623959-5643812	Ctr	Mut	3.10495	6.62457	1.0933	0.0003	0.01084	FBgn0035625
190	XLOC_006816	CG10725	3L:13431053-13432109	Ctr	Mut	520.684	1111.87	1.0945	0.0001	0.00448	FBgn0036362
191	XLOC_008667	CG6933	3L:20211599-20213520	Ctr	Mut	670.713	1436.17	1.0985	5.00E-05	0.00245	FBgn0036952
192	XLOC_003427	CG13325	2R:9003809-9013517	Ctr	Mut	4.007	8.58533	1.0994	0.0002	0.00793	FBgn0033792
193	XLOC_005579	Cht9	2R:16955486-16956813	Ctr	Mut	254.781	547.998	1.1049	5.00E-05	0.00245	FBgn0034582
194	XLOC_000887	CG31728	2L:13251508-13257426	Ctr	Mut	13.7833	29.7011	1.1076	0.00015	0.00635	FBgn0051728
195	XLOC_009511	tal-AA	3R:9638830-9640370	Ctr	Mut	40.0736	86.386	1.1081	0.0001	0.00448	FBgn0259731
196	XLOC_002035	CG9289	2L:8681313-8684350	Ctr	Mut	37.6492	81.2407	1.1096	0.00085	0.02603	FBgn0032058
197	XLOC_003363	CG30043	2R:8246940-8250730	Ctr	Mut	17.1834	37.2144	1.1149	0.00015	0.00635	FBgn0050043
198	XLOC_012234	Jon99Ci	3R:25750947-25751911	Ctr	Mut	250.699	543.716	1.1169	0.0001	0.00448	FBgn0003358
199	XLOC_000654	CG4017	2L:9767058-9768727	Ctr	Mut	2.65353	5.76484	1.1194	0.0012	0.03321	FBgn0032143
200	XLOC_012290	PH4alphaSG1	3R:26329386-26331383	Ctr	Mut	3.33517	7.29475	1.1291	0.0002	0.00793	FBgn0051014

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
201	XLOC_012203	CG11470	3R:25414757-25416761	Ctr	Mut	131.788	291.218	1.1439	0.00065	0.02068	FBgn0039671
202	XLOC_012203	CG31041	3R:25414757-25416761	Ctr	Mut	131.788	291.218	1.1439	0.00065	0.02068	FBgn0051041
203	XLOC_002557	CG10283	2L:18472963-18484573	Ctr	Mut	3.92372	8.69094	1.1473	0.0003	0.01084	FBgn0032681
204	XLOC_005213	Amy-p	2R:13006210-13013887	Ctr	Mut	134.272	298.336	1.1518	0.00165	0.04288	FBgn0000079
205	XLOC_012951	CG12239	X:5686179-5688576	Ctr	Mut	7.9618	17.7984	1.1606	5.00E-05	0.00245	FBgn0029810
206	XLOC_000504	CG7191	2L:7767623-7775040	Ctr	Mut	7.66408	17.3393	1.1779	5.00E-05	0.00245	FBgn0031945
207	XLOC_005287	CG10912	2R:13943567-13944564	Ctr	Mut	420.073	950.946	1.1787	5.00E-05	0.00245	FBgn0034296
208	XLOC_000049	CG12506	2L:773546-776918	Ctr	Mut	13.5498	30.6975	1.1799	0.0015	0.03966	FBgn0031276
209	XLOC_000049	CG13946	2L:773546-776918	Ctr	Mut	13.5498	30.6975	1.1799	0.0015	0.03966	FBgn0040725
210	XLOC_004724	CG12912	2R:6091595-6124853	Ctr	Mut	6.83549	15.5198	1.183	0.00075	0.02331	FBgn0000448
211	XLOC_004724	Hr46	2R:6091595-6124853	Ctr	Mut	6.83549	15.5198	1.183	0.00075	0.02331	FBgn0033497
212	XLOC_009442	CG14395	3R:8488552-8499686	Ctr	Mut	2.64887	6.01455	1.1831	5.00E-05	0.00245	FBgn0038073
213	XLOC_000973	CG34166	2L:14743252-14743859	Ctr	Mut	790.567	1799.38	1.1865	0.00015	0.00635	FBgn0085195
214	XLOC_006427	CG18417	3L:7393169-7394771	Ctr	Mut	23.2746	53.1451	1.1912	0.0001	0.00448	FBgn0035780
215	XLOC_008498	Jon74E	3L:17523180-17542227	Ctr	Mut	158.571	365.792	1.2059	0.00075	0.02331	FBgn0023197
216	XLOC_009183	CG8369	3R:4647729-4648560	Ctr	Mut	74.2016	171.812	1.2113	5.00E-05	0.00245	FBgn0040532
217	XLOC_002040	CG9463	2L:8765360-8772344	Ctr	Mut	7.25103	16.8242	1.2143	5.00E-05	0.00245	FBgn0032066
218	XLOC_002040	CG9465	2L:8765360-8772344	Ctr	Mut	7.25103	16.8242	1.2143	5.00E-05	0.00245	FBgn0032067
219	XLOC_009690	CG5225	3R:12854427-12856423	Ctr	Mut	24.1347	56.0932	1.2167	5.00E-05	0.00245	FBgn0038468
220	XLOC_003280	sha	2R:7211818-7223968	Ctr	Mut	0.887007	2.07106	1.2234	5.00E-05	0.00245	FBgn0003382
221	XLOC_011372	CG14872	3R:11333780-11335676	Ctr	Mut	128.748	301.003	1.2252	5.00E-05	0.00245	FBgn0038346
222	XLOC_000606	CG9568	2L:9010585-9011358	Ctr	Mut	202.832	477.651	1.2357	0.0001	0.00448	FBgn0032087
223	XLOC_011462	Mur89F	3R:12977587-13027935	Ctr	Mut	12.793	30.1621	1.2374	0.0011	0.03093	FBgn0038492
224	XLOC_007793	Jon65Ai	3L:6046683-6047543	Ctr	Mut	775.503	1833.99	1.2418	5.00E-05	0.00245	FBgn0035667
225	XLOC_012425	yellow-h	4:248549-251054	Ctr	Mut	13.8626	33.1027	1.2558	5.00E-05	0.00245	FBgn0039896
226	XLOC_004578	Obp44a	2R:4018937-4022588	Ctr	Mut	263.126	630.886	1.2616	0.00025	0.00942	FBgn0033268
227	XLOC_008527	CG5506	3L:17872195-17877476	Ctr	Mut	179.893	432.693	1.2662	0.00015	0.00635	FBgn0036766
228	XLOC_008253	CG10154	3L:13428779-13430236	Ctr	Mut	81.5316	196.683	1.2704	5.00E-05	0.00245	FBgn0036361
229	XLOC_010241	CG11892	3R:21107269-21108964	Ctr	Mut	34.8713	84.6199	1.279	5.00E-05	0.00245	FBgn0039313

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
230	XLOC_000757	CG17134	2L:10804268-10842603	Ctr	Mut	70.5108	171.507	1.2824	0.00125	0.03415	FBgn0032304
231	XLOC_012748	CG11382	X:1104002-1105886	Ctr	Mut	20.6205	50.7274	1.2987	5.00E-05	0.00245	FBgn0040367
232	XLOC_011968	Cad96Cb	3R:21049877-21058043	Ctr	Mut	0.623151	1.536	1.3015	0.00125	0.03415	FBgn0039294
233	XLOC_000547	CG14275	2L:8327455-8332911	Ctr	Mut	12.3333	30.5993	1.3109	5.00E-05	0.00245	FBgn0032022
234	XLOC_001570	CG42296	2L:2032634-2035936	Ctr	Mut	3.04095	7.54833	1.3116	0.00185	0.04626	FBgn0259192
235	XLOC_011913	CG5768	3R:20211824-20216089	Ctr	Mut	3.49094	8.66815	1.3121	0.0003	0.01084	FBgn0039198
236	XLOC_006133	CG13806	3L:2266393-2267640	Ctr	Mut	21.5374	53.6067	1.3156	5.00E-05	0.00245	FBgn0035325
237	XLOC_010696	CG2663	3R:1226315-1229851	Ctr	Mut	31.0395	77.9	1.3275	5.00E-05	0.00245	FBgn0037323
238	XLOC_008164	obst-G	3L:11945400-11946417	Ctr	Mut	15.2566	38.4208	1.3325	5.00E-05	0.00245	FBgn0036228
239	XLOC_010597	mey	3R:27324960-27332716	Ctr	Mut	6.69616	16.9632	1.341	5.00E-05	0.00245	FBgn0039851
240	XLOC_002682	CG16798	2L:20376982-20382355	Ctr	Mut	2.30975	5.87391	1.3466	5.00E-05	0.00245	FBgn0032856
241	XLOC_007681	mas	3L:4161665-4167464	Ctr	Mut	9.87177	25.175	1.3506	5.00E-05	0.00245	FBgn0011653
242	XLOC_009023	CG31560	3R:2135717-2136612	Ctr	Mut	9.67755	24.7251	1.3533	5.00E-05	0.00245	FBgn0051560
243	XLOC_010182	CG13631	3R:20594690-20596687	Ctr	Mut	10.1682	26.0224	1.3557	5.00E-05	0.00245	FBgn0040600
244	XLOC_000833	CG17211	2L:12114359-12129896	Ctr	Mut	1.59323	4.09729	1.3627	5.00E-05	0.00245	FBgn0032414
245	XLOC_008599	CG14089	3L:19273219-19273900	Ctr	Mut	74.8296	193.375	1.3697	5.00E-05	0.00245	FBgn0036861
246	XLOC_010503	neo	3R:25644833-25651522	Ctr	Mut	34.9169	90.2799	1.3705	5.00E-05	0.00245	FBgn0039704
247	XLOC_006406	CG15829	3L:7122002-7122480	Ctr	Mut	172.854	447.837	1.3734	5.00E-05	0.00245	FBgn0035743
248	XLOC_010935	CG11966	3R:4775923-4788760	Ctr	Mut	4.00914	10.4212	1.3782	5.00E-05	0.00245	FBgn0037645
249	XLOC_011220	Lip3	3R:9195390-9197626	Ctr	Mut	2.11262	5.50742	1.3823	0.0001	0.00448	FBgn0023495
250	XLOC_014660	wus	X:16832529-16834903	Ctr	Mut	8.82736	23.1449	1.3906	5.00E-05	0.00245	FBgn0030805
251	XLOC_010767	CG31556	3R:2132706-2135106	Ctr	Mut	12.0779	32.1914	1.4143	0.0002	0.00793	FBgn0051556
252	XLOC_004806	epsilonTry	2R:7232741-7233571	Ctr	Mut	934.289	2495.49	1.4174	5.00E-05	0.00245	FBgn0010425
253	XLOC_010550	Jon99Fi	3R:26314684-26315617	Ctr	Mut	1462.32	3910.72	1.4192	5.00E-05	0.00245	FBgn0039778
254	XLOC_008406	CG13048	3L:16275901-16277061	Ctr	Mut	35.901	96.1071	1.4206	0.0018	0.04544	FBgn0036593
255	XLOC_003886	CG15080	2R:14670805-14682105	Ctr	Mut	10.0405	27.0252	1.4285	0.0017	0.04354	FBgn0034391
256	XLOC_007712	NT1	3L:4501651-4508433	Ctr	Mut	6.52158	17.6027	1.4325	5.00E-05	0.00245	FBgn0261526
257	XLOC_014066	CG4666	X:5970105-5971587	Ctr	Mut	17.3974	47.1084	1.4371	5.00E-05	0.00245	FBgn0029838
258	XLOC_004605	Cyp4ad1	2R:4326899-4330311	Ctr	Mut	18.4917	50.2592	1.4425	5.00E-05	0.00245	FBgn0033292

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
259	XLOC_004978	CG42807	2R:9459028-9460045	Ctr	Mut	59.2631	162.158	1.4522	5.00E-05	0.00245	FBgn0261989
260	XLOC_010325	TwdlS	3R:22462435-22463182	Ctr	Mut	11.4598	31.454	1.4567	0.00195	0.04831	FBgn0039443
261	XLOC_010825	Gld	3R:3075520-3091423	Ctr	Mut	9.77219	26.8999	1.4609	5.00E-05	0.00245	FBgn0001112
262	XLOC_010825	Pbprp4	3R:3075520-3091423	Ctr	Mut	9.77219	26.8999	1.4609	5.00E-05	0.00245	FBgn0011282
263	XLOC_009312	CG4089	3R:6641012-6643214	Ctr	Mut	4.62652	12.8226	1.4707	5.00E-05	0.00245	FBgn0037838
264	XLOC_006509	Cpr66D	3L:8633538-8636684	Ctr	Mut	42.8269	119.543	1.4809	0.001	0.02928	FBgn0052029
265	XLOC_014919	lcs	X:21267753-21268326	Ctr	Mut	94.4453	263.702	1.4814	5.00E-05	0.00245	FBgn0028583
266	XLOC_010465	CG14515	3R:25023304-25024933	Ctr	Mut	5.27957	14.8611	1.4931	0.00185	0.04626	FBgn0039648
267	XLOC_002251	CG6785	2L:12041048-12045200	Ctr	Mut	4.38303	12.3881	1.499	0.00055	0.0181	FBgn0032399
268	XLOC_006189	CG32284	3L:3186663-3187245	Ctr	Mut	180.857	516.018	1.5126	5.00E-05	0.00245	FBgn0052284
269	XLOC_014737	upd2	X:18134688-18139206	Ctr	Mut	0.733816	2.09396	1.5127	0.0012	0.03321	FBgn0030904
270	XLOC_002367	CG43333	2L:14163504-14180681	Ctr	Mut	12.7529	36.4789	1.5162	0.001	0.02928	FBgn0263038
271	XLOC_005286	CG10911	2R:13940051-13941569	Ctr	Mut	290.371	833.345	1.521	5.00E-05	0.00245	FBgn0034295
272	XLOC_012686	tyn	X:39387-97296	Ctr	Mut	21.4836	61.6564	1.521	5.00E-05	0.00245	FBgn0029128
273	XLOC_003343	CG43315	2R:8128722-8129264	Ctr	Mut	7.8951	22.726	1.5253	0.00035	0.01228	FBgn0263020
274	XLOC_007202	CG6996	3L:20200633-20202700	Ctr	Mut	18.1567	52.392	1.5288	5.00E-05	0.00245	FBgn0036950
275	XLOC_008902	TwdlU	3R:69327-71262	Ctr	Mut	1.08298	3.15648	1.5433	0.00155	0.04068	FBgn0037223
276	XLOC_011369	CG6118	3R:11246825-11259906	Ctr	Mut	9.69039	28.4101	1.5518	0.00025	0.00942	FBgn0038339
277	XLOC_005640	CG4386	2R:17686834-17688405	Ctr	Mut	21.1043	62.2835	1.5613	5.00E-05	0.00245	FBgn0034661
278	XLOC_014536	CG9095	X:15033961-15057255	Ctr	Mut	15.0702	44.5009	1.5621	5.00E-05	0.00245	FBgn0030617
279	XLOC_004639	CG34350	2R:4880448-4894092	Ctr	Mut	8.70067	25.6983	1.5625	5.00E-05	0.00245	FBgn0085379
280	XLOC_006200	CG14960	3L:3271035-3276255	Ctr	Mut	11.8606	35.1304	1.5665	5.00E-05	0.00245	FBgn0035428
281	XLOC_008755	CG5656	3L:21435326-21437919	Ctr	Mut	2.55847	7.57973	1.5669	5.00E-05	0.00245	FBgn0037083
282	XLOC_007938	CG7201	3L:8283385-8288427	Ctr	Mut	3.81519	11.3153	1.5684	0.0005	0.01666	FBgn0035865
283	XLOC_008903	TwdlF	3R:72743-74040	Ctr	Mut	174.186	519.08	1.5753	0.00065	0.02068	FBgn0037224
284	XLOC_010630	TwdlG	3R:74433-77467	Ctr	Mut	55.912	168.186	1.5888	0.00105	0.02992	FBgn0037225
285	XLOC_006812	ImpL1	3L:13375864-13377099	Ctr	Mut	74.6957	225.777	1.5958	5.00E-05	0.00245	FBgn0001256
286	XLOC_002908	CG43366	2R:1781323-1818996	Ctr	Mut	5.53372	16.7677	1.5994	5.00E-05	0.00245	FBgn0263109
287	XLOC_009203	CG8420	3R:5072370-5077153	Ctr	Mut	33.3051	102.882	1.6272	5.00E-05	0.00245	FBgn0037664

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
288	XLOC_004358	spz6	2R:20644902-20647989	Ctr	Mut	16.3428	50.6848	1.6329	5.00E-05	0.00245	FBgn0035056
289	XLOC_009911	CG15695	3R:16667863-16670709	Ctr	Mut	2.88911	8.98311	1.6366	5.00E-05	0.00245	FBgn0038832
290	XLOC_003095	PGRP-SC1a	2R:4597267-4597825	Ctr	Mut	30.8452	95.9132	1.6367	5.00E-05	0.00245	FBgn0043576
291	XLOC_005557	CG43666	2R:16675821-16678640	Ctr	Mut	81.5211	253.514	1.6368	5.00E-05	0.00245	FBgn0263741
292	XLOC_005557	CG43667	2R:16675821-16678640	Ctr	Mut	81.5211	253.514	1.6368	5.00E-05	0.00245	FBgn0263760
293	XLOC_005557	CG43677	2R:16675821-16678640	Ctr	Mut	81.5211	253.514	1.6368	5.00E-05	0.00245	FBgn0263742
294	XLOC_010562	Spn100A	3R:26516365-26518490	Ctr	Mut	30.611	95.2712	1.638	5.00E-05	0.00245	FBgn0039795
295	XLOC_011426	Cad89D	3R:12310287-12321141	Ctr	Mut	0.578649	1.80098	1.638	0.0006	0.01944	FBgn0038439
296	XLOC_009817	CG6026	3R:14859255-14870350	Ctr	Mut	8.28078	26.0506	1.6535	5.00E-05	0.00245	FBgn0038676
297	XLOC_006201	CG12017	3L:3280238-3290896	Ctr	Mut	8.21501	26.1564	1.6708	5.00E-05	0.00245	FBgn0035429
298	XLOC_014753	CG15043	X:18337059-18399457	Ctr	Mut	243.575	780.028	1.6792	0.0003	0.01084	FBgn0030929
299	XLOC_014746	CG32548	X:18329747-18334614	Ctr	Mut	27.994	90.2723	1.6892	0.00035	0.01228	FBgn0052548
300	XLOC_009967	burs	3R:17584301-17596047	Ctr	Mut	4.28532	13.8578	1.6932	5.00E-05	0.00245	FBgn0038901
301	XLOC_009967	CG42335	3R:17584301-17596047	Ctr	Mut	4.28532	13.8578	1.6932	5.00E-05	0.00245	FBgn0259237
302	XLOC_007850	Cpr65Eb	3L:7080812-7081887	Ctr	Mut	29.6654	96.0684	1.6953	0.0005	0.01666	FBgn0035736
303	XLOC_006483	Cpr66Cb	3L:8329771-8331695	Ctr	Mut	4.82177	15.6745	1.7008	0.0002	0.00793	FBgn0035875
304	XLOC_005830	CG4324	2R:19902432-19904753	Ctr	Mut	3.14831	10.2921	1.7089	5.00E-05	0.00245	FBgn0034956
305	XLOC_009998	CG7080	3R:18155022-18156268	Ctr	Mut	13.0949	42.8661	1.7108	0.0003	0.01084	FBgn0038941
306	XLOC_007952	CG32354	3L:8404950-8415305	Ctr	Mut	2.07254	6.79244	1.7125	0.0005	0.01666	FBgn0052354
307	XLOC_006271	CG1273	3L:4263718-4277778	Ctr	Mut	3.48678	11.5591	1.7291	5.00E-05	0.00245	FBgn0035522
308	XLOC_010669	CG2016	3R:782720-787072	Ctr	Mut	19.9646	66.5106	1.7361	5.00E-05	0.00245	FBgn0250839
309	XLOC_013272	m	X:11648305-11663541	Ctr	Mut	2.85741	9.52379	1.7368	5.00E-05	0.00245	FBgn0002577
310	XLOC_004242	St1	2R:19569146-19570475	Ctr	Mut	6.45377	21.656	1.7466	0.0001	0.00448	FBgn0034887
311	XLOC_005554	CG43709	2R:16663398-16663907	Ctr	Mut	2.78933	9.38669	1.7507	0.0004	0.01385	FBgn0263848
312	XLOC_004014	CG15225	2R:16680471-16681191	Ctr	Mut	165.374	556.67	1.7511	5.00E-05	0.00245	FBgn0034551
313	XLOC_002027	Bace	2L:8491867-8506845	Ctr	Mut	28.6797	96.5775	1.7517	5.00E-05	0.00245	FBgn0032049
314	XLOC_000301	CG3355	2L:4651402-4652892	Ctr	Mut	135.625	456.924	1.7523	5.00E-05	0.00245	FBgn0031619
315	XLOC_014367	dy	X:11663801-11672800	Ctr	Mut	2.12608	7.20186	1.7602	5.00E-05	0.00245	FBgn0004511
316	XLOC_006481	CG34461	3L:8318958-8321407	Ctr	Mut	4.80596	16.3028	1.7622	5.00E-05	0.00245	FBgn0250833

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
317	XLOC_007855	CG14829	3L:7229899-7231358	Ctr	Mut	0.459196	1.56882	1.7725	0.00035	0.01228	FBgn0035751
318	XLOC_012060	TwdlQ	3R:22518355-22519245	Ctr	Mut	18.8223	64.5411	1.7778	0.0001	0.00448	FBgn0039448
319	XLOC_003380	CG13157	2R:8303959-8305647	Ctr	Mut	4.22345	14.7132	1.8006	5.00E-05	0.00245	FBgn0033732
320	XLOC_009009	Osi24	3R:2022993-2033403	Ctr	Mut	3.85093	13.5343	1.8133	5.00E-05	0.00245	FBgn0037409
321	XLOC_004009	CG13438	2R:16580374-16581114	Ctr	Mut	33.6896	118.484	1.8143	5.00E-05	0.00245	FBgn0034545
322	XLOC_014327	C901	X:10858151-10861167	Ctr	Mut	1.42079	5.00172	1.8157	5.00E-05	0.00245	FBgn0021742
323	XLOC_002170	Lip1	2L:10699407-10701458	Ctr	Mut	6.32754	22.2789	1.816	0.00025	0.00942	FBgn0023496
324	XLOC_003282	betaTry	2R:7233839-7234641	Ctr	Mut	1365.42	4828.03	1.8221	5.00E-05	0.00245	FBgn0010357
325	XLOC_009024	Osi17	3R:2137491-2153129	Ctr	Mut	3.12688	11.1302	1.8317	5.00E-05	0.00245	FBgn0037427
326	XLOC_010598	nyo	3R:27368753-27391887	Ctr	Mut	6.87153	24.5262	1.8356	5.00E-05	0.00245	FBgn0039852
327	XLOC_004012	CG34202	2R:16640109-16641244	Ctr	Mut	135.503	484.403	1.8379	5.00E-05	0.00245	FBgn0085231
328	XLOC_007387	CG13239	3L:22853886-22854590	Ctr	Mut	2.48204	8.95219	1.8507	0.0002	0.00793	FBgn0037197
329	XLOC_003484	CG6347	2R:9723925-9726869	Ctr	Mut	9.53212	34.4286	1.8527	5.00E-05	0.00245	FBgn0033874
330	XLOC_009264	Spn85F	3R:5822884-5827137	Ctr	Mut	0.889567	3.22888	1.8599	0.0011	0.03093	FBgn0037772
331	XLOC_000505	Uro	2L:7780084-7781415	Ctr	Mut	23.7836	86.5583	1.8637	5.00E-05	0.00245	FBgn0003961
332	XLOC_008664	CG7298	3L:20195169-20196787	Ctr	Mut	430.129	1565.75	1.864	5.00E-05	0.00245	FBgn0036948
333	XLOC_011909	CG42811	3R:20192248-20192942	Ctr	Mut	2.16645	7.88943	1.8646	5.00E-05	0.00245	FBgn0261993
334	XLOC_008060	ect	3L:10183374-10194937	Ctr	Mut	70.8772	261.366	1.8827	5.00E-05	0.00245	FBgn0000451
335	XLOC_014922	CG10918	X:21293169-21328761	Ctr	Mut	60.1141	221.707	1.8829	5.00E-05	0.00245	FBgn0031178
336	XLOC_004638	CG8213	2R:4856714-4873429	Ctr	Mut	8.39915	31.2643	1.8962	5.00E-05	0.00245	FBgn0033359
337	XLOC_011911	CG17782	3R:20196471-20198484	Ctr	Mut	1.92462	7.16462	1.8963	5.00E-05	0.00245	FBgn0039195
338	XLOC_010753	CG10280	3R:1836280-1838934	Ctr	Mut	3.03928	11.3184	1.8969	5.00E-05	0.00245	FBgn0037395
339	XLOC_013414	CG12540	X:14390169-14391275	Ctr	Mut	6.95944	26.0222	1.9027	0.0009	0.02717	FBgn0030570
340	XLOC_011144	CG4702	3R:7947299-7953919	Ctr	Mut	33.957	127.047	1.9036	5.00E-05	0.00245	FBgn0037992
341	XLOC_003809	CG10910	2R:13907753-13909911	Ctr	Mut	302.593	1137.01	1.9098	5.00E-05	0.00245	FBgn0034289
342	XLOC_009013	Osi6	3R:2057257-2062816	Ctr	Mut	406.454	1530.01	1.9124	5.00E-05	0.00245	FBgn0027527
343	XLOC_008765	CG7173	3L:21548472-21551326	Ctr	Mut	3.30272	12.4929	1.9194	5.00E-05	0.00245	FBgn0037099
344	XLOC_009021	Osi15	3R:2127340-2130291	Ctr	Mut	181.462	693.751	1.9348	5.00E-05	0.00245	FBgn0037424
345	XLOC_011216	CG15887	3R:9113370-9114532	Ctr	Mut	24.7212	94.5457	1.9353	5.00E-05	0.00245	FBgn0038132

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
346	XLOC_006853	CG43120	3L:14375012-14376399	Ctr	Mut	0.847401	3.24286	1.9362	5.00E-05	0.00245	FBgn0262580
347	XLOC_011653	CG17267	3R:16561341-16564567	Ctr	Mut	16.2011	62.2299	1.9415	5.00E-05	0.00245	FBgn0038821
348	XLOC_003687	CG5550	2R:12704095-12706745	Ctr	Mut	44.0338	169.153	1.9416	5.00E-05	0.00245	FBgn0034160
349	XLOC_001191	amd	2L:19110158-19114110	Ctr	Mut	8.38565	32.2991	1.9455	5.00E-05	0.00245	FBgn0000075
350	XLOC_014701	CG8568	X:17343017-17346576	Ctr	Mut	2.46059	9.49891	1.9488	0.0001	0.00448	FBgn0030841
351	XLOC_007808	CG13297	3L:6149691-6150477	Ctr	Mut	16.8005	65.3755	1.9602	5.00E-05	0.00245	FBgn0035685
352	XLOC_008529	CG16775	3L:17879047-17879745	Ctr	Mut	17.4425	68.5032	1.9736	5.00E-05	0.00245	FBgn0036767
353	XLOC_009463	Hsc70-2	3R:8870480-8873112	Ctr	Mut	0.467092	1.83548	1.9744	5.00E-05	0.00245	FBgn0001217
354	XLOC_003445	CG17047	2R:9187106-9189045	Ctr	Mut	7.41122	29.2453	1.9804	5.00E-05	0.00245	FBgn0033827
355	XLOC_009027	Osi20	3R:2165818-2167146	Ctr	Mut	121.812	481.068	1.9816	5.00E-05	0.00245	FBgn0037430
356	XLOC_012877	CG15239	X:3778204-3784520	Ctr	Mut	24.8507	98.1453	1.9816	5.00E-05	0.00245	FBgn0029681
357	XLOC_009018	Osi12	3R:2104252-2108217	Ctr	Mut	1.18754	4.70062	1.9849	5.00E-05	0.00245	FBgn0037419
358	XLOC_011628	CG4362	3R:16131937-16133258	Ctr	Mut	198.709	789.311	1.9899	5.00E-05	0.00245	FBgn0038784
359	XLOC_014437	CG12723	X:13083675-13085915	Ctr	Mut	15.6144	62.1549	1.993	0.0001	0.00448	FBgn0030459
360	XLOC_009017	Osi10	3R:2089437-2093262	Ctr	Mut	0.639865	2.55673	1.9985	0.0007	0.02195	FBgn0037417
361	XLOC_006784	CG17672	3L:13022061-13029604	Ctr	Mut	19.9941	80.3402	2.0066	5.00E-05	0.00245	FBgn0083978
362	XLOC_006205	dro2	3L:3314348-3314681	Ctr	Mut	48.5817	196.783	2.0181	5.00E-05	0.00245	FBgn0052279
363	XLOC_010208	CG11786	3R:20872046-20873131	Ctr	Mut	25.1621	102.008	2.0194	5.00E-05	0.00245	FBgn0039264
364	XLOC_009016	Osi9	3R:2085934-2087832	Ctr	Mut	89.1747	365.955	2.037	5.00E-05	0.00245	FBgn0037416
365	XLOC_004013	CG13443	2R:16661193-16662869	Ctr	Mut	8.61873	35.3796	2.0374	5.00E-05	0.00245	FBgn0034548
366	XLOC_009014	Osi7	3R:2074713-2077405	Ctr	Mut	121.334	498.155	2.0376	5.00E-05	0.00245	FBgn0037414
367	XLOC_003516	CG34444	2R:10256684-10258621	Ctr	Mut	0.985987	4.07991	2.0489	0.0009	0.02717	FBgn0085473
368	XLOC_004011	CG34201	2R:16638451-16639628	Ctr	Mut	30.3836	126.565	2.0585	5.00E-05	0.00245	FBgn0085230
369	XLOC_009026	Osi19	3R:2160998-2162919	Ctr	Mut	90.6602	385.629	2.0887	5.00E-05	0.00245	FBgn0037429
370	XLOC_006486	ImpE1	3L:8365504-8384271	Ctr	Mut	7.67726	32.9869	2.1032	5.00E-05	0.00245	FBgn0001253
371	XLOC_009025	Osi18	3R:2155971-2157536	Ctr	Mut	63.4048	273.256	2.1076	5.00E-05	0.00245	FBgn0037428
372	XLOC_013138	CG15370	X:9119488-9120601	Ctr	Mut	3.2796	14.2283	2.1172	5.00E-05	0.00245	FBgn0030107
373	XLOC_011840	CG33337	3R:19415666-19418502	Ctr	Mut	1.21027	5.41907	2.1627	5.00E-05	0.00245	FBgn0053337
374	XLOC_003096	PGRP-SC1b	2R:4600948-4601587	Ctr	Mut	19.3875	86.8636	2.1636	5.00E-05	0.00245	FBgn0033327

	gene_id	gene	locus	sample_1	sample_2	value_1	value_2	log2(fold_change)	p_value	q_value	oID_genes
375	XLOC_004842	Damm	2R:7751665-7753908	Ctr	Mut	4.37792	20.3545	2.217	5.00E-05	0.00245	FBgn0033659
376	XLOC_010088	CG31148	3R:19526359-19528324	Ctr	Mut	7.76002	36.6681	2.2404	5.00E-05	0.00245	FBgn0051148
377	XLOC_004988	CG6337	2R:9713879-9722593	Ctr	Mut	23.6537	113.218	2.259	0.00025	0.00942	FBgn0033873
378	XLOC_010280	MCO3	3R:21723197-21725466	Ctr	Mut	1.44111	7.14989	2.3107	5.00E-05	0.00245	FBgn0039387
379	XLOC_011912	CG17780	3R:20199175-20210770	Ctr	Mut	5.11759	25.3919	2.3108	5.00E-05	0.00245	FBgn0039197
380	XLOC_011912	CG17781	3R:20199175-20210770	Ctr	Mut	5.11759	25.3919	2.3108	5.00E-05	0.00245	FBgn0039196
381	XLOC_010154	CG13616	3R:20218631-20219502	Ctr	Mut	5.22073	26.6516	2.3519	5.00E-05	0.00245	FBgn0039200
382	XLOC_009373	CG14720	3R:7537500-7538076	Ctr	Mut	1.24108	6.47883	2.3841	0.0017	0.04354	FBgn0037940
383	XLOC_009022	Osi16	3R:2130867-2132036	Ctr	Mut	1.09495	5.74398	2.3912	5.00E-05	0.00245	FBgn0051561
384	XLOC_007762	CG10591	3L:5591656-5592792	Ctr	Mut	0.232809	1.23748	2.4102	0.0007	0.02195	FBgn0035621
385	XLOC_010183	CG42331	3R:20597902-20610284	Ctr	Mut	3.8151	20.7072	2.4403	5.00E-05	0.00245	FBgn0259233
386	XLOC_005578	Cht4	2R:16952884-16954592	Ctr	Mut	27.0575	147.77	2.4493	5.00E-05	0.00245	FBgn0022700
387	XLOC_009915	TotC	3R:16698709-16699310	Ctr	Mut	1.93065	10.5465	2.4496	5.00E-05	0.00245	FBgn0044812
388	XLOC_000087	CG42329	2L:1219317-1229802	Ctr	Mut	0.959453	5.37126	2.485	5.00E-05	0.00245	FBgn0259229
389	XLOC_009691	CG31268	3R:12857538-12859237	Ctr	Mut	11.5768	67.3905	2.5413	5.00E-05	0.00245	FBgn0051268
390	XLOC_011752	CG13857	3R:18215501-18219826	Ctr	Mut	2.15513	13.4446	2.6412	0.0006	0.01944	FBgn0038958
391	XLOC_008168	CG42397	3L:11966184-11966890	Ctr	Mut	3.15012	19.9335	2.6617	5.00E-05	0.00245	FBgn0259748
392	XLOC_000287	CG33003	2L:4400952-4403142	Ctr	Mut	3.38111	21.4192	2.6633	5.00E-05	0.00245	FBgn0053003
393	XLOC_009842	CG17751	3R:15432975-15435044	Ctr	Mut	5.46427	36.7634	2.7502	5.00E-05	0.00245	FBgn0038717
394	XLOC_006653	CG12522	3L:11101334-11101832	Ctr	Mut	70.7546	477.189	2.7537	5.00E-05	0.00245	FBgn0036131
395	XLOC_008037	CG6749	3L:9724337-9726907	Ctr	Mut	0.637892	4.33214	2.7637	0.0001	0.00448	FBgn0263541
396	XLOC_009223	CG8147	3R:5320033-5322675	Ctr	Mut	0.810923	5.54246	2.7729	5.00E-05	0.00245	FBgn0043791
397	XLOC_003686	CG15919	2R:12680161-12680585	Ctr	Mut	5.59236	42.4773	2.9252	5.00E-05	0.00245	FBgn0040743
398	XLOC_008095	CG32071	3L:11096520-11096973	Ctr	Mut	18.5365	144.989	2.9675	5.00E-05	0.00245	FBgn0052071
399	XLOC_001815	CG9021	2L:5903358-5904674	Ctr	Mut	0.262034	2.14408	3.0325	5.00E-05	0.00245	FBgn0031747
400	XLOC_010837	snRNA:7SK	3R:3300274-3300718	Ctr	Mut	1.69003	28.9889	4.1004	5.00E-05	0.00245	FBgn0065099

REFERENCE LIST

Reference List

1 Syvertsen, A., Haughton, V. M., Williams, A. L. & Cusick, J. F. The computed tomographic appearance of the normal pituitary gland and pituitary microadenomas. *Radiology* **133**, 385-391, (1979).

2 Swartz, J. D., Russell, K. B., Basile, B. A., O'Donnell, P. C. & Popky, G. L. High-resolution computed tomographic appearance of the intrasellar contents in women of childbearing age. *Radiology* **147**, 115-117, (1983).

3 Lechan, R. M. Neuroendocrinology of pituitary hormone regulation. *Endocrinology and metabolism clinics of North America* **16**, 475-501, (1987).

4 Rey-Dios, R., Payner, T. D. & Cohen-Gadol, A. A. Pituitary macroadenoma causing symptomatic internal carotid artery compression: surgical treatment through transsphenoidal tumor resection. *Journal of Clinical Neuroscience* **21**, 541-546, (2014).

5 Lechan, R. M. & Toni, R.*Endotext: Functional Anatomy of the Hypothalamus and Pituitary*, <http://www.endotext.org/chapter/functional-anatomy-of-the-hypothalamus-and-pituitary/2/>. Accessed 18.07 2015

6 Levy, M. J., Matharu, M. S., Meeran, K., Powell, M. & Goadsby, P. J. The clinical characteristics of headache in patients with pituitary tumours. *Brain: a Journal of Neurology* **128**, 1921-1930, (2005).

7 Abe, T., Matsumoto, K., Kuwazawa, J., Toyoda, I. & Sasaki, K. Headache associated with pituitary adenomas. *Headache* **38**, 782-786, (1998).

8 Suwanwela, N., Phanthumchinda, K. & Kaoropthum, S. Headache in brain tumor: a cross-sectional study. *Headache* **34**, 435-438, (1994).

9 Forsyth, P. A. & Posner, J. B. Headaches in patients with brain tumors: a study of 111 patients. *Neurology* **43**, 1678-1683, (1993).

10 Levy, M. J., Jager, H. R., Powell, M., Matharu, M. S., Meeran, K. & Goadsby, P. J. Pituitary volume and headache: size is not everything. *Archives of Neurology* **61**, 721-725, (2004).

11 Heaney, A. P. & Melmed, S. Molecular targets in pituitary tumours. *Nature Reviews. Cancer* **4**, 285-295, (2004).

12 Melmed, S. & Kleinberg, D. in *Williams Textbook of Endocrinology* eds H.M. Kronenberg, S. Melmed, K.S. Polonsky, & P.R. Larsen Ch. 8, 155-263 (2008).

13 Asa, S. L. & Kovacs, K. Functional morphology of the human fetal pituitary. *Pathology Annual* **19 Pt 1**, 275-315, (1984).

14 Scully, K. M. & Rosenfeld, M. G. Pituitary development: regulatory codes in mammalian organogenesis. *Science* **295**, 2231-2235, (2002).

15 Prince, K. L., Walvoord, E. C. & Rhodes, S. J. The role of homeodomain transcription factors in heritable pituitary disease. *Nature Reviews. Endocrinology* **7**, 727-737, (2011).

16 Takuma, N., Sheng, H. Z., Furuta, Y., Ward, J. M., Sharma, K., Hogan, B. L., Pfaff, S. L., Westphal, H., Kimura, S. & Mahon, K. A. Formation of Rathke's pouch requires dual induction from the diencephalon. *Development* **125**, 4835-4840, (1998).

17 Sheng, H. Z., Zhadanov, A. B., Mosinger, B., Jr., Fujii, T., Bertuzzi, S., Grinberg, A., Lee, E. J., Huang, S. P., Mahon, K. A. & Westphal, H. Specification of pituitary cell lineages by the LIM homeobox gene Lhx3. *Science* **272**, 1004-1007, (1996).

18 Alatzoglou, K. S., Kelberman, D. & Dattani, M. T. The role of SOX proteins in normal pituitary development. *The Journal of Endocrinology* **200**, 245-258, (2009).

19 Semina, E. V., Datson, N. A., Leysens, N. J., Zabel, B. U., Carey, J. C., Bell, G. I., Bitoun, P., Lindgren, C., Stevenson, T., Frants, R. R., van Ommen, G. & Murray, J. C. Exclusion of epidermal growth factor and high-resolution physical mapping across the Rieger syndrome locus. *American Journal of Human Genetics* **59**, 1288-1296, (1996).

20 Lanctot, C., Gauthier, Y. & Drouin, J. Pituitary homeobox 1 (Ptx1) is differentially expressed during pituitary development. *Endocrinology* **140**, 1416-1422, (1999).

21 Cohen, L. E. & Radovick, S. Molecular basis of combined pituitary hormone deficiencies. *Endocrine Reviews* **23**, 431-442, (2002).

22 Muscatelli, F., Strom, T. M., Walker, A. P., Zanaria, E., Recan, D., Meindl, A., Bardoni, B., Guioli, S., Zehetner, G., Rabl, W. & et al. Mutations in the DAX-1 gene give rise to both X-linked adrenal hypoplasia congenita and hypogonadotropic hypogonadism. *Nature* **372**, 672-676, (1994).

23 Lamolet, B., Pulichino, A. M., Lamonerie, T., Gauthier, Y., Brue, T., Enjalbert, A. & Drouin, J. A pituitary cell-restricted T box factor, Tpit, activates POMC transcription in cooperation with Pitx homeoproteins. *Cell* **104**, 849-859, (2001).

24 Ben-Shlomo, A. & Melmed, S. in *The Pituitary* ed Melmed S 21-45 (2011).

25 Bilezikjian, L. M., Blount, A. L., Leal, A. M., Donaldson, C. J., Fischer, W. H. & Vale, W. W. Autocrine/paracrine regulation of pituitary function by activin, inhibin and follistatin. *Molecular and Cellular Endocrinology* **225**, 29-36, (2004).

26 Mechenthaler, I. Galanin and the neuroendocrine axes. *Cellular and Molecular Life Sciences* **65**, 1826-1835, (2008).

27 Alexander, J. M., Biller, B. M., Bikkal, H., Zervas, N. T., Arnold, A. & Klibanski, A. Clinically nonfunctioning pituitary tumors are monoclonal in origin. *The Journal of Clinical Investigation* **86**, 336-340, (1990).

28 Meij, B. P., Lopes, M. B., Ellegala, D. B., Alden, T. D. & Laws, E. R., Jr. The long-term significance of microscopic dural invasion in 354 patients with pituitary adenomas treated with transsphenoidal surgery. *Journal of Neurosurgery* **96**, 195-208, (2002).

29 Pei, L., Melmed, S., Scheithauer, B., Kovacs, K. & Prager, D. H-ras mutations in human pituitary carcinoma metastases. *The Journal of Clinical Endocrinology & Metabolism* **78**, 842-846, (1994).

30 Scheithauer, B. W., Gaffey, T. A., Lloyd, R. V., Sebo, T. J., Kovacs, K. T., Horvath, E., Yapicier, O., Young, W. F., Jr., Meyer, F. B., Kuroki, T., Riehle, D. L. & Laws, E. R., Jr. Pathobiology of pituitary adenomas and carcinomas. *Neurosurgery* **59**, 341-353, (2006).

31 Costello, R. T. Subclinical Adenoma of the Pituitary Gland. *American Journal of Pathology* **12**, 205-216, (1936).

32 Ezzat, S., Asa, S. L., Couldwell, W. T., Barr, C. E., Dodge, W. E., Vance, M. L. & McCutcheon, I. E. The prevalence of pituitary adenomas: a systematic review. *Cancer* **101**, 613-619, (2004).

33 Daly, A. F., Rixhon, M., Adam, C., Dempegioti, A., Tichomirowa, M. A. & Beckers, A. High prevalence of pituitary adenomas: a cross-sectional study in the province of Liege, Belgium. *The Journal of Clinical Endocrinology and Metabolism* **91**, 4769-4775, (2006).

34 Fontana, E. & Gaillard, R. Epidemiology of pituitary adenoma: results of the first Swiss study. *Revue médicale suisse* **5**, 2172-2174, (2009).

35 Fernandez, A., Karavitaki, N. & Wass, J. A. Prevalence of pituitary adenomas: a community-based, cross-sectional study in Banbury (Oxfordshire, UK). *Clinical Endocrinology* **72**, 377-382, (2010).

36 Raappana, A., Koivukangas, J., Ebeling, T. & Pirila, T. Incidence of pituitary adenomas in Northern Finland in 1992-2007. *The Journal of Clinical Endocrinology & Metabolism* **95**, 4268-4275, (2010).

37 Gruppetta, M., Mercieca, C. & Vassallo, J. Prevalence and incidence of pituitary adenomas: a population based study in Malta. *Pituitary*, (2012).

38 Clayton, R. N. Sporadic pituitary tumours: from epidemiology to use of databases. *Best Practice & Research: Clinical Endocrinology & Metabolism* **13**, 451-460, (1999).

39 Agustsson, T. T., Baldvinsdottir, T., Jonasson, J. G., Olafsdottir, E., Steinthorsdottir, V., Sigurdsson, G., Thorsson, A. V., Carroll, P. V., Korbonits, M. & Benediktsson, R. The epidemiology of pituitary adenomas in Iceland, 1955-2012: a nationwide population-based study. *European Journal of Endocrinology* **173**, 655-664, (2015).

40 Freda, P. U., Beckers, A. M., Katznelson, L., Molitch, M. E., Montori, V. M., Post, K. D. & Vance, M. L. Pituitary incidentaloma: an endocrine society clinical practice guideline. *The Journal of Clinical Endocrinology and Metabolism* **96**, 894-904, (2011).

41 Aflorei, E. D. & Korbonits, M. Epidemiology and etiopathogenesis of pituitary adenomas. *Journal of Neuro-Oncology* **117**, 379-394, (2014).

42 Potter, V. R. Initiation and promotion in cancer formation: the importance of studies on intercellular communication. *Yale Journal of Biology and Medicine* **53**, 367-384, (1980).

43 Asa, S. L. & Ezzat, S. The pathogenesis of pituitary tumors. *Annual Review of Pathology* **4**, 97-126, (2009).

44 Clayton, R. N., Pfeifer, M., Atkinson, A. B., Belchetz, P., Wass, J. A., Kyrodimou, E., Vanderpump, M., Simpson, D., Bicknell, J. & Farrell, W. E. Different patterns of allelic loss (loss of heterozygosity) in recurrent human pituitary tumors provide evidence for multiclonal origins. *Clinical cancer research* **6**, 3973-3982, (2000).

45 Asa, S. L., Kovacs, K., Stefaneanu, L., Horvath, E., Billestrup, N., Gonzalez-Manchon, C. & Vale, W. Pituitary adenomas in mice transgenic for growth hormone-releasing hormone. *Endocrinology* **131**, 2083-2089, (1992).

46 Mayo, K. E., Hammer, R. E., Swanson, L. W., Brinster, R. L., Rosenfeld, M. G. & Evans, R. M. Dramatic pituitary hyperplasia in transgenic mice expressing a human growth hormone-releasing factor gene. *Molecular Endocrinology* **2**, 606-612, (1988).

47 Sano, T., Asa, S. L. & Kovacs, K. Growth hormone-releasing hormone-producing tumors: clinical, biochemical, and morphological manifestations. *Endocrine Reviews* **9**, 357-373, (1988).

48 Thorner, M. O., Perryman, R. L., Cronin, M. J., Rogol, A. D., Draznin, M., Johanson, A., Vale, W., Horvath, E. & Kovacs, K. Somatotroph hyperplasia. Successful treatment of acromegaly by removal of a pancreatic islet tumor secreting a growth hormone-releasing factor. *The Journal of Clinical Investigation* **70**, 965-977, (1982).

49 Sugiyama, K., Kimura, M., Abe, T., Ikezawa, Y., Manaka, H., Yamatani, K., Tominaga, M., Sasaki, H. & Misawa, T. Hyper-adrenocorticotropinemia in a patient with Addison's disease after treatment with corticosteroids. *Internal Medicine.* **35**, 555-559, (1996).

50 Hannaford, P. C., Selvaraj, S., Elliott, A. M., Angus, V., Iversen, L. & Lee, A. J. Cancer risk among users of oral contraceptives: cohort data from the Royal College of General Practitioner's oral contraception study. *BMJ.* **335**, 651, (2007).

51 Gorbman, A. Functional and structural changes consequent to high dosages of radioactive iodine. *The Journal of Clinical Endocrinology & Metabolism* **10**, 1177-1191, (1950).

52 Ueno, M., Inano, H., Onoda, M., Murase, H., Ikota, N., Kagiya, T. V. & Anzai, K. Modification of mortality and tumorigenesis by tocopherol-mono-glucoside (TMG) administered after X irradiation in mice and rats. *Radiation Research* **172**, 519-524, (2009).

53 Johnson, K. A., Gorzinski, S. J., Bodner, K. M., Campbell, R. A., Wolf, C. H., Friedman, M. A. & Mast, R. W. Chronic toxicity and oncogenicity study on acrylamide incorporated in the drinking water of Fischer 344 rats. *Toxicology and Applied Pharmacology* **85**, 154-168, (1986).

54 Svensson, K., Abramsson, L., Becker, W., Glynn, A., Hellenas, K. E., Lind, Y. & Rosen, J. Dietary intake of acrylamide in Sweden. *Food and Chemical Toxicology - Journal* **41**, 1581-1586, (2003).

55 Dirtu, A. C., Niessen, S. J., Jorens, P. G. & Covaci, A. Organohalogenated contaminants in domestic cats' plasma in relation to spontaneous acromegaly and type 2 diabetes mellitus: a clue for endocrine disruption in humans? *Environment International* **57-58**, 60-67, (2013).

56 Cannavo, S., Ferrau, F., Ragonese, M., Curto, L., Torre, M. L., Magistri, M., Marchese, A., Alibrandi, A. & Trimarchi, F. Increased prevalence of acromegaly in a highly polluted area. *European Journal of Endocrinology* **163**, 509-513, (2010).

57 Pesatori, A. C., Baccarelli, A., Consonni, D., Lania, A., Beck-Peccoz, P., Bertazzi, P. A. & Spada, A. Aryl hydrocarbon receptor-interacting protein and pituitary adenomas: a population-based study on subjects exposed to dioxin after the Seveso, Italy, accident. *European Journal of Endocrinology* **159**, 699-703, (2008).

58 Couldwell, W. T. & Cannon-Albright, L. A heritable predisposition to pituitary tumors. *Pituitary* **13**, 130-137, (2010).

59 Zhu, X., Lin, C. R., Prefontaine, G. G., Tollkuhn, J. & Rosenfeld, M. G. Genetic control of pituitary development and hypopituitarism. *Current opinion in genetics & development* **15**, 332-340, (2005).

60 Trivellin, G., Daly, A. F., Faucz, F. R., Yuan, B., Rostomyan, L., Larco, D. O., Schernthaner-Reiter, M. H., Szarek, E., Leal, L. F., Caberg, J. H., Castermans, E., Villa, C., Dimopoulos, A., Chittiboina, P., Xekouki, P., Shah, N., Metzger, D., Lysy, P. A., Ferrante, E., Strebkova, N., Mazerkina, N., Zatelli, M. C., Lodish, M., Horvath, A., de Alexandre, R. B., Manning, A. D., Levy, I., Keil, M. F., Sierra Mde, L., Palmeira, L., Coppieters, W., Georges, M., Naves, L. A., Jamar, M., Bours, V., Wu, T. J., Choong, C. S., Bertherat, J., Chanson, P., Kamenicky,

P., Farrell, W. E., Barlier, A., Quezado, M., Bjelobaba, I., Stojilkovic, S. S., Wess, J., Costanzi, S., Liu, P., Lupski, J. R., Beckers, A. & Stratakis, C. A. Gigantism and acromegaly due to Xq26 microduplications and GPR101 mutation. *The New England Journal of Medicine* **371**, 2363-2374, (2014).

61 McAndrew, J., Paterson, A. J., Asa, S. L., McCarthy, K. J. & Kudlow, J. E. Targeting of transforming growth factor-alpha expression to pituitary lactotrophs in transgenic mice results in selective lactotroph proliferation and adenomas. *Endocrinology* **136**, 4479-4488, (1995).

62 Ma, Z. Y., Song, Z. J., Chen, J. H., Wang, Y. F., Li, S. Q., Zhou, L. F., Mao, Y., Li, Y. M., Hu, R. G., Zhang, Z. Y., Ye, H. Y., Shen, M., Shou, X. F., Li, Z. Q., Peng, H., Wang, Q. Z., Zhou, D. Z., Qin, X. L., Ji, J., Zheng, J., Chen, H., Wang, Y., Geng, D. Y., Tang, W. J., Fu, C. W., Shi, Z. F., Zhang, Y. C., Ye, Z., He, W. Q., Zhang, Q. L., Tang, Q. S., Xie, R., Shen, J. W., Wen, Z. J., Zhou, J., Wang, T., Huang, S., Qiu, H. J., Qiao, N. D., Zhang, Y., Pan, L., Bao, W. M., Liu, Y. C., Huang, C. X., Shi, Y. Y. & Zhao, Y. Recurrent gain-of-function USP8 mutations in Cushing's disease. *Cell Res* **25**, 306-317, (2015).

63 Reincke, M., Sbiera, S., Hayakawa, A., Theodoropoulou, M., Osswald, A., Beuschlein, F., Meitinger, T., Mizuno-Yamasaki, E., Kawaguchi, K., Saeki, Y., Tanaka, K., Wieland, T., Graf, E., Saeger, W., Ronchi, C. L., Allolio, B., Buchfelder, M., Strom, T. M., Fassnacht, M. & Komada, M. Mutations in the deubiquitinase gene USP8 cause Cushing's disease. *Nature Genetics* **47**, 31-38, (2015).

64 Pellegrini, I., Barlier, A., Gunz, G., Figarella-Branger, D., Enjalbert, A., Grisoli, F. & Jaquet, P. Pit-1 gene expression in the human pituitary and pituitary adenomas. *The Journal of Clinical Endocrinology & Metabolism* **79**, 189-196, (1994).

65 Riss, D., Jin, L., Qian, X., Bayliss, J., Scheithauer, B. W., Young, W. F., Jr., Vidal, S., Kovacs, K., Raz, A. & Lloyd, R. V. Differential expression of galectin-3 in pituitary tumors. *Cancer Research* **63**, 2251-2255, (2003).

66 Galland, F., Lacroix, L., Saulnier, P., Dessen, P., Meduri, G., Bernier, M., Gaillard, S., Guibourdenche, J., Fournier, T., Evain-Brion, D., Bidart, J. M. & Chanson, P. Differential gene expression profiles of invasive and non-invasive non-functioning pituitary adenomas based on microarray analysis. *Endocrine-Related Cancer* **17**, 361-371, (2010).

67 Takino, H., Herman, V., Weiss, M. & Melmed, S. Purine-binding factor (nm23) gene expression in pituitary tumors: marker of adenoma invasiveness. *The Journal of Clinical Endocrinology & Metabolism* **80**, 1733-1738, (1995).

68 Gadelha, M. R., Trivellin, G., Hernández-Ramírez, L. C. & Korbonits, M. Genetics of pituitary adenomas. *Frontiers of hormone research* **41**, 111-140, (2013).

69 Spada, A. & Vallar, L. G-protein oncogenes in acromegaly. *Hormone Research* **38**, 90-93, (1992).

70 Levy, A. & Lightman, S. Molecular defects in the pathogenesis of pituitary tumours. *Frontiers in Neuroendocrinology* **24**, 94-127, (2003).

71 Tada, M., Kobayashi, H. & Moriuchi, T. Molecular basis of pituitary oncogenesis. *Journal of Neuro-Oncology* **45**, 83-96, (1999).

72 Cai, W. Y., Alexander, J. M., Hedley-Whyte, E. T., Scheithauer, B. W., Jameson, J. L., Zervas, N. T. & Klibanski, A. ras mutations in human prolactinomas and pituitary carcinomas. *The Journal of Clinical Endocrinology & Metabolism* **78**, 89-93, (1994).

73 Karga, H. J., Alexander, J. M., Hedley-Whyte, E. T., Klibanski, A. & Jameson, J. L. Ras mutations in human pituitary tumors. *The Journal of Clinical Endocrinology & Metabolism* **74**, 914-919, (1992).

74 Knudson, A. G., Jr. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 820-823, (1971).

75 Luzi, E., Marini, F., Giusti, F., Galli, G., Cavalli, L. & Brandi, M. L. The negative feedback-loop between the oncomir Mir-24-1 and menin modulates the Men1 tumorigenesis by mimicking the "Knudson's second hit". *PLoS One* **7**, e39767, (2012).

76 Woloschak, M., Yu, A., Xiao, J. & Post, K. D. Abundance and state of phosphorylation of the retinoblastoma gene product in human pituitary tumors. *International Journal of Cancer* **67**, 16-19, (1996).

77 Wierinckx, A., Roche, M., Raverot, G., Legras-Lachuer, C., Croze, S., Nazaret, N., Rey, C., Auger, C., Jouanneau, E., Chanson, P., Trouillas, J. & Lachuer, J. Integrated genomic profiling identifies loss of chromosome 11p impacting transcriptomic activity in aggressive pituitary PRL tumors. *Brain Pathology* **21**, 533-543, (2011).

78 Jacks, T., Fazeli, A., Schmitt, E. M., Bronson, R. T., Goodell, M. A. & Weinberg, R. A. Effects of an Rb mutation in the mouse. *Nature* **359**, 295-300, (1992).

79 Kiyokawa, H., Kineman, R. D., Manova-Todorova, K. O., Soares, V. C., Hoffman, E. S., Ono, M., Khanam, D., Hayday, A. C., Frohman, L. A. & Koff, A. Enhanced growth of mice lacking the cyclin-dependent kinase inhibitor function of p27(Kip1). *Cell* **85**, 721-732, (1996).

80 Moreno, C. S., Evans, C. O., Zhan, X., Okor, M., Desiderio, D. M. & Oyesiku, N. M. Novel molecular signaling and classification of human clinically nonfunctional pituitary adenomas identified by gene expression profiling and proteomic analyses. *Cancer Research* **65**, 10214-10222, (2005).

81 Bazina, M., Vukojevic, K., Roje, D. & Saraga-Babic, M. Influence of growth and transcriptional factors, and signaling molecules on early human pituitary development. *Journal of Molecular Histology* **40**, 277-286, (2009).

82 Ezzat, S., Zheng, L., Zhu, X. F., Wu, G. E. & Asa, S. L. Targeted expression of a human pituitary tumor-derived isoform of FGF receptor-4 recapitulates pituitary tumorigenesis. *The Journal of Clinical Investigation* **109**, 69-78, (2002).

83 Theodoropoulou, M., Arzberger, T., Gruebler, Y., Jaffrain-Rea, M. L., Schlegel, J., Schaaf, L., Petrangeli, E., Losa, M., Stalla, G. K. & Pagotto, U. Expression of epidermal growth factor receptor in neoplastic pituitary cells: evidence for a role in corticotropinoma cells. *Journal of Endocrinology* **183**, 385-394, (2004).

84 Missale, C., Boroni, F., Losa, M., Giovanelli, M., Zanellato, A., Dal, T. R., Balsari, A. & Spano, P. Nerve growth factor suppresses the transforming phenotype of human prolactinomas. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 7961-7965, (1993).

85 Shorts-Cary, L., Xu, M., Ertel, J., Kleinschmidt-Demasters, B. K., Lillehei, K., Matsuoka, I., Nielsen-Preiss, S. & Wierman, M. E. Bone morphogenetic protein and retinoic acid-inducible neural specific protein-3 is expressed in gonadotrope cell pituitary adenomas and induces proliferation, migration, and invasion. *Endocrinology* **148**, 967-975, (2007).

86 Sornson, M. W., Wu, W., Dasen, J. S., Flynn, S. E., Norman, D. J., O'Connell, S. M., Gukovsky, I., Carriere, C., Ryan, A. K., Miller, A. P., Zuo, L., Gleiberman, A. S., Andersen, B., Beamer, W. G. & Rosenfeld, M. G. Pituitary lineage determination by the Prophet of Pit-1 homeodomain factor defective in Ames dwarfism. *Nature* **384**, 327-333, (1996).

87 Pellegrini-Bouiller, I., Manrique, C., Gunz, G., Grino, M., Zamora, A. J., Figarella-Branger, D., Grisoli, F., Jaquet, P. & Enjalbert, A. Expression of the members of the Ptx family of transcription factors in human pituitary adenomas. *The Journal of Clinical Endocrinology and Metabolism* **84**, 2212-2220, (1999).

88 Nakamura, S., Ohtsuru, A., Takamura, N., Kitange, G., Tokunaga, Y., Yasunaga, A., Shibata, S. & Yamashita, S. Prop-1 gene expression in human pituitary tumors. *The Journal of Clinical Endocrinology and Metabolism* **84**, 2581-2584, (1999).

89 Leontiou, C. A., Gueorguiev, M., van der Spuy, J., Quinton, R., Lolli, F., Hassan, S., Chahal, H. S., Igreja, S. C., Jordan, S., Rowe, J., Stolbrink, M., Christian, H. C., Wray, J., Bishop-Bailey, D., Berney, D. M., Wass, J. A., Popovic, V., Ribeiro-Oliveira, A., Jr., Gadelha, M. R., Monson, J. P., Akker, S. A., Davis, J. R., Clayton, R. N., Yoshimoto, K., Iwata, T., Matsuno, A., Eguchi, K., Musat, M., Flanagan, D., Peters, G., Bolger, G. B., Chapple, J. P., Frohman, L. A., Grossman, A. B. & Korbonits, M. The role of the aryl hydrocarbon receptor-interacting protein gene in familial and sporadic pituitary adenomas. *Journal of Clinical Endocrinology and Metabolism.* **93**, 2390-2401, (2008).

90 Beckers, A. & Daly, A. F. The clinical, pathological, and genetic features of familial isolated pituitary adenomas. *European Journal of Endocrinology* **157**, 371-382, (2007).

91 Moy, J. & Lawson, D. Temporal effects of estradiol and diethylstilbestrol on pituitary and plasma prolactin levels in ovariectomized Fischer 344 and Holtzman rats: a comparison of radioimmunoassay and Nb2 lymphoma cell bioassay. *Proceedings of the Society for Experimental Biology and Medicine* **200**, 507-513, (1992).

92 Holtzman, S., Stone, J. P. & Shellabarger, C. J. Influence of diethylstilbestrol treatment on prolactin cells of female ACI and Sprague-Dawley rats. *Cancer Research* **39**, 779-784, (1979).

93 Heshmat, M. Y., Kovi, J., Simpson, C., Kennedy, J. & Fan, K. J. Neoplasms of the central nervous system. incidence and population selectivity in the Washington DC, metropolitan area. *Cancer* **38**, 2135-2142, (1976).

94 McDowell, B. D., Wallace, R. B., Carnahan, R. M., Chrischilles, E. A., Lynch, C. F. & Schlechte, J. A. Demographic differences in incidence for pituitary adenoma. *Pituitary* **14**, 23-30, (2011).

95 Goldson, A., Henschke, U., Leffall, L. D. & Schneider, R. L. Is there a genetic basis for the differences in cancer incidence between Afro-Americans and Euro-Americans? *Journal of the National Medical Association* **73**, 701-706, (1981).

96 Dwight, T., Mann, K., Benn, D. E., Robinson, B. G., McKelvie, P., Gill, A. J., Winship, I. & Clifton-Bligh, R. J. Familial SDHA mutation associated with pituitary adenoma and pheochromocytoma/paraganglioma. *The Journal of Clinical Endocrinology & Metabolism* **98**, E1103-1108, (2013).

97 Brahma, A., Heyburn, P. & Swords, F. Familial prolactinoma occuring in association with SDHB mutation positive paraganglioma. *Endocrine Abstracts Spring* **19**, 239-239, (2009).

98 Denes, J., Swords, F. M., Xekouki, P., Kumar, A. V., Maher, E. R., Ferscht, N., Grieve, J., Baldeweg, S. E., Stratakis, C. A. & Korbonits, M. Familial pituitary adenoma and paraganglioma syndrome - A novel type of multiple endocrine neoplasia. *Endocrine Reviews* **33**, OR41-42, (2012).

99 Xekouki, P., Pacak, K., Almeida, M., Wassif, C. A., Rustin, P., Nesterova, M., de la Luz, S. M., Matro, J., Ball, E., Azevedo, M., Horvath, A., Lyssikatos, C., Quezado, M., Patronas, N., Ferrando, B., Pasini, B., Lytras, A., Tolis, G. & Stratakis, C. A. Succinate dehydrogenase

(SDH) D subunit (SDHD) inactivation in a growth-hormone-producing pituitary tumor: a new association for SDH? *The Journal of Clinical Endocrinology & Metabolism* **97**, E357-E366, (2012).

100 Varsavsky, M., Sebastian-Ochoa, A. & Torres, V. E. Coexistence of a pituitary macroadenoma and multicentric paraganglioma: a strange coincidence. *Endocrinología y Nutrición* **60**, 154-156, (2013).

101 Wildi-Runge, S., Bahubeshi, A., Carret, A., Crevier, L. & Robitaille. New phenotype in the familial DICER1 tumor syndrome: pituitary blastoma presenting at age 9 months. *Endocrine Reviews* **32**, P1-777, (2011).

102 Korbonits, M., Storr, H. & Kumar, A. V. Familial pituitary adenomas - Who should be tested for AIP mutations? *Clinical Endocrinology* **77**, 351-356, (2012).

103 Coire, C. I., Smyth, H. S., Rosso, D., Horvath, E. & Kovacs, K. A double pituitary adenoma presenting as a prolactin-secreting tumor with partial response to medical therapy. Case report. *Endocrine Pathology* **21**, 135-138, (2010).

104 Kurozumi, K., Tabuchi, A., Ono, Y., Tamiya, T., Ohmoto, T., Furuta, T. & Hamasaki, S. Pituitary adenoma associated with neurofibromatosis type 1: case report. *Neurological surgery* **30**, 741-745, (2002).

105 Drimmie, F. M., MacLennan, A. C., Nicoll, J. A., Simpson, E., McNeill, E. & Donaldson, M. D. Gigantism due to growth hormone excess in a boy with optic glioma. *Clinical Endocrinology* **53**, 535-538, (2000).

106 Barberis, M., Gambacorta, M., Versari, P. & Filizzolo, F. About a case of Recklinghausen's disease associated with pituitary adenoma. *Pathologica* **71**, 265-272, (1979).

107 Boudin, G., Pepin, B. & Vernant, C. L. Multiple tumours of the nervous system in Recklinghausen's disease. An anatomo-clinical case with chromophobe adenoma of the pituitary gland. *La Presse medicale* **78**, 1427-1430, (1970).

108 Hartemann, P., Schmitt, J., Arnould, G. & Kissel, P. Acromegaly and Recklinghausen's Neuro-Fibromatosis. Apropos of 10 Cases. *Annales d'endocrinologie* **25**, 601-618, (1964).

109 Thakker, R. V. Multiple endocrine neoplasia type 1 (MEN1). *Best Practice & Research: Clinical Endocrinology & Metabolism* **24**, 355-370, (2010).

110 Lips, C. J., Vasen, H. F. & Lamers, C. B. Multiple endocrine neoplasia syndromes. *Critical Reviews in Oncology/Hematology* **2**, 117-184, (1984).

111 Pellegata, N. S., Quintanilla-Martinez, L., Siggelkow, H., Samson, E., Bink, K., Hofler, H., Fend, F., Graw, J. & Atkinson, M. J. Germ-line mutations in p27Kip1 cause a multiple endocrine neoplasia syndrome in rats and humans. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15558-15563, (2006).

112 Agarwal, S. K., Mateo, C. M. & Marx, S. J. Rare germline mutations in cyclin-dependent kinase inhibitor genes in multiple endocrine neoplasia type 1 and related states. *The Journal of Clinical Endocrinology & Metabolism* **94**, 1826-1834, (2009).

113 Adam MP Pagon RA, Bird TD, et al., editors. Carney Complex. in *Gene Reviews* (1993).

114 Daly, A. F., Jaffrain-Rea, M. L., Ciccarelli, A., Valdes-Socin, H., Rohmer, V., Tamburrano, G., Borson-Chazot, C., Estour, B., Ciccarelli, E., Brue, T., Ferolla, P., Emy, P., Colao, A., de, M. E., Lecomte, P., Penfornis, F., Delemer, B., Bertherat, J., Wemeau, J. L., De, H. W., Archambeaud, F., Stevenaert, A., Calender, A., Murat, A., Cavagnini, F. & Beckers, A. Clinical characterization of familial isolated pituitary adenomas. *The Journal of Clinical Endocrinology and Metabolism* **91**, 3316-3323, (2006).

115 Chahal, H. S., Chapple, J. P., Frohman, L. A., Grossman, A. B. & Korbonits, M. Clinical, genetic and molecular characterization of patients with familial isolated pituitary adenomas (FIPA). *Trends in Endocrinology & Metabolism* **21**, 419-427, (2010).

116 Vierimaa, O., Georgitsi, M., Lehtonen, R., Vahteristo, P., Kokko, A., Raitila, A., Tuppurainen, K., Ebeling, T. M., Salmela, P. I., Paschke, R., Gundogdu, S., de, M. E., Makinen, M. J., Launonen, V., Karhu, A. & Aaltonen, L. A. Pituitary adenoma predisposition caused by germline mutations in the AIP gene. *Science* **312**, 1228-1230, (2006).

117 Hernandez-Ramirez, L. C., Gabrovska, P., Denes, J., Stals, K., Trivellin, G., Tilley, D., Ferrau, F., Evanson, J., Ellard, S., Grossman, A. B., Roncaroli, F., Gadelha, M. R., Korbonits, M. & International, F. C. Landscape of familial isolated and young-onset pituitary adenomas: prospective diagnosis in AIP mutation carriers. *The Journal of Clinical Endocrinology & Metabolism* **100**, E1242–E1254, (2015).

118 Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C. & Kehrer-Sawatzki, H. Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Human Genetics* **132**, 1077-1130, (2013).

119 Daly, A. F., Tichomirowa, M. A., Petrossians, P., Heliovaara, E., Jaffrain-Rea, M. L., Barlier, A., Naves, L. A., Ebeling, T., Karhu, A., Raappana, A., Cazabat, L., de, M. E., Montanana, C. F., Raverot, G., Weil, R. J., Sane, T., Maiter, D., Neggers, S., Yaneva, M., Tabarin, A., Verrua, E., Eloranta, E., Murat, A., Vierimaa, O., Salmela, P. I., Emy, P., Toledo, R. A., Sabate, M. I., Villa, C., Popelier, M., Salvatori, R., Jennings, J., Longas, A. F., Labarta Aizpun, J. I., Georgitsi, M., Paschke, R., Ronchi, C., Valimaki, M., Saloranta, C., De, H. W., Cozzi, R., Guitelman, M., Magri, F., Lagonigro, M. S., Halaby, G., Corman, V., Hagelstein, M. T., Vanbellinghen, J. F., Barra, G. B., Gimenez-Roqueplo, A. P., Cameron, F. J., Borson-Chazot, F., Holdaway, I., Toledo, S. P., Stalla, G. K., Spada, A., Zacharieva, S., Bertherat, J., Brue, T., Bours, V., Chanson, P., Aaltonen, L. A. & Beckers, A. Clinical characteristics and therapeutic responses in patients with germ-line AIP mutations and pituitary adenomas: an international collaborative study. *The Journal of Clinical Endocrinology and Metabolism* **95**, E373-E383, (2010).

120 Beckers, A., Aaltonen, L. A., Daly, A. F. & Karhu, A. Familial isolated pituitary adenomas (FIPA) and the pituitary adenoma predisposition due to mutations in the aryl hydrocarbon receptor interacting protein (AIP) gene. *Endocrine Reviews* **34**, 239-277, (2013).

121 DiGiovanni, R., Serra, S., Ezzat, S. & Asa, S. L. AIP Mutations are not identified in patients with sporadic pituitary adenomas. *Endocrine Pathology* **18**, 76-78, (2007).

122 Cazabat, L., Bouligand, J., Salenave, S., Bernier, M., Gaillard, S., Parker, F., Young, J., Guiochon-Mantel, A. & Chanson, P. Germline AIP mutations in apparently sporadic pituitary adenomas: prevalence in a prospective single-center cohort of 443 patients. *The Journal of Clinical Endocrinology & Metabolism* **97**, E663-670, (2012).

123 Stratakis, C. A., Tichomirowa, M. A., Boikos, S., Azevedo, M. F., Lodish, M., Martari, M., Verma, S., Daly, A. F., Raygada, M., Keil, M. F., Papademetriou, J., Drori-Herishanu, L., Horvath, A., Tsang, K. M., Nesterova, M., Franklin, S., Vanbellinghen, J. F., Bours, V., Salvatori, R. & Beckers, A. The role of germline AIP, MEN1, PRKAR1A, CDKN1B and CDKN2C mutations in causing pituitary adenomas in a large cohort of children, adolescents, and patients with genetic syndromes. *Clinical Genetics* **78**, 457-463, (2010).

124 Gadelha, M. R., Prezant, T. R., Une, K. N., Glick, R. P., Moskal, S. F., 2nd, Vaisman, M., Melmed, S., Kineman, R. D. & Frohman, L. A. Loss of heterozygosity on chromosome 11q13 in two families with acromegaly/gigantism is independent of mutations of the multiple endocrine neoplasia type I gene. *The Journal of Clinical Endocrinology & Metabolism* **84**, 249-256, (1999).

125 Gadelha, M. R., Une, K. N., Rohde, K., Vaisman, M., Kineman, R. D. & Frohman, L. A. Isolated familial somatotropinomas: establishment of linkage to chromosome 11q13.1-11q13.3 and evidence for a potential second locus at chromosome 2p16-12. *The Journal of Clinical Endocrinology & Metabolism* **85**, 707-714, (2000).

126 Soares, B. S., Eguchi, K. & Frohman, L. A. Tumor deletion mapping on chromosome 11q13 in eight families with isolated familial somatotropinoma and in 15 sporadic somatotropinomas. *The Journal of Clinical Endocrinology & Metabolism* **90**, 6580-6587, (2005).

127 Linquette, M., Fossati, P., Lefebvre, J., Decoulx, M. & Dupont-Lecomte, J. Familial gynecomastia with hypospadias: Reifenstein's syndrome. *Annales d'endocrinologie* **28**, 381-390, (1967).

128 Levin, S. R., Hofeldt, F. D., Becker, N., Wilson, C. B., Seymour, R. & Forsham, P. H. Hypersomatotropism and acanthosis nigricans in two brothers. *Archives of Internal Medicine* **134**, 365-367, (1974).

129 Himuro, H., Kobayashi, E., Kono, H., Jinbo, M. & Kitamura, K. Familial occurrence of pituitary adenoma. *Neurological surgery* **4**, 371-377, (1976).

130 Salti, I. S. & Mufarrij, I. S. Familial Cushing Disease. *American Journal of Medical Genetics* **8**, 91-94, (1981).

131 Berezin, M. & Karasik, A. Familial prolactinoma. *Clinical Endocrinology* **42**, 483-486, (1995).

132 Benlian, P., Giraud, S., Lahlou, N., Roger, M., Blin, C., Holler, C., Lenoir, G., Sallandre, J., Calender, A. & Turpin, G. Familial acromegaly: a specific clinical entity--further evidence from the genetic study of a three-generation family. *European Journal of Endocrinology* **133**, 451-456, (1995).

133 Yamada, S., Yoshimoto, K., Sano, T., Takada, K., Itakura, M., Usui, M. & Teramoto, A. Inactivation of the tumor suppressor gene on 11q13 in brothers with familial acrogigantism without multiple endocrine neoplasia type 1. *The Journal of Clinical Endocrinology & Metabolism* **82**, 239-242, (1997).

134 Daly, A. F., Jaffrain-Rea, M. L. & Beckers, A. Clinical and genetic features of familial pituitary adenomas. *Hormone and metabolic research* **37**, 347-354, (2005).

135 Pestell, R. G., Alford, F. P. & Best, J. D. Familial acromegaly. *Acta Endocrinologica (Copenhagen)* **121**, 286-289, (1989).

136 Igreja, S., Chahal, H. S., King, P., Bolger, G. B., Srirangalingam, U., Guasti, L., Chapple, J. P., Trivellin, G., Gueorguiev, M., Guegan, K., Stals, K., Khoo, B., Kumar, A. V., Ellard, S., Grossman, A. B. & Korbonits, M. Characterization of aryl hydrocarbon receptor interacting protein (AIP) mutations in familial isolated pituitary adenoma families. *Human Mutations* **31**, 950-960, (2010).

137 Daly, A. F., Tichomirowa, M. A. & Beckers, A. Genetic, molecular and clinical features of familial isolated pituitary adenomas. *Hormone Research* **71 Suppl 2**, 116-122, (2009).

138 Iwata, T., Yamada, S., Mizusawa, N., Golam, H. M., Sano, T. & Yoshimoto, K. The aryl hydrocarbon receptor-interacting protein gene is rarely mutated in sporadic GH-secreting adenomas. *Clinical Endocrinology* **66**, 499-502, (2007).

139 Kuzhandaivelu, N., Cong, Y. S., Inouye, C., Yang, W. M. & Seto, E. XAP2, a novel hepatitis B virus X-associated protein that inhibits X transactivation. *Nucleic acids research* **24**, 4741-4750, (1996).

140 Carver, L. A. & Bradfield, C. A. Ligand-dependent interaction of the aryl hydrocarbon receptor with a novel immunophilin homolog in vivo. *The Journal of Biological Chemistry* **272**, 11452-11456, (1997).

141 Thakker, R. V., Pook, M. A., Wooding, C., Boscaro, M., Scanarini, M. & Clayton, R. N. Association of somatotrophinomas with loss of alleles on chromosome 11 and with gsp mutations. *The Journal of Clinical Investigation* **91**, 2815-2821, (1993).

142 Morgan, R. M. L., Hernandez Ramirez, L. C., Trivellin, G., Zhou, L., Roe, S. M., Korbonits, M. & Prodromou, C. Structure of the TPR domain of AIP: lack of client protein interaction with the C-terminal á-7 helix of the TPR domain of AIP is sufficient for pituitary adenoma predisposition. *PloS One* **7**, e53339, (2012).

143 Cunningham, F., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C. G., Gordon, L., Hourlier, T., Hunt, S. E., Janacek, S. H., Johnson, N., Juettemann, T., Kahari, A. K., Keenan, S., Martin, F. J., Maurel, T., McLaren, W., Murphy, D. N., Nag, R., Overduin, B., Parker, A., Patricio, M., Perry, E., Pignatelli, M., Riat, H. S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S. P., Zadissa, A., Aken, B. L., Birney, E., Harrow, J., Kinsella, R., Muffato, M., Ruffier, M., Searle, S. M., Spudich, G., Trevanion, S. J., Yates, A., Zerbino, D. R. & Flicek, P. Ensembl 2015. *Nucleic Acids Res* **43**, D662-669, (2015).

144 UniProt, C. UniProt: a hub for protein information. *Nucleic acids research* **43**, D204-212, (2015).

145 Linnert, M., Haupt, K., Lin, Y. J., Kissing, S., Paschke, A. K., Fischer, G., Weiwad, M. & Lucke, C. NMR assignments of the FKBP-type PPIase domain of the human aryl-hydrocarbon receptor-interacting protein (AIP). *Biomolecular NMR Assignments* **6**, 209-212, (2012).

146 Das, A. K., Cohen, P. W. & Barford, D. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *The EMBO Journal* **17**, 1192-1199, (1998).

147 Russell, L. C., Whitt, S. R., Chen, M. S. & Chinkers, M. Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. *The Journal of Biological Chemistry* **274**, 20060-20063, (1999).

148 Bell, D. R. & Poland, A. Binding of aryl hydrocarbon receptor (AhR) to AhR-interacting protein. The role of hsp90. *The Journal of Biological Chemistry* **275**, 36407-36414, (2000).

149 Meyer, B. K., Petrulis, J. R. & Perdew, G. H. Aryl hydrocarbon (Ah) receptor levels are selectively modulated by hsp90-associated immunophilin homolog XAP2. *Cell Stress & Chaperones* **5**, 243-254, (2000).

150 Laenger, A., Lang-Rollin, I., Kozany, C., Zschocke, J., Zimmermann, N., Ruegg, J., Holsboer, F., Hausch, F. & Rein, T. XAP2 inhibits glucocorticoid receptor activity in mammalian cells. *FEBS letters* **583**, 1493-1498, (2009).

151 Daly, A. F., Vanbellinghen, J. F., Khoo, S. K., Jaffrain-Rea, M. L., Naves, L. A., Guitelman, M. A., Murat, A., Emy, P., Gimenez-Roqueplo, A. P., Tamburrano, G., Raverot, G., Barlier, A., De Herder, W., Penfornis, A., Ciccarelli, E., Estour, B., Lecomte, P., Gatta, B., Chabre, O., Sabate, M. I., Bertagna, X., Garcia Basavilbaso, N., Stalldecker, G., Colao, A., Ferolla, P., Wemeau, J. L., Caron, P., Sadoul, J. L., Oneto, A., Archambeaud, F., Calender, A., Sinilnikova, O., Montanana, C. F., Cavagnini, F., Hana, V., Solano, A., Delettieres, D., Luccio-Camelo, D. C., Basso, A., Rohmer, V., Brue, T., Bours, V., Teh, B. T. & Beckers, A. Aryl hydrocarbon receptor-interacting protein gene mutations in familial isolated pituitary adenomas: analysis in 73 families. *The Journal of Clinical Endocrinology and Metabolism* **92**, 1891-1896, (2007).

152 Montanana, C. F., Daly, A. F., Tichomirowa, M. A., Vanbellinghen, J. F., Jaffrain-Rea, M. L., Trescoli Serrano, C., Riesgo Suares, P., Gomez Vela, J., Tenes, S. & Bours, V. TSH-secreting pituitary adenoma in a male patient with a novel missense AIP mutation. *Proceedings of the 91st Annual Meeting of the Endocrine*, P1668, (2009).

153 Trivellin, G. & Korbonits, M. AIP and its interacting partners. *Journal of Endocrinology* **210**, 137-155, (2011).

154 Ma, Q. & Whitlock, J. P., Jr. A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *The Journal of Biological Chemistry* **272**, 8878-8884, (1997).

155 Daly, A. F., Cogne, M., Tabarin, A., Murat, A., Delemer, B., Luger, A., Gaillard, R., Colao, A., Harris, A. G., Herlocu, M. C., Petrossians, P. & Beckers, A. The epidemiology of pituitary tumors: Results of an international collaborative study. in *The Endocrine Society Annual Meeting.*Toronto, Canada (2007).

156 Berg, P. & Pongratz, I. Two parallel pathways mediate cytoplasmic localization of the dioxin (aryl hydrocarbon) receptor. *The Journal of Biological Chemistry* **277**, 32310-32319, (2002).

157 Li, S., Wang, L., Fu, B., Berman, M. A., Diallo, A. & Dorf, M. E. TRIM65 regulates microRNA activity by ubiquitination of TNRC6. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 6970-6975, (2014).

158 Petrulis, J. R., Hord, N. G. & Perdew, G. H. Subcellular localization of the aryl hydrocarbon receptor is modulated by the immunophilin homolog hepatitis B virus X-associated protein 2. *The Journal of Biological Chemistry* **275**, 37448-37453, (2000).

159 Carver, L. A., LaPres, J. J., Jain, S., Dunham, E. E. & Bradfield, C. A. Characterization of the Ah receptor-associated protein, ARA9. *The Journal of Biological Chemistry* **273**, 33580-33587, (1998).

160 Dunham, E. E., Stevens, E. A., Glover, E. & Bradfield, C. A. The aryl hydrocarbon receptor signaling pathway is modified through interactions with a Kelch protein. *Molecular Pharmacology* **70**, 8-15, (2006).

161 Kazlauskas, A., Sundstrom, S., Poellinger, L. & Pongratz, I. The hsp90 chaperone complex regulates intracellular localization of the dioxin receptor. *Molecular and Cellular Biology* **21**, 2594-2607, (2001).

162 Vargiolu, M., Fusco, D., Kurelac, I., Dirnberger, D., Baumeister, R., Morra, I., Melcarne, A., Rimondini, R., Romeo, G. & Bonora, E. The tyrosine kinase receptor RET interacts in vivo with aryl hydrocarbon receptor-interacting protein to alter survivin availability. *The Journal of Clinical Endocrinology & Metabolism* **94**, 2571-2578, (2009).

163 Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z. Y., Larsen, B., Choi, H., Berger, B., Gingras, A. C. & Lindquist, S. A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* **158**, 434-448, (2014).

164 Varjosalo, M., Sacco, R., Stukalov, A., van Drogen, A., Planyavsky, M., Hauri, S., Aebersold, R., Bennett, K. L., Colinge, J., Gstaiger, M. & Superti-Furga, G. Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. *Nature Methods* **10**, 307-314, (2013).

165 Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M. P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., Dong, R., Guarani, V., Vaites, L. P., Ordureau, A., Rad, R., Erickson, B. K., Wuhr, M., Chick, J., Zhai, B., Kolippakkam, D., Mintseris, J., Obar, R. A., Harris, T., Artavanis-Tsakonas, S., Sowa, M. E., De Camilli, P., Paulo, J. A., Harper, J. W. & Gygi, S. P. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425-440, (2015).

166 Deribe, Y. L., Wild, P., Chandrashaker, A., Curak, J., Schmidt, M. H., Kalaidzidis, Y., Milutinovic, N., Kratchmarova, I., Buerkle, L., Fetchko, M. J., Schmidt, P., Kittanakom, S., Brown, K. R., Jurisica, I., Blagoev, B., Zerial, M., Stagljar, I. & Dikic, I. Regulation of epidermal growth factor receptor trafficking by lysine deacetylase HDAC6. *Science Signaling* **2**, ra84, (2009).

167 Kristensen, A. R., Gsponer, J. & Foster, L. J. A high-throughput approach for measuring temporal changes in the interactome. *Nature Methods* **9**, 907-909, (2012).

168 Nakata, A., Urano, D., Fujii-Kuriyama, Y., Mizuno, N., Tago, K. & Itoh, H. G-protein signalling negatively regulates the stability of aryl hydrocarbon receptor. *EMBO Reports* **10**, 622-628, (2009).

169 Schulke, J. P., Wochnik, G. M., Lang-Rollin, I., Gassen, N. C., Knapp, R. T., Berning, B., Yassouridis, A. & Rein, T. Differential impact of tetratricopeptide repeat proteins on the steroid hormone receptors. *PLoS One* **5**, e11717, (2010).

170 Li, J., Richter, K. & Buchner, J. Mixed Hsp90-cochaperone complexes are important for the progression of the reaction cycle. *Nature Structural & Molecular Biology* **18**, 61-66, (2011).

171 Yano, M., Terada, K. & Mori, M. AIP is a mitochondrial import mediator that binds to both import receptor Tom20 and preproteins. *The Journal of Cell Biology* **163**, 45-56, (2003).

172 de Oliveira, S. K., Hoffmeister, M., Gambaryan, S., Muller-Esterl, W., Guimaraes, J. A. & Smolenski, A. P. Phosphodiesterase 2A forms a complex with the co-chaperone XAP2 and regulates nuclear translocation of the aryl hydrocarbon receptor. *The Journal of Biological Chemistry* **282**, 13656-13663, (2007).

173 Bolger, G. B., Peden, A. H., Steele, M. R., MacKenzie, C., McEwan, D. G., Wallace, D. A., Huston, E., Baillie, G. S. & Houslay, M. D. Attenuation of the activity of the cAMP-specific phosphodiesterase PDE4A5 by interaction with the immunophilin XAP2. *The Journal of Biological Chemistry* **278**, 33351-33363, (2003).

174 Sumanasekera, W. K., Tien, E. S., Turpey, R., Vanden Heuvel, J. P. & Perdew, G. H. Evidence that peroxisome proliferator-activated receptor alpha is complexed with the 90-kDa heat shock protein and the hepatitis virus B X-associated protein 2. *The Journal of Biological Chemistry* **278**, 4467-4473, (2003).

175 Zhao, Y., Meng, X. M., Wei, Y. J., Zhao, X. W., Liu, D. Q., Cao, H. Q., Liew, C. C. & Ding, J. F. Cloning and characterization of a novel cardiac-specific kinase that interacts specifically with cardiac troponin I. *Journal of Molecular Medicine* **81**, 297-304, (2003).

176 Hollingshead, B. D., Petrulis, J. R. & Perdew, G. H. The aryl hydrocarbon (Ah) receptor transcriptional regulator hepatitis B virus X-associated protein 2 antagonizes p23 binding to Ah receptor-Hsp90 complexes and is dispensable for receptor function. *The Journal of Biological Chemistry* **279**, 45652-45661, (2004).

177 Petrulis, J. R., Kusnadi, A., Ramadoss, P., Hollingshead, B. & Perdew, G. H. The hsp90 Co-chaperone XAP2 alters importin beta recognition of the bipartite nuclear localization signal of the Ah receptor and represses transcriptional activity. *The Journal of Biological Chemistry* **278**, 2677-2685, (2003).

178 Murray, I. A., Patterson, A. D. & Perdew, G. H. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nature Reviews. Cancer* **14**, 801-814, (2014).

179 Ozfirat, Z. & Korbonits, M. AIP gene and familial isolated pituitary adenomas. *Molecular and Cellular Endocrinology* **326**, 71-79, (2010).

180 Lahvis, G. P., Lindell, S. L., Thomas, R. S., McCuskey, R. S., Murphy, C., Glover, E., Bentz, M., Southard, J. & Bradfield, C. A. Portosystemic shunting and persistent fetal vascular structures in aryl hydrocarbon receptor-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10442-10447, (2000).

181 Lin, B. C., Nguyen, L. P., Walisser, J. A. & Bradfield, C. A. A hypomorphic allele of aryl hydrocarbon receptor-associated protein-9 produces a phenocopy of the AHR-null mouse. *Molecular Pharmacology* **74**, 1367-1371, (2008).

182 Dull, A. B., Carlson, D. B., Petrulis, J. R. & Perdew, G. H. Characterization of the phosphorylation status of the hepatitis B virus X-associated protein 2. *Archives of Biochemistry and Biophysics* **406**, 209-221, (2002).

183 Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V. & Sullivan, M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Research* **40**, D261-270, (2012).

184 Formosa, R., Xuereb-Anastasi, A. & Vassallo, J. Aip regulates cAMP signalling and GH secretion in GH3 cells. *Endocrine-Related Cancer* **20**, 495-505, (2013).

185 Tuominen, I., Heliovaara, E., Raitila, A., Rautiainen, M. R., Mehine, M., Katainen, R., Donner, I., Aittomaki, V., Lehtonen, H. J., Ahlsten, M., Kivipelto, L., Schalin-Jantti, C., Arola, J., Hautaniemi, S. & Karhu, A. AIP inactivation leads to pituitary tumorigenesis through defective Galphai-cAMP signaling. *Oncogene* **34**, 1174-1184, (2015).

186 Romano, D., Magalon, K., Pertuit, M., Rasolonjanahary, R., Barlier, A., Enjalbert, A. & Gerard, C. Conditional overexpression of the wild-type Gs alpha as the gsp oncogene initiates chronic extracellularly regulated kinase 1/2 activation and hormone hypersecretion in pituitary cell lines. *Endocrinology* **148**, 2973-2983, (2007).

187 Yonehara, T., Kanasaki, H., Yamamoto, H., Fukunaga, K., Miyazaki, K. & Miyamoto, E. Involvement of mitogen-activated protein kinase in cyclic adenosine 3',5'-monophosphate-induced hormone gene expression in rat pituitary GH(3) cells. *Endocrinology* **142**, 2811-2819, (2001).

188 Freda, P. U., Chung, W. K., Matsuoka, N., Walsh, J. E., Kanibir, M. N., Kleinman, G., Wang, Y., Bruce, J. N. & Post, K. D. Analysis of GNAS mutations in 60 growth hormone secreting pituitary tumors: correlation with clinical and pathological characteristics and surgical

outcome based on highly sensitive GH and IGF-I criteria for remission. *Pituitary* **10**, 275-282, (2007).

189 Weinstein, L. S., Shenker, A., Gejman, P. V., Merino, M. J., Friedman, E. & Spiegel, A. M. Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. *The New England Journal of Medicine* **325**, 1688-1695, (1991).

190 Boikos, S. A. & Stratakis, C. A. Molecular genetics of the cAMP-dependent protein kinase pathway and of sporadic pituitary tumorigenesis. *Human Molecular Genetics* **16**, R80-R87, (2007).

191 Kirschner, L. S., Sandrini, F., Monbo, J., Lin, J. P., Carney, J. A. & Stratakis, C. A. Genetic heterogeneity and spectrum of mutations of the PRKAR1A gene in patients with the carney complex. *Human Molecular Genetics* **9**, 3037-3046, (2000).

192 Kirschner, L. S., Carney, J. A., Pack, S. D., Taymans, S. E., Giatzakis, C., Cho, Y. S., Cho-Chung, Y. S. & Stratakis, C. A. Mutations of the gene encoding the protein kinase A type I-alpha regulatory subunit in patients with the Carney complex. *Nature Genetics* **26**, 89-92, (2000).

193 Galigniana, M. D., Harrell, J. M., Murphy, P. J., Chinkers, M., Radanyi, C., Renoir, J. M., Zhang, M. & Pratt, W. B. Binding of hsp90-associated immunophilins to cytoplasmic dynein: direct binding and in vivo evidence that the peptidylprolyl isomerase domain is a dynein interaction domain. *Biochemistry* **41**, 13602-13610, (2002).

194 Kazlauskas, A., Poellinger, L. & Pongratz, I. The immunophilin-like protein XAP2 regulates ubiquitination and subcellular localization of the dioxin receptor. *The Journal of Biological Chemistry* **275**, 41317-41324, (2000).

195 LaPres, J. J., Glover, E., Dunham, E. E., Bunger, M. K. & Bradfield, C. A. ARA9 modifies agonist signaling through an increase in cytosolic aryl hydrocarbon receptor. *The Journal of Biological Chemistry* **275**, 6153-6159, (2000).

196 Munjal, A. & Lecuit, T. Actomyosin networks and tissue morphogenesis. *Development* **141**, 1789-1793, (2014).

197 Sahai, E. & Marshall, C. J. Differing modes of tumour cell invasion have distinct requirements for Rho/ROCK signalling and extracellular proteolysis. *Nature Cell Biology* **5**, 711-719, (2003).

198 Gillitzer, R. & Goebeler, M. Chemokines in cutaneous wound healing. *Journal of Leukocyte Biology* **69**, 513-521, (2001).

199 Charras, G. & Sahai, E. Physical influences of the extracellular environment on cell migration. *Nature reviews. Molecular Cell Biology* **15**, 813-824, (2014).

200 DeMali, K. A., Wennerberg, K. & Burridge, K. Integrin signaling to the actin cytoskeleton. *Cell Biol Current Opinion in Cell Biology* **15**, 572-582, (2003).

201 Kasuki Jomori de Pinho, L., Vieira Neto, L., Armondi Wildemberg, L. E., Gasparetto, E. L., Marcondes, J., de Almeida Nunes, B., Takiya, C. M. & Gadelha, M. R. Low aryl hydrocarbon receptor-interacting protein expression is a better marker of invasiveness in somatotropinomas than Ki-67 and p53. *Neuroendocrinology* **94**, 39-48, (2011).

202 Gadelha, M. R., Kasuki, L. & Korbonits, M. Novel pathway for somatostatin analogs in patients with acromegaly. *Trends in Endocrinology and Metabolism* **24**, 238–246, (2012).

203 Lovell, S. C., Li, X., Weerasinghe, N. R. & Hentges, K. E. Correlation of microsynteny conservation and disease gene distribution in mammalian genomes. *BMC Genomics* **10**, 521, (2009).

204 Lin, B. C., Sullivan, R., Lee, Y., Moran, S., Glover, E. & Bradfield, C. A. Deletion of the aryl hydrocarbon receptor-associated protein 9 leads to cardiac malformation and embryonic lethality. *The Journal of Biological Chemistry* **282**, 35924-35932, (2007).

205 Lahvis, G. P., Pyzalski, R. W., Glover, E., Pitot, H. C., McElwee, M. K. & Bradfield, C. A. The aryl hydrocarbon receptor is required for developmental closure of the ductus venosus in the neonatal mouse. *Molecular Pharmacology* **67**, 714-720, (2005).

206 Raitila, A., Lehtonen, H. J., Arola, J., Heliovaara, E., Ahlsten, M., Georgitsi, M., Jalanko, A., Paetau, A., Aaltonen, L. A. & Karhu, A. Mice with inactivation of aryl hydrocarbon receptor-interacting protein (Aip) display complete penetrance of pituitary adenomas with aberrant ARNT expression. *American Journal of Pathology* **177**, 1969-1976, (2010).

207 Rostomyan, L., Daly, A. F., Petrossians, P., Natchev, E., Lila, A. R., Lecoq, A. L., Lecumberri Santamaria, B., Trivellin, G., Salvatori, R., Moraitis, A., Holdaway, I., Kranenburg-Van Klaveren, D., Zatelli, M. C., Palacios, N., Nozieres, C., Zacharin, M., Ebeling, T. M., Ojaniemi, M., Rozhinskaya, L., Verrua, E., Jaffrain Rea, M. L., Filipponi, S., Guskova, D., Pronin, V., Bertherat, J., Belaya, Z., Ilovaiskaya, I., Sahnoun Fathallah, M., Sievers, C., Stalla, G. K., Castermans, E., Caberg, J. H., Sorkina, E., Auriemma, R., Mittal, S., Kareva, M., Lysy, P., Emy, P., de Menis, E., Choong, C., Mantovani, G., Bours, V., de Herder, W. W., Brue, T., Barlier, A., Neggers, S., Zacharieva, S., Chanson, P., Shah, N., Stratakis, C. A., Naves, L. A. & Beckers, A. Clinical and genetic characterization of pituitary gigantism: an international study in 208 patients. *Endocrine-Related Cancer* **22**, 745-757, (2015).

208 Tichomirowa, M. A., Barlier, A., Daly, A. F., Jaffrain-Rea, M. L., Ronchi, C. L., Yaneva, M., Urban, J. D., Petrossians, P., Elenkova, A. P., Tabarin, A., Desailloud, R., Maiter, D., Schurmeyer, T., Cozzi, R., Theodoropoulou, M., Sievers, C., Bernabeu, I., Naves, L. A., Chabre, O., Fajardo, M. C., Hana, V., Halaby, G., Delemer, B., Labarta, J. I., Sonnet, E., Ferrandez, A., Hagelstein, M. T., Caron, P., Stalla, G. K., Bours, V., Zacharieva, S., Spada, A., Brue, T. & Beckers, A. High prevalence of AIP gene mutations following focused screening in young patients with sporadic pituitary macroadenomas. *European Journal of Endocrinology* **165**, 509-515, (2011).

209 Cuny, T., Pertuit, M., Sahnoun-Fathallah, M., Daly, A., Occhi, G., Odou, M. F., Tabarin, A., Nunes, M. L., Delemer, B., Rohmer, V., Desailloud, R., Kerlan, V., Chabre, O., Sadoul, J. L., Cogne, M., Caron, P., Cortet-Rudelli, C., Lienhardt, A., Raingeard, I., Guedj, A. M., Brue, T., Beckers, A., Weryha, G., Enjalbert, A. & Barlier, A. Genetic analysis in young patients with sporadic pituitary macroadenomas: besides AIP don't forget MEN1 genetic analysis. *European Journal of Endocrinology* **168**, 533-541, (2013).

210 McGonnell, I. M. & Fowkes, R. C. Fishing for gene function--endocrine modelling in the zebrafish. *Journal of Endocrinology* **189**, 425-439, (2006).

211 Lohr, H. & Hammerschmidt, M. Zebrafish in endocrine systems: recent advances and implications for human disease. *Annual Review of Physiology* **73**, 183-211, (2011).

212 Aflorei, E. D., Chen, C., McGonnel, I. M., Foweks, R. C., Grossman, A. B., Tapon, N., Stanewsky, R. & Korbonits, M. Development of novel *AIP* (Aryl Hydrocarbon Receptor-Interacting Protein) gene study models using the fruitfly and the zebrafish. in *15th International Congress of Endocrinology, Florence, Italy.* (2012).

213 Stojanovic, M., Aflorei, E. D., McGonnel, I. M. & Korbonits, M. AIP inactivation leads to pituitary enlargement in the zebrafish. in *Aspiring to Excellence: Pituitary Expert Forum.*Barcelona, Spain (2015).

214 Arbouzova, N. I. & Zeidler, M. P. JAK/STAT signalling in Drosophila: insights into conserved regulatory and cellular functions. *Development* **133**, 2605-2616, (2006).

215 Huangfu, D. & Anderson, K. V. Signaling from Smo to Ci/Gli: conservation and divergence of Hedgehog pathways from Drosophila to vertebrates. *Development* **133**, 3-14, (2006).

216 St Johnston, D. The art and design of genetic screens: Drosophila melanogaster. *Nature Reviews. Genetics* **3**, 176-188, (2002).

217 Molnar, C., Resnik-Docampo, M., Organista, M. F., Martin, M., Hevia, C. F. & Celis, J. D. in *Human Genetic Diseases* Vol. 1 (ed Dijana Plaseska-Karanfilska) (INTECH Open Access, 2011).

218 Wang, S. H., Simcox, A. & Campbell, G. Dual role for Drosophila epidermal growth factor receptor signaling in early wing disc development. *Genes & Development* **14**, 2271-2276, (2000).

219 Martin, G. R. The roles of FGFs in the early development of vertebrate limbs. *Genes & Development* **12**, 1571-1586, (1998).

220 Herskowitz, I. MAP kinase pathways in yeast: for mating and more. *Cell* **80**, 187-197, (1995).

221 Morgan, T. H., Sturtevant, A. H. & Bridge, C. B. The Evidence for the Linear Order of the Genes. *Proceedings of the National Academy of Sciences of the United States of America* **6**, 162-164, (1920).

222 Sturtevant, A. H., Bridges, C. B. & Morgan, T. H. The Spatial Relations of Genes. *Proceedings of the National Academy of Sciences of the United States of America* **5**, 168-173, (1919).

223 Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor, G. L., Abril, J. F., Agbayani, A., An, H. J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., WoodageT, Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W.,

Rubin, G. M. & Venter, J. C. The genome sequence of Drosophila melanogaster. *Science* **287**, 2185-2195, (2000).

224 Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H. & Lewis, S. Comparative genomics of the eukaryotes. *Science* **287**, 2204-2215, (2000).

225 Rajan, A. & Perrimon, N. Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion. *Cell* **151**, 123-137, (2012).

226 Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D. & Sklar, J. TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* **66**, 649-661, (1991).

227 Barakat, M. T., Humke, E. W. & Scott, M. P. Learning from Jekyll to control Hyde: Hedgehog signaling in development and cancer. *Trends in Molecular Medicine* **16**, 337-348, (2010).

228 Klovstad, M., Abdu, U. & Schupbach, T. Drosophila brca2 is required for mitotic and meiotic DNA repair and efficient activation of the meiotic recombination checkpoint. *PLoS Genetics* **4**, e31, (2008).

229 Read, R. D., Cavenee, W. K., Furnari, F. B. & Thomas, J. B. A drosophila model for EGFR-Ras and PI3K-dependent human glioma. *PLoS Genetics* **5**, e1000374, (2009).

230 Ferner, R. E., Golding, J. F., Smith, M., Calonje, E., Jan, W., Sanjayanathan, V. & O'Doherty, M. [18F]2-fluoro-2-deoxy-D-glucose positron emission tomography (FDG PET) as a diagnostic tool for neurofibromatosis 1 (NF1) associated malignant peripheral nerve sheath tumours (MPNSTs): a long-term clinical study. *Annals of Oncology* **19**, 390-394, (2008).

231 Williams, J. A., Su, H. S., Bernards, A., Field, J. & Sehgal, A. A circadian output in Drosophila mediated by neurofibromatosis-1 and Ras/MAPK. *Science* **293**, 2251-2256, (2001).

232 Barkan, B., Starinsky, S., Friedman, E., Stein, R. & Kloog, Y. The Ras inhibitor farnesylthiosalicylic acid as a potential therapy for neurofibromatosis type 1. *Clinical Rancer Research* **12**, 5533-5542, (2006).

233 Rudrapatna, V. A., Cagan, R. L. & Das, T. K. Drosophila cancer models. *Developmental Dynamics* **241**, 107-118, (2012).

234 Committee on Developmental Toxicology. Using animal models to assess and undertand developmental toxicity. in *Scientific Frontiers in Developmental Toxicology and Risk Assessment* Vol. 1 151-196 The National Academies Press, (2000).

235 Hochman, B. Analysis of chromosome 4 in Drosophila melanogaster. II. Ethyl methanesulfonate induced lethals. *Genetics* **67**, 235-252, (1971).

236 Greenspan, R. J. Fly Pushing: The Theory and Practice of Drosophila Genetics. 2nd edn, (2004).

237 Childress, J., Behringer, R. & Halder, G. Learning to Fly: Phenotypic Markers in Drosophila. in *Genesis - The Journal of Genetics and Development*, (2005).

238 Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415, (1993).

239 Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research* **33**, D501-504, (2005).

240 Lindsley, D. L. & Zimm, G. G. The Genome of Drosophila melanogaster. Vol. 1 Academic Press, (1992).

241 Missirlis, F., Rahlfs, S., Dimopoulos, N., Bauer, H., Becker, K., Hilliker, A., Phillips, J. P. & Jackle, H. A putative glutathione peroxidase of Drosophila encodes a thioredoxin peroxidase that provides resistance against oxidative stress but fails to complement a lack of catalase activity. *Biological Chemistry* **384**, 463-472, (2003).

242 Luo, L., Liao, Y. J., Jan, L. Y. & Jan, Y. N. Distinct morphogenetic functions of similar small GTPases: Drosophila Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes & Development* **8**, 1787-1802, (1994).

243 Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in Drosophila. *Current Biology* **12**, 1293-1300, (2002).

244 Yang, J., McCart, C., Woods, D. J., Terhzaz, S., Greenwood, K. G., ffrench-Constant, R. H. & Dow, J. A. A Drosophila systems approach to xenobiotic metabolism. *Physiological Genomics* **30**, 223-231, (2007).

245 Lo, P. C. & Frasch, M. A role for the COUP-TF-related gene seven-up in the diversification of cardioblast identities in the dorsal vessel of Drosophila. *Mechanisms of Development* **104**, 49-60, (2001).

246 Klapholz, B., Herbert, S. L., Wellmann, J., Johnson, R., Parsons, M. & Brown, N. H. Alternative mechanisms for talin to mediate integrin function. *Current Biology* **25**, 847-857, (2015).

247 Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S., Couto, A., Marra, V., Keleman, K. & Dickson, B. J. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-156, (2007).

248 Bellen, H. J., Levis, R. W., He, Y., Carlson, J. W., Evans-Holm, M., Bae, E., Kim, J., Metaxakis, A., Savakis, C., Schulze, K. L., Hoskins, R. A. & Spradling, A. C. The Drosophila gene disruption project: progress using transposons with distinctive site specificities. *Genetics* **188**, 731-743, (2011).

249 Patel, H., Le, T., Liang, Z. G., Ray, S., Slovitt, M., Sivasubramaniam, G., Yu, M. & Beitel, G. A new family of direct-drive dfd::YFP balancers. *46th Annual Drosophila Research Conference, San Diego, CA*, 1035C, (2005).

250 Clark, J. B. & Kidwell, M. G. A phylogenetic perspective on P transposable element evolution in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 11428-11433, (1997).

251 Chou, T. B. & Perrimon, N. Use of a yeast site-specific recombinase to produce female germline chimeras in Drosophila. *Genetics* **131**, 643-653, (1992).

252 Zervas, C. G., Psarra, E., Williams, V., Solomon, E., Vakaloglou, K. M. & Brown, N. H. A central multifunctional role of integrin-linked kinase at muscle attachment sites. *Journal of Cell Science* **124**, 1316-1327, (2011).

253 Rubin, G. M. & Spradling, A. C. Genetic transformation of Drosophila with transposable element vectors. *Science* **218**, 348-353, (1982).

254 Thorpe, H. M., Wilson, S. E. & Smith, M. C. Control of directionality in the site-specific recombination system of the Streptomyces phage phiC31. *Molecular Microbiology* **38**, 232-241, (2000).

255 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675, (2012).

256 Polesello, C., Huelsmann, S., Brown, N. H. & Tapon, N. The Drosophila RASSF homolog antagonizes the hippo pathway. *Current Biology* **16**, 2459-2465, (2006).

257 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25, (2009).

258 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111, (2009).

259 Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8**, 175-185, (1998).

260 Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* **8**, 186-194, (1998).

261 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621-628, (2008).

262 Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L. & Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562-578, (2012).

263 Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C. & Jensen, L. J. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research* **41**, D808-815, (2013).

264 Linnert, M., Lin, Y. J., Manns, A., Haupt, K., Paschke, A. K., Fischer, G., Weiwad, M. & Lucke, C. The FKBP-type domain of the human aryl hydrocarbon receptor-interacting protein reveals an unusual Hsp90 interaction. *Biochemistry* **52**, 2097-2107, (2013).

265 Meyer, B. K., Pray-Grant, M. G., Vanden Heuvel, J. P. & Perdew, G. H. Hepatitis B virus X-associated protein 2 is a subunit of the unliganded aryl hydrocarbon receptor core complex and exhibits transcriptional enhancer activity. *Molecular and Cellular Biology* **18**, 978-988, (1998).

266 Duffy, J. B. GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis* **34**, 1-15, (2002).

267 Witsell, A., Kane, D. P., Rubin, S. & McVey, M. Removal of the bloom syndrome DNA helicase extends the utility of imprecise transposon excision for making null mutations in Drosophila. *Genetics* **183**, 1187-1193, (2009).

268 Ryder, E. & Russell, S. Transposable elements as tools for genomics and genetics in Drosophila. *Briefings in Functional Genomics & Proteomics* **2**, 57-71, (2003).

269 Venken, K. J. & Bellen, H. J. Emerging technologies for gene manipulation in Drosophila melanogaster. *Nature Reviews. Genetics* **6**, 167-178, (2005).

270 Engels, W. R., Johnson-Schlitz, D. M., Eggleston, W. B. & Sved, J. High-frequency P element loss in Drosophila is homolog dependent. *Cell* **62**, 515-525, (1990).

271 Daniels, S. B., McCarron, M., Love, C. & Chovnick, A. Dysgenesis-induced instability of rosy locus transformation in Drosophila melanogaster: analysis of excision events and the selective recovery of control element deletions. *Genetics* **109**, 95-117, (1985).

272 Takasu-Ishikawa, E., Yoshihara, M. & Hotta, Y. Extra sequences found at P element excision sites in Drosophila melanogaster. *Molecular & general genetics* **232**, 17-23, (1992).

273 Staveley, B. E., Heslip, T. R., Hodgetts, R. B. & Bell, J. B. Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. *Genetics* **139**, 1321-1329, (1995).

274 Thorpe, H. M. & Smith, M. C. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5505-5510, (1998).

275 Groth, A. C., Fish, M., Nusse, R. & Calos, M. P. Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775-1782, (2004).

276 Boguski, M. S., Tolstoshev, C. M. & Bassett, D. E., Jr. Gene discovery in dbEST. *Science* **265**, 1993-1994, (1994).

277 Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics* **10**, 57-63, (2009).

278 Dickerson, J. E., Zhu, A., Robertson, D. L. & Hentges, K. E. Defining the role of essential genes in human disease. *PLoS One* **6**, e27368, (2011).

279 Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539, (2011).

280 Stolc, V., Gauhar, Z., Mason, C., Halasz, G., van Batenburg, M. F., Rifkin, S. A., Hua, S., Herreman, T., Tongprasit, W., Barbano, P. E., Bussemaker, H. J. & White, K. P. A gene expression map for the euchromatic genome of Drosophila melanogaster. *Science* **306**, 655-660, (2004).

281 Arbeitman, M. N., Furlong, E. E., Imam, F., Johnson, E., Null, B. H., Baker, B. S., Krasnow, M. A., Scott, M. P., Davis, R. W. & White, K. P. Gene expression during the life cycle of Drosophila melanogaster. *Science* **297**, 2270-2275, (2002).

282 Chintapalli, V. R., Wang, J., Herzyk, P. & Dow, J. A. T. FlyAtlas: survey of adult and larval expression. *personal communication to FlyBase*, (2010), FlyAtlas (http://www.flyatlas.org/).

283 mod, E. C., Roy, S., Ernst, J., Kharchenko, P. V., Kheradpour, P., Negre, N., Eaton, M. L., Landolin, J. M., Bristow, C. A., Ma, L., Lin, M. F., Washietl, S., Arshinoff, B. I., Ay, F., Meyer, P. E., Robine, N., Washington, N. L., Di Stefano, L., Berezikov, E., Brown, C. D., Candeias, R., Carlson, J. W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M. Y., Will, S., Alekseyenko, A. A., Artieri, C., Booth, B. W., Brooks, A. N., Dai, Q., Davis, C. A., Duff,

M. O., Feng, X., Gorchakov, A. A., Gu, T., Henikoff, J. G., Kapranov, P., Li, R., MacAlpine, H. K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S. K., Riddle, N. C., Sakai, A., Samsonova, A., Sandler, J. E., Schwartz, Y. B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K. H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S. E., Brent, M. R., Cherbas, L., Elgin, S. C., Gingeras, T. R., Grossman, R., Hoskins, R. A., Kaufman, T. C., Kent, W., Kuroda, M. I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J. W., Ren, B., Russell, S., Cherbas, P., Graveley, B. R., Lewis, S., Micklem, G., Oliver, B., Park, P. J., Celniker, S. E., Henikoff, S., Karpen, G. H., Lai, E. C., MacAlpine, D. M., Stein, L. D., White, K. P. & Kellis, M. Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* **330**, 1787-1797, (2010).

284 Lessing, D. & Bonini, N. M. Maintaining the brain: insight into human neurodegeneration from Drosophila melanogaster mutants. *Nature Reviews. Genetics* **10**, 359-370, (2009).

285 Brand, A. H., Manoukian, A. S. & Perrimon, N. Ectopic expression in Drosophila. *Methods in Cell Biology* **44**, 635-654, (1994).

286 Hartenstein, V. Atlas of Drosophila Development. Cold Spring Harbor Laboratory Press, (1993).

287 Bross, T. G., Rogina, B. & Helfand, S. L. Behavioral, physical, and demographic changes in Drosophila populations through dietary restriction. *Aging Cell* **4**, 309-317, (2005).

288 Edgecomb, R. S., Harth, C. E. & Schneiderman, A. M. Regulation of feeding behavior in adult Drosophila melanogaster varies with feeding regime and nutritional state. *The Journal of Experimental Biology* **197**, 215-235, (1994).

289 Dorer, D. R., Rudnick, J. A., Moriyama, E. N. & Christensen, A. C. A family of genes clustered at the Triplo-lethal locus of Drosophila melanogaster has an unusual evolutionary history and significant synteny with Anopheles gambiae. *Genetics* **165**, 613-621, (2003).

290 Guan, X., Middlebrooks, B. W., Alexander, S. & Wasserman, S. A. Mutation of TweedleD, a member of an unconventional cuticle protein family, alters body shape in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16794-16799, (2006).

291 Kang, B. H., Xia, F., Pop, R., Dohi, T., Socolovsky, M. & Altieri, D. C. Developmental control of apoptosis by the immunophilin aryl hydrocarbon receptor-interacting protein (AIP) involves mitochondrial import of the survivin protein. *The Journal of Biological Chemistry* **286**, 16758-16767, (2011).

292 Fedorov, Y., Anderson, E. M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leake, D., Marshall, W. S. & Khvorova, A. Off-target effects by siRNA can induce toxic phenotype. *RNA* **12**, 1188-1196, (2006).

293 Moffat, J., Reiling, J. H. & Sabatini, D. M. Off-target effects associated with long dsRNAs in Drosophila RNAi screens. *Trends in Pharmacological Sciences* **28**, 149-151, (2007).

294 Kulkarni, M. M., Booker, M., Silver, S. J., Friedman, A., Hong, P., Perrimon, N. & Mathey-Prevot, B. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. *Nature Methods* **3**, 833-838, (2006).

295 Seinen, E., Burgerhof, J. G., Jansen, R. C. & Sibon, O. C. RNAi-induced off-target effects in Drosophila melanogaster: frequencies and solutions. *Briefings in Functional Genomics* **10**, 206-214, (2011).

296 Adams, M. D. & Sekelsky, J. J. From sequence to phenotype: reverse genetics in Drosophila melanogaster. *Nature Review Genetics* **3**, 189-198, (2002).

297 Levitt, P. S., Liu, H., Manning, C. & Weiss, R. S. Conditional inactivation of the mouse Hus1 cell cycle checkpoint gene. *Genomics* **86**, 212-224, (2005).

298 Gloor, G. B., Moretti, J., Mouyal, J. & Keeler, K. J. Distinct P-element excision products in somatic and germline cells of Drosophila melanogaster. *Genetics* **155**, 1821-1830, (2000).

299 Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **29**, e45, (2001).

300 Bustin, S. A. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology* **29**, 23-39, (2002).

301 Van Hiel, M. B., Van Wielendaele, P., Temmerman, L., Van Soest, S., Vuerinckx, K., Huybrechts, R., Broeck, J. V. & Simonet, G. Identification and validation of housekeeping genes in brains of the desert locust Schistocerca gregaria under different developmental conditions. *BMC Molecular Biology* **10**, 56, (2009).

302 Gentile, C., Lima, J. & Peixoto, A. Isolation of a fragment homologous to the rp49 constitutive gene of Drosophila in the Neotropical malaria vector Anopheles aquasalis (Diptera: Culicidae). *Memórias do Instituto Oswaldo Cruz* **100**, 545-547, (2005).

303 Mayer, C. & Grummt, I. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**, 6384-6391, (2006).

304 Rudra, D. & Warner, J. R. What better measure than ribosome synthesis? *Genes & Development* **18**, 2431-2436, (2004).

305 Boehlke, K. W. & Friesen, J. D. Cellular content of ribonucleic acid and protein in Saccharomyces cerevisiae as a function of exponential growth rate: calculation of the apparent peptide chain elongation rate. *Journal of Bacteriology* **121**, 429-433, (1975).

306 Xu, R., Deng, K., Zhu, Y., Wu, Y., Ren, J., Wan, M., Zhao, S., Wu, X., Han, M., Zhuang, Y. & Xu, T. A large-scale functional approach to uncover human genes and pathways in Drosophila. *Cell Research* **18**, 1114-1127, (2008).

307 de Moed, G. H., De Jong, G. & Scharloo, W. Environmental effects on body size variation in Drosophila melanogaster and its cellular basis. *Genetical Research* **70**, 35-43, (1997).

308 Schlichting, C. D. & Pigliucci, M. Phenotypic evolution: a reaction norm perspective. Sinauer Associates Incorporated, (1998).

309 Testa, N. D., Ghosh, S. M. & Shingleton, A. W. Sex-specific weight loss mediates sexual size dimorphism in Drosophila melanogaster. *PLoS One* **8**, e58936, (2013).

310 Wilhelm, E., Pellay, F. X., Benecke, A. & Bell, B. Determining the impact of alternative splicing events on transcriptome dynamics. *BMC Research Notes* **1**, 94, (2008).

311 Torales, S. L., Rivarola, M., Pomponio, M. F., Gonzalez, S., Acuna, C. V., Fernandez, P., Lauenstein, D. L., Verga, A. R., Hopp, H. E., Paniego, N. B. & Poltri, S. N. De novo assembly and characterization of leaf transcriptome for the development of functional molecular markers of the extremophile multipurpose tree species Prosopis alba. *BMC Genomics* **14**, 705, (2013).

312 Fan, X., Tang, X., Yan, J. & Xie, J. Identification of idiosyncratic Mycobacterium tuberculosis ribosomal protein subunits with implications in extraribosomal function,

persistence, and drug resistance based on transcriptome data. *Journal of Biomolecular Structure & Dynamics* **32**, 1546-1551, (2014).

313 Vidal, E. A., Moyano, T. C., Krouk, G., Katari, M. S., Tanurdzic, M., McCombie, W. R., Coruzzi, G. M. & Gutierrez, R. A. Integrated RNA-seq and sRNA-seq analysis identifies novel nitrate-responsive genes in Arabidopsis thaliana roots. *BMC Genomics* **14**, 701, (2013).

314 Graveley, B. R. Tips for preparing mRNA-Seq libraries from poly(A)+ mRNA for Illumina transcriptome high-throughput sequencing. *Cold Spring Harbor Protocols* **2013**, 482-484, (2013).

315 Kogenaru, S., Qing, Y., Guo, Y. & Wang, N. RNA-seq and microarray complement each other in transcriptome profiling. *BMC Genomics* **13**, 629, (2012).

316 Hitzemann, R., Bottomly, D., Darakjian, P., Walter, N., Iancu, O., Searles, R., Wilmot, B. & McWeeney, S. Genes, behavior and next-generation RNA sequencing. *Genes, Brain, and Behavior* **12**, 1-12, (2013).

317 Chintapalli, V. R., Wang, J. & Dow, J. A. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nature Genetics* **39**, 715-720, (2007).

318 Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A. & Rinn, J. L. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development* **25**, 1915-1927, (2011).

319 Cherbas, L., Willingham, A., Zhang, D., Yang, L., Zou, Y., Eads, B. D., Carlson, J. W., Landolin, J. M., Kapranov, P., Dumais, J., Samsonova, A., Choi, J. H., Roberts, J., Davis, C. A., Tang, H., van Baren, M. J., Ghosh, S., Dobin, A., Bell, K., Lin, W., Langton, L., Duff, M. O., Tenney, A. E., Zaleski, C., Brent, M. R., Hoskins, R. A., Kaufman, T. C., Andrews, J., Graveley, B. R., Perrimon, N., Celniker, S. E., Gingeras, T. R. & Cherbas, P. The transcriptional diversity of 25 Drosophila cell lines. *Genome Research* **21**, 301-314, (2011).

320 Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **27**, 2325-2329, (2011).

321 Young, R. S., Marques, A. C., Tibbit, C., Haerty, W., Bassett, A. R., Liu, J. L. & Ponting, C. P. Identification and properties of 1,119 candidate lincRNA loci in the Drosophila melanogaster genome. *Genome Biology and Evolution* **4**, 427-442, (2012).

322 Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. & Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* **28**, 511-515, (2010).

323 Graveley, B. R., Brooks, A. N., Carlson, J. W., Duff, M. O., Landolin, J. M., Yang, L., Artieri, C. G., van Baren, M. J., Boley, N., Booth, B. W., Brown, J. B., Cherbas, L., Davis, C. A., Dobin, A., Li, R., Lin, W., Malone, J. H., Mattiuzzo, N. R., Miller, D., Sturgill, D., Tuch, B. B., Zaleski, C., Zhang, D., Blanchette, M., Dudoit, S., Eads, B., Green, R. E., Hammonds, A., Jiang, L., Kapranov, P., Langton, L., Perrimon, N., Sandler, J. E., Wan, K. H., Willingham, A., Zhang, Y., Zou, Y., Andrews, J., Bickel, P. J., Brenner, S. E., Brent, M. R., Cherbas, P., Gingeras, T. R., Hoskins, R. A., Kaufman, T. C., Oliver, B. & Celniker, S. E. The developmental transcriptome of Drosophila melanogaster. *Nature* **471**, 473-479, (2011).

324 Meyer, B. K. & Perdew, G. H. Characterization of the AhR-hsp90-XAP2 core complex and the role of the immunophilin-related protein XAP2 in AhR stabilization. *Biochemistry* **38**, 8907-8917, (1999).

325 Jain, S., Maltepe, E., Lu, M. M., Simon, C. & Bradfield, C. A. Expression of ARNT, ARNT2, HIF1 alpha, HIF2 alpha and Ah receptor mRNAs in the developing mouse. *Mechanisms of Development* **73**, 117-123, (1998).

326 Motojima, K. & Hirai, T. Peroxisome proliferator-activated receptor alpha plays a vital role in inducing a detoxification system against plant compounds with crosstalk with other xenobiotic nuclear receptors. *The FEBS Journal* **273**, 292-300, (2006).

327 Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, J. & Rothberg, J. M. A protein interaction map of Drosophila melanogaster. *Science* **302**, 1727-1736, (2003).

328 Lindsley, D. L., Sandler, L., Baker, B. S., Carpenter, A. T., Denell, R. E., Hall, J. C., Jacobs, P. A., Miklos, G. L., Davis, B. K., Gethmann, R. C., Hardy, R. W., Steven, A. H., Miller, M., Nozawa, H., Parry, D. M., Gould-Somero, M. & Gould-Somero, M. Segmental aneuploidy and the genetic gross structure of the Drosophila genome. *Genetics* **71**, 157-184, (1972).

329 Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. & Bateman, A. The Pfam protein families database. *Nucleic acids research* **38**, D211-222, (2010).

330 Smoyer, L. K., Dorer, D. R., Nickerson, K. W. & Christensen, A. C. Phenotype of the Triplo-lethal locus of Drosophila melanogaster and its suppression by hyperoxia. *Genetical Research* **82**, 163-170, (2003).

331 Krebs, R. A. & Feder, M. E. Tissue-specific variation in Hsp70 expression and thermal damage in Drosophila melanogaster larvae. *The Journal of Experimental Biology* **200**, 2007-2015, (1997).

332 Roehrdanz, R. L. & Lucchesi, J. C. Mutational Events in the Triplo- and Haplo-Lethal Region (83de) of the Drosophila melanogaster Genome. *Genetics* **95**, 355-366, (1980).

333 Dorer, D. R., Ezekiel, D. H. & Christensen, A. C. The Triplo-lethal locus of Drosophila: reexamination of mutants and discovery of a second-site suppressor. *Genetics* **141**, 1037-1042, (1995).

334 Eissenberg, J. C., Ma, J., Gerber, M. A., Christensen, A., Kennison, J. A. & Shilatifard, A. dELL is an essential RNA polymerase II elongation factor with a general role in development. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9894-9899, (2002).

335 Reines, D., Conaway, R. C. & Conaway, J. W. Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Current Opinion in Cell Biology* **11**, 342-346, (1999).

336 Shilatifard, A. Identification and purification of the Holo-ELL complex. Evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL. *The Journal of Biological Chemistry* **273**, 11212-11217, (1998).

337 Lindquist, S. in *Translational Regulation of Gene Expression 2* ed Joseph Ilan Ch. 14, 279-320 (1993).

338 Rebers, J. E. & Riddiford, L. M. Structure and expression of a Manduca sexta larval cuticle gene homologous to Drosophila cuticle genes. *Journal of Molecular Biology* **203**, 411-423, (1988).

339 Ostrowski, S., Dierick, H. A. & Bejsovec, A. Genetic control of cuticle formation during embryonic development of Drosophila melanogaster. *Genetics* **161**, 171-182, (2002).

340 Vasilev, V., Daly, A. F., Petrossians, P., Zacharieva, S. & Beckers, A. Familial Pituitary Tumor Syndromes. *Endocrine Practice*, 1-16, (2011).

341 Jaffrain-Rea, M. L., Angelini, M., Gargano, D., Tichomirowa, M. A., Daly, A. F., Vanbellinghen, J. F., D'Innocenzo, E., Barlier, A., Giangaspero, F., Esposito, V., Ventura, L., Arcella, A., Theodoropoulou, M., Naves, L. A., Fajardo, C., Zacharieva, S., Rohmer, V., Brue, T., Gulino, A., Cantore, G., Alesse, E. & Beckers, A. Expression of aryl hydrocarbon receptor (AHR) and AHR-interacting protein in pituitary adenomas: pathological and clinical implications. *Endocrine-Related Cancer* **16**, 1029-1043, (2009).

342 Georgitsi, M., de, M. E., Cannavo, S., Makinen, M. J., Tuppurainen, K., Pauletto, P., Curto, L., Weil, R. J., Paschke, R., Zielinski, G., Wasik, A., Lubinski, J., Vahteristo, P., Karhu, A. & Aaltonen, L. A. Aryl hydrocarbon receptor interacting protein (AIP) gene mutation analysis in children and adolescents with sporadic pituitary adenomas. *Clinical Endocrinology* **69**, 621-627, (2008).

343 Toledo, R. A., Mendonca, B. B., Fragoso, M. C., Soares, I. C., Almeida, M. Q., Moraes, M. B., Lourenco-Jr, D. M., Alves, V. A., Bronstein, M. D. & Toledo, S. P. Isolated familial somatotropinoma: 11q13-loh and gene/protein expression analysis suggests a possible involvement of aip also in non-pituitary tumorigenesis. *Clinics (Sao Paulo)* **65**, 407-415, (2010).

344 Barlier, A., Vanbellinghen, J. F., Daly, A. F., Silvy, M., Jaffrain-Rea, M. L., Trouillas, J., Tamagno, G., Cazabat, L., Bours, V., Brue, T., Enjalbert, A. & Beckers, A. Mutations in the aryl hydrocarbon receptor interacting protein gene are not highly prevalent among subjects with sporadic pituitary adenomas. *The Journal of Clinical Endocrinology and Metabolism* **92**, 1952-1955, (2007).

345 Dominguez-Gimenez, P., Brown, N. H. & Martin-Bermudo, M. D. Integrin-ECM interactions regulate the changes in cell shape driving the morphogenesis of the Drosophila wing epithelium. *Journal of Cell Science* **120**, 1061-1071, (2007).

346 Fristrom, D., Wilcox, M. & Fristrom, J. The distribution of PS integrins, laminin A and F-actin during key stages in Drosophila wing development. *Development* **117**, 509-523, (1993).

347 Murray, M. A., Schubiger, M. & Palka, J. Neuron differentiation and axon growth in the developing wing of Drosophila melanogaster. *Developmental Biology* **104**, 259-273, (1984).

348 Fristrom, D. & Liebrich, W. The hormonal coordination of cuticulin deposition and morphogenesis in Drosophila imaginal discs in vivo and in vitro. *Developmental Biology* **114**, 1-11, (1986).

349 Bainbridge, S. P. & Bownes, M. Staging the metamorphosis of Drosophila melanogaster. *Journal of Embryology and Experimental Morphology* **66**, 57-80, (1981).

350 Johnson, S. A. & Milner, M. J. The final stages of wing development in Drosophila melanogaster. *Tissue Cell* **19**, 505-513, (1987).

351 Waddington, C. H. Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **25**, 299-307, (1939).

352 Classen, A. K., Aigouy, B., Giangrande, A. & Eaton, S. Imaging Drosophila pupal wing morphogenesis. *Methods in Molecular Biology* **420**, 265-275, (2008).

353 Wilcox, M., Brower, D. L. & Smith, R. J. A position-specific cell surface antigen in the drosophila wing imaginal disc. *Cell* **25**, 159-164, (1981).

354 Brower, D. L., Wilcox, M., Piovant, M., Smith, R. J. & Reger, L. A. Related cell-surface antigens expressed with positional specificity in Drosophila imaginal discs. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 7485-7489, (1984).

355 Tadokoro, S., Shattil, S. J., Eto, K., Tai, V., Liddington, R. C., de Pereda, J. M., Ginsberg, M. H. & Calderwood, D. A. Talin binding to integrin beta tails: a final common step in integrin activation. *Science* **302**, 103-106, (2003).

356 Golic, K. G. & Lindquist, S. The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. *Cell* **59**, 499-509, (1989).

357 Wong, A. M., Wang, J. W. & Axel, R. Spatial representation of the glomerular map in the Drosophila protocerebrum. *Cell* **109**, 229-241, (2002).

358 Theodosiou, N. A. & Xu, T. Use of FLP/FRT system to study Drosophila development. *Methods* **14**, 355-365, (1998).

359 Wehr, M. C., Holder, M. V., Gailite, I., Saunders, R. E., Maile, T. M., Ciirdaeva, E., Instrell, R., Jiang, M., Howell, M., Rossner, M. J. & Tapon, N. Salt-inducible kinases regulate growth through the Hippo signalling pathway in Drosophila. *Nature Cell Biology*, (2012).

360 Baker, S. E., Lorenzen, J. A., Miller, S. W., Bunch, T. A., Jannuzi, A. L., Ginsberg, M. H., Perkins, L. A. & Brower, D. L. Genetic interaction between integrins and moleskin, a gene encoding a Drosophila homolog of importin-7. *Genetics* **162**, 285-296, (2002).

361 Walsh, E. P. & Brown, N. H. A screen to identify Drosophila genes required for integrin-mediated adhesion. *Genetics* **150**, 791-805, (1998).

362 Lee, S. B., Cho, K. S., Kim, E. & Chung, J. blistery encodes Drosophila tensin protein and interacts with integrin and the JNK signaling pathway during wing development. *Development* **130**, 4001-4010, (2003).

363 Katanayeva, N., Kopein, D., Portmann, R., Hess, D. & Katanaev, V. L. Competing activities of heterotrimeric G proteins in Drosophila wing maturation. *PLoS One* **5**, e12331, (2010).

364 Prout, M., Damania, Z., Soong, J., Fristrom, D. & Fristrom, J. W. Autosomal mutations affecting adhesion between wing surfaces in Drosophila melanogaster. *Genetics* **146**, 275-285, (1997).

365 Fristrom, D., Gotwals, P., Eaton, S., Kornberg, T. B., Sturtevant, M., Bier, E. & Fristrom, J. W. Blistered: a gene required for vein/intervein formation in wings of Drosophila. *Development* **120**, 2661-2671, (1994).

366 Bilousov, O., Koval, A., Keshelava, A. & Katanaev, V. L. Identification of novel elements of the Drosophila blisterome sheds light on potential pathological mechanisms of several human diseases. *PLoS One* **9**, e101133, (2014).

367 Stark, K. A., Yee, G. H., Roote, C. E., Williams, E. L., Zusman, S. & Hynes, R. O. A novel alpha integrin subunit associates with betaPS and functions in tissue morphogenesis and movement during Drosophila development. *Development* **124**, 4583-4594, (1997).

368 Prokop, A., Martin-Bermudo, M. D., Bate, M. & Brown, N. H. Absence of PS integrins or laminin A affects extracellular adhesion, but not intracellular assembly, of hemiadherens and neuromuscular junctions in Drosophila embryos. *Developmental Biology* **196**, 58-76, (1998).

369 Bokel, C. & Brown, N. H. Integrins in development: moving on, responding to, and sticking to the extracellular matrix. *Developmental cell* **3**, 311-321, (2002).

370 Urbano, J. M., Torgler, C. N., Molnar, C., Tepass, U., Lopez-Varea, A., Brown, N. H., de Celis, J. F. & Martin-Bermudo, M. D. Drosophila laminins act as key regulators of basement membrane assembly and morphogenesis. *Development* **136**, 4165-4176, (2009).

371 Kimura, K., Kodama, A., Hayasaka, Y. & Ohta, T. Activation of the cAMP/PKA signaling pathway is required for post-ecdysial cell death in wing epidermal cells of Drosophila melanogaster. *Development* **131**, 1597-1606, (2004).

372 Lubarsky, B. & Krasnow, M. A. Tube morphogenesis: making and shaping biological tubes. *Cell* **112**, 19-28, (2003).

373 Schock, F. & Perrimon, N. Molecular mechanisms of epithelial morphogenesis. *Annual Review of Cell and Developmental Biology* **18**, 463-493, (2002).

374 Gumbiner, B. M. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell* **84**, 345-357, (1996).

375 Tepass, U., Tanentzapf, G., Ward, R. & Fehon, R. Epithelial cell polarity and cell junctions in Drosophila. *Annual Review of Genetics* **35**, 747-784, (2001).

376 Rozario, T. & DeSimone, D. W. The extracellular matrix in development and morphogenesis: a dynamic view. *Developmental Biology* **341**, 126-140, (2010).

377 Hynes, R. O. The extracellular matrix: not just pretty fibrils. *Science* **326**, 1216-1219, (2009).

378 Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *Journal of Cell Science* **123**, 4195-4200, (2010).

379 Hynes, R. O. Integrins: bidirectional, allosteric signaling machines. *Cell* **110**, 673-687, (2002).

380 Delon, I. & Brown, N. H. Integrins and the actin cytoskeleton. *Cell Biol Current Opinion in Cell Biology* **19**, 43-50, (2007).

381 Brown, N. H. Cell-cell adhesion via the ECM: integrin genetics in fly and worm. *Matrix biology* **19**, 191-201, (2000).

382 Levental, K. R., Yu, H., Kass, L., Lakins, J. N., Egeblad, M., Erler, J. T., Fong, S. F., Csiszar, K., Giaccia, A., Weninger, W., Yamauchi, M., Gasser, D. L. & Weaver, V. M. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* **139**, 891-906, (2009).

383 Paszek, M. J., Zahir, N., Johnson, K. R., Lakins, J. N., Rozenberg, G. I., Gefen, A., Reinhart-King, C. A., Margulies, S. S., Dembo, M., Boettiger, D., Hammer, D. A. & Weaver, V. M. Tensional homeostasis and the malignant phenotype. *Cancer Cell* **8**, 241-254, (2005).

384 Sternlicht, M. D., Lochter, A., Sympson, C. J., Huey, B., Rougier, J. P., Gray, J. W., Pinkel, D., Bissell, M. J. & Werb, Z. The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. *Cell* **98**, 137-146, (1999).

385 Lu, P., Weaver, V. M. & Werb, Z. The extracellular matrix: A dynamic niche in cancer progression. *The Journal of Cell Biology* **196**, 395-406, (2012).

386 Wozniak, M. A., Desai, R., Solski, P. A., Der, C. J. & Keely, P. J. ROCK-generated contractility regulates breast epithelial cell differentiation in response to the physical properties of a three-dimensional collagen matrix. *The Journal of Cell Biology* **163**, 583-595, (2003).

387 Zusman, S., Patel-King, R. S., Ffrench-Constant, C. & Hynes, R. O. Requirements for integrins during Drosophila development. *Development* **108**, 391-402, (1990).

388 MacKrell, A. J., Blumberg, B., Haynes, S. R. & Fessler, J. H. The lethal myospheroid gene of Drosophila encodes a membrane protein homologous to vertebrate integrin beta subunits. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 2633-2637, (1988).

389 Brabant, M. C. & Brower, D. L. PS2 integrin requirements in Drosophila embryo and wing morphogenesis. *Developmental Biology* **157**, 49-59, (1993).

390 Brown, N. H. Integrins hold Drosophila together. *Bioessays* **15**, 383-390, (1993).

391 Wilcox, M., DiAntonio, A. & Leptin, M. The function of PS integrins in Drosophila wing morphogenesis. *Development* **107**, 891-897, (1989).

392 Brower, D. L. & Jaffe, S. M. Requirement for integrins during Drosophila wing development. *Nature* **342**, 285-287, (1989).

393 Clark, K. A., McGrail, M. & Beckerle, M. C. Analysis of PINCH function in Drosophila demonstrates its requirement in integrin-dependent cellular processes. *Development* **130**, 2611-2621, (2003).

394 Zervas, C. G., Gregory, S. L. & Brown, N. H. Drosophila integrin-linked kinase is required at sites of integrin adhesion to link the cytoskeleton to the plasma membrane. *The Journal of Cell Biology* **152**, 1007-1018, (2001).

395 Wagner, C. R., Mahowald, A. P. & Miller, K. G. One of the two cytoplasmic actin isoforms in Drosophila is essential. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 8037-8042, (2002).

396 Anson, M., Drummond, D. R., Geeves, M. A., Hennessey, E. S., Ritchie, M. D. & Sparrow, J. C. Actomyosin kinetics and in vitro motility of wild-type Drosophila actin and the effects of two mutations in the Act88F gene. *Biophysical Journal* **68**, 1991-2003, (1995).

397 Nonaka, S., Nagaosa, K., Mori, T., Shiratsuchi, A. & Nakanishi, Y. Integrin alphaPS3/betanu-mediated phagocytosis of apoptotic cells and bacteria in Drosophila. *The Journal of Biological Chemistry* **288**, 10374-10380, (2013).

398 van der Flier, A. & Sonnenberg, A. Structural and functional aspects of filamins. *Biochimica et Biophysica Acta* **1538**, 99-117, (2001).

399 Zaidel-Bar, R., Milo, R., Kam, Z. & Geiger, B. A paxillin tyrosine phosphorylation switch regulates the assembly and form of cell-matrix adhesions. *Journal of Cell Science* **120**, 137-148, (2007).

400 Wolfenson, H., Henis, Y. I., Geiger, B. & Bershadsky, A. D. The heel and toe of the cell's foot: a multifaceted approach for understanding the structure and dynamics of focal adhesions. *Cell Motility and the Cytoskeleton* **66**, 1017-1029, (2009).

401 Miranti, C. K. & Brugge, J. S. Sensing the environment: a historical perspective on integrin signal transduction. *Nature Cell Biology* **4**, E83-90, (2002).

402 Juliano, R. L. & Haskill, S. Signal transduction from the extracellular matrix. *The Journal of Cell Biology* **120**, 577-585, (1993).

403 Ginsberg, M. H., Partridge, A. & Shattil, S. J. Integrin regulation. *Current Opinion in Cell Biology* **17**, 509-516, (2005).

404 Giannone, G. & Sheetz, M. P. Substrate rigidity and force define form through tyrosine phosphatase and kinase pathways. *Trends in Cell Biology* **16**, 213-223, (2006).

405 Welch, M. D. & Mullins, R. D. Cellular control of actin nucleation. *Annual Review of Cell and Developmental Biology* **18**, 247-288, (2002).

406 Revenu, C., Athman, R., Robine, S. & Louvard, D. The co-workers of actin filaments: from cell structures to signals. *Nature reviews. Molecular Cell Biology* **5**, 635-646, (2004).

407 Yamaguchi, H. & Condeelis, J. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochimica et Biophysica Acta* **1773**, 642-652, (2007).

408 Campbell, I. D. & Ginsberg, M. H. The talin-tail interaction places integrin activation on FERM ground. *Trends in Biochemical Sciences* **29**, 429-435, (2004).

409 Wu, C. & Dedhar, S. Integrin-linked kinase (ILK) and its interactors: a new paradigm for the coupling of extracellular matrix to actin cytoskeleton and signaling complexes. *The Journal of Cell Biology* **155**, 505-510, (2001).

410 Wickstrom, S. A., Lange, A., Montanez, E. & Fassler, R. The ILK/PINCH/parvin complex: the kinase is dead, long live the pseudokinase! *The EMBO Journal* **29**, 281-291, (2010).

411 Hannigan, G. E., Leung-Hagesteijn, C., Fitz-Gibbon, L., Coppolino, M. G., Radeva, G., Filmus, J., Bell, J. C. & Dedhar, S. Regulation of cell adhesion and anchorage-dependent growth by a new beta 1-integrin-linked protein kinase. *Nature* **379**, 91-96, (1996).

412 Legate, K. R., Montanez, E., Kudlacek, O. & Fassler, R. ILK, PINCH and parvin: the tIPP of integrin signalling. *Nature reviews. Molecular Cell Biology* **7**, 20-31, (2006).

413 Wu, C. The PINCH-ILK-parvin complexes: assembly, functions and regulation. *Biochimica et Biophysica Acta* **1692**, 55-62, (2004).

414 Vakaloglou, K. M., Chountala, M. & Zervas, C. G. Functional analysis of parvin and different modes of IPP-complex assembly at integrin sites during Drosophila development. *Journal of Cell Science* **125**, 3221-3232, (2012).

415 Zaidel-Bar, R. Evolution of complexity in the integrin adhesome. *The Journal of Cell Biology* **186**, 317-321, (2009).

416 Bloor, J. W. & Brown, N. H. Genetic analysis of the Drosophila alphaPS2 integrin subunit reveals discrete adhesive, morphogenetic and sarcomeric functions. *Genetics* **148**, 1127-1142, (1998).

417 Volk, T., Fessler, L. I. & Fessler, J. H. A role for integrin in the formation of sarcomeric cytoarchitecture. *Cell* **63**, 525-536, (1990).

418 Bunch, T. A., Graner, M. W., Fessler, L. I., Fessler, J. H., Schneider, K. D., Kerschen, A., Choy, L. P., Burgess, B. W. & Brower, D. L. The PS2 integrin ligand tiggrin is required for proper muscle function in Drosophila. *Development* **125**, 1679-1689, (1998).

419 Fogerty, F. J., Fessler, L. I., Bunch, T. A., Yaron, Y., Parker, C. G., Nelson, R. E., Brower, D. L., Gullberg, D. & Fessler, J. H. Tiggrin, a novel Drosophila extracellular matrix protein that functions as a ligand for Drosophila alpha PS2 beta PS integrins. *Development* **120**, 1747-1758, (1994).

420 Moerman, D. G., Hutter, H., Mullen, G. P. & Schnabel, R. Cell autonomous expression of perlecan and plasticity of cell shape in embryonic muscle of Caenorhabditis elegans. *Developmental Biology* **173**, 228-242, (1996).

421 Burridge, K., Chrzanowska-Wodnicka, M. & Zhong, C. Focal adhesion assembly. *Trends in Cell Biology* **7**, 342-347, (1997).

422 Newman, S. M., Jr. & Wright, T. R. A histological and ultrastructural analysis of developmental defects produced by the mutation, lethal(1)myospheroid, in Drosophila melanogaster. *Developmental Biology* **86**, 393-402, (1981).

423 Sturtevant, M. A. & Bier, E. Analysis of the genetic hierarchy guiding wing vein development in Drosophila. *Development* **121**, 785-801, (1995).

424 Noll, R., Sturtevant, M. A., Gollapudi, R. R. & Bier, E. New functions of the Drosophila rhomboid gene during embryonic and adult development are revealed by a novel genetic method, enhancer piracy. *Development* **120**, 2329-2338, (1994).

425 Sturtevant, M. A., Roark, M. & Bier, E. The Drosophila rhomboid gene mediates the localized formation of wing veins and interacts genetically with components of the EGF-R signaling pathway. *Genes & Development* **7**, 961-973, (1993).

426 O'Keefe, D. D., Prober, D. A., Moyle, P. S., Rickoll, W. L. & Edgar, B. A. Egfr/Ras signaling regulates DE-cadherin/Shotgun localization to control vein morphogenesis in the Drosophila wing. *Developmental Biology* **311**, 25-39, (2007).

427 D'Souza-Schorey, C. Disassembling adherens junctions: breaking up is hard to do. *Trends in Cell Biology* **15**, 19-26, (2005).

428 Dumstrei, K., Wang, F., Shy, D., Tepass, U. & Hartenstein, V. Interaction between EGFR signaling and DE-cadherin during nervous system morphogenesis. *Development* **129**, 3983-3994, (2002).

429 Kumar, A., Gupta, T., Berzsenyi, S. & Giangrande, A. N-cadherin negatively regulates collective Drosophila glial migration through actin cytoskeleton remodeling. *Journal of Cell Science* **128**, 900-912, (2015).

430 Vlotides, G., Siegel, E., Donangelo, I., Gutman, S., Ren, S. G. & Melmed, S. Rat prolactinoma cell growth regulation by epidermal growth factor receptor ligands. *Cancer Research* **68**, 6377-6386, (2008).

431 Guarino, M., Rubino, B. & Ballabio, G. The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* **39**, 305-318, (2007).

432 Lo, H. W. EGFR-targeted therapy in malignant glioma: novel aspects and mechanisms of drug resistance. *Current Molecular Pharmacology* **3**, 37-52, (2010).

433 Desgrosellier, J. S. & Cheresh, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nature Reviews Cancer* **10**, 9-22, (2010).

434 Stupp, R., Hegi, M. E., Gorlia, T., Erridge, S. C., Perry, J., Hong, Y. K., Aldape, K. D., Lhermitte, B., Pietsch, T., Grujicic, D., Steinbach, J. P., Wick, W., Tarnawski, R., Nam, D. H., Hau, P., Weyerbrock, A., Taphoorn, M. J., Shen, C. C., Rao, N., Thurzo, L., Herrlinger, U., Gupta, T., Kortmann, R. D., Adamska, K., McBain, C., Brandes, A. A., Tonn, J. C., Schnell, O., Wiegel, T., Kim, C. Y., Nabors, L. B., Reardon, D. A., van den Bent, M. J., Hicking, C., Markivskyy, A., Picard, M., Weller, M., European Organisation for, R., Treatment of, C., Canadian Brain Tumor, C. & team, C. s. Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *The Lancet. Oncology* **15**, 1100-1108, (2014).

435 Farnoud, M. R., Veirana, N., Derome, P., Peillon, F. & Li, J. Y. Adenomatous transformation of the human anterior pituitary is associated with alterations in integrin expression. *International Journal of Cancer* **67**, 45-53, (1996).

436 Tsutsumi, S., Beebe, K. & Neckers, L. Impact of heat-shock protein 90 on cancer metastasis. *Future Oncology* **5**, 679-688, (2009).

437 Murray, A. W. Cell-cycle control: turning on mitosis. *Current biology* **3**, 291-293, (1993).

438 Simanis, V., Carr, A. M., Goss, M., Lee, M. G., MacNeill, S. A. & Nurse, P. Cell cycle regulation in yeasts and man: towards a unifying mechanism. *Antonie Van Leeuwenhoek* **53**, 319-323, (1987).

439 Marshall, N. F., Peng, J., Xie, Z. & Price, D. H. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *The Journal of Biological Chemistry* **271**, 27176-27183, (1996).

440 Dahmus, M. E. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *The Journal of Biological Chemistry* **271**, 19009-19012, (1996).

441 Yang, Z., He, N. & Zhou, Q. Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression. *Molecular and Cellular Biology* **28**, 967-976, (2008).

442 Wang, S. & Fischer, P. M. Cyclin-dependent kinase 9: a key transcriptional regulator and potential drug target in oncology, virology and cardiology. *Trends in Pharmacological Sciences* **29**, 302-313, (2008).

443 Bettayeb, K., Baunbaek, D., Delehouze, C., Loaec, N., Hole, A. J., Baumli, S., Endicott, J. A., Douc-Rasy, S., Benard, J., Oumata, N., Galons, H. & Meijer, L. CDK Inhibitors Roscovitine and CR8 Trigger Mcl-1 Down-Regulation and Apoptotic Cell Death in Neuroblastoma Cells. *Genes & Cancer* **1**, 369-380, (2010).

444 Kang, B. H. & Altieri, D. C. Regulation of survivin stability by the aryl hydrocarbon receptor-interacting protein. *The Journal of Biological Chemistry* **281**, 24721-24727, (2006).

445 Altieri, D. C. New wirings in the survivin networks. *Oncogene* **27**, 6276-6284, (2008).

446 Zhu, Y., Li, Y., Zhu, S., Tang, R., Liu, Y. & Li, J. Association of survivin polymorphisms with tumor susceptibility: a meta-analysis. *PLoS One* **8**, e74778, (2013).

447 Guo, G., Zhang, Q., Yu, Z., Li, J., Ding, Z., Li, J. & Tan, W. Correlation between survivin genetic polymorphisms and lung cancer susceptibility. *International Journal of Clinical and Experimental Pathology* **8**, 7426-7430, (2015).

448 Mita, A. C., Mita, M. M., Nawrocki, S. T. & Giles, F. J. Survivin: key regulator of mitosis and apoptosis and novel target for cancer therapeutics. *Clinical cancer research: an official journal of the American Association for Cancer Research* **14**, 5000-5005, (2008).

449 Schofl, C., Honegger, J., Droste, M., Grussendorf, M., Finke, R., Plockinger, U., Berg, C., Willenberg, H. S., Lammert, A., Klingmuller, D., Jaursch-Hancke, C., Tonjes, A., Schneidewind, S., Flitsch, J., Bullmann, C., Dimopoulou, C., Stalla, G., Mayr, B., Hoeppner, W. & Schopohl, J. Frequency of AIP gene mutations in young patients with acromegaly: a registry-based study. *The Journal of Clinical Endocrinology & Metabolism* **99**, E2789-2793, (2014).

450 Krawczak, M., Ball, E. V. & Cooper, D. N. Neighboring-nucleotide effects on the rates of germ-line single-base-pair substitution in human genes. *American Journal of Human Genetics* **63**, 474-488, (1998).

451 Occhi, G., Jaffrain-Rea, M. L., Trivellin, G., Albiger, N., Ceccato, F., De Menis, E., Angelini, M., Ferasin, S., Beckers, A., Mantero, F. & Scaroni, C. The R304X mutation of the aryl hydrocarbon receptor interacting protein gene in familial isolated pituitary adenomas: Mutational hot-spot or founder effect? *Journal of Endocrinological Investigation* **33**, 800-805, (2010).

452 Chahal, H. S., Stals, K., Unterlander, M., Balding, D. J., Thomas, M. G., Kumar, A. V., Besser, G. M., Atkinson, A. B., Morrison, P. J., Howlett, T. A., Levy, M. J., Orme, S. M., Akker, S. A., Abel, R. L., Grossman, A. B., Burger, J., Ellard, S. & Korbonits, M. AIP mutation in pituitary adenomas in the 18th century and today. *The New England Journal of Medicine* **364**, 43-50, (2011).

453 Thusberg, J., Olatubosun, A. & Vihinen, M. Performance of mutation pathogenicity prediction methods on missense variants. *Human Mutations* **32**, 358-368, (2011).

454 Martucci, F., Trivellin, G. & Korbonits, M. Familial isolated pituitary adenomas: an emerging clinical entity. *Journal of Endocrinological Investigation* **35**, 1003-1014, (2012).

455 Yarman, S., Ogret, Y. D. & Oguz, F. S. Do the aryl hydrocarbon receptor interacting protein variants (Q228K and Q307R) play a role in patients with familial and sporadic hormone-secreting pituitary adenomas? *Genetic testing and molecular biomarkers* **19**, 394-398, (2015).

456 Cazabat, L., Libe, R., Perlemoine, K., Rene-Corail, F., Burnichon, N., Gimenez-Roqueplo, A. P., Dupasquier-Fediaevsky, L., Bertagna, X., Clauser, E., Chanson, P., Bertherat, J. & Raffin-Sanson, M. L. Germline inactivating mutations of the aryl hydrocarbon receptor-interacting protein gene in a large cohort of sporadic acromegaly: mutations are found in a subset of young patients with macroadenomas. *European Journal of Endocrinology* **157**, 1-8, (2007).

457 Rowlands, J. C., Urban, J. D., Wikoff, D. S. & Budinsky, R. A. An evaluation of single nucleotide polymorphisms in the human aryl hydrocarbon receptor-interacting protein (AIP) gene. *Drug Metabolism and Pharmacokinetics* **26**, 431-439, (2011).

458 Yamamoto, S., Jaiswal, M., Charng, W. L., Gambin, T., Karaca, E., Mirzaa, G., Wiszniewski, W., Sandoval, H., Haelterman, N. A., Xiong, B., Zhang, K., Bayat, V., David, G., Li, T., Chen, K., Gala, U., Harel, T., Pehlivan, D., Penney, S., Vissers, L. E., de Ligt, J., Jhangiani, S. N., Xie, Y., Tsang, S. H., Parman, Y., Sivaci, M., Battaloglu, E., Muzny, D., Wan, Y. W., Liu, Z., Lin-Moore, A. T., Clark, R. D., Curry, C. J., Link, N., Schulze, K. L., Boerwinkle, E., Dobyns, W. B., Allikmets, R., Gibbs, R. A., Chen, R., Lupski, J. R., Wangler, M. F. & Bellen, H. J. A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. *Cell* **159**, 200-214, (2014).

459 Kazlauskas, A., Poellinger, L. & Pongratz, I. Two distinct regions of the immunophilin-like protein XAP2 regulate dioxin receptor function and interaction with hsp90. *The Journal of Biological Chemistry* **277**, 11795-11801, (2002).

460 Guaraldi, F. & Salvatori, R. Familial isolated pituitary adenomas: from genetics to therapy. *Clinical and Translational Science* **4**, 55-62, (2011).

461 Georgitsi, M., Raitila, A., Karhu, A., Tuppurainen, K., Makinen, M. J., Vierimaa, O., Paschke, R., Saeger, W., van der Luijt, R. B., Sane, T., Robledo, M., de, M. E., Weil, R. J., Wasik, A., Zielinski, G., Lucewicz, O., Lubinski, J., Launonen, V., Vahteristo, P. & Aaltonen, L. A. Molecular diagnosis of pituitary adenoma predisposition caused by aryl hydrocarbon receptor-interacting protein gene mutations. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 4101-4105, (2007).

462 Occhi, G., Jaffrain-Rea, M. L., Trivellin, G., Albiger, N., Ceccato, F., de, M. E., Angelini, M., Ferasin, S., Beckers, A., Mantero, F. & Scaroni, C. The R304X mutation of the aryl hydrocarbon receptor interacting protein gene in familial isolated pituitary adenomas: Mutational hot-spot or founder effect? *Journal of Endocrinological Investigation* **33**, 800-805, (2010).

463 Pardi, E., Marcocci, C., Borsari, S., Saponaro, F., Torregrossa, L., Tancredi, M., Raspini, B., Basolo, F. & Cetani, F. Aryl hydrocarbon receptor interacting protein (AIP) mutations occur rarely in sporadic parathyroid adenomas. *The Journal of Clinical Endocrinology & Metabolism* **98**, 2800-2810, (2013).

464 Wang, J. W., Beck, E. S. & McCabe, B. D. A modular toolset for recombination transgenesis and neurogenetic analysis of Drosophila. *PLoS One* **7**, e42102, (2012).

465 Grossman, T. R., Gamliel, A., Wessells, R. J., Taghli-Lamallem, O., Jepsen, K., Ocorr, K., Korenberg, J. R., Peterson, K. L., Rosenfeld, M. G., Bodmer, R. & Bier, E. Over-expression of DSCAM and COL6A2 cooperatively generates congenital heart defects. *PLoS Genetics* **7**, e1002344, (2011).

466 Ikmi, A., Gaertner, B., Seidel, C., Srivastava, M., Zeitlinger, J. & Gibson, M. C. Molecular evolution of the Yap/Yorkie proto-oncogene and elucidation of its core transcriptional program. *Molecular Biology and Evolution* **31**, 1375-1390, (2014).

467 Kudumala, S., Freund, J., Hortsch, M. & Godenschwege, T. A. Differential effects of human L1CAM mutations on complementing guidance and synaptic defects in Drosophila melanogaster. *PLoS One* **8**, e76974, (2013).

468 Bergmiller, T., Ackermann, M. & Silander, O. K. Patterns of evolutionary conservation of essential genes correlate with their compensability. *PLoS Genetics* **8**, e1002803, (2012).

469 Rocha, E. P. & Danchin, A. An analysis of determinants of amino acids substitution rates in bacterial proteins. *Molecular Biology and Evolution* **21**, 108-116, (2004).

470 Gerdes, S. Y., Scholle, M. D., Campbell, J. W., Balazsi, G., Ravasz, E., Daugherty, M. D., Somera, A. L., Kyrpides, N. C., Anderson, I., Gelfand, M. S., Bhattacharya, A., Kapatral, V., D'Souza, M., Baev, M. V., Grechkin, Y., Mseeh, F., Fonstein, M. Y., Overbeek, R., Barabasi, A. L., Oltvai, Z. N. & Osterman, A. L. Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. *Journal of Bacteriology* **185**, 5673-5684, (2003).

471 Jordan, I. K., Rogozin, I. B., Wolf, Y. I. & Koonin, E. V. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Research* **12**, 962-968, (2002).

472 Makarova, K. S. & Koonin, E. V. Comparative genomics of Archaea: how much have we learned in six years, and what's next? *Genome Biology* **4**, 115, (2003).

473 Mimura, J. & Fujii-Kuriyama, Y. Functional role of AhR in the expression of toxic effects by TCDD. *Biochimica et Biophysica Acta* **1619**, 263-268, (2003).

474 Zatelli, M. C., Torre, M. L., Rossi, R., Ragonese, M., Trimarchi, F., degli Uberti, E. & Cannavo, S. Should aip gene screening be recommended in family members of FIPA patients with R16H variant? *Pituitary* **16**, 238-244, (2013).

475 Raitila, A., Georgitsi, M., Bonora, E., Vargiolu, M., Tuppurainen, K., Makinen, M. J., Vierimaa, O., Salmela, P. I., Launonen, V., Vahteristo, P., Aaltonen, L. A., Romeo, G. & Karhu, A. Aryl hydrocarbon receptor interacting protein mutations seem not to associate with familial non-medullary thyroid cancer. *Journal of Endocrinological Investigation* **32**, 426-429, (2009).

476 *dbSNP (Exome Sequencing Project)*, <http://www.ncbi.nlm.nih.gov/snp>. Accessed 27.08 2015

477 Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. & Sirotkin, K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research* **29**, 308-311, (2001).

478 NHLBI GO Exome Sequencing Project (ESP) *Exome Variant Server (EVS)*, <http://evs.gs.washington.edu/EVS>. Accessed 27.08 2015

479 Hernandez Ramirez, L. C., Martucci, F., Ferrau, F., Morgan, R. M. L., Trivellin, G., Begum, F., Tilley, D., Ramos-Guajardo, N., Iacovazzo, D., Prodromou, C. & Korbonits, M. The enhanced proteasomal degradation of AIP mutant proteins is a mechanism for AIP deficiency in AIP mutation-associated pituitary adenomas. in *The Endocrine Society Annual Meeting ENDO 2015,* .San Diego, USA, March 5-8 (2015).

480 *Exome Aggregation Consortium (ExAC)*, <exac.broadinstitute.org>. Accessed 27.08 2015

481 Williams, F., Hunter, S., Bradley, L., Chahal, H. S., Storr, H. L., Akker, S. A., Kumar, A. V., Orme, S. M., Evanson, J., Abid, N., Morrison, P. J., Korbonits, M. & Atkinson, A. B. Clinical experience in the screening and management of a large kindred with familial isolated pituitary adenoma due to an aryl hydrocarbon receptor interacting protein (AIP) mutation. *The Journal of Clinical Endocrinology & Metabolism* **99**, 1122-1131, (2014).

482 Goldberg, A. L. Protein degradation and protection against misfolded or damaged proteins. *Nature* **426**, 895-899, (2003).

483 Occhi, G., Trivellin, G., Ceccato, F., De Lazzari, P., Giorgi, G., Dematte, S., Grimaldi, F., Castello, R., Davi, M. V., Arnaldi, G., Salviati, L., Opocher, G., Mantero, F. & Scaroni, C. Prevalence of AIP mutations in a large series of sporadic Italian acromegalic patients and evaluation of CDKN1B status in acromegalic patients with multiple endocrine neoplasia. *European Journal of Endocrinology* **163**, 369-376, (2010).

484 Wang, J. W., Brent, J. R., Tomlinson, A., Shneider, N. A. & McCabe, B. D. The ALS-associated proteins FUS and TDP-43 function together to affect Drosophila locomotion and life span. *The Journal of Clinical Investigation* **121**, 4118-4126, (2011).

485 Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genetics* **40**, 476-483, (2008).

486 Bridges, C. B. Partial Sex-Linkage in the Pigeon. *Science* **37**, 112-113, (1913).

487 Bridges, C. B. Non-Disjunction as Proof of the Chromosome Theory of Heredity. *Genetics* **1**, 1-52, (1916).

488 Bridges, C. B. Non-Disjunction as Proof of the Chromosome Theory of Heredity (Concluded). *Genetics* **1**, 107-163, (1916).

489 Orr-Weaver, T. L. Meiosis in Drosophila: seeing is believing. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10443-10449, (1995).

490 McKim, K. S., Jang, J. K. & Manheim, E. A. Meiotic recombination and chromosome segregation in Drosophila females. *Annual Review of Genetics* **36**, 205-232, (2002).

491 Koehler, K. E., Hawley, R. S., Sherman, S. & Hassold, T. Recombination and nondisjunction in humans and flies. *Human Molecular Genetics* **5 Spec No**, 1495-1504, (1996).

492 Xiang, Y. & Hawley, R. S. The mechanism of secondary nondisjunction in Drosophila melanogaster females. *Genetics* **174**, 67-78, (2006).

493 Carpenter, A. T. & Sandler, L. On recombination-defective meiotic mutants in Drosophila melanogaster. *Genetics* **76**, 453-475, (1974).

494 Chahal, H. S., Trivellin, G., Leontiou, C. A., Alband, N., Fowkes, R. C., Tahir, A., Igreja, S. C., Chapple, J. P., Jordan, S., Lupp, A., Schulz, S., Ansorge, O., Karavitaki, N., Carlsen, E., Wass, J. A., Grossman, A. B. & Korbonits, M. Somatostatin analogs modulate AIP in somatotroph adenomas: the role of the ZAC1 pathway. *The Journal of Clinical Endocrinology & Metabolism* **97**, E1411-E1420, (2012).