Exploring biomarkers from the tumour and the microenvironment in Diffuse Large B-cell Lymphoma

Rita Ribeiro Coutinho

Submitted in partial fulfilment of the requirements of the

Degree of Doctor of Philosophy

November 2014

Centre for Haemato-Oncology,

Barts Cancer Institute,

Queen Mary School of Medicine and Dentistry,

Queen Mary University of London

Statement of originality

I, Rita Ribeiro Coutinho, confirm that the investigation described within this thesis is my own work. All work carried out in collaboration is fully acknowledged below. Previously published material is also acknowledged below.

I confirm that I have exerted sufficient care to ensure that the data 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 consent that plagiarism detection software is used to check the electronic version of the document.

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

This work was supported by the portuguese institution *Fundação para a Ciência e Tecnologia*, grant number SFRH / BD / 68462 / 2010.

Details of collaboration:

Andrew Clear and Andrew Owen constructed tissue microarrays from St Bartholomew's Hospital.

Dr Rute Alvarez and Dr Maria Gomes da Silva collected data and Dr José Cabeçadas constructed tissue microarrays from Instituto Português de Oncologia Francisco Gentil (IPO), Lisbon.

Andrew Clear optimized antibodies for immunohistochemistry studies.

Dr Abigail Lee and Dr Maria Calaminici assisted in data analysis for Chapter 3.

Dr Emanuele Mazzola and Dr Donna Neuberg performed statistical analysis for Chapter 4.

William Day and Dr Guglielmo Rosignoli performed flow sorting.

Tracy Chaplin performed the set-up and hybridisation of microarrays.

Jacek Marzec performed bioinformatics analysis of gene expression data.

Dr Aaron Newman and Dr Ash Alizadeh develop bioinformatics-derived macrophage signatures.

Timothy Farren assisted in cytokine bead array experiments.

Details of original publications:

1. Coutinho R, Clear AJ, Owen A, Wilson A, Matthews J, Lee A, Alvarez R, Gomes da Silva M, Cabeçadas J, Calaminici M, Gribben JG. Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma: implications for therapeutic strategies. Clin Cancer Res. 2013 Dec 15;19(24):6686-95. doi: 10.1158/1078-0432.CCR-13-1482. Epub 2013 Oct 11.
(URL: http://clincancerres.aacrjournals.org/content/19/24/6686.full.pdf+html)

2. Coutinho R, Clear AJ, Mazzola E, Owen A, Greaves P, Wilson A, Matthews J, Lee A, Alvarez R, Gomes da Silva M, Cabeçadas J, Neuberg D, Calaminici M and Gribben JG. Revisiting the immune microenvironment of diffuse large B-cell lymphoma using a tissue microarray and immunohistochemistry: robust semi-automated analysis reveals CD3 and FoxP3 as potential predictors of response to R-CHOP. Haematologica. 2015 Mar;100(3):363-9. doi: 10.3324/haematol.2014.110189.
(URL: http://www.haematologica.org/content/haematol/100/3/363.full.pdf)

3. Coutinho R, Pria AD, Gandhi S, Bailey K, Fields P, Cwynarski K, Wilson A, Papanastasopoulos P, Tenant-Flowers M, Webb A, Burns F, Marcus RE, Orkin C, Montoto S, Bower M. HIV status does not impair the outcome of patients diagnosed with diffuse large B-cell lymphoma treated with R-CHOP in the cART era. AIDS. 2014 Mar 13;28(5):689-97. doi: 10.1097/QAD.0000000000000133.
(URL: http://www.ncbi.nlm.nih.gov/pubmed/24418826)

Details of main oral presentations:

1. Coutinho R et al. A diffuse large B-cell lymphoma associated-macrophage transcriptome analysis reveals a unique M1-M2 polarization profile that challenges current concepts on the lymphoma microenvironment. 12-ICML, Lugano. 2013.

2. Coutinho R et al. The Diffuse Large B-Cell Lymphoma Infiltrating Macrophage Transcriptome Signature Is Enriched for Both M1 and M2 Genes and Provides an Excellent Platform for Functional Validation of Macrophage Biology in DLBCL. 54th ASH. 2012.

3. Coutinho R et al. Exploring the Immune Microenvironment of Diffuse Large B Cell Lymphoma in a Tissue Microarray: Predicting Survival with a Score That Incorporates Macrophages, Cytotoxic and Regulatory T Cells. 53rd ASH. 2011.

Signature: Date:

Abstract

In the last decade unprecedented improvement in cure rates and overall survival was achieved in diffuse large B-cell Lymphoma (DLBCL) through the introduction of rituximab and anthracyclin-based chemotherapy (R-CHOP) as first line treatment. However, 40% of patients are refractory or relapse after R-CHOP and are hardly salvaged. To date, only age, International Prognostic Index (IPI) stratification and genetic aberrations defining gray-zone lymphomas have been used in clinical trials to select high-risk patients for more aggressive regimens. However, these prognostic features do not take into account the full biological heterogeneity of DLBCL. This reflects our limited knowledge on comprehensive prognostication in this group of disorders and supports our choice to investigate old and new prognostic factors for DLBCL in this thesis.

Molecular characterization is generating opportunities for personalized therapy in poor-risk DLBCL. In order for targeted therapies to succeed in this disease, reliable and reproducible strategies that adequately segregate patients into distinct molecular groups are needed. While gene expression profiling (GEP) is the gold standard method, there is presently a lack of standardized methodology for array analysis, which can lead to variable results. The lack of a routine methodology for GEP has led investigators to develop immunohistochemistry (IHC) based approaches for the molecular classification in DLBCL. In fact, the Hans algorithm is being used to identify non-GCB DLBCLs in clinical trials offering NF-kB targeting agents to patients with this subtype. By performing a systematic comparison of nine IHC algorithms for molecular classification in a new large dataset of diagnostic DLBCL, we document an extremely low concordance across all classifiers (<21%) when classifying each individual patient, and a lack of outcome impact of all strategies, demonstrating that IHC is not a reliable alternative to molecular-based methods to be used for clinical decisions in DLBCL.

GEP studies also suggested that the microenvironment could provide prognostic biomarkers in DLBCL in the R-CHOP era. Most authors have focused on the use of IHC to enumerate and functionally characterize the microenvironment in DLBCL. In our second study, by comparing two methods of semi-automated analysis for IHC staining

of the microenvironment, we demonstrate that the computerized results are highly reproducible, add the required robustness to IHC studies and should be used in the future instead of manual analysis. By applying comprehensive statistical analysis we propose that CD3 and FoxP3 should be validated as predictors of response to R-CHOP in clinical trials.

Whereas a number of mechanisms by which cancer cells influence macrophage function have been described, currently there is very limited understanding of the macrophage polarisation status and effector function in human DLBCL. In our third study we analysed the GEP of macrophages sorted from human DLBCL samples. Unsupervised hierarchical clustering does not resolve DLBCL macrophage samples from reactive macrophage samples, indicating that macrophage heterogeneity in DLBCL should be considered. 202 genes are differentially expressed in DLBCL relative to controls. Functional annotation supports that these genes are macrophage-specific. We demonstrate that DLBCL macrophages have a bidirectional M1 and M2 functional activation, challenging the concept, widespread in the literature, that macrophages in tumours have a predominant M2 transcriptome.

In our fifth study we used a two-cell co-culture model in an attempt to demonstrate that DLBCL cells influence macrophage transcriptome and proteome. The heterogeneity of the results, which precludes the confirmation of our hypothesis, is fully discussed.

In our last study we tease out the DLBCL macrophage GEP heterogeneity and propose IFN-γ as a culprit B-cell derived molecule influencing macrophage activation status. Finally, using immunofluorescence we demonstrate that both M1 and M2 proteins are expressed in DLBCL macrophages.

Dedication

To my family, who led me here.

To Nuno, for unconditional love and support.

To Clara, my new wonder.

Acknowledgements

The work presented in this thesis was possible due to funding provided by Fundação para a Ciência e Tecnologia (reference SFRH / BD / 68462 / 2010).

I would sincerely like to thank Professor John Gribben for accepting to act as my primary supervisor and for helping to delineate the conceptual framework of our grant proposal. During these four years John provided constant support. He also encouraged me to participate in the Translational Research Training in Haematology program, one of the best experiences I had during my PhD.

I would like to thank a number of people who provided invaluable help with my work: Sameena Iqbal, Catherine Durance and Tom Dowe for retrieving patient samples; Dave Williamson for keeping a tidy and organized lab; Janet Matthews and Andrew Wilson for providing the patient information; Andrew Clear for his unparalleled expertise in immunohistochemistry; Andrew Owen who constructed the tissue microarrays; William Day and Guglielmo Rosignoli for performing flow sorting; Jacek Marzec who performed the bioinformatics analysis; Tim Farren for his help with the cytokine bead arrays; Robert Petty for his guidance in performing and interpreting gene expression data; Eleni Kotsiou, Cristina Ghirelli, Melania Capasso and Jude Fitzgibbon for their willingness to help in interpreting my data.

I am grateful to Abi Lee and Maria Calaminici for finding time to help me analyse data and giving input in the two manuscripts published.

I am particularly indebted to Farideh Miraki-Moud for her permanent support in the lab. Her capacity to share expertise with a serene and friendly posture helped me to believe I was capable of everything.

Finally, I would like to thank everyone in the Haemato-Oncology department and Barts Cancer Institute, in particular my dear friends Silvia Montoto, Áine McCarthy, Lauren Wallis, Lenushka Maharaj, Katherine Hodby, Andrew Clear, Timothy Farren, Robert Petty and Paul Greaves. I never imagined I was going to meet such supportive, kind and truthful friends – thank you. I will miss you all.

Table of Contents

STATEMENT OF ORIGINALITY 2

ABSTRACT 5

DEDICATION 7

ACKNOWLEDGEMENTS 8

LIST OF FIGURES 14

LIST OF TABLES 16

LIST OF ABBREVIATIONS 17

CHAPTER 1 INTRODUCTION AND OBJECTIVES 23

1.1 Diffuse Large B-cell Lymphoma: background 23
1.2 Prognostication in DLBCL 26
1.2.1 Clinical prognostic factors 28
1.2.2. Prognostic impact of Fluorodeoxyglucose (FDG)-PET-CT 31
1.2.3. Morphology and Immunophenotype prognostic factors 31
1.3 B-cell Ontogeny in DLBCL 32
1.3.1 The GC reaction 33
1.3.2 The GC reaction and B-cell Lymphomas 35
1.3.4 The Cell-of-Origin (COO) Classification in DLBCL 36
1.4 Main oncogenic mechanisms in DLBCL 38
1.5 GEP-based prognostic models suggests that a host inflammatory/immune response plays a role in the biology of DLBCL 46
1.6 Limitations of GEP-based predictive algorithms 50
1.7 Cancer and Immunity 52
1.8 Evolving models of macrophage ontogeny 53
1.9 Macrophage functions in normal physiology 56
1.10 Shifting the paradigm of macrophage polarisation 58
1.10.1 Classical and alternative macrophage activation 58
1.10.2 M1 and M2 activation 59
1.10.3 Problems of macrophage activation nomenclature 60
1.10.4 Transcriptional analysis of macrophages 62
1.10.5 Obstacles to the understanding of human macrophage biology 66
1.11 Tumour-associated macrophages (TAM) 67
1.11.1 Protumoural features in the context of a macrophage classical activation 70
1.11.2 Protumoural features in the context of a macrophage alternative activation 70
1.11.3 "Gray areas" of TAM activation 72
1.12 Crosstalk between B-cells and macrophages and their relevance for antibody therapies 73

1.13 Correlative IHC studies of the DLBCL-associated macrophages 75
1.14 Functional studies exploring the role of macrophages in lymphoma 78
1.15 Objectives 81

CHAPTER 2 MATERIALS AND METHODS 82

2.1 Patient Samples 82
2.1.1 Formalin-Fixed Paraffin-Embedded Tissue 82
2.1.2 Frozen Single Cell Suspensions 83
2.1.3 Healthy Donor Buffy Cones 84
2.1.4 Cell Lines 84
2.2 Immunohistochemistry 84
2.2.1 Tissue Microarray Assembly 84
2.2.2 Principles of Immunohistochemistry 85
2.2.3. Primary antibodies for Immunohistochemistry 86
2.2.4 Staining Protocol using the Dako Autostainer System 86
2.2.5. Immunohistochemistry analysis 87
2.2.6 Cutpoint Determination and Survival Analysis 89
2.3 Immunofluorescence Staining and analysis 90
2.4 Multicolour flow cytometry 92
2.4.1 Single cell sorting 92
2.4.2 Immunophenotyping of macrophage cell surface markers 96
2.5 RNA extraction 97
2.6 RNA Quantity and Quality Assessment 97
2.6.1 NanoDrop Spectrophotometer 97
2.6.2 Agilent Bioanalyzer 98
2.7 Gene Expression Analysis 99
2.7.1 Gene Expression Profiling by Microarray technology 99
2.7.2 Gene Expression Analysis by quantitative Real-Time PCR 103
2.8 Co-culture systems involving primary human macrophages and malignant and reactive B-cells 105
2.8.1 Positive selection of monocytes from using magnetic microbeads and macrophage maturation *in vitro* 106
2.8.2 Positive selection of reactive B-cells from tonsils using magnetic microbeads 106
2.8.3 Macrophage harvesting and replatting for co-culture experiments 107
2.8.4 Co-culture set-up 107
2.8.5 Macrophage harvesting after co-culture 109
2.9 Cytokine profiling of co-culture supernatants using cytometric bead arrays 109
2.9.1 FCAP Array analysis of cytokine secretion 111
2.10 Statistical analysis 111

CHAPTER 3 RELIABILITY OF IMMUNOHISTOCHEMISTRY CLASSIFIERS OF CELL-OF-ORIGIN FOR DIFFUSE LARGE B-CELL LYMPHOMA 112

3.1 INTRODUCTION 112
3.2 OBJECTIVES 114
3.3 MATERIALS AND METHODS 116
3.3.1 PATIENTS 116
3.3.2 TMA AND IHC 119
3.3.3. SLIDE SCANNING, SCORING AND ANALYSIS 120
3.3.4 STATISTICAL ANALYSIS 121
3.4 RESULTS 124
3.4.1. ANALYSIS OF INDIVIDUAL MARKERS 124
3.4.2 ALGORITHM CLASSIFICATION: DISTRIBUTION AND CONSISTENCY 127
3.4.3 SURVIVAL ANALYSIS 131
3.5 DISCUSSION 133

CHAPTER 4 REVISITING THE IMMUNE MICROENVIRONMENT OF DLBCL USING A TISSUE MICROARRAY AND IMMUNOHISTOCHEMISTRY: ROBUST SEMI-AUTOMATED ANALYSIS REVEALS CD3 AND FOXP3 AS PREDICTORS OF RESPONSE TO R-CHOP 138

4.1 INTRODUCTION 138
4.2 OBJECTIVES 139
4.3 METHODS 139
4.3.1 PATIENT CHARACTERISTICS 139
4.3.2 TISSUE MICROARRAY AND IMMUNOHISTOCHEMISTRY 139
4.3.3. SEMI-AUTOMATED IMAGE ANALYSIS 141
4.3.4. CUTPOINT DETERMINATION 141
4.3.5 STATISTICAL ANALYSIS 142
4.4 RESULTS 144
4.4.1 HETEROGENEOUS DENSITY OF IMMUNE CELLS IN THE MICROENVIRONMENT OF DLBCL 144
4.4.2 EXCELLENT CONCORDANCE BETWEEN TWO SYSTEMS OF SEMI-AUTOMATED IMAGE ANALYSIS FOR THE CHARACTERIZATION OF THE MICROENVIRONMENT 147
4.4.3 CD3 AND FOXP3 ARE POTENTIAL PREDICTORS OF RESPONSE TO R-CHOP 150
4.5 DISCUSSION 156

CHAPTER 5 DEFINING THE DIFFUSE LARGE B-CELL LYMPHOMA ASSOCIATED-MACROPHAGE TRANSCRIPTOME 161

5.1 INTRODUCTION 161
5.2 AIMS 162
5.3 MATERIALS AND METHODS 163
5.3.1 SAMPLES 163
5.3.2 MICROARRAY QUALITY CONTROL ASSESSMENT 163

5.3.3 DATA NORMALISATION 165
5.3.4 DATA ANALYSIS 165
5.3.5 TARGETED GENE EXPRESSION VALIDATION BY QRT-PCR 166
5.3.6 HIERARCHICAL CLUSTERING OF RQ VALUES 167
5.3.7 GENE ENRICHMENT ANALYSIS 167
5.3.8 GENERATION OF RESTING AND POLARISED MACROPHAGE GENE SIGNATURES 167
5.4 RESULTS 169
5.4.1 CD36 IS EXPRESSED BY THE MAJORITY OF TONSIL CD68+ CELLS, REPRESENTING A GOOD SINGLE MARKER FOR MACROPHAGE CELL SORTING 169
5.4.2 LN SINGLE CELL SUSPENSIONS OF DLBCL ARE NOT REPRESENTATIVE OF THE TUMOUR CONTENT WHEN ASSESSED BY FLOW CYTOMETRY 170
5.4.3 QUALITY ASSESSMENT OF THE EXPERIMENTAL WORKFLOW EMPLOYED IN THIS STUDY 171
5.4.4 UNSUPERVISED HIERARCHICAL CLUSTERING 174
5.4.5 STATISTICAL ANALYSIS OF DIFFERENTIALLY EXPRESSED GENES BETWEEN DLBCL TAM AND REACTIVE LN MACROPHAGES 175
5.4.6 CONFIRMATION OF DIFFERENTIALLY EXPRESSED GENES BY QRT-PCR 178
5.4.7 HIERARCHICAL CLUSTERING OF TARGETED GENE EXPRESSION BY QRT-PCR 180
5.4.8 GENE ENRICHMENT ANALYSIS ESTABLISHES LINKS BETWEEN OUR GENE SET, THE "LN SIGNATURE" OF DLBCL AND M1 AND M2 POLARISED MACROPHAGE FUNCTIONS 181
5.4.9 COMPARATIVE ANALYSIS WITH MACROPHAGE GENE SIGNATURES 185
5.5 DISCUSSION 189

CHAPTER 6 ATTEMPTS FOR FUNCTIONAL VALIDATION OF TRANSCRIPTOMIC ANALYSIS 199

6.1 INTRODUCTION 199
6.2 AIMS 200
6.3 MATERIALS AND METHODS 200
6.3.1 SAMPLES 200
6.3.2 CO-CULTURE EXPERIMENTS 200
6.3.3 STATISTICAL ANALYSIS 202
6.4 RESULTS 202
6.4.1 INVESTIGATING CHANGES IN THE EXPRESSION OF MACROPHAGE ACTIVATION MARKERS AFTER CO-CULTURE 202
6.4.2 EXPLORING WHETHER CHANGES IN GENE EXPRESSION FOUND BY GEP CAN BE MIMICKED USING THE CO-CULTURE SYSTEM 205
6.4.3 EXPRESSION DYNAMICS OF THE IMMUNOREGULATORY MOLECULE PD-L1 209
6.4.4 INVESTIGATING CHANGES IN CD36 EXPRESSION IN MACROPHAGES AFTER CO-CULTURE 210
6.5 DISCUSSION 211

CHAPTER 7 EXPLORING MACROPHAGE HETEROGENEITY IN DIFFUSE LARGE B-CELL LYMPHOMA 217

7.1 INTRODUCTION 217
7.2 AIMS 218
7.3 MATERIALS AND METHODS 218
7.3.1 SAMPLES 218
7.3.2 CYTOKINE STUDIES 219
7.3.3 IHC STAINING AND IF ANALYSIS 219
7.3.4 STATISTICAL ANALYSIS 219
7.4 RESULTS 220
7.4.1 TRANSCRIPTOME HETEROGENEITY IN DLBCL TAM 220
7.4.2 INVESTIGATING SOLUBLE FACTORS DIFFERENTIALLY EXPRESSED BETWEEN DLBCL TAM GROUPS 221
7.4.3 INVESTIGATING TH1 AND TH2 CYTOKINES IN CO-CULTURE SUPERNATANT 226
7.4.4 USING IMMUNOHISTOCHEMISTRY AND IMMUNOFLUORESCENCE TO DISCOVER SPECIFIC MACROPHAGE SUBSETS IN THE MICROENVIRONMENT 227
7.5 DISCUSSION 233

CHAPTER 8 FINAL DISCUSSION AND FURTHER WORK 238

CHAPTER 9 REFERENCES 245

APPENDIX 267

List of Figures

Figure 1.1 Recurrent genetic aberrations and oncogenic pathways in DLBCL according to the cell-of-origin molecular profile. 39
Figure 1.2 Origins of tumour-associated macrophages and crosstalk with tumour cells and other cells of the microenvironment. 69
Figure 2.1 Training Optimization using the Ariol System 88
Figure 2.2 Sorting strategy for isolation of macrophages, B-cells and T-cells from single cell suspensions of DLBCL and reactive LNs 95
Figure 2.3 Representative Electropherograms generated by the Agilent 2100 Bioanalyser 98
Figure 2.4 Ribo-SPIA® Technology used for cDNA synthesis and amplification 100
Figure 2.5 Macrophage and B-cells co-culture system employed in this study 108
Figure 2.6 Outline of the cytokine profiling using the Cytokine Bead Array (CBA) System (BD Biosciences) 110
Figure 3.1 Algorithms applied in the current study. 115
Figure 3.2 Kaplan Meier curves according to treatment regimens and treating institution. 117
Figure 3.3 Immunohistochemistry results for each antibody. 123
Figure 3.4 Pair wise agreement according to kappa statistics. 131
Figure 4.1 Selection of tumour representative regions. 143
Figure 4.2 Biomarker distribution. 146
Figure 4.3 Expression correlation for CD3, FoxP3 and TIA1. 148
Figure 4.4 Expression correlation for CD4, CD8 and CD68. 149
Figure 4.5 CD3 Expression and Outcome. 152
Figure 4.6 FoxP3 Expression and Outcome. 154
Figure 5.1 Development of macrophage gene signatures. 168
Figure 5.2 CD68+ cells co-express the membrane scavenger receptor CD36. 169
Figure 5.3 Proportion of macrophages, T-cells and B-cells in SCSs of DLBCL and reactive conditions. 170
Figure 5.4 Excellent RNA quality obtained from macrophage populations. 172
Figure 5.5 cDNA fragment size distribution analysis using the Agilent RNA 6000 Nano kit. 173
Figure 5.6 Correlation of intensity signals for the samples hybridised in the two experimental batches 174
Figure 5.7 Unsupervised hierarchical clustering of DLBCL and reactive LN-associated macrophages. 175
Figure 5.8 DLBCL TAM and reactive LN macrophages differ in expression of 202 genes. 176
Figure 5.9 Expression heterogeneity of most differentially expressed genes in DLBCL and control macrophages. 177
Figure 5.10 Validation of targeted differentially expressed genes by qRT-PCR. 179
Figure 5.11 Hierarchical cluster analysis of RQ values of seven transcripts resolves DLBCL TAM from controls. 180
Figure 5.12 GO enrichment analysis. 182
Figure 5.13 Heatmap of “resting” macrophage-enriched genes. 186

Figure 5.14 Comparative analyses of our DLBCL macrophage associated gene signature with other macrophage gene signatures. 187
Figure 6.1 Expression dynamics of activation markers CD80 (A), CD86 (B) and HLA-DR (C) in macrophages harvested from co-culture. 204
Figure 6.2 Targeted mRNA expression is affected by sorting after co-culture. 206
Figure 6.3 Gene expression changes of targeted genes in macrophages harvested from co-culture. 208
Figure 6.4 Expression of PD-L1 protein in macrophages after co-culture. 209
Figure 6.5 CD36 expression dynamics in macrophages harvested from co-culture. . 210
Figure 7.1 GEP analysis suggesting two groups of DLBCL TAM. 222
Figure 7.2 Differentially expressed soluble factors between B-cells. 223
Figure 7.3 CCL5 is detected in the cell culture media used in co-culture experiments. 225
Figure 7.4 Expression of CD80 and PD-L1 in macrophages treated with soluble factors. 226
Figure 7.5 Detection of Th1 and Th2 cytokines in co-culture supernatants. 227
Figure 7.6 Analysis of CD68 and CD163 immunostainings. 228
Figure 7.7 Expression of ALOX15 in reactive LNs and DLBCL. 230
Figure 7.8 Expression of IDO1 in reactive LNs and DLBCL. 232

List of Tables

Table 1.1 The NCCN enhanced IPI.....29
Table 2.1 Staining strategy for cell sorting 94
Table 2.2 Components of the 2x reverse transcription (RT) master mix used for cDNA synthesis.....103
Table 2.3 Components of Reaction Mix for qRT-PCR 104
Table 3.1 Clinical features of the R-CHOP series 118
Table 3.2 Primary antibodies and conditions of use 120
Table 3.3 Single marker analysis of the R-CHOP cohort 125
Table 3.4 Distribution of R-CHOP treated patients according to the nine IHC classifiers 129
Table 3.5 Survival analysis according to clinical characteristics, IPI and COO classifiers 132
Table 4.1 Primary antibodies and conditions of use 140
Table 4.2 Descriptive statistics showing heterogeneity of expression of biomarkers145
Table 4.3 Univariate and multivariate survival analysis 155
Table 5.1 Samples used for transcriptomic analysis: patients characteristics and partial results 164
Table 6.1 Antibodies used in this study 202
Table 7.1 Antibodies used in this study 220

List of Abbreviations

ABC	Activated B-cell–like
ADCC	Antibody-dependent cytotoxicity
ALOX15	Arachidonate 15-Lipoxygenase
AQP	Aquaporins
AID	Activation-induced cytidine deaminase
ASCT	Autologous stem-cell transplantation
AP-1	Activator protein 1
APC	Antigen presenting cells
APRIL	A proliferation inducing ligand
BAFF	B-cell activating factor
BCR	B-cell receptor
BCL-2	B-cell lymphoma/leukaemia-2
BCL-6	B-cell lymphoma/leukaemia-6
BCL10	B-cell lymphoma-10
BH	Benjamini and Hochberg
BSA	Bovine serum albumin
BTK	Bruton's tyrosine kinase
BL	Burkitt's Lymphoma
Blimp-1	B-lymphocyte-induced maturation protein 1
BM	Bone marrow
CARD11	Caspase recruitment domain-containing protein 11
CBA	Cytometric Bead Arrays
cDNA	Complementary DNA
CDH1	E-cadherin
C/EBPβ	CCAAT-enhancer-binding protein beta
CGH	Chromosomal genomic hybridisation
cHL	Classical Hodgkin's lymphoma
CI	Confidence intervals
CR	Complete response

DMEM	Dulbecco's Modified Eagle's
COO	Cell-of-Origin
CREBBP	cAMP response element-binding protein
CSF-1	Colony stimulating factor-1
CSF-1R	Csf-1 receptor
CSR	Class switch recombination
CT	Cycle threshold
CTGF	Connective-tissue growth factor
DAB	Diaminobenzidine
DAPI	4’6-diamidino-2-phenylindole
DLBCL	Diffuse Large B-cell Lymphoma
DMSO	Di-Methyl Sulphoxide
DNA	Deoxyribonucleic acid
EBV	Epstein Barr virus
ECM	Extracellular matrix
ECOG	Eastern Cooperative Oncology Group
EFS	Event-free survival
EGR	Early growth response
EP300	E1A binding protein p300
ETS	E-twenty six
FC	Fold change
FCS	Foetal calph serum
FcγR	Fcγ receptors
FDC	Follicular dendritic cells
FDG-PET-CT	Fluorodeoxyglucose positron emission tomography-computed tomography
FDR	False discovery rate
FFPE	Formalin-fixed paraffin embedded
FMO	Fluorescence-minus-one
FOXO1	Forkhead box protein O1
FSC-H	Forward scatter

GC	Germinal centre
GCB	Germinal centre B-cell–like
GCET1	Centerin
GEP	Gene expression profiling
GFP	Green fluorescent protein
GM-CSF	Granulocyte macrophage colony-stimulating factor
H	Heavy
H&E	Haematoxylin and Eosin
HIF-1α	Hypoxia-inducible factor-1α
HLA	Human leukocyte antigen
HR	Hazard ratios
HRP	Horseradish peroxidase
Id	Idiotype
IDO1	Indoleamine 2,3-dioxygenase 1
IF	Immunofluorescence
Ig	Immunoglobulin
IHC	Immunohistochemistry
IKBα	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
IL	Interleukin
IMDM	Iscove's Modified Dulbecco's Medium
IMS	Industrial Methylated Spirits
IPI	International Prognostic Index
IPO	Instituto Português de Oncologia
IRAK	Interleukin-1 receptor-associated kinase
IRF	Interferon regulatory factor
JAK	Janus tyrosine kinase
JMJD3	Jumonji domain containing-3
LDH	Lactate dehydrogenase
LN	Lymph node
LMO2	LIM domain only 2

LPS	Lipopolysaccharide
LYN	Tyrosine-protein kinase
MALT1	Mucosa-associated lymphoid tissue lymphoma translocation protein 1
MAP	Mitogen-activated protein
MHC	Major histocompatibility complex
mAb	Monoclonal antibody
MFI	Median fluorescence intensity
miRNA	microRNAs
MLL2	mixed-lineage leukemia protein 2
MPS	Mononuclear-phagocyte system
MSF	Migration-stimulating factor
MR	Mannose receptor
MT	Metallothionein
NCCN	National Comprehensive Cancer Network
NF-kB	Nuclear factor-kB
NGS	Next-generation sequencing
NHL	Non-Hodgkin lymphoma
NLPHL	Nodular lymphocyte predominant Hodgkin Lymphoma
NO	Nitric oxide
NOS	Not otherwise specified
NSB	Non-specific binding
OD	Optical density
OS	Overall survival
qRT-PCR	Quantitative real-time PCR
PAX5	Paired box protein 5
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PD-L1	Programmed death-ligand 1
PGE2	Prostaglandin E2
PFS	Progression-free survival
PI3K	Phosphoinositide 3-kinase

PK	Pharmacokinetic
PKCβ	Protein Kinase Cβ
PM	Perfect match
PPAR-γ	Peroxisome proliferator-activated receptor-γ
PRC2	Polycomb repression complex-2
PRDM1	PR domain zinc finger protein 1
PTEN	Phosphatase and tensin homolog
RAG	Recombination-activating gene enzymes
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone
RCT	Randomised clinical trials
RIN	RNA Integrity Number
RLE	Relative log expression
RMA	Robust multi-array average
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RQ	Relative quantities
RPMI	Roswell Park Memorial Institute
SCARB1	Scavenger receptor class B member 1
SCID	Severe combined immunodeficiency
SCSs	Single cell suspensions
SD	Standard deviation
SEM	Standard error of the mean
SHM	Somatic hypermutation
SPARC	Secreted protein, acidic and rich in cysteine
SSC-A	Side scatter
STAT	Signal Transducers and Activators of Transcription
SYK	Spleen tyrosine kinase
TAM	Tumour-associated macrophages
THRBCL	T-cell/histiocyte-rich B-cell lymphomas
TCR	T-cell receptor
TFH	T follicular helper cells

TGM2	Transglutaminase-2
TLR	Toll-like receptors
TMA	Tissue microarray
TNF	Tumour necrosis factor
TP53	Tumour protein 53
V(D)J	V(ariable), D(iversity) and J(oining)
VEGF	Vascular endothelial growth factor
WHO	World Health Organization

Chapter 1 Introduction and Objectives

PART ONE

1.1 Diffuse Large B-cell Lymphoma: background

Diffuse Large B-cell Lymphoma (DLBCL) is the most common aggressive lymphoma in the western world, accounting for 30-40% of all cases.[1] The incidence of DLBCL has risen slowly over the last 20 years. An estimated age-adjusted incidence of 9.8/100.000 cases in Europe in 2012 has been recently reported, with predominance for males.[2] However disease-related mortality has significantly decreased, recognisably due to the introduction of rituximab chemo-immunotherapy into first-line treatment, together with better supportive care and intention to treat in the elderly population. DLBCL is in fact a disease of the elderly, with a median age at diagnosis in population-based studies of 71 years.

The aetiology of DLBCL remains largely unknown. Although most commonly arising *de novo*, DLBCL can be the result of transformation from indolent non-Hodgkin lymphomas (NHL).[3] Potential etiological factors include: a family history of lymphoproliferative diseases, prior history of malignancy, primary or treatment-related immunodeficiency, auto-immune diseases or viral infections including human immunodeficiency virus[4], human herpes virus 8 and Epstein Barr virus (EBV).[5]

Patients generally present with a rapidly growing lymph node (LN), accompanied in a third of the cases by B-symptoms (fever, night sweats, and weight loss). Extranodal involvement at presentation occurs in 40% of the cases.[6] Half of the patients are diagnosed in advanced stage disease as defined by the Ann Arbor system,[7] with bone marrow involvement reported in 11-27% of the cases.[8] A revised staging system has been recently proposed incorporating results from the diagnostic positron emission tomography-computed tomography (PET-CT) that will likely change the distribution of patients according to stage.[9]

A tissue-based histopathology examination remains the standard tool to establish a diagnosis of DLBCL. This disease is described microscopically as a neoplasm of large B lymphoid cells with a diffuse pattern often effacing the normal LN architecture. The malignant B-cells exhibit a nuclear size at least twice that of normal lymphocytes.[1] The malignant cells express pan–B-cell markers including CD19, CD79a, CD20, PAX5, and CD22, although some of these markers can be lost in individual cases.[10] The proliferation fraction, as detected by Ki-67 staining, is generally high.

However, the diagnosis of DLBCL encompasses a large number of sub-diagnoses, reflecting the heterogeneity of this disease.[1] What has accelerated the sub-classification of DLBCL has undoubtedly been the use of gene expression profiling (GEP). This technology generated evidence that certain conditions amalgamated within the DLBCL diagnosis and previously recognized by their particular clinical behaviour, morphology or immunophenotype had a distinct molecular background and hence likely a distinct cell of origin. For example, T-cell/histiocyte-rich B-cell lymphomas (THRBCL) constitute a rare aggressive subtype of DLBCL with abundant T-cell and macrophage infiltration in the microenvironment.[11] Only a minority of the cells are large CD20+ B-cells. Although not pathognomonic, most of these cases carry a GEP characterized by a host immune response that differs from the majority of DLBCL cases.[12] Another example is the provisional entities of unclassified large cell lymphomas with intermediate features between DLBCL and Burkitt's Lymphoma (BL) or classical Hodgkin's lymphoma (cHL). For many years, pathologists encountered cases with intermediate morphologic features between DLBCL and BL or cHL that posed diagnostic dilemma. Microarray technology again helped recognizing that these cases constitute a real biological gray zone of lymphomas.[13-16]

Taking into consideration the prolific number of publications characterizing new DLBCL "entities" it is envisaged that the next World Health Organization (WHO) lymphoma classification will incorporate changes in how we define what is now termed DLBCL. The issue of misdiagnosis in DLBCL is still debated, highlighting the

importance of expert central review. A large-scale assessment in the United Kingdom highlighted that there is still up to a 25% rate of discordance between primary and revised diagnosis of DLBCL.[17]

DLBCL has an aggressive natural history, with a median survival of less than one year if left untreated, but is readily treatable. The anthracycline-based combination CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) was introduced in 1976 and has been the standard of care for 30 years.[18] However survival rates were suboptimal, with 10-year progression-free survival (PFS) of 30% and overall survival (OS) rates of 35%. Attempts to add treatment efficacy with more intensified chemotherapy regimens failed.[19]

In the last decade unprecedented advances were made in the standard of care of patients with DLBCL that translated in an improvement in cure rates and OS. With modern therapeutic combinations, approximately 60% of patients with this disease are now being cured. This progress is largely due to the addition of the anti-CD20 antibody rituximab to CHOP chemotherapy (R-CHOP).[20-26] Long-term follow-up of the first randomised clinical trial (RCT) comparing R-CHOP to CHOP demonstrated a 16% absolute improvement in the 10-year OS rate in favour of R-CHOP.[21] The higher costs associated with the use of rituximab are offset by the significant advantage in survival and by the decrease in the number of patients requiring expensive salvage therapies.

Currently there are a number of research priorities in the treatment of patients with DLBCL.

A substantial 40% of patients will be primary refractory or relapse after R-CHOP, and this constitutes a real management challenge. The estimate of PFS at 3 years for this group is only 23%, even with subsequent high-dose chemotherapy and autologous stem-cell transplantation (ASCT), proving that salvage regimens remain largely ineffective for this sub-group of patients.[27] In fact, in keeping with the findings of other groups, we have confirmed in our patient cohort that R-CHOP refractoriness is a strong predictor of a worse OS. Identification of these high-risk patients at diagnosis, incorporation of novel therapies into RCTs where they are enrolled and recognition of

the mechanisms for rituximab-refractoriness are of paramount priority. On the other hand it is hypothesized that some patients with better prognosis disease might do as well with four rather than six or eight cycles of R-CHOP, an issue that is being addressed by the German high-grade non-Hodgkin's Lymphoma study group.

To date, only age, the International Prognostic Index (IPI) stratification and genetic aberrations defining gray-zone lymphomas have been used in clinical trials to select high-risk DLBCL patients for more aggressive regimens, most of them incorporating front-line consolidation with ASCT. However, these prognostic features do not take into account the full biological heterogeneity partially responsible for disease aggressiveness. Moreover, despite different approaches taken in RCTs, there is no current evidence for offering treatments other than R-CHOP according to risk factors in DLBCL.

This reflects our limited knowledge on comprehensive prognostication in this heterogeneous group of disorders and supports our choice to investigate old and new prognostic factors for DLBCL in this thesis. The treatment scenario is, however, changing rapidly owing to a deeper knowledge of the molecular biology and the drivers of oncogenic transformation of DLBCL, leading to the development of rational targeted therapeutic approaches (see section 1.4).

1.2 Prognostication in DLBCL

Despite the number of scientific publications describing potential prognostic markers in medicine, the field of outcome prediction is complex. This is even more so in heterogeneous disorders such as DLBCL. It is argued that only modest outcome predictive accuracy can be achieved with clinical data or single biochemical or molecular markers.[28]

The statistical validation of prognostic variables is dependent on the number of patients enrolled in a study. If the objective of a particular study is to validate a new prognostic marker, then other formerly validated markers need to be taken into account. With a growing amount of markers to be incorporated in multivariate analyses, a large number of patients are required to accomplish statistical validity. In practice, however, the results in highly cited biomarker studies often significantly overestimated their findings.

To ascertain the impact of a clinical or biologic feature on patient's survival or treatment response, robust clinical validation under prospective RCTs or large multi-institutional cohorts is needed. However, this is rarely performed. It has been reported that only around 3% of studies describing potential clinical applications in genomic medicine moved to assessment of clinical utility under RCT.[29] Moreover, studies using patients enrolled in RCT are essential for predictive marker validation as it guarantees uniformity in clinical characteristics from both the experimental and the control arm.

Only after scrutiny in this manner can a clinical or biologic marker predictive of response to R-CHOP help us to individualize treatment in DLBCL. The final and essential step in this validation process is developing robust, simple and reproducible methodologies for assessment of these biomarkers so as to guarantee their implementation in the clinical practice.

GEP and genomic studies in DLBCL were of paramount importance in deepening our understanding of the biology of DLBCL. These studies were the starting point for many others detailing the classification of the heterogeneous group of DLBCL, not otherwise specified (NOS), and searching for the functional roles of specific aberrations. More importantly they provided new prognostic biomarkers and treatment approaches. Details on the molecular biology of DLBCL will be covered in section 1.4.

1.2.1 Clinical prognostic factors

The **IPI** is a prognostic score incorporating five factors reflecting patients' clinical and biological characteristics. Based on the number of negative factors detected at the time of diagnosis (age >60 years, stage III/IV disease, elevated serum lactate dehydrogenase [LDH] level, Eastern Cooperative Oncology Group [ECOG] performance status ≥2, >1 extranodal site of disease), four groups were identified, with 5-year OS ranging from 26%-73%.[30] The IPI was devised in the pre-rituximab era and its power to discriminate outcome has declined after introduction of R-CHOP, particularly for higher risk patients.[31-33]

The prognostic value of the individual factors incorporated in the IPI, however, remains unchanged. The National Comprehensive Cancer Network (NCCN)-IPI[33] was recently developed based on DLBCL patients treated in academic centres and in the community with R-CHOP (Table 1.1). It incorporates the same markers as the IPI and is able to better discriminate patients with low and high risk disease that might merit different treatment approaches.

Table 1.1 The NCCN enhanced IPI

Variables		**Score**
Age, years	>40 to ≤60	1
	>60 to ≤75	2
	>75	3
LDH, normalised	>1 to ≤3	1
	>3	2
Ann Arbor stage III-IV		1
Extranodal disease		1
Performance status ≥2		1

Low risk: 0 – 1 High-intermediate risk: 4 - 5

Low-intermediate risk: 2 – 3 High risk: ≥ 6

However, many authors acknowledge that patients belonging to the same IPI group have significant survival differences, likely reflecting different biological backgrounds. Moreover a number of papers suggest that some biomarkers, such as the GEP-based stromal signatures[34] or chromosomal translocations involving *MYC*,[35] B-cell lymphoma/leukaemia-2 *(BCL2)*,[36] or B-cell lymphoma/leukaemia-6 *(BCL6),*[37] have an IPI-independent prognostic impact. These data demonstrate that there is opportunity to improve upon the IPI with new biomarkers. Although not yet available due to the reasons discussed above, it is likely that a clinico-biological index will be more comprehensive for prognostic stratification than the NCCN-IPI.

Zhou et al. recognize **advanced age** as being associated with incremental risk even with R-CHOP treatment. In fact, aging seems to be a determinant of the molecular biology of DLBCL and a surrogate of aggressive disease.[38]

Other clinical variables not incorporated in the IPI have been suggested as prognostic markers in DLBCL:

1. **Male gender** has been associated with worse outcome in independent studies in DLBCL.[39] In fact, pharmacokinetic (PK) studies of rituximab demonstrated that elderly males have a faster clearance of rituximab and hence might be sub-optimally treated with the standard dose of the monoclonal antibody. This hypothesis formed the basis of two German clinical trials: the SMART-E-R-CHOP-14 trial demonstrated that a PK-guided rituximab schedule benefited older patients with high tumour burden, regardless of the patient's gender;[40] the SEXIE-R-CHOP-14 trial showed that increasing rituximab dose eliminates the increased risk of elderly males as well as young male and female patients with low rituximab serum levels.

2. A high **body mass index** has provocatively been stated as a favourable outcome predictor after R-CHOP treatment in DLBCL.[41] However others have disputed this association.[42,43]

3. The **maximum tumour diameter** has an adverse impact on OS in R-CHOP-treated patients,[44,45] and is helpful for stratifying patients with low-risk IPIs.

4. **Bone marrow** (BM) involvement is associated with a poor outcome, particularly with concordant rather than discordant low-grade histology. Those patients who have diffuse large cell infiltrates in the BM have a particularly bad outcome, with reported 10% OS at 5 years.[8,46]

5. Primary involvement of the **Waldeyer ring** appears to confer a better outcome.[6,47] In fact, the role of rituximab in the treatment of primary extranodal DLBCL is a matter of debate.[48]

1.2.2. Prognostic impact of Fluorodeoxyglucose (FDG)-PET-CT

FDG-PET-CT has currently replaced standard CT scan for staging and response assessment of FDG-avid lymphomas.[9] A negative FDG-PET-CT at the end of treatment is an excellent predictor of good outcome in DLBCL. The role of interim FDG-PET scan is, however, controversial. Recent reports show an increase in PFS (and OS in some studies) in patients with a negative PET after 2 or 4 cycles of R-CHOP. Yet others have reported contrary results. In an attempt to validate the biological significance of mid-treatment PET results, Moskowitz et al.[49] re-biopsied all lesions deemed positive and found that 87% of PET positive were indeed false positive. The absence of strict scoring criteria, best standard scoring method[50] and a high inter-observer variability are given as potential explanations for this controversy. Using PET-CT to guide optimal treatment is not established in DLBCL.

1.2.3. Morphology and Immunophenotype prognostic factors

Although morphology has been disregarded as a method to gather prognostic insight in DLBCL, a thorough pathological report at diagnosis is required to appropriately allocate patients according to the WHO classification and might in fact provide some information on patient's outcome.

A recent analysis from a large cohort of patients enrolled in the Ricover-60 trial showed that immunoblastic morphology[51] is an adverse prognostic factor at diagnosis. Similarly, a plasmablastic phenotype[52] in cases of DLBCL-NOS was associated with shorter survival after R-CHOP.

CD5 expressing DLBCLs constitute 10% of cases and seem to represent an independent entity with a distinct genomic and transcriptomic profile and to be associated with a poorer outcome.[53-55] The intensity of CD20 expression is heterogeneous in patients with DLBCL. Cases showing decreased CD20 expression appear to have a worse survival, independently of clinical prognostic factors.[56]

1.3 B-cell Ontogeny in DLBCL

B-cell malignancies arise from a subversion of the differentiation and proliferation patterns of a normal B-cell. B-cells are specialized in the recognition and elimination of foreign antigens, with these functions mediated through the B-cell receptor (BCR). In order for the BCR to recognize non-self antigens it has to go through molecular modifications, which lead to increased epitope specificity and affinity. This dependency on physiological genetic changes in the BCR coding region potentiate the acquisition of oncogenic aberrations and help to explain why B-cells are particularly prone to malignant transformation.[57]

The process of B-cell differentiation initiates in the BM where a precursor cell undergoes sequential rearrangements of the heavy and light chain locus of the Ig gene. This molecular process involves a modulated series of genetic events which ultimately leads to the expression of a functional BCR at the surface of the B-cell.[58] Assembly of the Ig gene segments, required to produce a functional BCR, occurs through three recombination events catalysed by the enzymatic complex recombinase.[59] If in any event a non-productive rearrangement occurs, cells undergo apoptosis.

The recombinase enzymatic complex starts this process by introducing double strand DNA (deoxyribonucleic acid) breaks.[60] Repair mechanisms introduce a high potential for oncogenic chromosomal translocations to occur. The element responsible for regulating the expression of a specific gene is replaced by a regulatory element that is constitutively active and hence leads to aberrant gene expression. An illustrative example is the t(14;18)/IgH-BCL2, which brings the *BCL2* gene under the control of the active regulatory element of the Ig heavy (H) locus, ultimately leading to BCL2 protein overexpression.[61,62] The occurrence of a potential initial oncogenic event does not undermine the differential potential of the malignant cell, allowing for further genetic events to occur later on the living pathway of the lymphoma cell.

Precursor B-cells expressing a functional BCR are positively selected into the antigen inexperienced, B-cell compartment in the peripheral blood. Once they encounter foreign antigens, naïve B-cells follow one of three pathways:[59] (1) travel to the peripheral lymphoid tissues and initiate the germinal centre (GC) reaction; (2) differentiate into short lived plasmablasts outside the germinal centre; or (3) enter into an anergic state.

1.3.1 The GC reaction

The GC reaction occurs in the follicles of lymphoid tissues and is a physiological process required for final B-cell maturation.[60] Through molecular mechanisms known as somatic hypermutation (SHM) and class switch recombination (CSR), an efficient humoral response against the antigen is carried out with the production of high-affinity antibodies.

After contacting with the antigen presenting cells (APC), naïve B-cells uptake, internalize and expose the antigen through the major histocompatibility complex (MHC)-class II molecule. B-cells are then driven to the T-cell rich areas of the lymphoid organs, where they become activated through interaction with CD4+ helper T-cells. This activation requires not only the interaction between the MHC class-II and the T-cell receptor (TCR) but also ligation of the co-stimulatory molecules CD28 and CD40L with their ligands on APCs. This is essential for the initial formation of the GC reaction, as has been demonstrated in murine models of genetic or pharmacological disturbance of CD40 or CD40L function.[58]

Within the dark zone of the GC, B-cells exhibit an impressively high proliferation rate and can be histologically recognized as centroblasts. It has been demonstrated that the major gene expression changes between a resting naïve B-cell and a centroblast are regulated by the BCR signalling and nuclear factor-kB (NF-kB) and result on the over-expression of the proliferation machinery and down-regulation of the DNA damage sensing machinery.[63] Centroblasts producing antibodies with improved

antigen binding capacity progress through the GC reaction, stop proliferating (now denominated centrocytes) and undergo further selection. Centrocytes engage with antigen-exposing follicular dendritic cells (FDCs) and CD40L-expressing T follicular helper cells (TFH) and are selected based on their antigen affinity. Finally, centrocytes exit the GC either as plasma cells or memory B-cells.

The SHM[64] molecular event involves the introduction of single nucleotides, small deletions and duplications into the Ig variable locus or genes other than the Ig genes, such as the *BCL-6* proto-oncogene.[65] CSR is the molecular mechanism by which Ig isotype switching occurs, leading to a modification of the effector functions of an antibody while retaining specificity. Both the SHM and CSR events require the activity of the enzyme activation-induced cytidine deaminase (AID). AID deaminates cytosines leading to the initiation of DNA repair mechanisms that ultimately allow the introduction of somatic mutations and the generation of IgV clonal diversity.[66]

The transcriptome of the B-cells encompassing the GC is tightly regulated and involves many molecules. However, BCL6 is regarded as the master regulator of the GC reaction. *BCL6* knock-in mice are unable to mount a GC reaction upon antigen triggering. These mice develop lymphoma that recapitulates the biology of human DLBCL.[67]

BCL6 is a transcriptional repressor that targets genes involved in cellular processes fundamental for the GC reaction. BCL6 is known to repress apoptosis[68] and cell-cycle arrest responses as well as DNA damage sensor related genes,[69] enabling persistent proliferation and tolerance to genetic transformation required for B-cell antigen affinity maturation. Furthermore, BCL6 down-regulates the expression of genes involved in B-cell activation,[70] enabling centroblasts to complete clonal expansion and undergo SHM. Finally, BCL6 has been shown to repress GC B-cell differentiation into plasma cells by direct or indirect (via targeting Interferon regulatory factor 4, IRF4) repression of a master regulator of B-cell terminal differentiation, Blimp-1 (B-lymphocyte-induced maturation protein 1).[71]

Overall BCL6 functions to sustain the GC B-cell phenotype and hence needs to be repressed for conclusion of B-cell maturation. Its repression in centrocytes is mediated by IRF4. Once *BCL6* is "turned-off", Blimp-1 establishes the characteristic plasma cell phenotype by inactivating PAX5 (paired box protein 5). Continued signalling through CD40 is crucial for driving centrocytes towards memory B-cell differentiation, and continued PAX5 expression maintains B-cell identity in memory B cells.

1.3.2 The GC reaction and B-cell Lymphomas

Nearly all B-cell NHLs harbour SHM in the IgV genes, indicating that the malignant B-cell is blocked within or has passed through the GC. Moreover the commonest oncogenic events in B-cell neoplasms are chromosomal translocations and aberrant SHM, which represent errors in the physiological process of B cell differentiation and maturation. With the exception of the t(14;18), and a subset of t(8;14) involving the Ig and *MYC* loci in endemic-type BL, most breakpoints in chromosomal translocations involving the Ig genes arise in the switch region or in the target regions for SHM.[16] AID is absolutely required for the SHM to occur and has been detected in exceedingly high levels in some DLBCLs. This again indicates the occurrence of malignant transformation during the GC reaction.

The consequence of these genetic abnormalities is the transcriptional deregulation of genes involved in GC B-cell development or the aberrant expression of genes restricted to other B-cell developmental stages. As an example, chromosomal translocations involving *BCL6* prevent silencing of its expression at the end of the GC reaction, leading to a block in post-GC differentiation.[72] Cells harbouring *BCL6* translocations maintain a centroblast-specific functional program, with typical pro-proliferative, DNA-damage-tolerant phenotypes, with a high potential to acquire further genetic alterations, eventually leading to lymphomagenesis.

The understanding of B-cell development, the physiological process of the GC reaction and the mechanisms of B-cell transformation were instrumental for the seminal discovery of the molecular subtypes of DLBCL described in the following section.

1.3.4 The Cell-of-Origin (COO) Classification in DLBCL

Up to the year 2000 molecular heterogeneity in DLBCL was recognized only by the detection of chromosomal translocations, numerical chromosomal abnormalities, mutations and deletions in tumour suppressor genes. It was by then also accepted that these molecular changes arose in B-cells in different developmental stages. However the normal cellular counterpart of DLBCL was still unknown. GEP, a high throughput technique for interrogation of the whole transcriptome of cells or tissues was by then starting to be used to study human cancer.[73]

In this scenario, Alizadeh et al. performed GEP in normal human B-cells in different stages of development or stimulated in vitro, in parallel with a variety of snap-frozen single-cell suspensions derived from B-cell malignancies.[74] They wanted to test the hypothesis that the concerted gene transcription in lymphomas is similar to the one from the normal counterpart from which they derive; and to deepen the current understanding of the lymphoma biological phenotypes. The Lymphochip array was purposely designed to detect transcripts related to the GC biology.[75] By visually comparing, in a completely unsupervised fashion, how all samples clustered according to similarities between them (hierarchical clustering), the authors indeed recognized that around half of the DLBCL cases studied expressed genes that are hallmarks of normal centroblasts (e.g. *BCL6*). By contrast, another DLBCL subgroup lacked the expression of GC B-cell–restricted genes and instead had a transcriptome similar to the one of a B-cell triggered to divide (e.g. NF-kB, *IRF4*). To consolidate these findings, patient division according to resemblances with normal B-cell transcriptomes was later confirmed in larger datasets of more than 200 patients, either before[76] or after the introduction of R-CHOP.[34]

Importantly all GEP studies recognized that the two subgroups had disparate survival rates independently of the treatment given. The prognostic significance of this classification of DLBCL was also corroborated by GEP studies done on formalin-fixed paraffin embedded (FFPE) tissue using different systems of nucleic acid amplification and done using low input, partially degraded RNA (ribonucleic acid).[77-79]

GEP studies in DLBCL[12,74,76,80] could also recognize groups of genes involved in different biological processes, which had synchronous expression. These gene signatures were differentially expressed by the lymphomas profiled, and helped to elucidate the biological processes that are being exploited by the malignant cells and to predict outcome independently of the IPI.

These analyses led to the COO classification, recognizing two molecularly distinct entities in DLBCL: the germinal centre B-cell–like (GCB) DLBCL and the activated B-cell–like (ABC) DLBCL. Mutational analysis showed that the GCB-DLBCLs have ongoing SHM of the IgV genes,[81] suggesting that the malignant B-cells result from a block within the centroblast stage in the GC reaction; whereas ABC-DLBCLs have a high mutational burden in the IgV genes, supporting that the malignant B-cells have passed through the GC but fail to fully mature.

Additional work helped to allocate previously known genetic aberrations (such as the t(14;18) translocation) into each molecular subgroup, and to describe novel genetic lesions that are mutually exclusive to each subgroup.[82] Importantly, the COO classification accelerated functional studies exploring the role of cytokines, signalling pathways and transcription factors differentially involved in the pathogenesis of these lymphomas.

The COO classification led to the design of RCTs incorporating targeted therapies, with the perspective of improving survival of the high-risk ABC-DLBCLs. These trials

might demonstrate that the COO is predictive of response to R-CHOP, but as yet there is no controlled data supporting it.

1.4 Main oncogenic mechanisms in DLBCL

Despite having genetic and epigenetic alterations in common, GCB and ABC-DLBCLs utilize mutually exclusive oncogenic pathways, providing compelling evidence for their different origin. Pathway activation depends on gain-of-function mutations in signalling effectors, loss-of-function mutations in negative signalling regulators, or autocrine receptor activation.[83] Here we will first describe the most relevant genetic alterations, followed by the commonest oncogenic pathways utilized by the DLBCL cells and the epigenetic changes in this disease, paying particular attention to the changes with potential for therapeutic targeting. For a comprehensive list of genetic aberrations see Figure 1.1.

Genetic aberrations	proportion of patients (%)	GCB	ABC
t(14;18)/BCL2	50	x	
EZH2 mutations	20	x	
MEF2B mutations	~10	x	
TP53 mutations/deletions	20	x	x
MLL2 mutations	40	x	x
CREBBP/EP300 mutations	40	x	x
MYD88 L265P mutations	30		x
CD79 A/B mutations	20		x
t(3;x)/BCL6	10 - 25	x	x
SPIB translocations/gains/amplifications	20 - 50		x
CDKN2A deletions	20 - 50		x
INK4a/ARF deletion	20 - 50		x
A20 deletions/mutations	20 - 50		x
PRDM1	20 - 50		
BCL2 amplifications	20		x
MYC translocations	5 - 20	x	x
MDM2 amplifications	< 20	x	
IRF4 rearrangements	< 20	x	
CARD11 mutations	10	x	x
MYD88 mutations (other)	10	x	x
B2M mutations	< 20	x	x
CD58 mutations	< 20	x	x
PTEN deletions	< 20	x	
TNFRSF14 mutations	5 - 20	x	
FOXO1 mutations	~8	x	x
FAS mutations/deletions	< 20	x	
GNA13 mutations	~25	x	
SGK1 mutations		x	
miR-17-92 amplifications	12	x	

Figure 1.1 Recurrent genetic aberrations and oncogenic pathways in DLBCL according to the cell-of-origin molecular profile.

The most common changes of each molecular group are highlighted.

Next-generation sequencing (NGS) technologies have been used to deepen our understanding of the coding genome of DLBCL (reviewed by Jardin[84]). More than 200 cases have been reported to date and the findings support that DLBCL has a high degree of genomic complexity, harbouring on average between 30-100 changes/case. However the modest overlap of lesions described among the four studies published[85] and the limited depth of detection of the technologies used, suggests that the full genomic landscape is yet to be unravelled. Apart from confirming previously acknowledged genetic aberrations, NGS uncovered novel targeted genes, some of which have prognostic impact in the R-CHOP era (e.g. Forkhead box protein O1 (*FOXO1*) mutations). Some of these somatic mutations conduct to changes in signalling pathways that play crucial roles in B-cell function, immunity, cell death, or epigenetic regulation, as will be detailed.

The ***TP53*** (tumour protein 53) gene is targeted by mutations (20% of cases),[86] and by deletions irrespective of the COO of DLBCL. This gene encodes for the tumour-suppressor protein p53. Loss-of-function mutations in *TP53* impair the regulation of many biological processes controlled by p53: cell cycle, apoptosis, cell differentiation, DNA repair, angiogenesis, and genomic stability. All aberrations were associated with worse survival in DLBCL.[87-89] Importantly, it was suggested that *TP53* mutations might help to stratify GCB patients into different prognostic subgroups.[89,90]

Chromosomal abnormalities can be detected by karyotype, targeted in-situ hybridisation (fluorescence or colorimetric) or, in a more detailed and comprehensive fashion, by chromosomal genomic hybridisation (CGH) arrays. CGH-arrays demonstrated that GCB and ABC-DLBCLs harbour distinct chromosomal aberrations.[91] As an example, copy number alterations on chromosome 3 are exclusively diagnosed in ABC-DLBCL, whereas amplifications of *C-REL* are exclusively detected in GCB cases.

Around half of GCB-DLBCLs exhibit the t(14;18) translocation, leading to ***BCL2*** overexpression and subsequent escape from apoptosis. *BCL2* overexpression is also detected in the majority of ABC-DLBCLs, driven by transcriptional deregulation and

gene amplifications rather than translocations.[76,92] BCL2 protein is highly expressed in both molecular groups (70% in our series[93]), and is accepted as a predictor of poor outcome.[94-97] Hence BCL2 represents an excellent candidate for therapeutic targeting. Several molecules targeting it at the DNA and protein level, including ABT-199 are under study and represent an opportunity to improve the outcome of GCB-cases.

A number of gene **transcriptional regulators** exhibit chromosomal changes in DLBCL. Translocations involving the transcription factor ***MYC*** are detected in 5-10% of DLBCLs irrespective of the COO and are usually associated with complex karyotypes.[13,16] Amplifications and somatic mutations have also been described. These genetic aberrations lead to MYC protein overexpression and activation of a proliferative phenotype. In RCTs, the presence of *MYC* aberrations was associated with poorer OS and EFS, independently of the IPI and the COO classification.[98] Other studies correlated *MYC* aberrations with worse survival.[35,99-101] The presence of MYC staining as detected by IHC was correlated with *MYC* rearrangements in two independent studies.[99,102] MYC+ and BCL2 or BCL6+ double-hit lymphomas[103] have a very poor outcome, which cannot be solely explained by the presence of a *MYC* breakpoint, hence suggesting a synergism between these genetic events. Applying an IHC BCL2 and MYC double-hit score[102] might help to recognize patients with worse OS, and PFS, independently of the IPI and COO. Targeting DNA secondary structures within the *MYC* promoter region is emerging as a potential intervention in DLBCL.[104]

As already stated, **BCL6** is a transcriptional repressor essential for the formation of the GC reaction, and, hence, a key molecule involved in the lymphomagenesis of GCB-DLBCLs. Most genetic aberrations involving *BCL6* lead to its overexpression. Although these are more common in ABC-DLBCLs, translocations involving *BCL6* can occur in both molecular subtypes.[83] *BCL6* mutations have also been described.[105] It has been shown that BCL6 inhibitors are toxic for GCB-DLBCL cell lines even in the absence of chromosomal translocations.[106]

A common event in ABC-DLBCLs is the acquisition of genetic or epigenetic events in ***PRDM1*** (PR domain zinc finger protein 1) leading to repression of Blimp-1, the key transcription factor involved in the terminal differentiation of B-cells.

Similarly, **IRF4** is a transcription factor that has an important role in plasma cell differentiation and in the survival of post-GC neoplasms such as the ABC-DLBCLs. In GCB-DLBCLs, IRF4 was found to be involved in rearrangements with the Ig genes. These cases exhibited a specific GEP and presented a favourable outcome.[107]

Addiction to IRF4 in ABC-DLBCLs is independent of genetic aberrations. Its overexpression is attributed to NF-kB constitutional activation. On the other hand, **SPIB**, its essential partner for DNA binding is targeted by chromosomal translocations, gains and amplifications in ABC-DLBCLs.[82]

The most commonly exploited signalling pathway by malignant B-cells is the **NF-kB pathway**.[83] Crosslinking of different B-cell membrane receptors, including the BCR, CD40 and Toll-like receptors (TLR), leads to phosphorylation of IKBα (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha), unleashing the NF-kB complex of transcription factors from the cytoplasm to the nucleus, where they alter gene expression. The NF-kB pathway activates the transcription of an anti-apoptotic module highly convenient for lymphoma cell survival and chemotherapy resistance. NF-kB constitutive activation constitutes a hallmark for the poor risk ABC-DLBCLs[108] and is a promising candidate for targeted therapies (see below). The mechanisms by which NF-kB activation occurs in DLBCL are diverse.[83] The CBM protein complex, constituted by CARD11 (Caspase recruitment domain-containing protein 11), BCL10 (B-cell lymphoma-10), MALT1 (Mucosa-associated lymphoid tissue lymphoma translocation protein 1) and casein kinase 1α, is activated downstream of Bruton's tyrosine kinase (BTK) and Protein Kinase Cβ (PKCβ) after BCR ligation and is required for activation of NF-kB, either directly or indirectly by inactivation of negative regulators such as A20.[109,110] Gain-of-function ***CARD11*** somatic mutations,

which are detected in 10% of the ABC-DLBCLs,[111] or loss of A20[112] both lead to NF-kB constitutive signalling.

Targeting the NF-kB and other pathways activated via BCR is under scrutiny in clinical trials for DLBCL. Bortezomib is known to block NF-kB by avoiding degradation of IκBα. It has been combined with chemotherapy and immunochemotherapy in phase 2 trials and, as expected, seems more effective in ABC-DLBCLs, improving CR rates and OS.[113,114] Although phase 3 RCTs currently recruiting are trying to demonstrate whether this proteasome inhibitor is really selective against the ABC-DLBCLs, this is not entirely expected given that the drug has a broad activity against protein function. However more selective signalling inhibitors are available.

Despite harbouring molecular aberrations in the Ig genes, B-cell malignancies retain the BCR in the cell surface to profit from its downstream survival and proliferation signals. Mouse models support the role for **BCR signalling** for normal B-cell survival and malignant transformation.[115]

As already mentioned, the CD79A/B molecules sustain the BCR receptor assembly and transmit signals to a variety of downstream pathways. Activation of spleen tyrosine kinase (SYK) and its downstream targets leads to engagement of NF-κB, Phosphoinositide 3-kinase (PI3K), Mitogen-activated protein (MAP) kinase, and RAS signalling pathways, leading to cell survival and proliferation. Around 20% of ABC-DLBCL tumours harbour gain-of-function mutations in ***CD79A/B***, which are rarely, if ever, found in GCB-DLBCLs.[116] Theoretically, the mutant isoforms are selected as this may allow premalignant B-cells to escape anergy and expand in the GC microenvironment, where they can engage in Ig affinity maturation and further expand. *CD79A/B* mutations increase the expression of the BCR in the surface membrane and reduce activation of tyrosine-protein kinase (LYN), a negative regulator of the BCR signalling.

ABC-DLBCLs with wild type *CARD11* exhibit a chronic "antigen-like" BCR engagement able to sustain their survival.[116] In concordance, knockdown of any of the components of the BCR or of its downstream effector molecules, including BTK, or CARD11 itself, leads to cell death. Most of the GCB cell lines are unresponsive to the knockdown of BCR or related molecules.

Ibrutinib, a BTK inhibitor, showed exclusive in vitro cytotoxicity against ABC-DLBCL cell lines, and it has recently been suggested that it has synergistic activity with lenalidomide in blocking the NF-κB pathway in this subset of DLBCLs.[117] In a phase 2 RCT enrolling relapsed/refractory patients, ibrutinib induced a high overall response rate, showing preferential activity against the ABC-DLBCLs.[118] Activity against the GCB subtype is, however, not depreciated and the mechanisms behind this are not entirely clear.

Preclinical studies with Fostamatinib, a SYK inhibitor, revealed its ability to inhibit BCR signalling and induce cell-cycle arrest in cases relying on tonic BCR signalling for survival.[119,120] Fostamatinib showed activity against relapsed DLBCL in early phase trials.[121]

PKCβ inhibitors have been effective in preclinical studies against ABC-DLBCL cell lines with *CD79A/B* mutations[122] and are being used in monotherapy or in combination with Everolimus in patients with the CD79-mutant or the ABC subtype of DLBCL.

PI3K inhibitors also have prominent activity against ABC-DLBCLs with *CD79B* mutations,[123] as well as in cases with *FOXO1* mutations[85] or *PTEN* (Phosphatase and tensin homolog) deletions.[124]

RNA interference screening in ABC cell lines hinted that **innate immune signalling** is involved in the biology of this DLBCL subgroup.[117] Nearly 30% of ABC-DLBCLs have L265P mutations and an additional 10% have other gain-of-function mutations in the ***MYD88*** (Myeloid differentiation primary response gene 88) gene.[125] These occur with

CD79A/B mutations, suggesting oncogenic cooperation, as these pathways are non-redundant. MYD88 is an adaptor molecule that in coordination with IRAK (Interleukin-1 receptor-associated kinase) kinases, is able to couple receptors like TLRs with downstream signalling pathways such as NF-kB and type-I interferon (IFN) circuit. The oncogenic *MYD88* forms express IL-6 and IL-10, which in turn activate JAK-STAT3 (Janus tyrosine kinase-Signal Transducers and Activators of Transcription 3) and NF-kB pathways.[126] Type-I IFN might be beneficial for tumour cells, since IFN-β secretion might dampen the immune microenvironment. Inhibitors of IRAK kinases selectively kill ABC cell lines harbouring *MYD88* mutations.[125] IRAK kinases inhibitors targeting the TLR signalling, probably in combination with other agents with efficacy against the BCR signalling, are promising strategies for these patients with *MYD88* mutations.[127]

Recent cancer genomics studies in lymphoma unveiled recurrent events in molecules that induce **epigenetic changes** in the DNA. Histone modifications can have a permissive or repressive effect in gene transcription by altering chromatin structure.

Mutations in histone modifying enzymes are common in DLBCL. Genes coding for the histone acetyltransferases cAMP response element-binding protein (*CREBBP*), E1A binding protein p300 (*EP300*) and the hystone H3K4 methyl transferase mixed-lineage leukemia protein 2 (*MLL2*) are mutated in up to 40% of DLBCLs, irrespective of the molecular subtypes,[128] suggesting an important role for epigenetic changes in lymphomagenesis. The precise oncogenic mechanism behind these mutations is, however, unclear.

EZH2 is a member of the epigenetic regulator polycomb repression complex-2 (PRC2), responsible for restraining transcription through the methylation of the histones H3K27. Experiments inhibiting *EZH2* proved its role in Ig affinity maturation in the GC by transiently suppressing B-cell differentiation.[129] In concordance, heterozygous gain-of-function mutations in *EZH2* are found exclusively in 21% of GCB-DLBCLs.[130] Hence, targeting *EZH2* represents an opportunity to personalize treatment based on

the COO classification. In fact EZH2 inhibitors were found to be effective in halting proliferation of GCB-DLBCL cell lines and *EZH2* mutant DLBCL xenografts in mouse models.[131] A clinical trial in the United Kingdom will explore the clinical activity of EZH2 inhibitors (GSK2816126) in relapsed/refractory GCB-DLBCLs and will study the EZH2 mutational status irrespective of the molecular subtype (clinicaltrial.gov number NCT02082977).

microRNAs are also altered in DLBCL. It has been shown that the locus that encodes for the miR-17-92 cluster is amplified in 12,5% of GCB-DLBCL tumours, while it is never targeted in ABC cases.[82] miR-17-92 works in collaboration with MYC and induces a proliferative and anti-apoptotic phenotype in cancer cells, as well as promoting escape from senescence and cell cycle checkpoints.[83]
The expression of certain miRNAs was associated with prognosis together with the IPI and the COO in RCTs.[132,133]

1.5 GEP-based prognostic models suggests that a host inflammatory/immune response plays a role in the biology of DLBCL

GEP studies performed in whole LNs infiltrated with DLBCL captured not only the transcriptome of the malignant B-cells and its resemblance with the putative COO but also other relevant aspects of the biology of the disease that could be implicated in the outcome of patients. One such biological feature is the host immune response to the tumour.

Using unsupervised hierarchical clustering, Alizadeh et al.[74] firstly recognized that DLBCL samples had a coherent overexpression of genes known to be expressed by natural killer (NK)-cells and macrophages (including *CD14, CD105, CSF-1R*) and of transcripts involved in matrix remodelling (e.g. *MMP9* and *TIMP-3*). This LN-signature was also highly expressed in normal LNs, but not in other B-cell malignancies. The

intensity of this signature varied, possibly reflecting the relative proportion of tumour and host cells in the biopsy.

A larger study exploring the use of GEP to define prognostic algorithms demonstrated that the molecular COO classification explained only part of the survival variability in DLBCL.[76] Indeed, prognostically significant genes selected by supervised analysis could be grouped into signatures according to their involvement in physiological processes: GC B-cell, LN, MHC class II, and proliferation.

Expectedly, a proliferative signature was associated with a worse outcome, whereas the expression of the MHC class II signature genes by the malignant B-cells improved survival independently of the COO. It was speculated that malignant cells devoid of surface MHC class II molecules could escape T-cell immune checkpoints that would help them to survive, emphasizing the role of the immune system in controlling malignancy.

Genes from the LN-signature incorporated in this model, however, were expressed by the non-malignant cells in the tumours.[134] Importantly, a higher transcriptional activity of these genes was associated with improved outcome even in the poor risk ABC-DLBCLs. Whether this was due to a mere heavier infiltration of immune cells in the samples that would increase the signal intensity of those transcripts in the arrays, or a qualitative difference was unclear.

Another GEP study applied different unsupervised clustering algorithms to model outcome in DLBCL.[12] Comparison of these methods allowed devising robust subsets of patients again highlighting tumour biological features. The “host response” cluster had high expression of T-cell molecules, IFN-induced genes, cytokine receptors, TNF ligands/receptors and extracellular matrix (ECM) component transcripts. Additionally, some of the transcripts of this signature codify for proteins with established roles in macrophage development and function, including *CSF-1R*, *CD14* and *CD163*. This cluster was enriched for genes from the LN-signature previously described. Finally,

samples allocated to this group had a significantly higher numbers of T-cells and DCs by IHC.

The relevance of the stromal compartment in DLBCL was later substantiated by a large study using samples of R-CHOP treated cases.[34] Their research used GEP of sorted malignant and non-malignant cells from a few DLBCL samples to help clarify the relative input of these two compartments on the prognostic algorithm designed. Transcripts that had a differential signal value in either the malignant or non-malignant cells were used to build multivariate survival models that were validated in whole GEP data from almost 400 patients. A malignant-derived GCB signature, together with two others derived from the non-malignant cells were able to predict OS and PFS in the R-CHOP treated patients. Additionally, the IPI and the GEP-based model added to the predictive power of each other, suggesting a shared role for clinical and biological features contributing to patient outcome. The stromal-1-signature was enriched for genes derived from macrophages and ECM components and was predictive of a good outcome. The stromal-2-signature was enriched for genes involved in angiogenesis and conferred an adverse outcome. Importantly, the two stromal signatures were strongly synergic. The relative expression of each of the stromal signatures in each sample was what most predicted length of survival, highlighting the power of GEP to unveil quantitative differences in immune responses that are prognostically significant. These data partially supports the model derived from solid tumours, in which mononuclear phagocytic cells are involved in stromal remodelling, drive the angiogenic switch and contribute to disease aggressiveness. The authors also demonstrated that the LN- signature, which shares a large amount of transcripts with the stromal-1-signature remained a good survival predictor in the R-CHOP era.

In an attempt to satisfy critical standpoints regarding the complexity of the models discovered and its impossible applicability to the clinical practice, other studies tried to define simplified predictive models. Similarly, those models incorporated genes encoding for microenvironment components. A model comprising four genes of the

GCB/ABC signature and two genes related to immune response (*APOBEC3G* and *RAB33A*), showed to be predictive of outcome in patients receiving R-CHOP.[135] Two of the genes included in the six-gene model of Lossos et al.[136] (*SCYA3* and *FN1*), reflect the tumour microenvironment. *FN1* was expressed at very low levels by B-cells, supporting that the transcript is being expressed by the accessory cells. Its expression has been associated with a better outcome, highlighting again the positive prognostic impact of a stromal response in DLBCL.

Alizadeh's group explored bioinformatics methods to analyse the extensive paired clinical and GEP data available online. By univariate analysis, *LMO2* was one of the best genes at predicting OS in DLBCL. Bivariate survival predictor models incorporating *LMO2* and a second gene more highly expressed in non-malignant cells were explored. *TNFRSF9/CD137* was also a strong predictor of good outcome on univariate analysis and the best in bivariate combination with *LMO2.* This bivariate model synergizes with the IPI for predicting outcome in DLBCL.[137] *TNFRSF9* expression was restricted to a minority of infiltrating T-cells of DLBCL.

The above studies consistently highlight that biological features of DLBCL are derived from the stromal microenvironment. Interestingly, in contrast to what is generally accepted in solid tumours and other lymphoid malignancies, in DLBCL the expression of genes derived from macrophages and from the ECM components of the malignant LNs confers an improved outcome.

Functional studies approaching this phenomenon are lacking, but potential hypotheses can be made. The most compelling is that the stromal cells are contributing to an immune response against the tumour that, although not fully efficient, might control some features associated with tumour aggressiveness. Within this hypothesis it can be speculated that the malignant cells of high-risk cases have found ways to dampen and escape this microenvironmental control. Chromosomal copy number changes are associated with differential expression of the LN-

signature,[134] suggesting that oncogenic changes might lead to dependence from tumour microenvironmental signals.

1.6 Limitations of GEP-based predictive algorithms

High throughput strategies such as GEP, NGS or proteome analysis are extremely useful to generate hypotheses for subsequent studies where simpler, affordable and highly reproducible methodologies are applied in real world patient's samples. The COO and other GEP-based strategies have the potential to improve prognostication in DLBCL and offer an unforeseen opportunity for tailored treatment in this disease. However a large body of criticism is out there regarding the implementation of these strategies in the context of RCTs and in the clinical practice.

The studies here described, although proving the importance of the molecular subgroups or the tumour microenvironment in the biology of DLBCL, have used different gene lists for their survival algorithms. After more than 10 years after the first study proposing a strategy for classifying patients according to the COO, other GEP-based algorithms are still being proposed.[79] This failure to implement and validate previously proposed methodologies suggests difficulties in standardization and arbitrary analysis approaches. In fact there is a lack of standard operating procedures for the analysis of GEP.

Another limitation is the number of samples explored in the original studies. Statisticians highlighted that inadequate sample sizes compromise any correlation of high throughput data with survival information, most times generating results that overestimate the prognostic value of a certain gene or algorithm. Additionally, there is some criticism of the use of cohort splitting into training and validation sets, which further reduces the size of the already small training group that is used for the development of the risk prediction model thus increasing problems of overfitting. In consequence results might not be replicated in subsequent validation studies.

Alternative strategies to sample splitting are proposed, such as the bootstrap cross-validation approach.[138] Since there are GEP results of over 700 patients with DLBCL available for the scientific community, techniques that can integrate this information, such probabilistic Bayesian modelling,[139] would almost certainly generate more coherent strategies.

The contribution of the microenvironment in algorithms developed using whole tumour cells is going to be variable. It will probably depend not only on the extent of non-malignant cells infiltrating the primary samples, but also on the differential sensitivity of particular cell types to the harvesting, RNA extraction and other techniques used prior to sample arraying.

Finally the lack of controlled prospective data is one of the main hindrances of gene-based prognostic models that might be overcome in the near future once results from RCTs applying GEP are reported.

The aspects exposed above hamper the implementation of the findings of GEP into the clinical arena. This contributed to the search for other methods that could be predictive of the COO, such as IHC classifiers. We will challenge the current use of COO IHC classifiers in Chapter 3.

PART TWO

1.7 Cancer and Immunity

Cancer cells are equipped with the ability to subvert the physiological processes that control the progression from a normal to a neoplastic state (reviewed by Hanahan and Weinberg[140]). One component of such physiological mechanisms is their close interaction with the immune microenvironment.

Whereas the original theory of a vigorous immune cell infiltration being an indicator of the host antitumoural response against the tumour still holds true for certain immune cell subsets in specific cancer types,[141] a more recent theory recognizes a double-faced immune system with both host-protecting and tumour-promoting features.[142]

As early as 1863, Virchow recognized the infiltration of cancer tissues with cells of the innate immune system, including macrophages, and hypothesised that mechanisms of immune activation such as chronic inflammation enhanced cell proliferation, creating the soil for cancer development.[143] Following Virchow's original hypothesis, a large body of work published over the last 20 years supports that tumour-associated inflammation could indeed be permissive to cancer formation.[144-146]

An inflammatory response can give rise to a number of molecules that are fundamental in cancer development. For example, reactive oxygen species (ROS) released in the context of an inflammatory response enhance the acquisition of oncogenic events in mammalian cells.[147] Additionally, inflammatory cells produce an array of molecules that have tumour-promoting functions: growth factors that sustain cancer proliferation; survival factors that limit cancer cell death; proangiogenic and matrix remodelling molecules that facilitate angiogenesis, invasion, and metastasis; cytokines and chemokines that amplify the inflammatory response; and immune suppressive molecules that are crucial for tumour cell escape. Cancer associated-inflammation closely resembles wound healing and subsequent

tissue remodelling responses.[148] Importantly, B-cells have been implicated in the recruitment, activation,
and persistence of such wound-healing and tumour-promoting innate immune cells.[149,150]

However the scenario is more complex and highly dependent on the type of tumour, on where the inflammatory cells lie within the tumour, and on the temporal evolution of the tumour, which endangers a general view and therapeutic approach of inflammation in cancer.

1.8 Evolving models of macrophage ontogeny

Macrophages are an obligatory component of all tissues in humans, constituting 5-20% of their cell content.[151] Macrophages were recognized as phagocytic cells involved in tissue function and homeostasis in the late 19th century, by Elie Metchnikoff. Only 100 years later van Furth et al.[152] described the mononuclear-phagocyte system (MPS) as a linear model where macrophages were terminally differentiated cells derived from monocytes arising from a common progenitor in the BM. The discovery of a common progenitor to macrophages and DCs led to the inclusion of the latter in this system.[153]

Colony stimulating factor-1 (CSF-1 or M-CSF) is the main cytokine involved in the development and function of cells of the MPS. This has been demonstrated in mice lacking functional Csf-1 or Csf-1 receptor (CSF-1R) genes.[154] The intriguingly worse phenotype of CSF-1R null mice led to implication of an alternative CSF-1R ligand, IL-34, in maintaining normal phagocytes, especially epidermal macrophages and microglia.[155] The E-twenty six (ETS) family transcription factor PU.1 controls the expression of CSF-1R. It has been demonstrated that by introducing PU.1 in PU.1-/- cells it is possible to rescue differentiation of CSF-1R deficient cells.[156]

Over the last decade studies on the dynamic evolution of monocytes and macrophages conducted to major breakthroughs in the field of macrophage ontogeny and previewed opportunities to selectively target recruited macrophages in the context of chronic diseases such as cancer without disturbance of local homeostatic macrophages. We will now highlight some of those findings.

Monocyte heterogeneity has been acknowledged. Two monocyte subsets exist in mice that show different phenotype, namely the expression of Ly6C, and function.[157,158] The Ly6C+ "inflammatory" monocytes have a short life span and are bound to traffic to sites of infection and inflammation, whereas another population crawls along the luminal surface of the vessels and patrols for integrity of the endothelial cells. These two populations might have a differential contribution to the resident macrophage pool.[159] What remains to be clarified is whether theses subsets derive from a common progenitor or whether the minor population of "patrolling" monocytes originates from the "inflammatory" subset.[153]

Importantly, there is evidence for functional homologues of mouse monocyte subsets in humans.[160] A small population of monocytes with dimmer expression of CD14 compared to the classical monocytes was described in humans which has a similar cytokine profile and antigen presentation capacity that the mouse "patrolling" monocytes. The impact of human monocyte subsets in pathological conditions is yet to be defined.

Another seminal finding was that embryonic phagocytes derived from non-haematopoietic cells in the yolk-sac and the foetal liver can persist after birth in most tissues and can repopulate them in the absence of a functional haematopoietic system.[161,162] These are self-maintained by slow but steady proliferation in homeostatic conditions[163,164] under the regulation of the transcription factor MAFB/cMAF.[165] The suggested contribution of this proliferation is to populate tissues with macrophages during organogenesis and at a lower rate thereafter to maintain tissue integrity. Such phenomenon would explain why haematopoietic stem-cell transplantation leads to inefficient replacement of tissue macrophages. Nonetheless

adult organs are infiltrated, with the exception of the gut,[166] by a mixed population of embryonic and bone marrow-derived macrophages. Macrophages from distinct progeny both rely on CSF-1.[167]

The balance of pre-natal and bone marrow-derived macrophages has been shown to change in pathological conditions.
In acute inflammation,[168] including IL-4-driven parasitic infections[169] the dynamics of resident macrophages changes: the proliferation capacity of resident macrophages is boosted, whereas monocytes recruited from circulation only transiently repopulate tissues and differentiate in situ to macrophages.[170]

The discovery of proliferation enhancement by IL-4[171] raises the possibility that accumulation of macrophages in tumours might also be partly due to local macrophage proliferation. Recent evidence, however partially disproves this hypothesis. Using a model of mammary cancer genetically driven by the expression of an oncogene, Franklin et al.[172] demonstrated that two populations of macrophages co-inhabited the tumour, which were phenotypically and transcriptionally diverse. One subset accumulated upon tumour growth and was required for tumour support by impeding infiltration of cytotoxic T-cells and by promoting tumour growth. These tumour-associated macrophages (TAM) had origin mostly on attracted peripheral blood inflammatory monocytes and exhibited high proliferative capacity. Targeted genetic knockout experiments implicated Notch signalling in the differentiation of TAM. On the other hand, another subset of macrophages relied solely on the input from circulating monocytes and seemed oblivious to the presence of the tumour cells, showing less proliferation and a transcriptome that equipped them to maintain tissue homeostasis. Importantly, this population persisted despite inhibition of Notch signalling. This study teasing out the origins of TAM in mice was able to describe a new pathway dependency for these cells. Moreover, the formal demonstration of the existence of “non-malignant” macrophages independent of the same signalling pathway suggests that TAM targeting strategies that do not compromise innate immunity and tissue homeostasis might be possible.

Although studies such as the one described change current paradigms in the understanding of macrophage biology, their translational potential has, however, to be questioned. The lack of any data in humans paralleling elegant findings in the murine model hinders the relevance of such findings. Additionally, the expansion of innate cells necessary for pathogen control or wound repair independently of monocyte recruitment puts into question standard methodology used to obtain human macrophages in the laboratory. The reasons for limited validation of macrophage studies in humans will be discussed later to highlight the relevance and novelty of the studies we undertook in the context of DLBCL.

1.9 Macrophage functions in normal physiology

Understanding the functions in which macrophages take part in normal physiology is important, as in the context of cancer those attributes are subverted via permanent contact with the tumour cells or products produced by the tumour cells, generally to favour tumour progression. We will next expose some of the fundamental functions of macrophages in healthy conditions. TAM functional diversity will be discussed in the next section.

Macrophages are cells from the innate immune system that play important roles in the immediate response to pathogens and in the regulation of humoral and cell-mediated immune responses.[149]

The recognition of moieties such as lipopolysaccharide (LPS) induces functional activation of macrophages towards a pro-inflammatory and pro-apoptotic phenotype. While pro-inflammatory molecules induce oxidative processes that contribute to the killing of invading organisms,[173] macrophages release soluble pattern-recognition molecules that activate the complement system and amplify the killing process. Additionally macrophages skew the differentiation of T-cells towards Th1 and Th17 phenotypes, boosting the whole inflammatory process.[174] During the resolution phase and on subsequent encounters with bacterial products macrophages

should skew towards wound healing and immune-regulatory properties that minimize tissue damage and re-establish equilibrium. A failure of this control mechanism leads to chronic inflammation and auto-immunity.

Besides host defence, macrophages are involved in many other aspects of tissue homeostasis (reviewed by Wynn, Chaula and Pollard[175]). Specialized tissue macrophages including Kuppfer cells, osteoclasts, microglia, and alveolar macrophages, all display functions that help keeping the integrity of the organs where they reside.

Macrophages contribute to tissue remodelling and repair under homeostatic and damage conditions. This role is observed in the prenatal stage and in adult animals. During embryogenesis macrophages are localized in areas of active tissue reconstruction, such as the inter-digit areas during limb formation.[155] This has been demonstrated by using the MacGreen mouse model[176] in which macrophages express a green fluorescent protein (GFP) from the Csf-1r promoter. In this site, macrophages scavenge cell debris and produce a myriad of molecules responsible for ECM degradation, reconstruction and angiogenesis. Additionally, macrophages reduce immune responses against auto-antigens that naturally occur during this process.[177] Studies of Csf-1 null mutant mice, which lack many macrophage populations, highlighted the role of macrophages in morphogenesis. These mice had tissue-remodelling defects involving the bone, mammary gland and pancreas. The same macrophage homeostatic and reparative attributes are recapitulated in the process of wound healing.

Macrophages ensure homeostasis in haematopoiesis in a number of ways, from controlling the egress of haematopoietic cells from the bone marrow to engulfing apoptotic neutrophils and red blood cells in the spleen. Bone marrow macrophages ingest the extruded erythrocyte nuclei from erythroblasts, a process fundamental to ensure normal erythropoiesis.[175]

The main iron supply for erythropoiesis derives from recycling after phagocytosis of

senescent red blood cells by macrophages.[178] These sequester the iron intra-cellularly in the context of infection in an attempt to withhold it from the invading pathogens. Recent evidence suggests, however, that iron storage homeostasis is influenced by the activation pattern of macrophages.[179] Macrophages also contribute to maintaining metabolic homeostasis,[180] such as promoting peripheral insulin resistance to increase nutrient supply to innate immune cells during bacterial infection.

1.10 Shifting the paradigm of macrophage polarisation

Macrophages are highly plastic cells, assuming unique phenotypes and functions consequent to their diverse ontogeny and contextualized in their ever-changing microenvironment. The complexity of macrophage polarisation is only recently being recognized. For ease of understanding immunologists attempted to categorize the spectrum of macrophage functional activation.

1.10.1 Classical and alternative macrophage activation

Using murine models of infection, Mackaness and collaborators observed reproducible changes in the morphology and antibacterial activation of macrophages that were induced by cellular factors upon second exposure to the pathogens.[181] Those cells were later identified as T-cells and the key soluble factor IFN-γ.[182] Under the influence of this cytokine and the bacterial product LPS, macrophages were potently activated to respond to foreign antigens by increasing antigen presentation and phagocytosis and by producing chemokines and proinflammatory cytokines, such as IFN-γ, IL-12, tumour necrosis factor (TNF), IL-6, and IL-1β. This reproducible phenomenon induced by intracellular pathogens and tumour cells[183] was coined macrophage classical activation. Granulocyte macrophage colony-stimulating factor (GM-CSF) was later found to induce a secretory pattern and function similar to IFN-γ and LPS.

In 1984 Mossman and Coffman[184] made an important discovery in the field of T-cell immunology that would help to understand macrophage polarisation even further. The authors found that IFN-γ and its then recently discovered opposing cytokine, IL-4, were produced by mutually exclusive CD4+ helper T-cell populations after TCR stimulation and hence were able to inhibit each other's functions. These were named Th1 and Th2 helper cells. Meanwhile other Th2 cytokines were beginning to be discovered, including IL-10 and IL-13.[185] Th2-mediated immune responses were found to be required to control infections by extracellular parasites, including helminths and protozoa.

Similar to the findings with IFN-γ, it was then hypothesized that Th2 cytokines such as IL-4 could also influence macrophage behaviour. Investigations recognized that IL-4 inhibited the production of pro-inflammatory cytokines and induced the expression of MHC class II and mannose receptor (MR) in macrophages. IL-4 induced macrophage fusion and decreased phagocytosis. This consistent response led to the recognition of the macrophage alternative activation response.[186] The alternative macrophage phenotype was entirely different from the IFN-γ-induced classical phenotype and served different purposes in an immune response.

It was later recognized that this dichotomic classification represented the extremes of a variety of macrophage activation patterns that could be elicited by different triggers. To acknowledge that, Mantovani et al.[187] proposed that the alternative pole should be ramified according to the stimuli used to trigger macrophage activation other than IL-4, including IL-10, glucocorticoids and immune complexes.

1.10.2 M1 and M2 activation

Hsieh and collaborators[188] recognized that Th1 and Th2 T-cell responses in mice were influenced by their genetic background. T-cells from C57BL/6 mice preferentially produced IFN-γ, whereas those from BALB/c mice favoured Th2 responses by producing high levels of IL-4. These polarised T-cell responses were independent of the function of antigen-presenting cells in the different strains and were linked to the

differences in susceptibility to Leishmaniasis that were observed in the mice.
While studying arginine metabolism, Mills et al.[189] recognized that macrophages from C57BL/6 and BALB/c mice metabolized arginine in different ways, but also had qualitative differences in the macrophage responses to the classical stimuli IFN-γ and LPS. The researchers hence proposed a new macrophage polarisation classification: M1 and M2 macrophage responses. M1 paralleled the classical whereas M2 paralleled the alternative macrophage responses. This simplified classification suffered ramifications to appreciate the combined effect of IFN-γ and LPS or the effect of the multiple M2 stimuli just described.

The polarised macrophage subsets were and still are defined according to responses induced by in vitro cytokine stimulation. The read-out of such experiments is usually limited to morphology, intracellular or surface protein detection and molecule release to the culture media.

Classical/M1 macrophages produce high levels of pro-inflammatory cytokines, including IL-12 and IL-23 and TNF-α; have an IL-12^{high}, IL-23^{high}, IL-10^{low} phenotype; express high levels of MHC molecules and co-stimulatory molecules, and Th1 cell–attracting chemokines such as CXCL9 and CXCL10. M1 cells metabolize arginine to nitric oxide and are proficient at engulfing pathogens, presenting antigen, and killing tumour cells.[190] On the other hand alternative/M2 macrophages have high expression of IL-10, IL-1 decoy receptor and IL-1RA and low expression of IL-12; express the chemokines CCL17, CCL22 and CCL24, have high levels of scavenger, mannose and galactose receptors, and the arginine metabolism is shifted towards ornithine and polyamines. Functionally, M2 macrophages induce Th2 responses, clear intracellular parasites and remodel the extracellular-matrix.

1.10.3 Problems of macrophage activation nomenclature

We should at this point acknowledge the chief problems of the classification systems just described, which galvanized seminal advances on the understanding of

macrophage polarisation.

The number of molecules known to activate macrophages has grown beyond LPS, IFN-γ, IL-4 and other alternative stimuli contemplated in the current classifications and include tissue oxygen tension[191] or pH.

Macrophage activation is influenced by factors other than the activator itself that are disregarded in the current classifications, including: the macrophage origin, the source of the stimulus and the multitude of concentrations and longevities of exposure to the stimulus, among others.

The lack of robust markers for macrophage polarisation hinders the evaluation of studies exploring this subject.

Macrophages do not form clear-cut activation subsets. Indeed they can develop mixed M1/M2 phenotypes in vitro and certain pathological conditions, including the resolution phase of bacterial infections.[192] By stimulating macrophages with a panel of six triggers, Stout et al. documented that macrophages can reversibly shift their functional phenotype through a multitude of patterns very different from the categorized classifications.[193] Sequential treatment of macrophages with multiple cytokines led to a progression through multiple functional phenotypes, suggesting that macrophage functional adaptivity is retained and can be therapeutically manipulated.

In the same line, a polarised macrophage phenotype can be reversed by applying triggers of the opposite pole. As Stout et al.[193] have shown treatment of macrophages with the Th2 cytokine IL-4, before LPS stimulation strongly enhanced production of inflammatory cytokines (TNF-α and IL-12 production), whereas inhibited anti-inflammatory IL-10 production. The impact of intermediate stages of activation during this process is unclear.

Complex in vivo systems where numerous cytokines and growth factors interact to determine the function of macrophages is difficult to mimic in vitro. It is unreasonable to accept that such simplified classifications replicate different tissue-specific and pathological scenarios.

1.10.4 Transcriptional analysis of macrophages

Until recently, macrophage classifications were hampered by the lack of methodologies for analysis of the influence of a certain stimulus administered to macrophages. Since the advent of GEP, a body of work has acknowledged that macrophages harbour specific gene expression signatures according to the animal species, the location where they reside within the organism and the different environmental signals they received. Moreover this technology helped predicting novel transcriptional regulators implicated in discrete activation conditions.

Due to their seminal role in maintaining tissue homeostasis, it is accepted that macrophages are kept in a restrained functional mode, which is unleashed once the cell receives “danger” signals from the microenvironment. Since macrophages play a large number of functions in tissues and are instructed by a multitude of molecules and other immune cells, plasticity in gene expression makes sense and leads to rapid functional adaptation according to their ever changing surroundings. Ultimately there is no macrophage identical to another as minute changes in the microenvironment have a great potential to influence the cell transcriptome and ultimately the function.

Martinez et al.[194] pioneered the field by describing the cardinal changes in the transcriptome during monocyte-to-macrophage differentiation and by comparing the GEP of resting and polarised macrophages in humans. Besides providing an extensive list of novel genes associated with maturation and activation, this study demonstrated that CSF-1, the standard cytokine used for maturation of macrophages in vitro and known to circulate at high levels in normal blood, inclined macrophages

towards an M2 transcriptome. Thus it could be suggested that an M2 GEP is a standard pathway in macrophage differentiation. This in our opinion also highlights the need for well-standardized cell maturation protocols.

Other GEP studies of IL-4 stimulated macrophages had been undertaken to illuminate the biology of alternative activation in humans.[195,196] Some prototypic mouse M2 genes have no homologues in humans or do not undergo transcriptional regulation by the standard cytokines and hence are irrelevant in pre-clinical models. By analyzing the 1000 most highly expressed genes, Martinez et al.[195] could recognize 87 with concordant protein expression that constituted conserved markers in mice and humans, and as such could be useful for translational research. One example is the enzyme transglutaminase-2 (TGM2). This gene signature helped predicting the extent of macrophage infiltration in human lungs. Finally, this study recognized that the G-protein coupled receptor signalling had the greatest number of genes affected by IL-4 activation in humans. Pello et al.[197] also reported novel human markers of alternative activation, including Scavenger receptor class B member 1 (SCARB1) and Arachidonate 15-Lipoxygenase (ALOX15), which the authors validated at the protein level in physiological and pathological conditions.

A study recently published[198] undertook the extraordinary task of analysing GEP data from human macrophages obtained in a standardized manner and activated by 28 different signals (single or in combination), including pattern recognition receptor ligands, cytokines, and metabolic factors. This comprehensive approach addressed the criticism directed against oversimplified in vitro procedures to mirror in vivo complex systems, and demonstrated what has long been suggested but not formally proven: that macrophage activation in humans works in a spectrum much broader than the current M1 versus M2-polarisation model. Nine distinct macrophage activation programs were identified and correlated to transcriptional regulators. IFN-γ selectively induced *FEM1C* gene expression, which could constitute a potential specific marker of classical activation.

In a similar comprehensive way, the "Immunological Genome Project" has assessed the transcriptional profiling of resident mouse macrophages.[199] Supporting the relevance of the anatomical context on functional diversity, macrophages sorted from different organs exhibited remarkable transcriptional diversity with minimal overlap. The authors also suggested which transcriptional regulators could be implicated in universal and specific macrophage programs and recognized well-characterized and novel protein markers, including CD14, CD64, and MerTK that help identifying all tissue macrophages. Importantly, proteins previously predicted to distinguish macrophages from other cell types, such as CD68 and CSF-1R, did not emerge as the most powerful markers of macrophages.

Studies such as these require substantial financial investment, alongside complex analysis done by researchers fully dedicated to the field of bioinformatics. However, they form an unprecedented platform for future validations in the setting of human diseases where macrophages are implicated, such as cancer. Moreover, they open the perspective of selectively targeting macrophages in diseased organs without affecting others cell types.

The molecular mechanisms that trigger and sustain macrophage transcriptional patterns of activation are being unravelled. By signalling through TLR4 and IFN receptors, the classical stimulants IFN-γ/LPS induce activation of the NF-kB pathway,[200] together with activator protein 1 (AP-1), IRF3, STAT1, and EGR (early growth response) family members, ultimately leading to the M1 transcriptome. Studies demonstrated that inhibition of the NF-kB pathway by conditional gene knockout[201] leads to a switch from M2 to M1 phenotypes, providing a proof-of-concept for the plasticity of mature activated macrophage phenotype.

By contrast, signalling downstream of IL-4 involves the activation of JAK kinases and STAT6, the master regulator of the M2 transcriptome.[202] Other transcription factors have been implicated in the regulation of subsets of M2-induced genes, including IRF4, peroxisome proliferator-activated receptor-γ (PPAR-γ),[203] MYC,[196] or CCAAT-

enhancer-binding protein beta (C/EBPβ).[204]

Recently the role of epigenetic mechanisms in the control of the macrophage functional repertoire has been described (reviewed by Ivashkic[205]). Mice genetically engineered to have a non-functional NF-kB displayed impaired chromatin remodelling, emphasizing the role of NF-kB not only as a transcription factor, but also as a mediator of genetic transcription. The histone demethylase jumonji domain containing-3 (JMJD3) has a potential role in the modulation of macrophage activation, as has been shown to upregulate a restricted subset of M2 markers, whereas downregulate M1 genes.[206]

Finally, PU.1 has been described as the master regulator of gene expression in macrophages. Genome-scale chromatin immunoprecipitation analysis to identify PU.1 binding sites has shown that this transcription factor determines the macrophage GEP repertoire by generating open chromatin around cell-specific enhancers.[207] This mechanism currently explains functional heterogeneity.

The confirmation of the functional influence of a gene or genetic program activated by a transcription factor would ultimately require animal models where the activity of the gene under study could be manipulated specifically in macrophages. This is enabled by enzymatic knock-in experiments on gene promoters or by macrophage ablation through the expression of the Diphtheria toxin receptor. However, there is no specific promoter of macrophage genes that can be exploited. These are also expressed in most macrophage types, making it difficult to discriminate the functions of sub-classes of macrophages. Investigators acknowledge the limited availability of these models that could specifically dissect functionality in mature tissue macrophages in steady state or under activation.

1.10.5 Obstacles to the understanding of human macrophage biology

Much work published in high impact factor journals using mouse models of macrophage biology strikingly lacks human correlations. This hinders a good understanding of human macrophage biology, and is related to a number of barriers.

The first obstacle is obtaining human tissue samples. Macrophage functional studies should be done using single cell suspensions, ideally fresh but acceptably stored in liquid nitrogen under well-standardized conditions. This implies the existence of tissue banks, dedicated staff and ethical approval in place.

The use of cell sorting allows obtaining highly pure cell populations. However, difficulties in obtaining sufficient, highly purified cells from tissue specific populations, such as macrophages from LNs are well accepted. Although some work has been done using laser captured macrophages from FFPE tissue,[208] issues related to purity and sample quality compromise results. Hence this methodology has not been generally pursued, although this may change with technical improvements that allow characterisation of smaller samples.

A number of central differences exist between mouse and human macrophages[209] that put into question much of the translational relevance of murine data. One difference relates to nitric oxide (NO),[210] an indispensable product of macrophage defence responses against pathogens in mice. Although IHC studies suggest that iNOS, the source of NO, is expressed in human macrophages,[211] in vitro experiments failed to demonstrate production of NO. Similarly, human macrophages also have restricted activation markers that consequently cannot be functionally validated in mouse models.[195] The reasons underlying these discrepancies, either technical or truly biological, remain unclear.

Comparative analyses of mouse and human macrophages have been performed using peripheral blood or BM-derived cells subjected to artificial stimulation in vitro; studies using other tissues are lacking. GEP comparative studies have highlighted

important differences.[212] Similarly, there is limited comparison of results obtained from tissue macrophages in a disease context and the results obtained from in vitro matured monocytes, the latter being usually accepted as adequate to predict the macrophage activation status of in vivo. This is particularly true in the human setting.

Finally, as genomic and proteomic data becomes available underlining macrophage heterogeneity between samples and within the same sample, novel techniques to label several proteins and RNA transcripts simultaneously in situ will be required to validate potential prognostic subsets in large datasets of patients. Only a combination of markers can ascribe activation outcomes.[213] Moreover, single-cell studies would be the only way to tease out whether subsets of macrophages within the same tissue have different activation profiles or whether a macrophage can acquire a unique and hybrid activation pattern. Novel techniques for single-cell studies are becoming available and could be exploited.[214,215]

In our studies we set out to compare the transcriptome of human DLBCL-associated macrophages with their "normal" counterparts. During this endeavour we faced the same obstacles highlighted in the literature to justify the scarcity of human macrophage data. However we feel we have provided data with enough quality to allow the recognition of a lymphoma-specific macrophage GEP and the generation of hypotheses for functional studies.

1.11 Tumour-associated macrophages (TAM)

Macrophages infiltrate all tumours in varying densities and constitute a hallmark of tumour-associated inflammation. A recent meta-analysis[216] of studies evaluating the prognostic impact of macrophage infiltration in human cancer confirmed that, with a few exceptions[217,218] TAM density inversely correlates with OS. Our own research has shown the same association in cHL.[219] However these data are correlative and the functional role of TAM in human cancer is not well established. As a consequence the

data discussed here is extrapolated from murine models and needs to be interpreted with care.

Macrophages in tumours mainly originate from recruited monocytes by CCL2 and CSF-1. In a model of breast cancer, stromal and tumour-derived CCL2 instructed inflammatory monocytes to migrate to metastatic niches.[220] Blocking of the chemokine halted monocyte recruitment and metastasis formation. CCL2-mediated monocyte recruitment has also been observed in primary follicular lymphoma.[221] In Csf-1 null mice susceptible to development of spontaneous mammary tumours, transgenic expression of the cytokine in the mammary epithelium led to an accelerated infiltration of macrophages into the primary tumour.[222] Our own research has shown that CSF-1R inhibition reduces TAM infiltration and tumour growth in a transplantable model of MYC-driven mature aggressive B-lymphoma (Hallam et al., under submission). In human cancer, CSF-1 protein expression at the tumour invasion edge correlated with macrophage infiltration.[222]

Although a few functional studies suggest that macrophages retain tumouricidal function,[223,224] clinical and experimental data more compellingly supports the opposite: that a strong inflammatory response and macrophage infiltration enhances tumour formation and progression. Indeed, there are a number of well-recognized tumour-promoting features of macrophages that are reminiscent of their functional diversity in health. It is accepted that tumour cells find direct or indirect ways of exploiting the repertoire of functions with which macrophages are equipped in physiological conditions to help them thrive.

In Figure 1.2 we provide a list of tumour-derived molecules that have been implicated in cancer facilitating features of TAM. Data on the crosstalk between B-cells and macrophages will be discussed in a separate section.

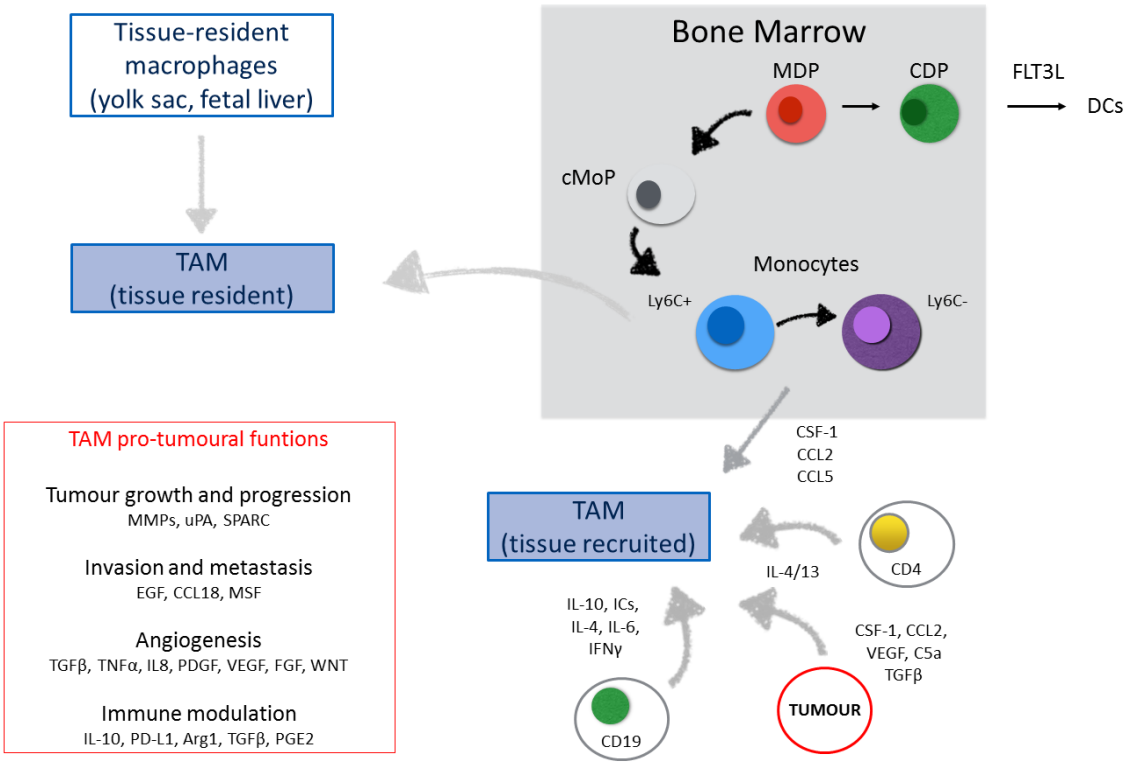

Figure 1.2 Origins of tumour-associated macrophages and crosstalk with tumour cells and other cells of the microenvironment.

TAMs can originate from the tissue resident pool, but mainly derive from inflammatory (Ly6C+) monocytes recruited from peripheral blood by CSF-1, CCL2 or CCL5. Within the microenvironment TAM can be influenced by a number of tumour-derived molecules, including CSF-1, CCL2, VEGF, TGF-β and others. T and B-cells in the tumour stroma can also influence macrophage behaviour through a variety of molecules. TAM play an number of pro-tumoural functions; examples and respective instrumental molecules are detailed in the bottom left box.

MDP, macrophage and DC precursor; CDC, common DC precursors; FLT3L, FMS-like tyrosine kinase 3 ligand; Ly6C/G, lymphocyte antigen 6C/G; cMoP, common monocyte progenitor; uPA, urokinase; SPARC, secreted protein acid and rich in cysteine; EGF, epidermal growth factor; MSF, mitochondrial important stimulation factor; PDGF, platelet derived growth factor; VEGF, vascular endothelial growth factor; FGF, fibroblast growth factor; WNT, wingless-type MMTV integration site family; PD-L1, programmed death-ligand 1; Arg, arginine; PGE2, Prostaglandin E2.

The protumoural attributes of TAM stem from their inability to switch-off activated phenotypes as described below.

1.11.1 Protumoural features in the context of a macrophage classical activation

Although an M1 phenotype can be efficient at eliciting tumour killing, the release of classical pro-inflammatory cytokines such as IFN-γ, TNF-α or IL-6, contributes to sustained inflammation and creates a mutagenic environment. A paradigmatic example of this mechanism is colitis-induced carcinogenesis.[225]

Macrophage-derived IL-6 can lead to STAT3/NF-kB activation and subsequent proliferation of tumour cells. Using a mouse model of B-cell lymphoma, Gilbert and Hemann propose that stromal derived IL-6 sustains malignant B-cells within a chemo-resistant niche. In gastric MALT lymphoma, it has been suggested that persistent infection with Helicobacter pylori induces macrophages to produce a proliferation inducing ligand (APRIL) that in turn promotes the survival and proliferation of neoplastic B-lymphocytes.[226]

Furthermore, it has been demonstrated that M1-derived cytokines affect the dynamic of T-cell infiltration and proliferation in tumours. Genetic deletion or antibody-mediated elimination of IL-23 increased infiltration of cytotoxic T-cells into the transformed tissue, rendering a protective effect against chemically induced carcinogenesis.[227]

Additionally, Kryczek et al.[228] hypothesized that CD80+ TAM could suppress cancer-associated T-cell immunity. After co-stimulatory blockade with an anti-CD80 antibody, macrophages regained the capacity to stimulate T-cells, contributing to regression of ovarian cancer in mice.

1.11.2 Protumoural features in the context of a macrophage alternative activation

An M2-like activation pattern specialized in resolving inflammation through tissue remodelling and immunoregulation, is more common in the context of cancer mainly

in the later stages of the disease. This phenotype promotes immunossupression,[229] tumour invasion and metastasis.[220]

Tumours invade the surrounding ECM by exploiting macrophages trophic ability, a functional attribute of alternatively/M2-activated macrophages. Destruction of the tumour edges facilitates metastasis formation by allowing tumour cells to escape into the circulatory or lymphatic system. A number of macrophage-derived molecules can reshape the ECM and promote cancer invasion, including metalloproteases,[230] cathepsin proteases[231,232] and EGF.[233,234] Macrophage-specific depletion of cathepsins resulted in reduced tumour cell invasion in mammary cancer models. The fibronectin isoform MSF (migration-stimulating factor) has also been shown to promote invasion at the tumour front by enhancing tumour cell motility.[235] Importantly a crosstalk between adaptive and innate immunity has been implicated in this process, as CD4+ T-cells are the required secretors of IL-4 that polarises macrophages towards this invasion enabling phenotype.[231,234]

Macrophage alternative activation has been shown to impair T-cell activation and effector functions by secreting a variety of immunosuppressive factors including IL-10, Prostaglandin E2[236] (PGE2), and arginase. M2 macrophages expressing programmed death-ligand 1 (PD-L1) have also been implicated in the suppression of cytotoxic T-cell responses against tumours.[237,238] Furthermore, M2-like TAM enhance recruitment of suppressive regulatory T-cells through production of the chemokine CCL22.[239]

Tumours also exploit M2 TAM-mediated angiogenesis. This function is triggered in avascular regions in tumours, where hypoxia[240] induces the expression of macrophage attractant molecules, such as vascular endothelial growth factor (VEGF),[241,242] CSF-1, TNF-α, IL-1β or IL-8.[146] Macrophages are then shifted towards a pro-angiogenic phenotype, and hypoxia-inducible factor-1α (HIF-1α) has a key role in controlling this response.[243] In addition, HIF-1α activation also potentiates macrophage-mediated T-cell suppression in vitro.[244] In concordance, a unique

population of angiogenic TAM has been characterized that also displays an immunossupressive function.[245]

By blocking distinct mechanisms of macrophage-induced angiogenesis, different groups substantiated the relevance for this process in tumourigenesis. Ablation of specialized angiogenic macrophages inhibited tumour angiogenesis, growth and metastasis[246] in murine models. On the other hand, blocking CSF-1 in tumours has also been shown to inhibit angiogenesis and decrease tumour growth in mammary cancer model. Additionally, blockade of the Semaphorin 3A/neuropilin-1 signalling pathway in macrophages impeded their migration to hypoxic tumour regions and relented angiogenesis.[247] Importantly, this mechanism is currently being tested in clinical trials. Finally, low-doses of anti-VEGFR2 have also been demonstrated to normalise angiogenesis and reprogram TAM towards a tumour surveillance phenotype.[248]

1.11.3 "Gray areas" of TAM activation

However this clear-cut scenario does not completely hold true, as TAM exhibit substantial heterogeneity depending on the type of tumour; and undergo dynamic changes in phenotype and function according to the temporal evolution of the tumours[175] or to their location within the tumours.[224,229,247,249-251] This might account for the detection of independent subsets of macrophages with M1 or M2-like features,[252] or "hybrid" M1/M2 TAM populations with both tumour permissive and opponent functions in tumours.[172,236,250,253] The role of macrophages in mediating efficacy of monoclonal antibody therapies is of particular relevance and will be discussed in a separate section.

These results caution against the overestimation of studies on the basis of whole TAM populations and compel to independently studies of TAM in each cancer type. Unravelling the network of signalling molecules and transcription factors that underlie these different activation states in TAM will help to devise conservative

strategies targeting the "malignant" macrophage population. In fact, important studies have demonstrated that it is possible to revert TAM towards an anti-tumour phenotype.[201,254-256]

1.12 Crosstalk between B-cells and macrophages and their relevance for antibody therapies

By studying the secretory phenotype of macrophages after crosslinking of the Fcγ receptors (FcγR), researchers documented a shift towards an M2-like phenotype with production of IL-10[257] and PGE2.[258]

B1 cells, initially identified in mice and recently proved to exist in humans,[259] are a subset of B-cells that constitutively produce IL-10. B-1-derived IL-10, in the presence of LPS, has been shown to also upregulate IL-10 production by macrophages.[260]

These studies provide evidence for two mechanisms whereby B-cells influence macrophage effector functions towards immunossupression. It has been recently suggested that these mechanisms can regulate the recruitment and effector function of TAM to induce cancer-related inflammation and tumour progression.

Using a mouse model of multistage skin carcinogenesis, Andreu et al.[261] demonstrated that macrophages were skewed towards a M2 pro-tumoural phenotype after crosslinking of the FcγR with autoantibodies against the ECM. The production of autoantibodies was T-cell dependent.

Similarly, in a transplanted tumour model of melanoma, B1 cells induced M2 polarisation of TAM, likely due to IL-10 production. On the other hand, in B cell–deficient mice transplanted with melanoma, TAM polarisation occurred towards a classical phenotype, suggesting that the acquisition of pro-tumoural features by macrophages in this model is entirely dependent on B-cells.[260]

The efficacy of rituximab in T cell-mediated autoimmune diseases,[262] together with studies on T-cell function upon antigen stimulation in mice transiently depleted of B-cells demonstrated that B-cells are involved in both suppressing and enhancing T-cell mediated immunity independently of antibody production (reviewed by Lund and Randall[263]). The mechanisms by which effector B-cells can affect T-cell responses are multiple: by presenting antigen,[264] by expressing co-stimulatory molecules,[265] and by producing cytokines.[266] Aside from IL-10 production, T-cell primed B-cells can in fact produce other Th1 and Th2 cytokines,[267] including IFN-γ, IL-12, IL-4 and IL-13. Their impact on macrophage activation can be hypothesized, but has not been demonstrated.

One of the mechanisms of action of rituximab and other monoclonal antibody (mAb) is by engaging activatory FcRs and by boosting antibody-dependent cytotoxicity (ADCC) of immune effector cells such as macrophages.[268] The relevance of macrophage-mediated ADCC comes from two animal models whereby the activity of mAb's was halted by genetically abolishing the common FcRγ[269] chain or by ablating macrophages.[270] Recognizing the importance of this mechanism of action, Roche has developed Obinutuzumab (GA101), an anti-CD20 mAb with improved binding affinity to FcγRIII on effector cells and hence improved ADCC.

Some research has explored other mechanisms by which macrophages mediate the efficacy of rituximab against B-cell malignancies. Using a murine model of NHL, Cittera et al.[271] showed that rituximab induced the expression of the CCL3 in lymphoma cells and blocking the chemokine's activity abrogated rituximab activity in this model. The authors demonstrated that CCL3 activates cytotoxic cells towards tumour killing.

The same group proposed that the macrophage phenotype has an impact on the cell's cooperation with rituximab. Indeed the authors demonstrated that human M2 polarised macrophages had greater capacity for killing lymphoma cells opsonised with rituximab compared with M1 macrophages, which was further up-regulated by

IL-10.[272] The authors suggest that, although M2 macrophages enable tumourigenesis, their improved phagocytic might be beneficial if therapeutic mAbs are to be used.

As already mentioned, rituximab-refractoriness constitutes one of the most challenging clinical situations in DLBCL. The mechanisms of resistance are not entirely clear, but one such mechanism is the acquisition of expression of surface CD47 by cancer cells. CD47+ cancer cells evade macrophage-induced phagocytosis, even when coated with targeting antibodies. Weiskopf et al.[273] exploited this phenomenon to demonstrate that human NHL can be eradicated in mice by sole stimulation of the immune system, with macrophages playing a central role. Treatment with blocking anti-CD47 antibodies synergized with rituximab, allowing elimination of lymphoma in engrafted mice. Macrophage depletion abrogated the therapeutic effect. The authors demonstrated that this synergism involved FcR-independent enabling of phagocytosis by anti-CD47 antibody and FcR-dependent stimulation of phagocytosis by rituximab.

1.13 Correlative IHC studies of the DLBCL-associated macrophages

The demand for comprehensive studies approaching macrophage biology in DLBCL is supported by the GEP studies already discussed in section 1.5. In brief, prognostic models of transcriptomic analysis performed in whole DLBCL tumours incorporate genes recognized to be expressed by macrophages. Contrary to most cancers and cHL,[274] the expression of those genes identifies patients with better outcome. However, these interesting findings are merely correlative and a coherent explanation for the association between mRNA levels and DLBCL biology should follow. Protein expression is generally the starting point. With very few exceptions, protein validation studies of the microenvironment genes discovered are surprisingly rare in the literature. We suspect this is due to the lack of specificity of the transcripts for macrophages and to difficulties that are posed to IHC analysis of macrophages.

Translational studies evaluating macrophage biology in human NHL are mostly limited to the interrogation of FFPE tissue for individual proteins. Immunostaining for the lysosomal glycoprotein CD68 is taken as a surrogate of the extent of infiltration of macrophages in tissues.[275] However this protein can be detected in other cells, including dendritic cells,[276] rendering the staining unspecific and an overestimate of macrophage content. Moreover, as a cytoplasmic marker, CD68 is intrinsically difficult to analyse.

Contrary to most cancers and cHL,[219] the published literature does not provide a reproducible association between macrophage numbers and survival in patients with DLBCL. Importantly, some studies were performed in the pre-rituximab era[277-280] and are hence irrelevant for the current prognostic scenario of DLBCL.

Two studies merit further discussion, as they were performed with samples from R-CHOP treated patients and tried to explore the potential of using IHC to recognize macrophage subsets.

Secreted protein, acidic and rich in cysteine (SPARC),[281] a glycoprotein involved in matrix remodelling, adhesion, cytokine signalling and apoptosis, has emerged as a potential marker for individualizing macrophages subsets in DLBCL with potential prognostic impact. In order to validate some of the markers depicted in the "stromal-1" as representative of macrophage subsets, Lenz[34] and collaborators performed double immunofluorescence staining for CD68 and SPARC and connective-tissue growth factor (CTGF). While being expressed in other cells in the microenvironment, SPARC and CTGF were indeed confined to only a proportion of the CD68+ cells, substantiating macrophage phenotypic heterogeneity that certainly could not be unravelled by morphology or immunostaining with a prototypic marker. However, the additional prognostic examination for SPARC expression is misleading. Firstly, the analysis was performed in a CHOP treated cohort. Moreover SPARC density is based the protein's single staining, scored in a fairly crude way: four cohorts of 0 to 4 according to density and then dichotomized in high versus low expression. This

methodology does not take advantage of the elegant findings of the double staining and the simplified analysis is hardly replicable in validation studies.

Following this report, Meyer et al.[282] revisited the prognostic value of CD68 and SPARC in a 262 R-CHOP treated patient dataset. CD68 staining was not predictive of any survival measure. The presence of any SPARC positive cells in the microenvironment correlated with longer OS, whereas those in a "high" SPARC category had longer event free survival (EFS) compared to the remaining. This positive impact of SPARC was restricted to ABC-DLBCLs as defined by GEP but retained its value on multivariate analysis together with the IPI and the COO classification.

The shortcomings of these data arise again from the methodology: the number of positive stromal cells was estimated visually as a percentage of all cells present within the tumour area, and graded in 5 or 10% increments; cases were then divided in "negative", "low" and "high" categories. Moreover SPARC staining on endothelial cells was recognized but ignored for any further analysis. We believe this strategy leads to data reduction, turns validation problematic and encapsulates the limitations of IHC analysis when automated systems are not applied.

The Osaka Lymphoma Study Group[283] has also studied a small cohort of 101 R-CHOP cases and reported a direct correlation of macrophage density and worse OS. Additionally the authors explored whether fairly accepted M1 (human leukocyte antigen (HLA)-DR) and M2 (CD163) markers could help establishing a relationship between outcome and macrophage subsets. The authors demonstrated that a large number of CD163+/CD68+ correlated with a worse OS, whereas the presence of M1 macrophages portrayed no relevant impact. As already discussed, the validation of in vitro defined markers of macrophage polarisation has limited value when the disease idiosyncrasies are ignored; and single markers are rarely faithful of a whole population of polarised macrophages. All in all, the Japanese data adds little value to our understanding of macrophage biology in DLBCL.

1.14 Functional studies exploring the role of macrophages in lymphoma

Due to the fundamental differences between murine and human macrophages (discussed in section 1.10.5), research hypotheses on the interactions between macrophages and B-cell lymphoma should be generated using human samples. The GEP studies on DLBCL are an excellent example of hypotheses generating subsequent macrophage research. In vitro studies using primary lymphoma cells are feasible[221,284] and should follow, but it should be acknowledged that a potential allogeneic reaction and the lack of other immune cells in the systems could deviate the results from the tumour context in situ. Murine models should finally be used to confirm the hypotheses, as they provide, with all the limitations, the closest approximation of an in vivo tumour microenvironment. The challenges posed to macrophage research have also been discussed previously.

We shall here debate data suggesting a heterogeneous and less accepted role for macrophages in lymphomagenesis. Their role in the context of treatment with rituximab has already been discussed. Further studies involving co-culture systems of macrophages and primary lymphoma samples will be approached in experimental chapter 6.

Murine macrophages have been shown to exert phagocytosis of lymphoma cells in the early 70's. Comprehensive work by Evans[285,286] and others appreciated that this was a highly regulated process, requiring for maximum efficacy preliminary priming of the effector cells and the presence of T-cells and soluble factors. Intuitively tumour cells could overcome immune surveillance by influencing any of these factors. Indeed, it was simultaneously reported that macrophages from lymphoma-bearing mice could produce the immunossupressive enzyme PGE2 to halt T-cell function and favour tumour growth.[287]

A murine model recurrently used as a surrogate of human DLBCL is the Eμ-myc transgenic model, with constitutive expression of the MYC oncogene in lymphocytes and inexorable progression to a highly proliferative lymphoma. As a transplantable

lymphoma, it constitutes a useful model to study the microenvironment. In this model macrophages are highly engaged in the clearance of lymphoma apoptotic cells, which has both protumoural and antitumoural consequences.[288,289] We believe that the mechanisms described below highlight the complexity of potential roles for macrophages in mediating lymphomagenesis and would apply to all DLBCL cases with high proliferative index.

Cell-free supernatant of lymphoma dying cells induces macrophage chemotaxis. Lymphocyte-derived CX3CL1[290] appears to be the culprit macrophage attractant. Once in the tumour microenvironment macrophages actively phagocyte apoptotic debris and amplify macrophage recruitment by producing CX3CR4. On the other hand, macrophages produce cytokines, including TGF-β and IL-10 that are implicated in shifting other macrophages towards an immunossupressive and pro-proliferative lymphoma permissive phenotype.[291] On the other hand, Reimann and colleagues[292] have shown that macrophage-derived TGF-β arising in the context of active apoptosis can limit MYC-driven lymphomagenesis by feedback induction of terminal cell-cycle arrest. By genetic inactivation of senescence or neutralization of TGF-β the authors observed an acceleration lymphomagenesis.

Perhaps the most intriguing data comes from Haabeth et al.,[293] who proposed a mechanism by which inflammation, if developed in the context of an efficient T-cell response against the tumour cells, protects against B-cell lymphomas. To clarify the process of CD4+ T-cell-mediated immune surveillance against B-cell lymphoma, the authors used an idiotype (Id)-specific TCR transgenic mouse model, which was made homozygous for the severe combined immunodeficiency (SCID) mutation to prevent rearrangement of endogenous TCR chains. In these mice, CD4+ T cells exclusively recognize an Id peptide from the IgV chain of a plasmacytoma cell line, rendering them resistant against inoculation with syngeneic plasmacytoma cell lines or with Id-transfected B-cell lymphoma cell lines.

The authors developed a strategy whereby tumour cells are embedded in a collagen gel prior to injection in mice. The collagen can then be studied ex vivo for immune cell infiltration and cytokine production. Using this method, the authors uncovered that, in the presence of Id-specific CD4+ T-cells, macrophages infiltrate the gel and capture the lymphoma-specific antigen. On recognition of tumour peptides presented by macrophages, T-cells were shown to secrete IFN-γ. This cytokine in turn activated macrophages to efficiently kill lymphoma cells. When the host was capable of initiating an antigen specific Th1 response, tumouricidal macrophages produced the proinflammatory cytokines IL-1β and IL-6.

Macrophage-induced inflammation accepted by experts as tumour permissive and here dissonantly associated with tumour surveillance, highlights in our opinion one main aspect that encapsulates the complexity of designing immunology studies: the functional spectrum of one immune cell cannot be fully understood in the absence of the other cells of the immune system. This paradigm renders in vitro studies quite simple and translationally meaningless and explains why a body of data developed in murine models does not move forward to the human setting.

1.15 Objectives

DLBCL is an aggressive lymphoma curable with R-CHOP in a sizable proportion of patients. The remaining patients have a dismal outcome and ideally should be selected at diagnosis for experimental therapies. Devising robust prognostic markers is hence a priority for this disease.

This study was set out to review prognostic biomarkers derived from the GEP studies performed in DLBCL. Using samples from patients with DLBCL, we aimed to:

- Construct tissue microarray (TMAs) of patients with known clinical outcome and follow up;

- Explore the applicability of IHC-based algorithm for molecular stratification of DLBCL;

- Reconsider the role of IHC-based studies for the immune microenvironment by applying novel strategies of analysis;

- Explore the macrophage heterogeneity in DLBCL based upon transcriptomic studies of highly pure cell populations;

- Confirm the existence of macrophage subsets in DLBCL by using novel proteins derived from the transcriptomic analysis;

- Develop in vitro co-culture systems to test whether the malignant B-cells are directly influencing macrophage polarisation in the microenvironment.

Chapter 2 Materials and Methods

2.1 Patient Samples

Patient samples were obtained after informed consent and stored under conditions compliant with the Human Tissue Act 2008. All sample collection was done under current regulatory permission from the local research ethics committee and according to the Declaration of Helsinki.

2.1.1 Formalin-Fixed Paraffin-Embedded Tissue

Immunohistochemistry studies were performed using FFPE. Samples were selected based on the availability of stored good quality paraffin blocks of the diagnostic biopsy and clinical and extended clinical and follow-up data. Only cases with de novo DLBCL were included. Patients with an immunodeficiency-associated lymphoma, central nervous system or primary mediastinal lymphomas were excluded from the study. December 2009 was selected as the cut-off date of diagnosis to include patients in the study to guarantee an acceptable follow-up.

Two patient cohorts were used for this study. The first cohort was diagnosed and treated at St. Bartholomew's Hospital (Bart's), London. Patients were identified by consultation of the clinical database, created in the 60's and used since then to document clinical details of all patients presenting initially to, or referred from local hospitals to Bart's for treatment. Subsequently, clinical information was compared with the tissue database for final confirmation of FFPE material available for research. Between 1977 and end of 2009, 651 patients were diagnosed with de novo DLBCL at Bart's. Finally, 225 patients with FFPE material stored amenable to be arrayed were identified. From these, 71 were treated in the rituximab era with R-CHOP. The remaining patients were treated with different approaches, from anthracycline-based therapy to palliative care.

The second cohort was diagnosed and treated at the Instituto Português de Oncologia Francisco Gentil (IPO), Lisbon. Patients were selected by Dr Maria Gomes da Silva, the lead haematologist responsible for clinical review, and Dr José Cabeçadas (JC), the lead haematopathologist responsible for research material review. All Portuguese patients were treated with R-CHOP.

Tissue sections were made and Haematoxylin and Eosin (H&E) staining was used to carry out tumour area selection for TMA.

2.1.2 Frozen Single Cell Suspensions

Single cell suspensions (SCSs) from DLBCL, tonsil and reactive LNs were obtained from the centre for Haemato-Oncology tissue bank storage. Surplus material from samples for diagnostic or treatment purposes was collected under sterile conditions in media. Samples were dissected on a pre-cooled tray (Biocision) and homogenised by passage through a 70µm filter (BD biosciences) under gravity. Cell suspensions were washed twice in media by centrifugation at 1500rpm for 5min. After assessing cell count and viability in an automated cell counter by trypan blue exclusion (ViCell, Beckman Coulter), samples were cryopreserved in 10% Di-Methyl Sulphoxide (DMSO, Sigma-Aldrich) in foetal calph serum (FCS, PAA Laboratories Ltd) for two hours at -80^{o}C in adapted containers (CoolCell, BioCision) and transferred to liquid nitrogen tanks (Custom Biogenic Systems (CBS) Isothermal V-1500 series) for long-term storage.

All studies described here were performed on previously cryopreserved SCSs with the purpose of normalising for storage effects. SCSs were thawed in a water bath at 37^{o}C, treated with 0.5mg/ml of DNAse (DNAse I from bovine pancreas, Sigma-Aldrich) for 5min at room temperature (RT), washed by centrifugation at 1400rpm for 5min in phosphate buffered saline (PBS, Sigma-Aldrich) with 100 U/ml penicillin and 100µg/ml streptomycin (both from Invitrogen) and filtered through a 70µm mesh to exclude cell clumps. Viability and cell count was assessed using either an automated or a manual cell counter.

2.1.3 Healthy Donor Buffy Cones

Peripheral blood mononuclear cells (PBMCs) from healthy donors were isolated from buffy cones supplied by the National Blood Service. After a 5-fold dilution in sterile PBS, 10ml of this solution was gently layered with an automated pipette (Pipetboy, Integra) in a 30^{o} angle over 5ml of Ficoll-Paque (Lymphoprep, Axis Shield) in a 15ml conical centrifugation tube (Corning). Samples were centrifuged at 1500rpm for 25min at 22^{o}C with slow acceleration and brake off. The PBMC layer was carefully removed using a pastette, diluted to a maximum of 40ml in PBS and centrifuged at 1200rpm for 10min at 4^{o}C for platelet removal. A second wash by centrifugation at 1800rpm for 10 min was performed after which cell pellets were resuspended in 10ml of diluted red blood cell lysis buffer (Pharm Lyse, Beckton Dickson) for 7min at RT. Finally, cells were washed and resuspended in PBS prior to cell counting.

2.1.4 Cell Lines

The DLBCL cell lines used in this study were maintained at 37°C in a 5% CO2 humidified incubator in sterile flasks in Roswell Park Memorial Institute (RPMI)-1640 (Sigma-Aldrich) or Iscove's Modified Dulbecco's Medium (IMDM, Sigma-Aldrich) as recommended. The GCB cell line Su-DHL4 was gifted from Dr. A Letai and the ABC cell line Ri1 was gifted from Dr. M Capasso. All culture medium was supplemented with 10% heat inactivated FCS, 100U/ml penicillin and 100µg/ml streptomycin. Cells were passaged every 2-3 days to maintain optimal cell concentration and viability.

2.2 Immunohistochemistry

2.2.1 Tissue Microarray Assembly

TMAs consist of tissue cylinders, extracted from FFPE tissue by an arraying machine that are aligned and embedded in a new paraffin block.[294] About 200 sections 2-5µm thick can be cut from each block and stained for large-scale protein expression profiles, allowing the study of large patient cohorts under controlled experimental

conditions. It has been previously shown that this technology can reproduce lymphoma tissue heterogeneity with the same accuracy as conventional sections when duplicate or triplicate cylinders are applied.[295,296]

After patient selection, each original FFPE block was cut and stained with H&E. Well represented tumour areas free of necrosis and fibrosis were marked with a pen on the slide. Donor blocks were then aligned with the marked slides.

A semi-automated arrayer (Beecher Scientific) was used as previously described.[294] Each patient block was cored in duplicate (1.5mm^2 diameter) or triplicate (1mm^2 diameter) and these inserted in a new recipient paraffin block. Reactive tonsils were used as internal staining controls and myocardium or pancreas as orientation. The British patient set was separated into eight blocks and the Portuguese into two blocks. Before sectioning each block was placed over an ice block for 10-15min. A standard microtome technique was used for sectioning 3µm sections into slides.

2.2.2 Principles of Immunohistochemistry

IHC uses direct labelling of a target protein antigen with an antibody or, alternatively, with a more sensitive secondary antibody labelling system. The detection of the antigen-antibody interaction requires an amplification step that is achieved by the use of multimeric molecules that are able to link to multiple proteins. The avidin-biotin complex (ABC) and the polymer-based systems (e.g. Biogenix Supersensitive Polymer horseradish peroxidase, HRP) are the most widely used amplification methods. The first relies on the use of biotinilated secondary antibodies to which multiple avidin-enzyme complexes bind strongly, providing antigen signal amplification. The second utilizes a unique technology based on a polymer sugar backbone to which multiple antibodies and enzyme molecules are conjugated. A Catalyzed Signal Amplification method is being incorporated in this polymer-based technique. Final visualisation of an antibody-antigen interaction is achieved by the use of enzyme modifiable chromogens.[297] The chromogen most commonly used is Diaminobenzidine (DAB) due to its crisp brown staining.

2.2.3. Primary antibodies for Immunohistochemistry

IHC staining was carried out using commercial available primary antibodies (clone, dilution and source detailed in each experimental chapter in Table 3.2, Table 4.1 and Table 7.1). Primary antibodies against GCET1 (clone Ram341, Novocastra), FOXP1 (clone JC12, Abcam), and LMO2 (clone SP51, Abcam) were first used in our laboratory for this study and required optimisation. Appropriate controls for titration and antigen retrieval were provided by the manufacturer. Serial 1/50, 1/100 and 1/200 antibody dilutions were used and three antigen retrieval techniques tried: no pre-treatment, pronase enzyme digestion for 15min and pressure-cooking. The primary antibody was diluted with 1% bovine serum albumin (BSA, Sigma A-70906) and sodium azide (Sigma – S8032). After first assessment, further dilution titrations were performed.

2.2.4 Staining Protocol using the Dako Autostainer System

TMA slides were placed in plastic racks at 60^{o}C overnight. For paraffin removal slides were incubated in xylene (VWR Chemicals – 28975.325) for 2 consecutive periods of 5min. Subsequently, slides were incubated for 3 consecutive periods of 2min in Industrial Methylated Spirits (IMS, Fisher Chemical – 11482874) and further 2 periods of 2min in hydrogen peroxide (BDH – 101284N) in order to dehydrate tissue and reduce non-specific staining from the action of endogenous peroxidises on the chromogen. A final incubation of 2min in IMS is required prior to antigen retrieval. While performing first incubation steps, 3000ml of a working solution of antigen unmasking solution (Citrate buffer, pH 6, Vector Laboratories – CA94010) was warmed up in a pressure cooker. When boiling, the plastic racks with slides were immersed and left for 10min at high heat (120-130^{o}C). When finished, the pressure cooker was left to cool down under cold tap water for 5min and the slides quickly transferred to washing buffer (DAKO – S3006).

Slides were marked using hydrophobic pen around the edge of the array field and kept wet with wash buffer throughout remaining procedure.

The DAKO Autostainer System is composed of a robot arm with nozzles and a pump system, which allows for timed dispensing of reagents into the slides. Before using,

the software (Dako Autostainer Plus) is programmed for the number of slides, reagents and incubation times and rinse steps. The Super Sensitive™ Polymer-HRP IHC Detection System (Biogenix - QD430-XAKE) was used for signal detection. The Autostainer was run for 2-3 hours as specified. After finishing all slides were replaced in plastic racks and rinsed in tap water for 5min. As a counterstain, the slides were suspended in haematoxylin solution for 5min, rinsed for 2min in running water and plunged into acid alcohol solution (1% hydrochloric acid in 70% IMS), quickly, for 5 times, after which were transferred into tap water wash. Finally the tissue is re-hydrated using IMS and clarified by incubation in xylene baths. Using DPX mounting media (VWR – 360294H), which provides a high quality durable mounting with refractive properties, cover slips were applied leaving no trapped air bubbles. Finally, slides were left to dry and labelled appropriately.

2.2.5. Immunohistochemistry analysis

2.2.5.1 Automated Image analysis using the Ariol SL-50 visual analysis software

Slides were scanned using an Olympus BX61 microscope with an automated platform (Prior). A review of all cores was performed manually. Whole cores with less than 50% of tumour representation were excluded from analysis. Fibrotic and necrotic areas were also excluded. Representative regions were selected for training. Positive stained cells acquire a brown/black colour characteristic of DAB. Colour hue, saturation and intensity were manipulated to allow contrast with the background. This was achieved by selection of individual pixels from positive events and avoiding the negative cells and the non-specific stained areas. Training is improved by limiting the size and the shape of the areas considered positively stained with DAB. In order to estimate areas of representative tumour tissue, a second colour training for non-cellular areas was included (Figure 2.1). The area of viable tissue was calculated by subtracting the non-cellular areas from the total area analysed. Using this system we calculated the number of positive cells as well as the area of DAB stained per area of lymphoma tissue. The values obtained were corrected to a $1mm^2$ area and a mean for each patient was calculated.

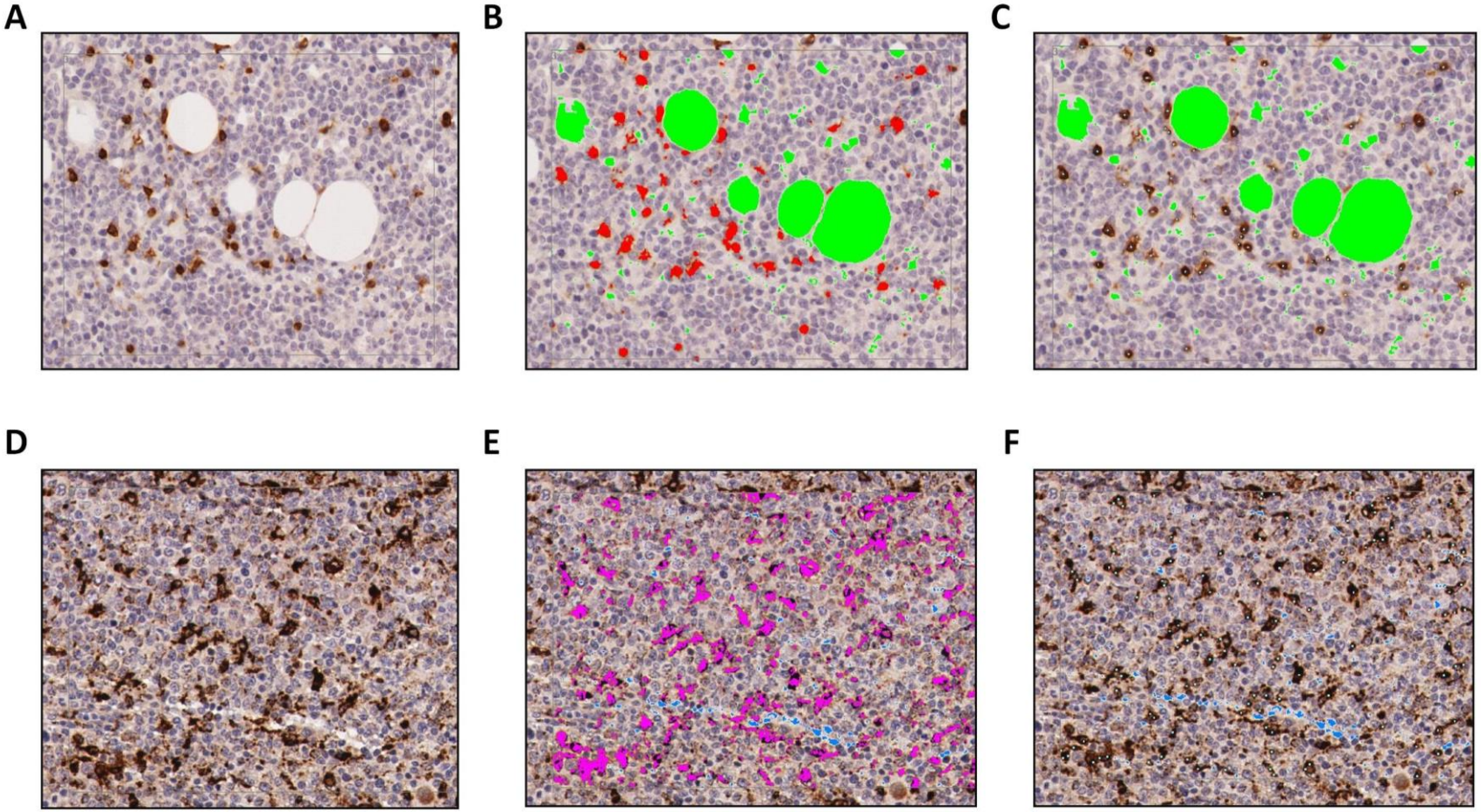

Figure 2.1 Training Optimization using the Ariol System

A&D: Representative example of training protocol, showing the DLBCL areas selected for training for CD3 (A) and CD68 (D). Magnification x20. B&C: Optimal results for % stained area of CD3 (B), where DAB staining is represented in red and non-cellular areas are represented in green; and for CD3+ cells (C), where individual cells are represented by white dots. E&F: Optimal results for % stained area of CD68 (E), where DAB staining is represented in pink and non-cellular areas are represented in blue; and for CD68+ cells (F), where individual cells are represented by white dots.

2.2.5.2 Automated Image analysis using the Pannoramic Viewer System

Slides were scanned using the Pannoramic 250 Flash II scanner (3DHISTECH). Each core was observed on a computer screen using the Pannoramic Viewer computer interface. Meticulous marking of representative tumour areas was done and areas quantified. After selection of representative tumour areas, the DensitoQuant module was used to quantify the number of DAB stained pixels. This module distributes pixels to 3 grades of positive classes using their RGB values. We used only the top red and orange levels for identification of stained areas. After adjusting the brown tolerance and the score levels an optimal script was saved for each antibody and applied for analysis in all areas. Finally, the number of brown pixels/selected area was calculated and a mean value was estimated for each patient.

2.2.6 Cutpoint Determination and Survival Analysis

Cut-point discrimination was assessed using a recursive partitioning algorithm in the rpart package (http://cran.r-project.org/web/packages/rpart/index.html) within the R statistical software version 3.0.2.

Although data validation in independent datasets is a more robust method for confirming a certain IHC marker as a prognostic biomarker, this is still seldom done.

The recursive splitting method with cross-validation is a statistical approach well accepted to develop outcome prediction models from novel variables. This algorithm involves finding a cutpoint in the variable under study that best divides the dataset in two groups with different outcome. The data is separated accordingly and novel splits are applied to each sub-group, in a recursive manner, until no improvement can be made in outcome prediction. This method requires that some limitations to the stepwise procedure are established in order to avoid developing too complex or suboptimal models for the problem under study. A “bucket size” approach that excludes data splitting where only a minority of patients are included in a given group is commonly used to increase robustness of results. The second stage of this method consists of using cross-validation to estimate the performance of cutpoints

considered predictive of survival.

For every quantified measurement, survival has been estimated using Kaplan-Meier estimators, and differences between groups of the same measurement have been assessed with the Log-rank test. To accommodate for the optimization method within the splitting algorithm we considered as significant only Log-rank p-values <0.01.

2.3 Immunofluorescence Staining and analysis

Three colour Immunofluorescence (IF) was used in this study to detect co-expression of two proteins of interest in macrophages in FFPE tissue. Nuclei counterstaining is performed by incubation with a third colour, 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI, Sigma – D9564).

This technique requires the use of primary antibodies against a target antigen epitope, ideally raised in different species; and secondary antibodies raised against Igs of the primary antibody species. The last are conjugated to a fluorochrome that allows visualisation of the proteins of interest. The secondary fluorescent antibodies must emit colour at different and non-overlapping wavelengths. The most commonly used wavelengths for this purpose are 488 (excitation: 495/emission: 519), which produces a green fluorescence and 546 (excitation: 556/emission: 573) which produces a red fluorescence. Optimization for these experiments involved changing the order of the primary antibodies incubation as well as the fluorochrome combinations. If both primary antibodies were raised in mouse species, an intermediate incubation step with mouse-on-mouse block reagent (Vector Labs) was required. Primary antibodies and conditions of use and secondary reagents are described in Chapter 6.

Slide deparaffinisation and antigen retrieval steps were performed as described in section 2.2.4. After placing the slides in the DAKO Autostainer the software was programmed. A template of the staining procedure is provided below:

1. Primary antibody (mouse) (40’ incubation)
2. Rinse with buffer
3. Secondary reagent (anti-mouse Texas red secondary antibody) (10’ incubation)
4. Rinse with buffer
5. Mouse-on-mouse blocking step
6. Primary antibody (mouse/rabbit) (40’ incubation)
7. Rinse with buffer
8. Secondary reagent (anti-rabbit FITC secondary antibody) (5’ incubation)
9. Rinse with buffer

Slides were then transferred to plastic racks and immersed in buffer solution. Immersion for 2min in 3 pots of graded ethanol (70/80/96%) was used for slide rehydration. Finally slides were mounted using 5 μL of Vectashield mounting medium (Vector Labs) with DAPI fluorochrome and kept at -20^{o}C.

Slides were scanned using an Olympus BX61 microscope and analysed using the Ariol SL-50 visual analysis software. The fluorescence wizard-training module was used to optimize capturing conditions. The DAPI channel capturing nuclei staining was used to focus at 5x amplification. All channels were subsequently observed at 40x amplification and adjustments on background capturing and exposure times were done. Selected cores were finally scanned and observed later on a computer screen. Single channel and overlapping images were used for final analysis quantifying single and double positive cell counts/core.

2.4 Multicolour flow cytometry

Flow cytometry is a single cell analysis technique, which enables the recognition and quantification of expression of multiple molecules simultaneously. Cells are stained with fluorochrome-conjugated monoclonal antibodies against the molecules of interest; moved through a fluidics system and forced to pass one by one through an integrated laser bean which results in fluorochrome excitation and light emission. The light side and forward scattering properties provides information on the size and intracellular organelle complexity of each cell, whereas the fluorescence emission and intensity delivers qualitative and quantitative information on molecule previously labelled. Besides non-specific fluorescence and background noise, the emission spectral overlap of different fluorochromes is frequently involved in interference with the final read-out and can be overcome by careful combination of fluorochromes with minimal overlap, adjustment of voltages of the light detectors and compensation of the detected data. The information acquired is studied using specific software packages. Data can be displayed as one-dimensional histograms showing detected fluorescence distribution and intensity, or multidimensional histograms combining different parameter intensities, each cell being represented by a dot.

This technology was used for single cell sorting (Chapters 5 and 6), carried out using a BD FACSAria II sorter, and macrophage surface marker staining (Chapter 6), carried out using a four laser LSR Fortessa (both from BD Biosciences).

2.4.1 Single cell sorting

Flow sorting is a technique that allows separating different cell populations from heterogeneous samples, based on the physical (such as the size) or chemical (such as expression of cell specific surface antigens) properties of the cells. The FACSAria sorter uses the electrostatic method in which the cells are ejected through a vibrating nozzle and broken up into a stream of regular droplets. A charge is applied to the droplets that contain particles of interest, deviating them from the main stream into plates at high voltage. Finally, cells are collected under sterile conditions as they move to the waste stream and can then be used for a variety of studies. The

efficiency of the sorting is tested by different parameters. Decisions regarding priority for recovery or purity of the population are important when programming the sorting experiment.

This technique is particularly useful for isolation of rare cell populations with high purity, such as macrophages in previously stored LN SCSs from patients with lymphoma or reactive conditions. When sorting a rare population that needs to be highly pure, a second sorting procedure might be needed at the expense of loss of cell yield.

Primary LN SCSs (2-3 vials/case) were processed according to section 2.1.2. Cell count and viability were confirmed using an automated cell counter by trypan blue exclusion (ViCell, Beckman Coulter). Cells were split into $1x10^7$ aliquots, resuspended in PBS with penicillin, streptomycin and 2% FCS (wash buffer) and centrifuged at 1400 rpm for 5min at RT. Smaller aliquots ($0.25x10^6$ cells/tube) were used for optimization controls. Excess supernatant was discarded and cell pellets incubated at 4^oC for 20min in 2% human anti-γ-globulins (Sigma) with the purpose of reducing non-specific antibody binding to Ig FcR. Cells were then incubated with the appropriate antibodies (Table 2.1) for 30min in the dark, at 4^oC, washed and resuspended in wash buffer with DAPI (1µL/2ml), the fluorochrome used for dead cell exclusion. Polypropylene tubes with 1ml of 50% filtered FCS in wash buffer were used for cell collection.

An example of the sorting strategy is illustrated in Figure 2.2. Firstly, DAPI negative, live cells were selected, followed by doublet exclusion and positive selection of total leukocytes according to CD45 expression. Total T-cells were separated according to the expression of CD3. B-cells were sorted based on the expression of the pan-B-cell marker CD20. Back gating confirmed these were viable and negative for CD3. Finally macrophages were isolated based on the expression of CD36, a membrane scavenger receptor. As CD36 is also expressed on endothelial cells, CD45 positivity was confirmed by back gating of this population prior to sorting. Purity was assessed by centrifuging and re-suspending cells in wash buffer/DAPI prior to re-analysis. If cell

purity was <90%, re-sorting was performed. Finally, B-cells were stored in liquid nitrogen in FCS/20% DMSO after centrifugation at 3000 rpm for 7min. T-cells and macrophages were pellet by centrifugation at 8000 rpm for 10min and subjected to RNA extraction immediately.

Table 2.1 Staining strategy for cell sorting

	CD45 PE	**CD3 FITC**	**CD36 APC**	**CD20 APC-H7**
Unstained sample				
CD45 PE	5 µL			
CD3 APC		5 µL		
CD20 APC-H7				5 µL
CD36 FITC			5 µL	
FMO APC	5 µL	5 µL		5 µL
FMO PE		5 µL	5 µL	5 µL
FMO FITC	5 µL		5 µL	5 µL
FMO APC-H7	5 µL	5 µL	5 µL	
Sample (/10^7 cells)	15 µL	15 µL	30 µL	30 µL

SS: single staining; FMO: fluorescence-minus-one

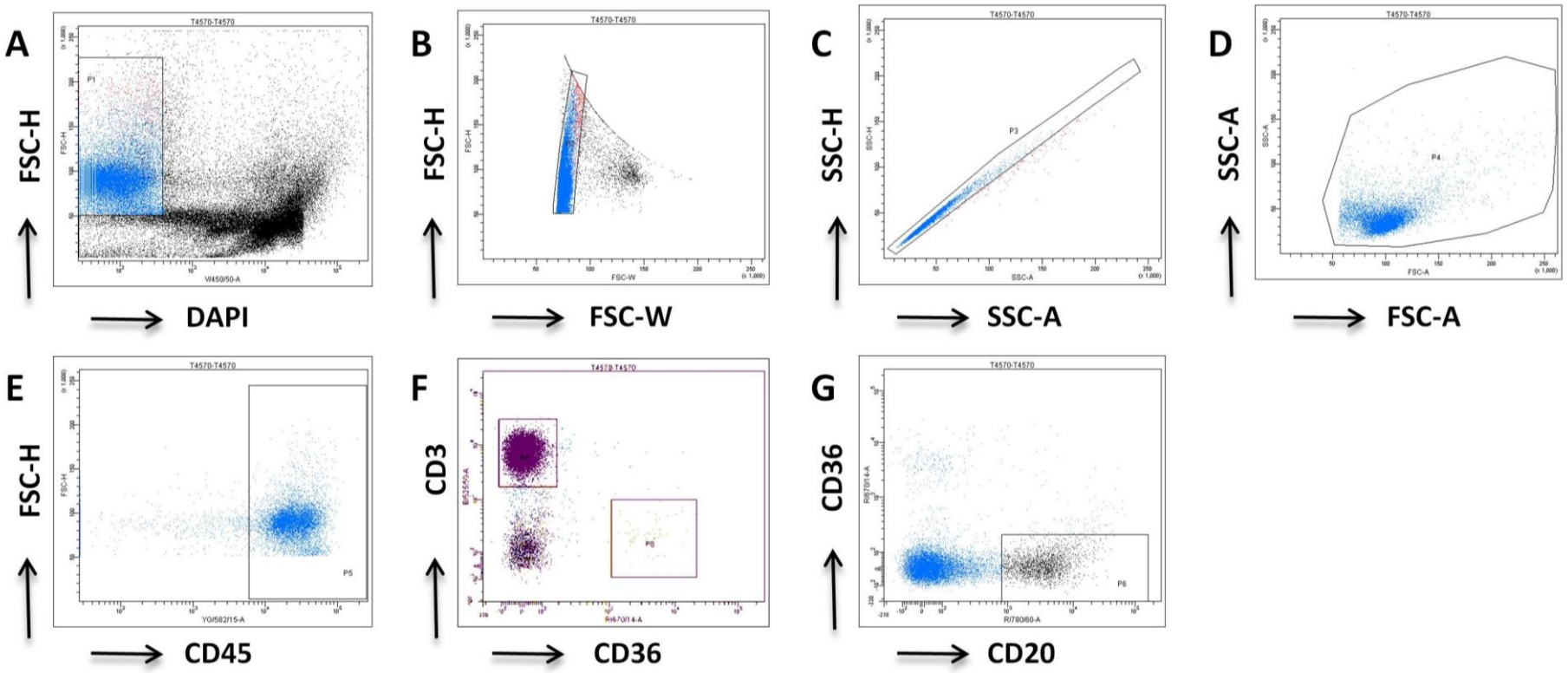

Figure 2.2 Sorting strategy for isolation of macrophages, B-cells and T-cells from single cell suspensions of DLBCL and reactive LNs

Viable cells were selected based on the lack of DAPI staining (A); doublets were excluded by two consecutive gatings (B,C); Leukocytes were selected based on the expression of CD45 (E); T-cells and macrophages were sorted based on the expression of CD3 and CD36, respectively (F); B-cells were isolated according to the expression of CD20 (G).

2.4.2 Immunophenotyping of macrophage cell surface markers

These experiments were carried out in V-bottomed 96 well plates and staining was performed after antibody cocktail preparation. This increased efficiency of protocols and reduced preparation times and costs.

Staining cocktails including all antibodies were prepared, with adequate volumes for control and experimental conditions (+10% for pipetting errors). The cocktails were pipetted into a 96 well plate prior to multichannel staining of the samples.

Unstained cells were used to measure intrinsic autofluorescence in each wavelength prior to acquiring stained cells. Fluoresce-minus-one (FMO) tubes allowed discriminating non-specific fluorescence and apparent fluorescence resulting from spectral overlap and setting the gating for positivity for each marker/channel.

Macrophages were harvested from culture plates (see section 2.8.6), re-platted into 96 well plates and centrifuged at 3000 rpm for 5min. Supernatant was discarded by turning plates upside down and cells were resuspended in 50µl of wash buffer. Staining cocktails (antibodies described in Chapter 6) were applied in a volume of 50µl/well, mixed by partial volume pipetting and left to incubate for 30min on ice in the dark. Wash buffer was then added (200µl/well) and the plate centrifuged at 3000rpm for 5min. Samples were transferred to 1ml labelled polypropylene adaptor tubes in 100µl of wash buffer.

Data was acquired using the four laser BD Fortessa flow cytometer. A minimum of 5000-10000 target events gated on compensated viable-singlet cells were acquired. The main read-out was the median fluorescence intensity (MFI), which was measured in all experimental conditions and retrieved as the difference compared to the un-manipulated sample in each individual experiment. FlowJo (Tree Star Inc) software was used for analysis.

2.5 RNA extraction

QIAGEN provides well-standardized protocols that were followed in these experiments with minimal changes after optimization.

RNA was extracted from sorted macrophages and B cells using the QIAGEN RNeasy® micro kit recommended for samples with a low cell number. The QIAGEN RNeasy® mini kit was used for RNA extraction in samples with higher cell yield. The protocol here described was used for low cell number samples.

Cell pellets were mixed twice by pipetting with 75+75µl of denaturing buffer (RLT+β-mercaptoethanol), transferred to a QIAshredder spin column and centrifuged at full speed (~14000rpm) for 2min. The homogenised lysate was mixed with 150µL of 70% ethanol, transferred to an RNeasyMinElute spin column and centrifuged for 15sec at 10000 rpm. This facilitates binding of the RNA to the column. The Buffer RW1 (350µl) was added to the column after discarding the flowthrough and a centrifugation step of 15sec at 13000 rpm was done. RW1 was re-added and an on-column DNase digestion step was also performed with the purpose of eliminating genomic DNA. 350µl of buffer RW1 was then added to the membrane and centrifuged for 15sec at 13000rpm. The membrane was washed 2× by incubation with 500µl buffer RPE followed by 500µl of 80% ethanol with intermediate centrifugations at high speed. The column was dried by centrifugation at 13300 for 5min with the lid open. Finally, the RNA was eluted by incubation of the membrane with 14µl RNase-free water for 3-5min at RT followed by centrifugation at 13300rpm for 1min; and stored at -80^{o}C.

2.6 RNA Quantity and Quality Assessment

2.6.1 NanoDrop Spectrophotometer

RNA quantity and purity was determined using the NanoDrop ND-1000 spectrophotometer (Nanodrop technologies). Concentration was calculated by determining the optical density (OD) or ultraviolet light absorbed at 260nm of 1µl of RNA, knowing that 1 OD unit is equal to 50ng/µl of RNA. The ratio of absorbance at 260/280nm indicates purity and should be ~2 for good quality. Lower ratios suggest

contamination with protein or other impurities, which absorb strongly at 280nm. The 260/230nm ratio is often taken as a secondary measurement of purity and should be in the range of 1.8 - 2.2, although cut-offs are not well standardized. A lower ratio may suggest contamination by compounds such as guanidine thiocyanate,[298] which are known to have limited effect on downstream analysis.

2.6.2 Agilent Bioanalyzer

The Agilent 2100 Bioanalyzer system enables researchers to determine the quality and integrity of an RNA sample.[299] The Agilent RNA 6000 Pico Kit is particularly useful to assess low concentrated RNA samples (<50pg/µL), for which NanoDrop readings are inaccurate. Results are based on the electrophoretic trace of the sample (Figure 2.3). The RNA Integrity Number (RIN) generated is calculated based on the identification of degraded RNA products and the height of the 18S peak.

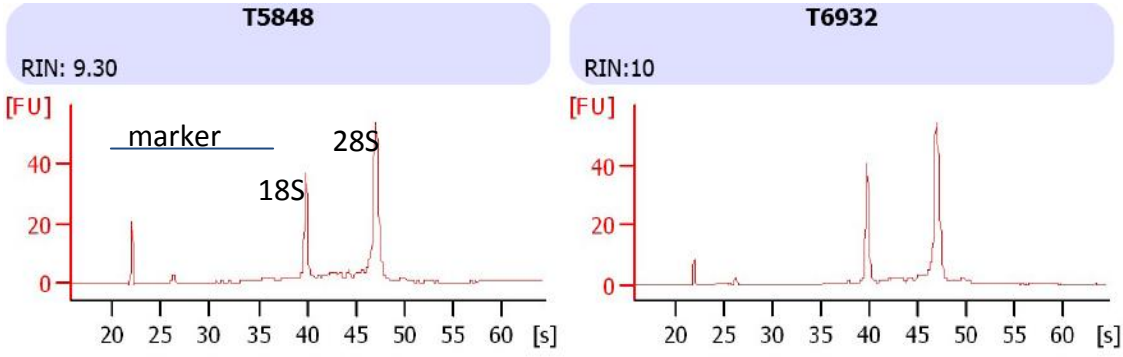

Figure 2.3 Representative Electropherograms generated by the Agilent 2100 Bioanalyser

The samples illustrated have excellent quality, the electropherogram traces showing a single marker peak and two peaks of ribosomal RNA (18S subunit, 28S subunit) migrating in the correct timing. RINs near 1 indicate poor quality, while values near 10 suggest intact RNA.

The kit reagents were equilibrated to RT prior to use. RNA concentrations were estimated using the NanoDrop and samples with higher RNA concentrations were diluted with RNAse free water. The RNA gel matrix (550µl) was transferred to a spin filter and centrifuged at 4000 rpm for 10min. A gel-dye mix (65µl of gel + 1µl of dye) was prepared and centrifuged for 10min at 11800 rpm. The gel-dye mix was loaded to RNA Pico chips in 9µl aliquots pipetted onto the wells marked 'G', followed by the conditioning solution. Subsequently, 5µl of RNA marker was added to each well including the ladder well. Sample and control RNA aliquots and the RNA 6000 Pico ladder were heat denatured and added to each well (1µl). After vortexing the chip was run on the Agilent 2100 Bioanalyzer using the Eukaryote RNA Pico chip assay.

2.7 Gene Expression Analysis

2.7.1 Gene Expression Profiling by Microarray technology

Microarray technology, developed in the mid-1990's, allows examining the expression of thousands of genes simultaneously. It exploits the ability of complementary strands of nucleic acids to base pair with each other and bind. The DNA samples under study are fluorescently labelled and put in contact with DNA copies corresponding to different mRNAs known to codify for a particular gene. If hybridisation occurs, fluorescence can be identified, indicating that a particular gene is being transcribed in that sample. Differences in fluorescence, and hence gene expression can be measured between "disease" and "control" samples.

This technology was used in our studies to compare the transcriptome of macrophages selected from DLBCL lymph nodes with macrophages extracted from reactive lymph nodes. The clinical features of the respective patients are described in the experimental chapter 5.

2.7.1.1 cDNA Synthesis and Amplification using the Nugen Ovation® Pico WTA System V2

RNA is unstable and readily degraded by ubiquitous RNases, making it unsuitable for direct analysis. Complementary DNA (cDNA), a stable template for RNA-based assays can be generated from RNA. The Ovation® Pico WTA System V2 enables preparation of amplified complementary cDNA from low concentrated RNA samples for microarray or quantitative real-time PCR (qRT-PCR) applications. Amplification occurs through a highly reproducible method using DNA/RNA chimeric primers, DNA polymerase and RNase H in an isothermal assay (Figure 2.4).

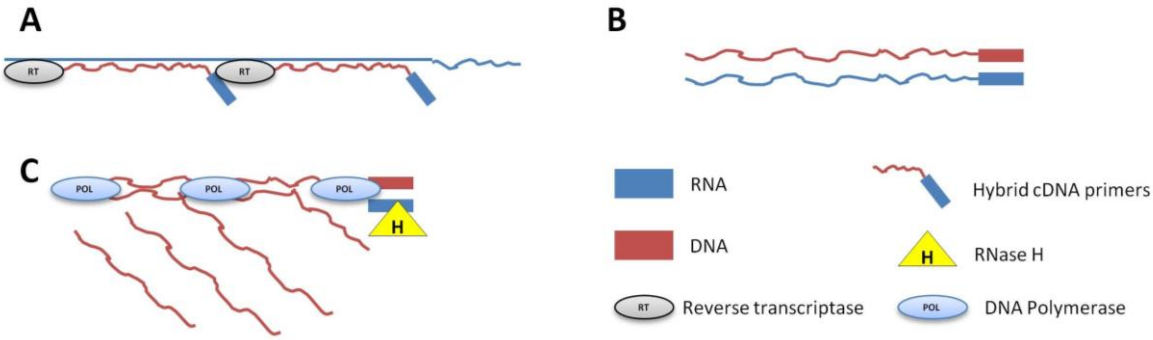

Figure 2.4 Ribo-SPIA® Technology used for cDNA synthesis and amplification

A: Generation of First Strand cDNA. Total RNA is mixed with random and oligo dT DNA/RNA chimeric primers, allowing for priming to occur across the whole transcript. Reverse transcriptase (RT) promoted extension occurs from the 3' DNA end of each primer generating first strand cDNA. B: Generation of a DNA/RNA Heteroduplex Double Strand cDNA. The cDNA is then fragmented to generate priming sites for DNA polymerase to synthesize a second cDNA strand containing DNA complementary to the RNA tag sequence localized at 5' on the first strand cDNA. The final product is a double-stranded cDNA with a DNA/RNA heteroduplex. C: Single Primer Isothermal Amplification (SPIA®). The RNA portion of the heteroduplex is firstly removed by RNase H. DNA polymerase then synthesizes cDNA starting at the 3' end of the SPIA primer, displacing the existing forward strand. The process of DNA/RNA primer binding, DNA replication, strand displacement and RNA cleavage is repeated, resulting in accumulation of amplified cDNA.

All samples were lyophilized in a 60Hz speed vac instrument and resuspended in 5µl nuclease-free water. This allowed decreasing sample volumes without compromising RNA concentration.

In a 0.2mL PCR, 2µL of first strand primer mix and 5µL of each sample were mixed and subsequently placed in a pre-warmed thermal cycler programmed to run Program 1 (Appendix, last section). Once primer annealing was complete, 2.5µL first strand buffer mix and 0.5µL first strand enzyme mix were added and samples were run on the thermal cycler, Program 2 (Appendix, last section). In order to perform Second Strand cDNA Synthesis, 10 µL of the second strand master mix (9.7µL buffer + 0.3µL enzyme mix) were mixed to each first strand reaction tube and Program 3 was run on the pre-cooled thermal cycler.

cDNA was purified using Agencourt RNAClean XP beads. At RT, 32µL of resuspended beads were admixed with each reaction and left to incubate for 10min. Samples were transferred to a magnet (SPRIplate® 96-ring, Agencourt Biosciences Corporation) with a strong magnetic field for 5min to completely clear the solution of beads. The solution was carefully discarded by pipetting and the beads washed with 200µL of 70% ethanol in triplicate. After completely removing the ethanol, the beads were left to air dry on the magnet for 15-20min. 100µL of the SPIA amplification mix (50µL buffer + 25µL primer + 25 µL enzyme) was mixed to each tube containing the double-stranded cDNA bound to the dried beads. Samples were then place on a pre-cooled thermal cycler programmed to run Program 4 (Appendix, last section). Tubes were transferred back to the magnet for 5min. The eluted amplified cDNA was removed and subsequently purified using the QIAGEN QIAquick PCR Purification Kit according to the company's instructions. The amplified cDNA was mixed with 500µL of Buffer PB, added to QIAquick spin columns and centrifuged for 1min at 13300 rpm. After discarding the flow-through, 700µL of 80% ethanol were added twice to the column, with intermediate 1min 13,300 rpm centrifugations. Columns were blotted onto absorbent paper to remove any residual wash buffer and placed into new 1.5mL eppendorfs. 30µL of nuclease-free water was added to the column, left to incubate for 5min at RT and forced through by spinning at maximum speed for 1min.

cDNA yield and purity were measured using the Nanodrop (single stranded DNA setting, where 1 OD unit at 260nm= 33µg/mL). cDNA quality was assessed using the Agilent Bioanalyzer RNA 6000 Nano Kit (protocol described in section 2.6.2, with minimal changes as required by manufacturer).

2.7.1.2 cDNA fragmentation and biotin labelling using the Nugen Encore Biotin Module

5µg of amplified cDNA in 25µL volume were fragmented by admixing with 7µL of fragmentation master mix (5µL buffer + 2µL enzyme mix) and placing in a pre-warmed thermal cycler programmed to run Program 5 (Appendix, last section).

Subsequently, 18µL of biotin labelling master mix (15µL buffer + 1.5µL labelling reagent + 1.5µL enzyme mix) was added to each fragmented cDNA sample and tubes were placed in a pre-warmed thermal cycler programmed to run Program 6 (Appendix, last section). The fragmented and labelled cDNA was processed immediately for array hybridisation. Fragmentation efficiency was assessed using the Bioanalyzer.

2.7.1.3 cDNA hybridisation to Affymetrix GeneChip® Human Gene 1.0 ST Arrays

The GeneChip® Hybridisation, Wash and Stain kit was used. The hybridisation cocktail was prepared according to the specifications for Mini Arrays (labelled cDNA, 5µg/25µl; control oligonucleotide B2 (3nM), 1.9µL; 20x hybridisation controls, 5.5µL; 2x hybridisation buffer, 55µL; DMSO, 11.6µL) and heated at 99^{o}C for 5min. The probe cartridges were meanwhile washed with 80µL of pre-hybridisation mix and incubated in rotation for 10min at 45^{o}C. The hybridisation cocktail was transferred to a 45°C heat block for 5min and spun at maximum speed in a microcentrifuge for 5min.

Finally the arrays were refilled with 90µL of the warmed hybridisation cocktail, left to hybridise for 16h at 45^{o}C at 60rpm and analysed using the GeneChip Fluidics station 450. Details on the approach taken to analyse the data will be detailed in the experimental Chapter 5.

2.7.2 Gene Expression Analysis by quantitative Real-Time PCR

qRT-PCR was used in these studies for validation of the GEP results and for interrogation of expression changes of target genes in macrophages after co-culture with reactive or malignant B-cells.

2.7.2.1 cDNA Synthesis

cDNA prepared as in section 2.7.1.1 was used for the array validation studies. RNA extracted from macrophages after co-culture was converted to cDNA using the high capacity reverse transcription kit (Applied Biosystems). RNA (0.5-1.5µg) in a volume of 10µl was mixed with the reaction master mix (Table 2.2) in a 0.2ml PCR reaction tube (final volume: 20µl). Reverse transcription took place on an automated thermal cycler with initial annealing for 10min at 25°C, followed by extension for 120min at 37°C, reverse transcriptase inactivation for 5min at 85°C and cooling at 4°C.

Table 2.2 Components of the 2x reverse transcription (RT) master mix used for cDNA synthesis

Components	**Volume per reaction (µL)**
10x RT buffer	2
25x dNTP mix (100mM)	0.8
10x RT random primers	2
Multiscribe reverse transcriptase	1
Nuclease free water	4.2

2.7.2.2 Quantitative Real Time quantitative PCR

This technique is used to quantify specific RNA transcripts in real-time. The reaction happens in a fluid mixture containing the template cDNA, an enzyme (Taq DNA polymerase), fluorescently-tagged primers-probes and water. The enzymatic cleavage of cDNA-bound probes in each round of PCR reaction leads to the releasing the fluorescent dye (FAM). The quantity of the detectable fluorescence is proportional to the amount of PCR product and the abundance of the RNA species of interest.

RNA transcripts were assayed in triplicate for each sample using TaqMan gene expression assays (Applied Biosystems). The housekeeping genes chosen were *B2M* for the microarray validation studies and *GAPDH* for the gene expression analysis performed in Chapter 6.

The reactions were set up in 386 well optical plates. Taqman 2x universal PCR Master Mix (Applied Biosystems) and expression assays were mixed first and placed into a loading reservoir and from here 11µl transferred to the designated well. cDNA were added in a 15µl final volume per reaction. The plate was sealed, centrifuged for 1min at 3300rpm to collect the reaction at the bottom of the well and loaded on the ABI HT-7900 system (Applied Biosystems). An initial 95^{o}C incubation for 10min enabled enzyme activation and cDNA denaturation. The reaction was then cycled for 40 cycles at 95^{o}C for 15 seconds and 60^{o}C for 1 minute.

Table 2.3 Components of Reaction Mix for qRT-PCR

Components	**Volume per 15µl reaction**
20x TaqMan gene expression assay	0.75
2x TaqMan gene expression master mix	7.5
RNase-free Water	2.75
cDNA template	4

Data was converted to relative quantities (RQ) for comparison between samples. A cycle threshold (CT), representing the number of PCR cycles at which the fluorescence signal for each target and control gene crosses a user defined threshold, was calculated. CT values are inversely proportional to mRNA abundance. An average of CT values for each sample was calculated, where replicate reactions with CT standard deviation (SD) ≥0.5 were discarded. Wells with a CT value of >35 were excluded from further analysis.

Target CT values were normalised to the housekeeping gene, generating a delta CT value (ΔCT= CT Target Gene – CT Endogenous Control Gene). The delta ΔCT (ΔΔCT) value, which represents the quantity of mRNA present in each condition, was determined by subtracting the ΔCT of a user-defined calibrator sample from the ΔCT of test samples. The Log^2 scaled data was finally transformed to a linear scale using the formula: $RQ=2^{-\Delta\Delta CT}$. RQ values were compared by T-test and fold changes (FC) calculated from the mean of each group as described in each chapter.

2.8 Co-culture systems involving primary human macrophages and malignant and reactive B-cells

Functional studies are the only ones able to provide a definitive biological explanation towards the impact of cell-cell interactions happening in vivo. In vitro cell systems are used in an attempt to mimic the cell-cell interactions and their functional effects. A dual-cell system such as the one used here, although limited to understand the relevance of the whole microenvironment in the biology of DLBCL, would help to explore one of our hypothesis for this study: the malignant DLBCL cells are inducing functional changes in the lymphoma-associated macrophages.

2.8.1 Positive selection of monocytes from using magnetic microbeads and macrophage maturation *in vitro*

Positive cell selection was performed using the magnetic cell sorting system developed by Miltenyi Biotec. PBMCs from healthy donors were isolated from buffy cones using the Ficoll method described in section 2.1.3. After cell counting, samples were divided in $1.5x10^8$ aliquots, pelleted by centrifugation and resuspended in 1000µl of cold buffer (PBS/0.5M EDTA/2% FCS) and 100µl CD14 antibody conjugated microbeads. After incubating for 15-20min in the dark at 4°C, cells were washed with 10ml of buffer, centrifuged at 1200rpm for 10min at 4°C and resuspended in 500µl of buffer. The LS columns were adapted to the QuadroMACS® magnet, rinsed with 3ml of buffer and loaded with the cell suspensions. Each column was washed 3 x 3ml of cold buffer. The positive cell fraction was forced through the column in 5ml of buffer by plunging, resuspended in Dulbecco's Modified Eagle's (DMEM) medium with high glucose and sodium pyruvate (Sigma-Aldrich), 10% filtered human AB serum (Sigma-Aldrich) and penicillin/streptomycin and centrifuged for 5min at 1500rpm at RT. Aliquots were taken for analysis of cell purity by flow cytometry (described in experimental Chapter 6). After automated cell counting, monocytes were plated in 90mm Petri dishes (Sterilin) in a 15ml volume at $4.5x10^6$/ml concentration and place at 37°C in a 5% CO2 humidified incubator for 7 days.

2.8.2 Positive selection of reactive B-cells from tonsils using magnetic microbeads

Tonsil SCSs were prepared, stored and resuscitated according to section 2.1.2. Viability and cell count was assessed using an automated cell counter.

Cell separation was performed according to the protocol described in section 2.8.1 with minor adjustment. The concentration of CD19 antibody conjugated microbeads used was kept as per manufacturers recommendations (20µl/10^7 cells).

After sorting, cell counting was assessed on a haemocytometer by mixing 5µl cell suspension with 5µl trypan blue dye 0.4% (Cell Viability Inc). The trypan blue is absorbed through the cell membrane of dead cells but not the intact membrane of live cells. For each sample 4x1mm^2 squares were counted on the haemocytometer

noting the number of live and dead cells. The average number of live and dead cells was multiplied by the dilution factor (2), giving a cell number x10^4/ml.

Finally, B-cells needed for co-culture were resuspended in DMEM medium with 10% filtered human AB serum and penicillin/streptomycin. Aliquots were kept for flow cytometry purity assessment.

2.8.3 Macrophage harvesting and replatting for co-culture experiments

After 1 week of incubation macrophages adhere to the bottom of the Petri dishes and need to be harvested by manual scrapping. The petri dishes were taken out of the incubator, transferred to the hood and sat on ice. Culture medium with unattached cells was gently harvested to 50ml centrifuge tubes using an automated pipette, leaving the bottom of the dish untouched. The adherent fraction was washed twice by stirring with 10ml of cold PBS and harvested together with the non-adherent cell fraction. 5ml cold PBS was added to each dish. A 1.8cm blade cell scraper (BD Falcon) was then used to detach the macrophages, with particular attention not to use circular movements and to keep the scrapper in a 45^o angle with the dish. The harvesting was repeated once to increase cell yield and a final wash with cold PBS was done. After centrifuging at 1400rpm for 10min at 4^{o}C, macrophages were resuspended in DMEM with 10% human serum and penicillin/streptomycin and counted on an automated counter. Finally, cells were adhered at a concentration of 2-5 x10^4/ml at 37^{o}C into 24-well plates for 24h prior to co-culture set-up.

2.8.4 Co-culture set-up

The following day, the DLBCL cell lines Su-DHL4 and Ri1 were washed and resuspended in DMEM with 10% human serum and penicillin/streptomycin in the concentration required for co-culture. Reactive B-cells were prepared according to section 2.8.2.

The 24-well plates were taken from the incubator and adequate adherence confirmed using the microscope (10x amplification). All work was then performed under the hood.

Reactive and malignant B-cells were layered in triplicate over the adherent macrophages at a 1:1 concentration. Transwell inserts with 0.4µm pore size and $1x10^4/cm^2$ pore density (Costar, Corning) were also used, allowing cell separation but passage of small molecules across the membrane. After transwell insertion, cells were added to the upper chamber at a 1:1 concentration in a 100µl volume. As a positive control, macrophages were treated with LPS at a concentration of 100ng/ml. Finally, as negative internal controls for each biological experiment, macrophages were cultured alone for an additional 24h. Figure 2.5 below illustrates the experimental set-up.

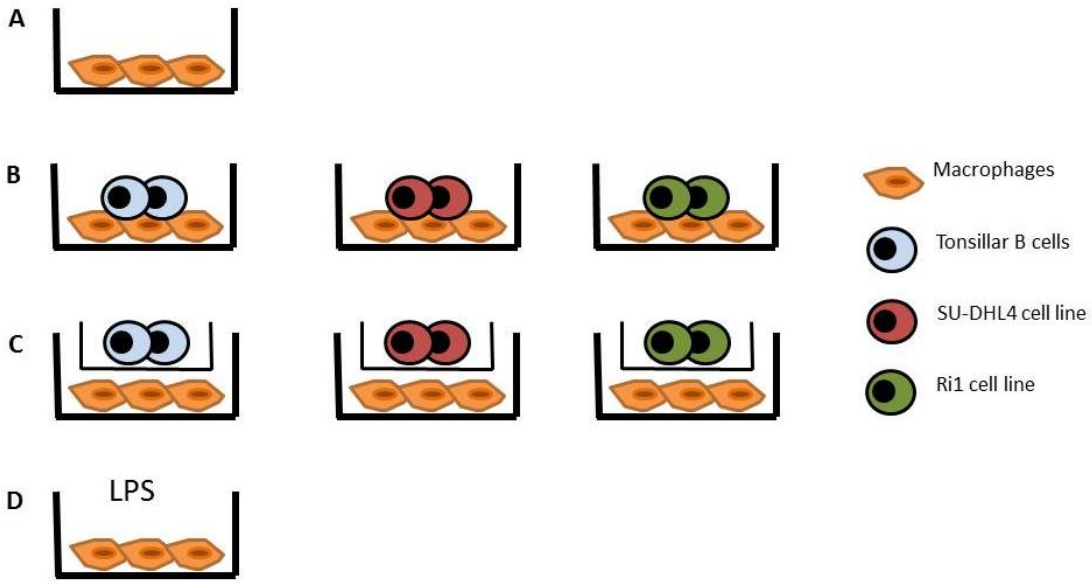

Figure 2.5 Macrophage and B-cells co-culture system employed in this study
Reactive tonsillar B-cells and DLBCL cell lines were plated in a 1:1 ratio in contact (B) or on the top layer of a transwell insert (C) in 24 microwell plates where macrophages were allowed to adhere for 24 hours. As negative internal controls for each biological experiment macrophages were cultured alone for additional 24h (A). Finally, macrophages were treated with 100ng/ml of LPS (D).

2.8.5 Macrophage harvesting after co-culture

After 24h, plates were placed over an icebox. 1ml of co-culture supernatant from all experimental conditions was collected in 1.5ml eppendorfs and centrifuged at 4000rpm for 5min at 4^{o}C in a microcentrifuge. The top layer was aspirated carefully (~0.5ml) and stored at -20^{o}C for cytokine quantification. The remaining supernatant was replaced by 200µl of cold fresh PBS with 2%FCS. Adherent cells were scrapped with a cell scraper with the blade cut on both sides, avoiding circular movements. Once the cell suspension was collected to 1.5ml eppendorfs, additional scrapping was done if significant residual adherent cells were seen under the microscope. Finally, all tubes were centrifuged at 4000rpm for 5min at 4^{o}C. The cell pellets were resuspended immediately for flow cytometry (section 2.4.2). In order to guarantee a pure population after co-culture, macrophages were sorted using size and intracellular complexity differences compared to B-cells and stored at -80^{o}C for future RNA extraction.

2.9 Cytokine profiling of co-culture supernatants using cytometric bead arrays

The human Th1/Th2/Th17 Cytometric Bead Arrays (CBA, Beckton Dickinson) were used to quantify IL-2, IL-4, IL-6, IL-10, IL-17, TNF, and IFN-γ in supernatants from the co-culture experiments just described. The CBAs are a flow cytometry based assay, which enables highly sensitive quantification of multiple proteins simultaneously in a 50µl sample volume. As a result, this method significantly reduces sample requirements and time to results in comparison with ELISA and Western blotting. The system relies on the addition of multiple antibody-coated beads, which capture the specific proteins in solution. These beads have unique fluorescence intensity emitted on the PE-FL2 channel of the flow cytometer and hence can be analysed simultaneously in a single tube.

The protocol recommended by the manufacturer was strictly followed.

The cytokine lyophilized standards were reconstituted in 2ml of Assay Diluent. After incubation at RT for 15min, 9 serial dilutions (1:1 - 1:256) were prepared. Diluent represented the negative control. An aliquot of 10µl/sample of each cytokine Capture Bead were mixed in a single tube and 50µl of mixed beads were distributed to the appropriate assay tubes. 50µl of the standard dilutions were added to the control assay tubes. Finally, 50µl of each test sample and 50µl of the PE Detection Reagent were added to the test assay tubes. All tubes were left to incubate for 3 hours at RT and protected from direct exposure to light. After incubation all samples were washed with 1ml of Wash Buffer and centrifuged at 1000rpm for 5min. Supernatant was discarded and pellets resuspended in 300µl of Wash Buffer. Data was acquired on a BD Fortessa II flow cytometer with FACS Diva Software. The beads were excited off the 488nm laser and detected in the PE (FL-2) detector (Figure 2.6). Approximately 2000 events were acquired per sample.

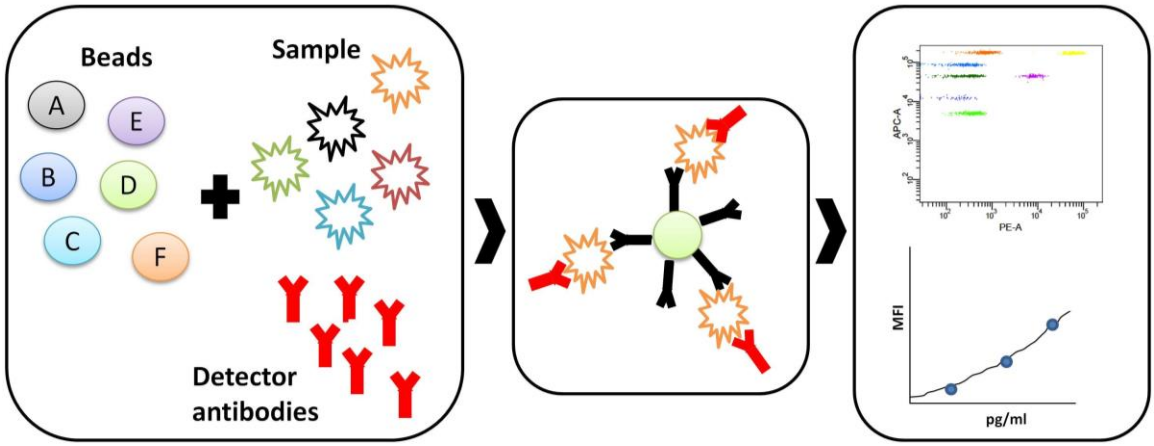

Figure 2.6 Outline of the cytokine profiling using the Cytokine Bead Array (CBA) System (BD Biosciences)

Cytokine specific beads are incubated with the sample and detector antibodies conjugated to the PE fluorochrome. After incubation, cytokine/antibody/detector antibody complexes are detected by flow cytometry. The concentration of each cytokine in a given sample is directly proportional to its fluorescence intensity on the PE-FL2 channel and is calculated based on the standard curve fitted from the concentration on the control assay tubes.

2.9.1 FCAP Array analysis of cytokine secretion

Data analysis was performed using FCAP Array software (Becton Dickinson). The software allows determination of the analyte concentrations in the samples based on know concentrations from the standards run in the same conditions.

The CBA beads were gated based on the forward scatter versus side scatter features, after which individual bead sets representing each cytokine were gated individually. Finally the MFI of each gated population on the PE detector was measured. A standard curve based on a 4-parameter logistic mathematical model was used to calculate the concentration of measured proteins in each condition. The results were expressed as a pattern code representing the MFI of the sample population and finally in pg/ml (Figure 2.6).

2.10 Statistical analysis

Details on statistical analyses applied are given within the relevant experimental chapter. Statistical tests were performed using Prism software Version 5.03, SPSS Version 19.0 and Excel.

Chapter 3 Reliability of immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma

3.1 Introduction

As discussed in Chapter 1, DLBCL represents a heterogeneous group of lymphoid malignancies with distinct oncogenic events and clinical behaviour that cannot be distinguished only by morphology and immunophenotype.[34,74,92] This biological diversity explains the heterogeneous responses to the current standard treatment, R-CHOP, and provides a rational for investigation of novel targeted therapies.

Through the use of microarray technology, Alizadeh et al.[74] described the two molecularly distinct forms of DLBCL, which replicate GEPs typical of different stages of B-cell differentiation. GCB-DLBCLs assume a transcriptome superimposable to that of normal GC B-cells. In concordance with this, GCB cases exhibit immunoglobulin gene ongoing somatic hypermutation.[81] ABC-DLBCLs express genes more characteristic of plasma cells,[76] but are blocked in their differentiation capacity. The two entities are very distinct in their genetic changes, and signalling pathway deregulation. Numerous studies are trying to pinpoint the biological features that explain outcome differences and could be used as markers for targeted therapy. Importantly, it has been shown that patients with GCB-DLBCLs have an improved OS compared with ABC-DLBCLs after treatment with R-CHOP.[34] In fact, the molecular classification is regarded the most robust biological prognostic tool available for DLBCL.

Molecular characterization is opening up opportunities for personalized therapy in poor-risk DLBCL. Emerging in vitro and clinical data support that the two main DLBCL molecular subtypes, the ABC and GCB-DLBCLs may benefit from different treatment approaches, with agents including bortezomib,[300] lenalidomide[117,301] or ibrutinib[118] appearing particularly active against the worse prognosis ABC subtype.

In order for targeted therapies to succeed in this disease, reliable and reproducible

strategies that adequately segregate patients into distinct molecular groups are needed. While GEP is the gold standard method for molecular stratification of DLBCL in the laboratory, this technique has only recently been incorporated into clinical trials for treatment stratification. The REMoDLB phase 3 clinical trial (NCT01324596) aims at determining whether the addition of bortezomib to standard R-CHOP improves EFS and if that benefit is related to the molecular features of the tumour cells, which is being characterized by GEP in the FFPE tissue. However, since the application of GEP is still restricted to research purposes, there is presently a lack of standardized methodology for array analysis, which can lead to variable results both at the inter- and intra-laboratory level. This issue, which may impact on its results and on patient care, is generally unreported.

The lack of a routine methodology for GEP based COO assessment has led investigators to develop IHC based approaches for the molecular classification in DLBCL, using proteins that either were already known to be expressed in GC and post-GC cells or that were unravelled by the GEP studies. In 2005 Hans and co-workers[302] established the first IHC algorithm, allegedly with high sensitivity for GEP classification. Subsequently, eight further strategies[303-308] have been published, most of which reported a better concordance with molecular-based classification than the Hans algorithm and the ability to segregate two groups with significantly different outcome. However, many investigators continue to question the clinical applicability of these algorithms.[26,51,94,309-313] Results are typically inconsistent and are generally poorly reproducible by independent groups. Nevertheless, IHC is attractive as a surrogate for molecular stratification in DLBCL and the Hans algorithm is being used to define DLBCL of the ABC-type in clinical trials offering NF-kB targeting agents to patients with this subtype.

3.2 Objectives

This study set out to provide an up-to-date systematic comparison of nine IHC scores (Figure 3.1) for molecular classification in a new large dataset of diagnostic DLBCL. Our primary objective was to test the reliability of these methodologies in classifying individual cases of this cohort. IHC profiles for single proteins and each algorithm were assessed and agreed among three expert observers. A consensus matrix based on all scoring combinations and the number of subjects for each combination was constructed in order to assess reliability. As a secondary aim, the survival impact of individual markers and the nine classifiers was evaluated.

We hypothesized that the IHC algorithms are not reliable predictors of the molecular classification of DLBCL. The approach used in this study addresses the important question of whether IHC is or not a reliable alternative to molecular-based methods to be used for clinical decisions in DLBCL.

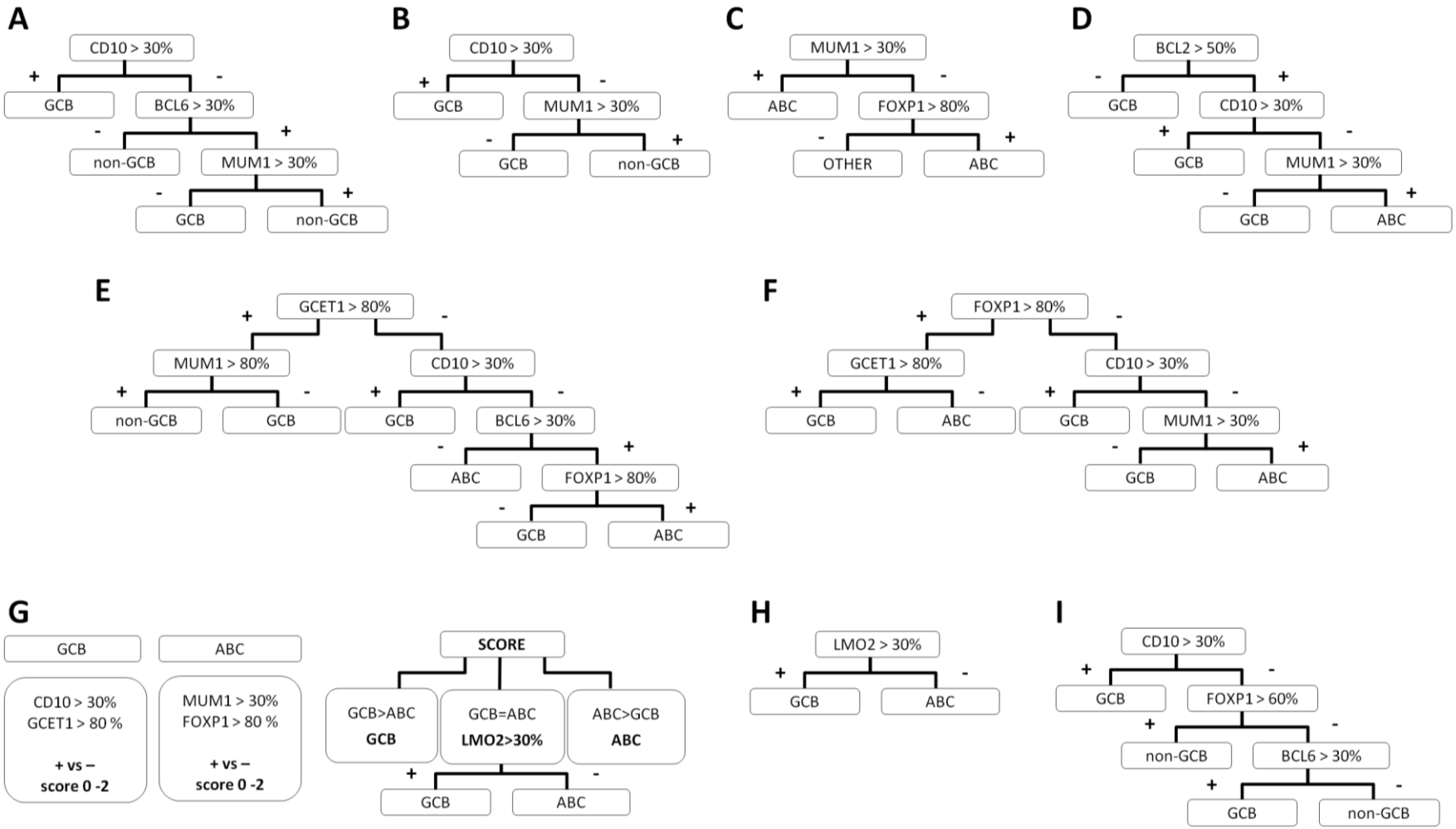

Figure 3.1 Algorithms applied in the current study.

(A) Hans; (B) Hans modified; (C) Nyman; (D) Muris; (E) Choi; (F) Choi modified; (G) Tally; (H) Natkunam; (I) Visco-Young.

3.3 Materials and Methods

3.3.1 Patients

Patients with de novo DLBCL included in the studies detailed here were treated in two institutions, as already described in Chapter 2. Patient selection was dependent on the availability of good quality FFPE tissue of the diagnostic biopsy and extended clinical and follow-up data. December 2009 was selected as the cut-off date of diagnosis to include patients in the study to guarantee an acceptable follow-up.

Overall, 161 R-CHOP chemo-immunotherapy treated patients and 148 patients treated with other regimens were included in these studies. Survival according to treatment era and institution is detailed in Figure 3.2. Whenever no prognostic studies were planned, data from all patients were included. Clinical data for the R-CHOP cohort, including response to chemotherapy is detailed in Table 3.1. The median follow-up of the R-CHOP cohort is of 55 months.

Due to TMA tissue loss and technical obstacles a variable number of cases were included in the final analysis. A comparison of clinical features and outcome between cases included and the ones lost was done in parallel and will be presented in the results.

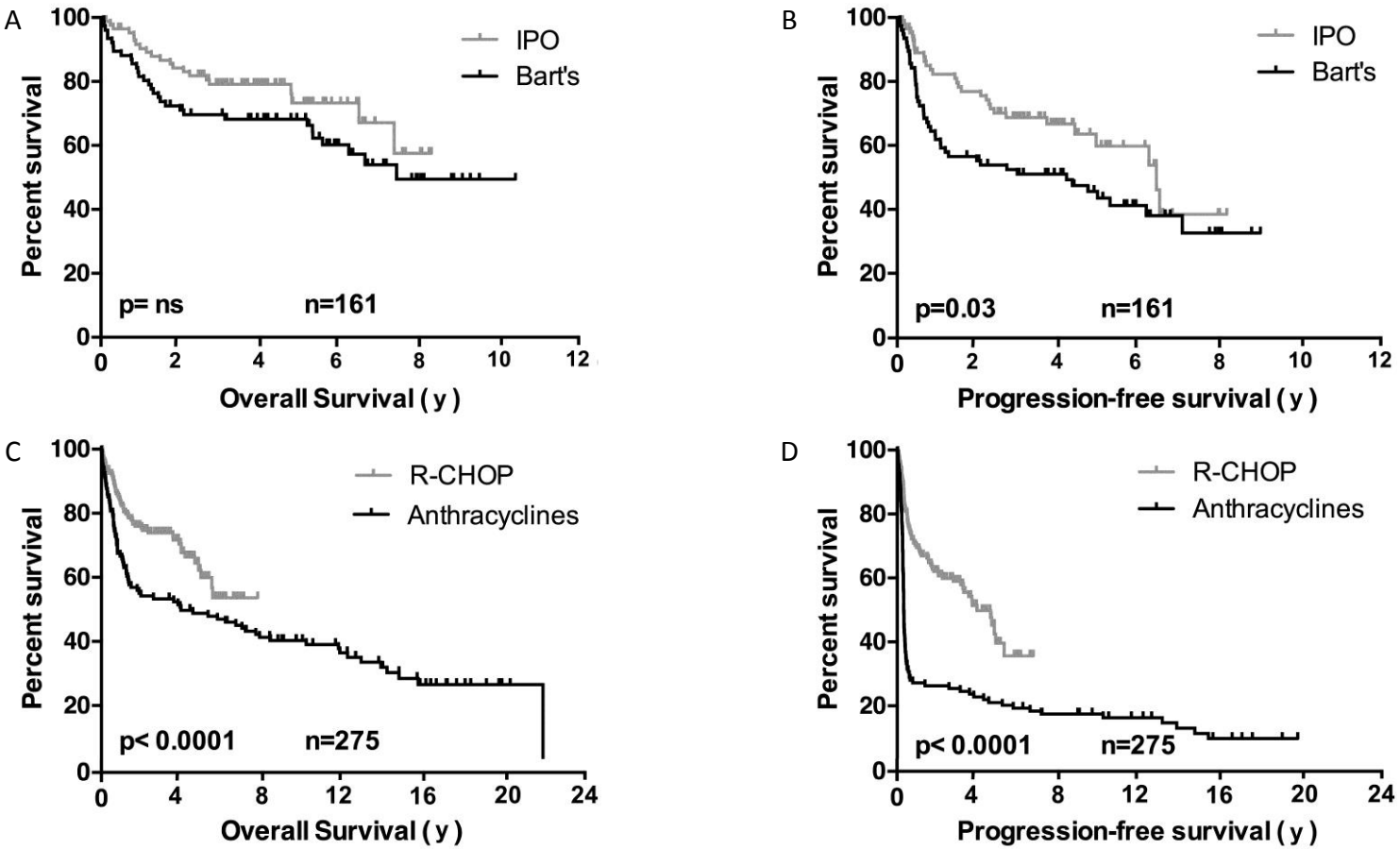

Figure 3.2 Kaplan Meier curves according to treatment regimens and treating institution.

A&B: OS and PFS of patients according to treatment regimen. 114 of 148 patients received anthracycline based regimens (CHOP, MACOP-B, VAPEC-B) and are highlighted in the curve. The remaining were managed palliatively or did not receive treatment and were excluded from analysis;

C&D: OS and PFS of patients stratified by treatment institution. Differences between groups were determined using the chi-square method with significance defined as $p<0.05$

Table 3.1 Clinical features of the R-CHOP series

	Barts (77 patients)		**IPO (84 patients)**		
Clinical characteristics	number	%	number	%	p-value
Male	53	68,8	39	46,4	0,004
Age, years (Median, range)	57 (19 - 86)		64,5 (16 - 85)		NS
Age > 60 years	34	44,2	50	59,5	
Stage III-IV	44	57,1	44	52,4	NS
"B-symptoms"	18	26,5	21	25,3	NS
≥ 2 Extranodal sites	13	16,9	16	18,8	NS
High LDH	38	51,4	67	79,8	< 0.001
ECOG PS ≥2	8	10,5	15	17,9	NS
IPI					0,01
Low	39	50,6	26	31,0	
Low-intermediate	11	14,3	22	26,2	
High-intermediate	22	28,6	21	25,0	
High	5	6,5	15	17,9	
IPI ≥3	27	35,1	36	42,9	NS
Response					< 0.001
CR/CRu	47	61,0	68	81,0	
PR	16	20,8	11	13,1	
SD	2	2,6	0	0,0	
PD/failure	11	14,3	4	4,8	
Relapse rate (from CR/CRu)	11	23,4	16	23,5	NS
Death rate	26	33,8	21	25,0	NS
Causes of death					NS
Lymphoma	19	73,1	12	57,1	
Toxicity	2	7,7	6	28,6	
Follow-up (median, range)	65.5 (16.3 - 110.4)		47.3 (0.6-94)		

ECOG: Eastern Cooperative Oncology Group; LDH: lactate dehydrogenase; IPI: International Prognostic Index; Low-int: low-intermediate; High-int: high-intermediate; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease; NS: not significant

3.3.2 TMA and IHC

Sample collection followed informed consent in accordance to the declaration of Helsinki. Ethical approval for this study was obtained from local ethics regional committees.

For the Bart's TMAs triplicate 1 mm^2 cores were taken from regions of biopsy material rich in malignant cells identified on H&E stained sections. Representative tumour regions had been previously marked by me and an expert Haematopathologist [Dr Maria Calaminici (MC)]. Tonsils were cored in all TMAs as internal controls. For the IPO TMA, duplicate 1.5 mm^2 cores were taken after selection of representative tumour areas by JC. Staining for the pan B-cell marker CD20 was performed on TMA sections to confirm adequate tumour representation. TMAs were sectioned and transferred onto glass slides. After dewaxing, blocking in hydrogen peroxide/methanol solution, rehydration, and pressure-cooker antigen retrieval, the slides were subjected to immunostaining. Primary antibody reaction for CD10, BCL6, BCL2, MUM1, FOXP1, GCET1 and LMO2 was detected using a peroxidase-labelled system (Super-Sensitive Polymer-HRP IHC Detection System, BioGenex). An immunological amplification method (CSA II, Catalyzed Amplification System, Dako) was used exclusively for BCL6. Heat induced antigen retrieval using a pressure cooker was used for all antibodies. In each batch of staining, tonsil sections were analyzed simultaneously for all markers. All IHC was performed in the same laboratory. Primary antibodies, conditions of use and source are provided in Table 3.2. Representative negative, positive and control cases are shown in Figure 3.3.

Table 3.2 Primary antibodies and conditions of use

Antibody	Clone	Species	Dilution	Source
CD10	56c6	Mouse	1/250	Novocastra
BCL2	124	Mouse	1/200	Dako
BCL6	LN22	Mouse	1/500	Novocastra
MUM1	MUM1p	Mouse	1/400	Dako
FOXP1	JC12	Mouse	1/500	Abcam
LMO2	SP51	Rabbit	1/500	Abcam
GCET1	Ram341	Mouse	1/100	Abcam
CD20	L26	Mouse	1/2000	Dako

3.3.3. Slide scanning, scoring and analysis

Slides were scanned using the Hammamatsu Virtual Slide Scanner NanoZoomer 2.0 (Hammamatsu), and viewed using the NDP.scan software.

All cores were jointly visualized on a computer screen at low and high magnification by MC and Dr Abigail Lee and myself to standardize the scoring criteria for each antibody applied. Each case was scored as positive or negative according to the cut-points defined in the original publications, as detailed in Figure 3.1. Thus, some antibodies were scored using more than one cut-point, as detailed in Table 3.3. As most publications report only scant methodological details, negative cases with absent internal controls were deemed unclassifiable, as recommended by the Lunenburg Lymphoma Biomarker Consortium (LLBC) guidelines.[314] This and the absence of whole cores in the TMA were the primary causes for the inability to score (unclassifiable cases are detailed for each antibody in Table 3.3). Whenever individual cores of a given case showed non-concordant results, the core with highest large cell infiltration was used. After all observers had assessed each staining, a meeting was organized to reach consensus on discordant cases. In almost all cases a 3:0 decision was reached, but in less than 5% of the cases a 2:1 score was accepted.

As expected, consensus was higher for the antibodies in use in the diagnostic setting, such as CD10, BCL2 and MUM1. LMO2 shows a nuclear distribution and is also expressed by T-cells so analysis also included nuclear size. FOXP1 scoring was difficult in some cases due to the background staining and inter-patient differences in staining intensity.

3.3.4 Statistical analysis

Correlations between the expression of specific proteins or allocation to a particular algorithm and clinical parameters were tested using the Pearson's test.

3.3.4.1 Outcome analysis

Results regarding single marker expression and algorithm distribution refer only to the more recently diagnosed and R-CHOP treated cohort, given that biomarker distribution was under prognostic scrutiny. Differences in clinical characteristics between the two R-CHOP series were tested using chi-square or Fisher's exact test, when appropriate.

In the univariate analysis, log-rank tests were performed. All parameters of the IPI were assessed for prognostic impact. Relevant interactions of the markers studied with the IPI factors were assessed. Cox proportional hazards models were used to obtain estimates of the hazard ratios (HR) with 95% confidence intervals (CI). Models included all variables with a p-value ≥ 0.2 on univariate analysis. Both backward and forward selection methods were used to test independent significance of each variable included in the models.

The outcomes, measured from date of diagnosis to occurrence of event or date of last follow-up, were: OS, the event being death from any cause; and PFS, the event being failure of treatment (including not achieving complete response (CR) or progressing after achieving CR) or death of any cause. Median follow-up was calculated for patients alive at last follow-up. Statistical analysis was performed using SPSS version 19.0 (SPSS, Chicago, IL) and Prism version 5.03 (GraphPad Software, La

Jolla, CA).

3.3.4.2 Comparison of IHC COO algorithms

All cases included in the arrays at the time of the current study were used in the comparison of algorithm performance, independently of the treatment received. Firstly data compiling the classification of all cases according to each of the IHC classifiers was gathered. Subsequently, a consensus matrix based on the nine classifiers and number of subjects for each combination was built. Concordance across all methods was measured using the kappa statistics.

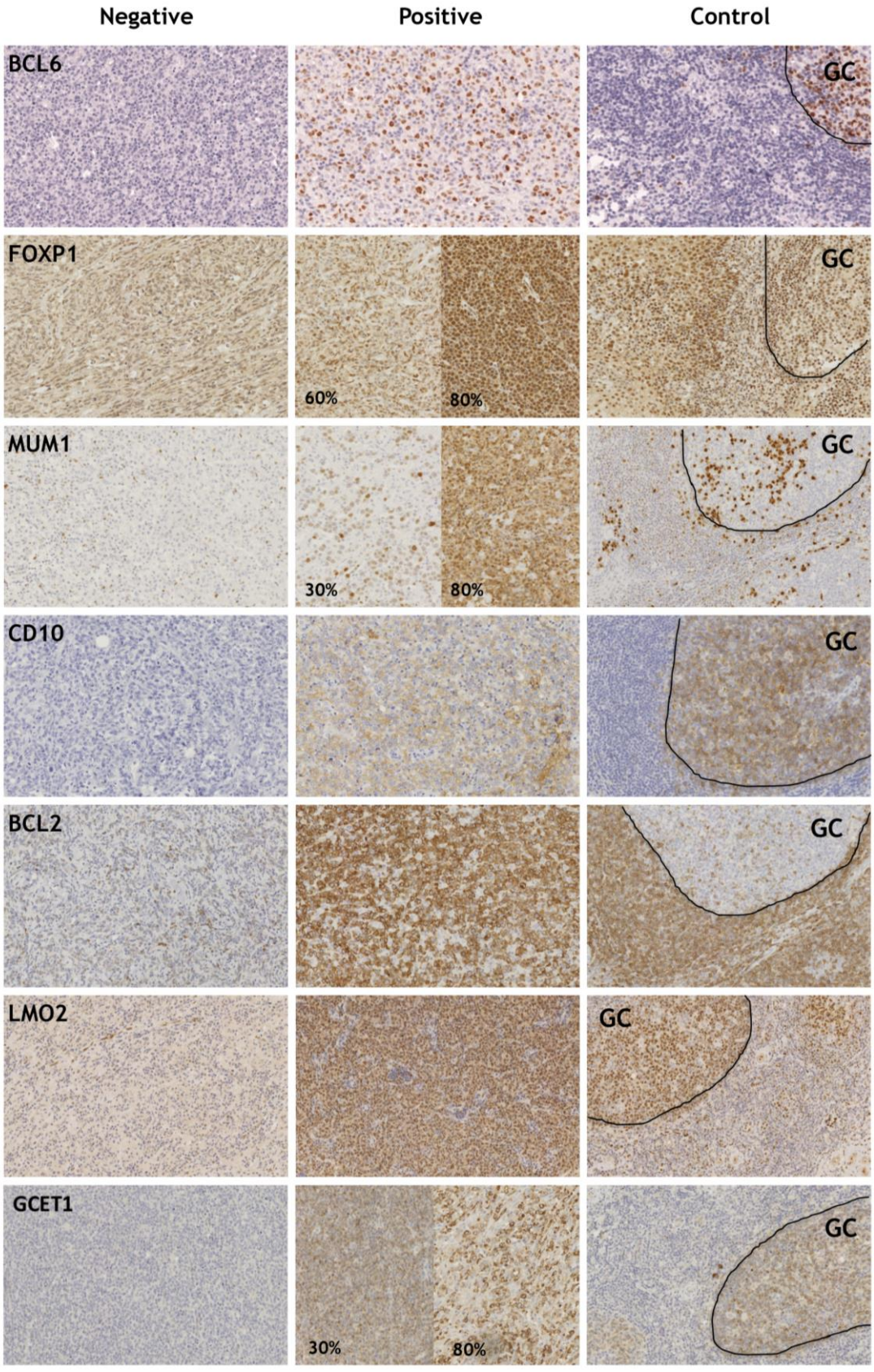

Figure 3.3 Immunohistochemistry results for each antibody.

Original magnification: x20. Cases were considered CD10 negative if CD10 staining fibroblasts were encountered in the tumour. In LMO2 negative cases was common to observe positive T-cells in the microenvironment, constituting an ideal internal control. GC, germinal centre.

3.4 Results

3.4.1. Analysis of individual markers

With the exception of the Natkunam algorithm, all classifiers tested use a combination of antigens for final allocation. In fact, it is generally accepted that no single protein expression appears able to mirror the GEP classification.

Although our main aim was to compare the results from the different algorithms, we hypothesized that the comparison of single antigen expression patterns from our series with those from others reported in the literature would give us an indication of the reliability of IHC for protein studies in this setting. Results from single protein expression of the R-CHOP treated cases (positive, negative and unclassifiable cases) are illustrated in Table 3.3. We could not document a survival impact of the expression of any the proteins studied, including BCL2 or BCL6.

CD10 is a glycoprotein detected in normal lymphoid ontogeny on pro-B cells and mature GC B-cells. In normal and malignant LNs, CD10 expression is also seen in cells from the stromal compartment. In the context of lymphoma it helps identifying cases of GC derivation, such as follicular lymphoma or a subset of GCB-DLBCLs. In our series its expression was detected in 28%, in keeping with previous results.[51,302,303,306,311] As has been previously suggested, CD10 positive cases might have a better OS, although the data available is not entirely concordant. For the purposes of ascertaining true outcome impact of each marker, we tested potential interactions with clinical parameters. In our series, the CD10 expressing cohort was enriched for younger patients ($p=0.03$), whereas negative cases had higher LDH ($p=0.02$). However, no differences in the IPI distribution were detected. We observed a positive correlation between CD10 expression and other GCB-"specific" proteins, including BCL6 (Pearson´s 0.311, $p<0.001$) and GCET1 (Pearson´s 0.45, $p<0.001$), and a negative correlation with the post-GC marker MUM1 (Pearson´s -0.164, $p=0.05$).

Table 3.3 Single marker analysis of the R-CHOP cohort

Antigens (cut-off)	positive [n (%)]	negative [n (%)]	unclassifiable (n)
CD10 (30%)	41 (28)	106 (72)	4
BCL6 (30%)	86 (61)	54 (39)	11
MUM1 (30%)	109 (76)	35 (24)	7
MUM1 (80%)	60 (42)	84 (58)	7
GCET1 (30%)[a]	29 (20)	118 (80)	4
GCET1 (80%)	12 (8)	135 (92)	4
FOXP1 (30%)	116 (82)	25 (18)	10
FOXP1 (60%)	102 (72)	39 (28)	10
FOXP1 (80%)[b]	77 (55)	64 (45)	10
LMO2 (30%)	79 (56)	62 (44)	10
BCL2 (50%)	96 (69)	43 (31)	12

Absolute number and percentage of classifiable cases is highlighted for each antigen. Staining and analysis were performed centrally. [a] a significantly higher proportion of GCET1+ cases (30% cut-off) was detected in patients diagnosed at Bart's Hospital (p=0.006); [b] a significantly higher number of patients diagnosed in IPO were considered positive for FOXP1 at the cutoff of 80% (p=0.001). n: number; NS: not significant

Both GCB and ABC-DLBCLs can haorbour genetic aberrations involving BCL6, a transcription repressor molecule essential for the formation of the GC reaction. Most genetic aberrations involving BCL6 lead to protein overexpression. In consequence, B-cells cannot differentiate into plasma cells but continue to divide and proliferate. GCB-DLBCLs harbour mutations within the *BCL6* auto-regulatory domain, whereas ABC-DLBCLs exhibit translocations deregulating *BCL6*. Nevertheless, BCL6 protein is considered a GC marker.

BCL6 expression in our series was comparable to previous reports.[302,311,313,315] A weak positive correlation was observed with other GCB-related proteins (CD10; GCET1, Pearson´s 0.39, p<0,001; and LMO2, Pearson's 0.51, p<0.001). BCL6+ patients were more likely to have low-risk IPI (p=0.05), and in this IPI subgroup BCL6 expression alone conferred a better 3-year OS (79 vs 93%, p=0.04). When the whole dataset was

analysed, BCL6 was unable to differentiate patients with distinct outcome. This contrasts with recent data suggesting that BCL6 protein overexpression alone is associated with an adverse prognosis, independently of the IPI score.[316]

The *MUM1* gene codifies for a lymphoid-specific transcription factor. MUM1 was introduced in the first IHC algorithm for prediction of the molecular stratification as a post-GC marker. Its expression at a cut-off of 30% was detected in 109 (76%) of our subjects, higher than previously reported.[302,305] However, using an 80% cut-off, 42% of the patients were deemed positive for MUM1, which is in keeping with Choi's data.[306] Moreover, this higher cut-off improved correlation with other post-GC markers, such as FOXP1 (Pearson´s 0.19 to 0.37, $p<0.01$). A significant correlation between BCL2 and MUM1 expression was also documented (Pearson´s 0.36 to 0.37, $p<0.01$).

Centerin (GCET1) protein expression is restricted to a subset of GCB cells[317] and should specifically identify GCB DLBCLs. However, expression of this marker was detected in only 8% of cases when assessed at a cut-point of 80%, and this only increased to 20% at the lower cut-off of 30%. At both cut-points, although positive patients were more likely to be younger ($p=0.003$), there was no association with IPI. At the 30% cut-off for expression, weak positive correlations with other GC markers, including CD10 and LMO2 (Pearson's 0.34, $p<0.001$), and negative correlation with MUM1 (Pearson's -0.19, $p=0.03$) were detected, with weaker correlations at the higher cut-point.

FOXP1 is a transcription factor, which has been detected at high levels in cases lacking GCB markers and expressing BCL6 and MUM1.[318] In our cohort, 82% of the patients had >30% and 72% had >60% of FOXP1+ cells. If a higher cut-off of 80% was used, a higher proportion of patients treated in the Portuguese institution were allocated to the FOXP1+ group (67% vs 40%, p=0,001).

LMO2 emerged from GEP studies as a strong prognostic marker in DLBCL. The LMO2 protein is expressed in LN solely by the nucleus of normal GC B-cells and by a subset of GCB-DLBCLs.[319] Natkunam et al.[304] proposed that LMO2 alone has a high predictive power for GCB allocation and consequently a positive impact on patient survival, even after the introduction of rituximab. We detected LMO2 staining in 79 (56%) cases, which is in keeping with the original data. However we were not able to detect any prognostic value in our cohort.

BCL2 is an oncogene commonly targeted in DLBCL, activating an anti-apoptotic program in the malignant cells. Forty-five per cent of GCB-DLBCLs are associated with t(14;18) translocations and consequently have BCL2 overexpression. This oncogenic event was divergently correlated with outcome. The majority of ABC-DLBCLs have BCL2 overexpression due to transcriptional deregulation. Studies have reported a negative outcome impact in these cases. The expression of BCL2 by the malignant B-cells has been associated with poor outcome in most studies.[94,96,320,321] However, its survival impact appears to be dependent on the COO allocation, rituximab treatment[322,323] and by the coexistence of MYC genetic aberrations. In our series, BCL2 was expressed in 96 (69%) patients, with no association with survival being detected.

3.4.2 Algorithm classification: distribution and consistency

According to the original and most widely utilized Hans method[302] we classified 53 of our subjects as GCB (Table 3.4). CD10 expression was determinant for this allocation, as 41 of these patients were positive for this antigen. This is in keeping with the original publication. The remaining 12 cases were all BCL6+ and MUM1-. Within the non-GCB cohort, BCL6 expression was detected in half of the cases. Hans's allocation was similar between the two R-CHOP cohorts and no association with clinical characteristics was found.

The modified Hans was proposed to decrease inconsistency introduced by the anti-BCL6 antibody.[307] According to modified Hans criteria, 61 (42%) patients were classified as GCB, including all GCB patients from the Hans method and a further eight patients scored CD10-/MUM1-.

The Choi classifier[306] relies on the expression of three antigens (GCET1, MUM1 and CD10) for initial allocation and on a further two (BCL6 and FOXP1) for final decision. Cut-off points were adopted according to appropriate detection of a single series of molecularly profiled patients. Using the cut-offs defined in the original paper, allocation into GCB and non-GCB groups in our cohort was 65 (45%) and 79 (55%) patients, respectively.

Acknowledging the complexity of the Choi algorithm, Meyer et al.[307] proposed the modified Choi. In our hands, only 35 (25%) of patients were classified as GCB. Compared to the Choi algorithm, the modified version re-classified both GCB (64 cases) and non-GCB cases (8 cases). Due to the different proportion of FOXP1+ patients between the two R-CHOP series, the Portuguese series was enriched for ABC cases using the Choi ($p=0.05$) and the modified Choi ($p=0.01$) criteria. The Choi ABC cohort was enriched for older patients ($p=0.01$).

From the clinical point of view it would be more relevant to use a method that is highly sensitive at identifying ABC patients, the subgroup that may be more amenable to targeting with new agents. Nyman et al.[305] proposed a method for this purpose, using only post-GC antibodies (MUM1 and FOXP1). Due to a high proportion of positive cases for both markers, in our series only 18 patients (13%) were considered non-ABC using this approach.

Table 3.4 Distribution of R-CHOP treated patients according to the nine IHC classifiers

Classifiers	Bart's [n (%)]		IPO [n (%)]		original data (%)		
	GCB	non-GCB	GCB	non-GCB	GCB	non-GCB	p-value
Hans	29 (44)	37 (56)	24 (32)	50 (68)	42	58	NS
Hans*	35 (51)	34 (49)	26 (34)	50 (66)	54	56	NS
Choi	37 (54)	32 (46)	28 (37)	47 (63)	58	42	0.05
Choi*	22 (34)	43 (66)	13 (17)	62 (83)	49	51	0.01
Natkunam	37 (57)	28 (43)	42 (55)	34 (45)	55	45	NS
Nyman	13 (19)	55 (81)	5 (7)	71 (93)	45	55	0.03
Muris	36 (56)	28 (44)	35 (45)	42 (55)	55	45	0.03
Tally	17 (26)	48 (74)	13 (17)	62 (83)	45	55	NS
Visco-Young	29 (43)	38 (57)	24 (32)	52 (68)	53	47	NS

* modified; p-values refer to differences between patients treated in the different institutions.

The Muris classifier[303] is the only method using BCL2. In our series 71 (50%) patients were classified as GCB, 43 of which were BCL2-, suggesting that this marker plays a predominant role at defining this COO subgroup.

The Tally algorithm[307] has the unique feature of attributing similar weight to GCB and post-GCB specific markers for allocation. There was a predominance of ABC cases (110, 78%) in our cohort. In 34 cases LMO2 expression had to be used for decision, being positive in half of the cases.

Finally Visco et al.[308] recently launched another method with an increased overlap with GEP data. Similarly to Hans, CD10 plays the central role for GCB allocation (41 of the 53 GCB patients were CD10+). Regarding non-GCB allocation, BCL6 plays less of a role than in the Hans classifier as in only 15 CD10-/FOXP1- patients was BCL6 expression taken into account for allocation. As with the Choi classifier, there was enrichment for older patients in the Visco ABC subset (p=0.02).

To address the question of how an individual case is classified across all methods, we then performed a parallel classification of all tumours using the nine COO algorithms.

Results for all classifiers were available for 242/298 cases (81%). Surprisingly, only 4.1% of the tumours were classified as GCB by all methods. The degree of agreement in allocation of patients to the non-GCB group was significantly higher, with 21% of patients being allocated to this group by all methods and 20.6% being classified as non-GCB by all methods except one - either the Choi (2 cases), the modified Choi (2 cases), the Natkunam (33 cases), or the Muris (13 cases) algorithms. Of note, the last two are the only methods in which allocation to the GCB subset was higher.

We then sought to assess pair wise agreement using the general kappa statistics, a method that tests for inter-scoring reliability. The kappa is considered a robust statistical method since it takes into account the agreement occurring by chance. Fig. 3.2 illustrates the strength of agreement among all scoring systems. Poor and fair kappa values were detected in 44.4% on pair wise concordance assessment; and in only 20% was kappa excellent or good. The Natkunam algorithm is the least concordant with the remaining, showing a poor agreement with four algorithms and only fair agreement with the other four. The highest level of agreement was found between the Choi and the Visco algorithms (Kappa=0.85). From all the methods investigated, the Hans and the Hans modified exhibited the highest degree of consistency with other algorithms.

Kappa	Hans	Hans*	Nyman	Choi	Choi*	Natkunam	Tally	Muris	Visco
Hans									
Hans*									
Nyman									
Choi									
Choi*									
Natkunam									
Tally									
Muris									
Visco									

	Poor	Fair	Moderate	Excellent	Very good
K	< 0.20	0.21 - 0.40	0.41 - 0.60	0.61 - 0.80	0.81 - 1.00

Figure 3.4 Pair wise agreement according to kappa statistics.
*modified

3.4.3 Survival Analysis

As the IPI remains the most robust prognostic discriminator in DLBCL, we assessed whether its individual variables or subgroups had value in predicting outcome in the R-CHOP cohort. On univariate analysis, age, stage, performance status and IPI groups were significant in predicting OS, whereas number of extranodal sites, staging, performance status and IPI groups were significant in predicting EFS.

No immunohistochemical marker alone achieved significance for outcome prediction in R-CHOP treated patients. Although patients expressing FOXP1 (60% cut-point) had a lower OS (72% vs 82%, p=0.09), and patients expressing BCL2 had a lower EFS (57% vs 77%, p=0.06), none reached significance on forward stepwise multivariate analysis together with either the IPI factors or the IPI subgroups.

As shown in Table 3.5, none of the algorithms rival the IPI for OS or PFS prediction in R-CHOP treated patients. We also looked at survival differences between patients classified as either GCB or non-GCB by all methods versus the remaining patients with heterogeneous classification. Although OS was similar among groups, 3-year EFS was significantly better for the GCB set (100%) compared to the ABC set (78%) or the

remaining patients (60%)(p=0.004).

As survival was similar across all classifiers, we sought to determine whether outcome stratification could be improved by analysing the expression of additional proteins not included in the original algorithms. If this was demonstrated, it would suggest that IHC classifiers are over-simplified methods for the purpose of outcome stratification. As an example, BCL2 expression was associated with worse PFS in GCB cases only, when incorporated into the Hans (54% vs 88% p=0.006), Hans modified (52% vs 83%, p=0.009), Visco (54 vs 86%, p=0.01), Natkunam (45% vs 78% p=0.02) and Choi modified (53% vs 86% p=0.02) methods. Similarly, expression of the post-GC markers FOXP1 and MUM1 was associated with worse survival in those cases defined as GCB using the Hans and Natkunam algorithms (data not shown).

Table 3.5 Survival analysis according to clinical characteristics, IPI and COO classifiers

Variables	**3-year OS**		**3-year PFS**	
	%	**p-value**	**%**	**p-value**
Sex (male vs female)	74 vs 77	NS	58 vs 72	NS
Age (<60 vs >60 y)	83 vs 68	0.03	66 vs 64	NS
Number extranodal sites (<2 vs ≥2)	79 vs 53	0.09	71 vs 33	0.003
Ann Arbor stage (I-II vs III-IV)	87 vs 63	0.001	83 vs 49	<0.001
ECOG performance status (<2 vs ≥2)	81 vs 48	<0.001	73 vs 23	<0.001
LDH (low vs high)	79 vs 72	NS	73 vs 60	NS
International Prognostic Index (Low/ Low-int/ High-int/High)	90 vs 72 vs 73 vs 45	<0.001	79 vs 80 vs 53 vs 21	<0.001
Algorithms (GCB vs non-GCB/ABC)				
Hans	77 vs 74	NS	66 vs 66	NS
modified Hans	75 vs 75	NS	63 vs 66	NS
Choi	75 vs 75	NS	63 vs 62	NS
modified Choi	74 vs 75	NS	63 vs 62	NS
Muris	79 vs 71	NS	67 vs 64	NS
Nyman	76 vs 75	NS	75 vs 64	NS
Tally	78 vs 73	NS	72 vs 64	NS
Natkunam	71 vs 78	NS	59 vs 74	NS
Visco-Young	75 vs 75	NS	62 vs 68	NS

OS: overall survival; PFS: progression-free survival; ECOG: Eastern Cooperative Oncology Group; LDH: lactate dehydrogenase; IPI: International Prognostic Index; Low-int: low-intermediate; High-int: high-intermediate; NS: not significant

3.5 Discussion

The tumour molecular profiling is for the first time opening prospects for personalized therapy in DLBCL. Clinical trials utilizing GEP for COO allocation are underway and will clarify the utility of targeted therapies in the poor risk ABC subtype. However, the applicability of molecular classification into clinical practice will require a robust, affordable and reproducible methodology for designation of molecular sub-types with clinical and prognostic relevance.

It was hypothesized that IHC approaches would be useful surrogates for classification of DLBCL subsets, would be readily applicable in clinical practice and would be incorporated into diagnostic work-up within haematopathology clinical laboratories. However, based upon previous work and the data presented here, we suggest that much work needs to be done to standardize IHC methods, which currently should be considered unreliable surrogates for molecular classification in DLBCL.

We provide for the first time a systematic analysis of the nine IHC COO classification algorithms in a representative dataset of diagnostic DLBCLs. This study compares single marker and algorithm distribution with previous reports and analyses the survival impact of these data in conjunction with the well-established clinical prognostic score IPI. However, the main objective was to describe how each individual sample scored by all classifiers. This analysis does not imply any comparison across samples from different tumours (with inherent differences in the quality of the material and consequently in the results obtained), but only how each method performs within the same tumour sample to assign the sample to a specific sub-type. Whereas this study would have been enhanced by the availability of GEP as the “gold-standard”, the methodology employed here does not require such a comparison, since we sought to examine the robustness of the more commonly used IHC algorithms and their ability to classify DLBCL compared with each other.

Using the kappa test, a statistical method that takes into account the agreement occurring by chance, we document an extremely low concordance across all classifiers, especially for cases more likely to represent the GCB subtype (only 6%).

Scoring allocation appeared more consistent across all methods for the non-GCB group (but still as low as 25%), confirming the data from previous groups using paired GEP and IHC.[79,307,311]

Classifier distribution was heterogeneous. Allocation to the GCB subtype was less frequent compared to the non-GCB, in line with the findings of others.[310,311] Using the most common Hans classifier, a proportion of molecularly defined GCB cases are likely being allocated to the non-GCB group due to higher expression of MUM1. The Nyman method appears to have too low specificity for detection ABC cases as only 18 patients were considered non-ABC using this method. Similarly, a predominant number of cases were classified as ABC using the Tally method, driven by the high expression of FOXP1 and MUM1 in our series. The Choi classifier is too complex, as both GCET1 and MUM1 had almost no impact on patient allocation. This method is in our experience very similar to that of Visco.

We report low but significant correlations of expression of GCB markers, including CD10, BCL6 and LMO2. BCL6 expression was associated with both GCB and ABC markers, supporting the evidence that BCL6 expression is not entirely restricted to GCB or ABC B-cells as both can harbour genetic aberrations that drive protein overexpression. However BCL6 detection required an amplification step, making it difficult to standardize across laboratories.

We identified a smaller proportion of cases expressing GCET1 than previously reported. In the original study[306] an amplification method was used to enhance GCET1 staining, while others used a different antigen retrieval strategy.[307] This, together with the staining pattern of GCET1, might explain our results. However, as this antigen has been studied by relatively few groups, we propose that more experience has to be gathered on patterns of expression and optimal staining procedures for GCET1 before this is widely incorporated into DLBCL classification. The expression of the post-GC marker MUM1 is incorporated in most of the algorithms, in spite of the fact that IHC assessment for MUM1 is highly sensitive to laboratory variations and inter-interpreter scoring. We noted >30% expression of MUM1 in a higher proportion of patients than previously reported and this cannot be explained

by common reasons such as non-specific cytoplasmic background staining and target cell artefacts.[314] Choi et al. claimed that a higher cut-off level of expression of 80% was required to achieve high specificity for ABC cases. Using this cut-point, 42% of our patients were deemed positive for MUM1 expression, in keeping with Choi's data. This, however, highlights the difficulties of standardizing results based upon arbitrary cut-offs.

We also documented significant differences in the proportion of cases expressing FOXP1 in 80% of malignant B-cells between the Portuguese and the English datasets. Although the biopsies were obtained in the two countries, staining and analysis were performed in the same laboratory. Whether this is a reflection of different fixation and storing methods, the use of arbitrary cut-offs or a true ethnic difference still has to be demonstrated and further population studies will be required to address this question.

As has already been demonstrated in other cohorts[324], we detected associations between clinical factors and protein expression. In both Choi and Visco classifiers, the ABC subset was enriched for older patients. This suggests that clinical prognostic factors might interact with biological predictors such as the COO classification for DLBCL.

Many authors have questioned the prognostic impact of IHC classifiers. Thus, our secondary aim was to analyse survival according to these algorithms. It is recognized that samples collected over a long period of time have differences in quality that might impact adequate interpretation of immunostaining results. This is particularly important when survival analysis is undertaken. As only recently diagnosed R-CHOP treated cases were included, this problem is overcome in our studies. None of the IHC classifiers was able to predict outcome in this series. Others, particularly using the Hans method, previously demonstrated this. We acknowledge that the R-CHOP series analysed (151 cases) is limited and the number of events registered during the study period render it underpowered to detect differences in survival between the two groups. However, we believe this fact supports our hypothesis that IHC classifiers are

inadequate to recognize the molecularly defined DLBCLs. Analysing individual markers not included in the algorithms can improve outcome prediction, as has been demonstrated by us and others[97] using BCL2.

Our study challenges the use of any IHC COO classifier in DLBCL, and it is important to resolve this, as this biological feature provides not only prognostic information but also offers the window for targeted therapies to improve outcome in this disease.

Further studies approaching the COO molecular classification of DLBCL should only be made in the context of prospective clinical trials. A large number of retrospective studies, including ours, have been published that had so far no impact on changing patient's management apart from questioning the utility of IHC for this purpose. There is, together with our data, enough evidence to suggest that at the present time IHC is not a reliable surrogate of the molecular classification of DLBCL. Taking our data into consideration, it isn't rational to recommend any specific classifier for further use in the clinical practice. Despite the advances in reagents and automation, the use of IHC is hampered by variable consistency, reduced reproducibility, and quality assurance disparities, resulting in poor concordance, validation and verification. Once a concerted multi-institutional effort towards a precise standardization of every single procedural step known to impact on IHC results is made, this technology might reunite potential for use in surrogate studies such as the COO. This would be ideal, as IHC is an inexpensive technology that can be entirely manually performed and that requires no batching of samples. The turnaround of the results is also extremely acceptable.

Meanwhile simpler and more affordable molecular technologies are being explored for the molecular classification in DLBCL. FFPE samples from patients enrolled in the REMoDLB trial are being studied using the Illumina whole-genome DASL technology, which enables performing GEP in low yield, partially degraded RNA samples. Validation of GEP in FFPE tumour biopsies is in fact crucial for the clinical setting and has been previously done. More recently the Nanostring nCounter technology has been applied in FFPE samples from the LLMPP.[79] The authors report a 20-gene GEP-based assay (Lymph2Cx) developed with the objective of accurately assigning samples

to the same COO group as with Affymetrix-based original method. The new assay is robust, accurate and results can be retrieved in less than 36 hours. Importantly, both the DASL and the nCounter assays require low nucleic acid input and guarantee highly reproducible results for a much affordable cost than the original array technologies. However it needs to be taken into account that such assays demand sample batching, which might not be feasible in the clinical setting. Profiling individual patients might hence be more costly than suggested.

Having in mind that the protocols recommended by the companies selling molecular biology solutions are very detailed and precise, it is not surprising that independent groups obtain similar end products. The problem lies on the analysis. Scott et al. describe in the supplementary data one phenomenon that illustrates this issue. As is well known, a number of samples, generally described as "unclassified", fail to be allocated to either the GCB or the ABC groups. The authors report that around 15% of samples "migrated" in and out of the "unclassified" group between the Lymph2Cx and "gold standard" GEP method. This significantly decreases the concordance between the 2 methods to around 80%, similar to the concordance reported with IHC algorithms. This phenomenon highlights for misclassification even with GEP and has fundamental consequences in the context of clinical trials testing agents with selective activity to a given molecular subtype.

In conclusion, there is a lack of standardized methodology for array analysis, which can lead to variable results both at the inter- and intra-laboratory level. This issue, which may impact on results, but most importantly, patient care, is generally unreported and should be a matter of future debate.

Chapter 4 Revisiting the immune microenvironment of DLBCL using a tissue microarray and immunohistochemistry: robust semi-automated analysis reveals CD3 and FoxP3 as predictors of response to R-CHOP

4.1 Introduction

As highlighted in previous chapters, rituximab primary refractory DLBCL has emerged as a particularly difficult group to cure. Furthermore, approximately half of the patients will eventually relapse after R-CHOP. As such, there is an urgent need for novel therapeutic approaches in these patients.

The role of the microenvironment in DLBCL biology and outcome gained relevance when independent GEP studies defined distinct biological traits that were driven by the non-malignant cells in the tumours.[12,34,74,76] Data supervised analyses delineated prognostic signatures that are enriched in genes encoding for ECM components, T-cell and macrophage markers, and angiogenic mediators. Importantly, all studies suggest that a high expression of genes characteristic from the non-malignant cells in the microenvironment confers a better outcome in DLBCL. Functional validation of recurrent mutations compellingly relates DLBCL biology with inflammation and immune surveillance.[325] These studies identify the lymphoma microenvironment and host inflammatory response as defining features in DLBCL.

However, GEP data requires further validation and needs to be made simpler in order to be useful for clinical trial design and for clinical practice. Many methods can be used to validate GEP information. Most authors have focused on the use of IHC to enumerate and functionally characterize the microenvironment in DLBCL and other lymphomas. IHC can be extended to clinical practice, which makes it highly attractive as a diagnostic and prognostic tool. Nevertheless the results published regarding IHC analysis of the immune microenvironment in DLBCL are often contradictory.[278,279,282,283,326,327] The use of inconsistent methodology likely explains

these results. Moreover, it is known that it is difficult to count reproducibly cells across large tumour areas manually. Categorization of the density of cell infiltration is used to overcome this problem, but results might be misleading and there is a lack of validation of cutpoints.

4.2 Objectives

The main scope of our study was to revisit the immune microenvironment of diagnostic samples of 309 patients with DLBCL by two different methods of semi-automated image analysis. The semi-automated analyses have the advantage of being capable of analysing large tumour areas, making them ideal for prognostic exploration of IHC studies. We expected to detect a high degree of inconsistency between the results of the two systems, similar to what is reported when manual and automated analysis is compared. Finally, we aimed at describing the prognostic role of different immune biomarkers in 161 R-CHOP treated cases. Total T-lymphocytes and their functional subsets and macrophages were studied.

4.3 Methods

4.3.1 Patient characteristics

Patient selection, clinical characteristics and outcome of the patients included in this study were already described. As in the previous chapter, cutpoint and outcome analysis was based on the R-CHOP dataset.

4.3.2 Tissue Microarray and Immunohistochemistry

Ethical approval for this study was obtained from Local Regional Ethics Boards. Sample collection followed informed consent in accordance to the declaration of Helsinki. Triplicate or duplicate 1-1.5 mm^2 cores were taken from representative tumour regions identified on H&E stained sections and confirmed by CD20 staining.

Dewaxed paraffin sections were rehydrated, blocked and subjected to heat antigen retrieval. CD3, CD4, CD8, CD68, TIA1 and FOXP3 were stained using a peroxidase-labelled system (Super-Sensitive Polymer-HRP IHC Detection System, BioGenex) on the Dako-autostainer. Tonsil sections were simultaneously stained for all antibodies. Finally, sections were counterstained with haematoxylin. Detailed information on the staining protocol is provided in the Materials and Methods section. Primary antibodies and conditions of use are provided in Table 4.1. Heat induced antigen retrieval was used for all antibodies.

Table 4.1 Primary antibodies and conditions of use

Antibody	**Clone**	**Species**	**Source**	**Dilution**
CD20	L26	mouse	Dako	1/2000
CD3	SP7	rabbit	Lab Vision, Thermo Scientific	1/500
CD4	4B12	mouse	Novocastra, Leica Biosystems	1/500
CD8	C8/144B	mouse	Dako	2,5/1000
CD68	KP1	mouse	Dako	1/8000
FOXP3	263A/E7	mouse	Abcam	1/100
TIA1	2G9A10F5	mouse	Immunotech, Beckman Coulter	1/400

4.3.3. Semi-automated Image analysis

4.3.3.1 Ariol SL-50 visual analysis software

Slides were scanned using an Olympus BX61 microscope. Representative regions were selected manually using the tools available on the Ariol Software. Fibrotic and necrotic areas were excluded so that only representative tumour areas were brought forward for analysis (Figure 4.1). Whole cores with less than 50% of tumour representation were also completely excluded from analysis.

Using the methodology detailed in Chapter 2 and illustrated in Figure 2.1, the number of positive cells as well as the area of DAB stained per area of lymphoma tissue was estimated. The values obtained were corrected to a 1mm2 area and a mean for each patient was calculated.

4.3.3.2 Pannoramic Viewer System

Slides were scanned using the Pannoramic 250 Flash II scanner (3DHISTECH, Hungary). Meticulous marking of representative tumour areas was done on the Pannoramic Viewer computer interface and individual areas quantified. Using the DensitoQuant module the number of DAB stained pixels per area selected was quantified. This module distributes pixels to 3 grades of positive classes by their RGB values. We used only the top red and orange levels for identification of stained areas. After adjusting the brown tolerance and the score levels an optimal script was saved for each antibody. A preliminary analysis was undertaken to confirm the capacity of the script to identify DAB stained cells in independent cores. This optimal script was then applied for analysis in all areas.

4.3.4. Cutpoint Determination

Cutpoint discrimination was assessed using the recursive splitting algorithm in the rpart package (http://cran.r-project.org/web/packages/rpart/index.html) within the R statistical software.[328]

To achieve a simple subdivision into two classes, the algorithm has been forced to

obtain a single split for each considered measurement, with a splitting criterion based on the Gini index. For robustness, only splits yielding ≥30 individuals in the smallest group by the recursive algorithm were considered. Two to ten cross validations have been performed for each measurement to check for consistency of each cutpoint, showing no modifications in the outcomes.

A more comprehensive explanation of this method is provided in Chapter 2. Dr Emanuele Mazolla and Dr Donna Neuberg performed this analysis.

4.3.5 Statistical Analysis

Differences between patient's groups were tested using χ^2 or Fisher's exact test, when appropriate. Association between pairs of biomarkers has been tested using the Fisher's exact test. Pearson's correlation was calculated for all IHC variables and clinical parameters.

The outcomes measured were: OS, defined as the time from diagnosis to death from any cause, with surviving patients censored at last follow-up; and PFS, defined as time from diagnosis to failure of treatment (including not achieving CR or relapse after CR) or death from any cause. Median follow-up was calculated for patients alive at last follow-up. For every quantified measurement survival has been estimated using Kaplan-Meier estimators, and differences between groups of the same measurement have been assessed with the Log-rank test. To accommodate for the optimization method within the splitting algorithm we considered as significant only Log-rank p-values <0.01.

Multivariate analysis was performed using a Cox proportional hazards model (stepwise backward and forward methods) including IHC parameters together with the clinical factors included in the IPI with prognostic significance on univariate analysis. Statistical significance was set as $p<0.05$. Statistical analysis was performed using SPSS version 19.0 (SPSS) and Prism version 5.03 (GraphPad Software) and R version 3.0.2.

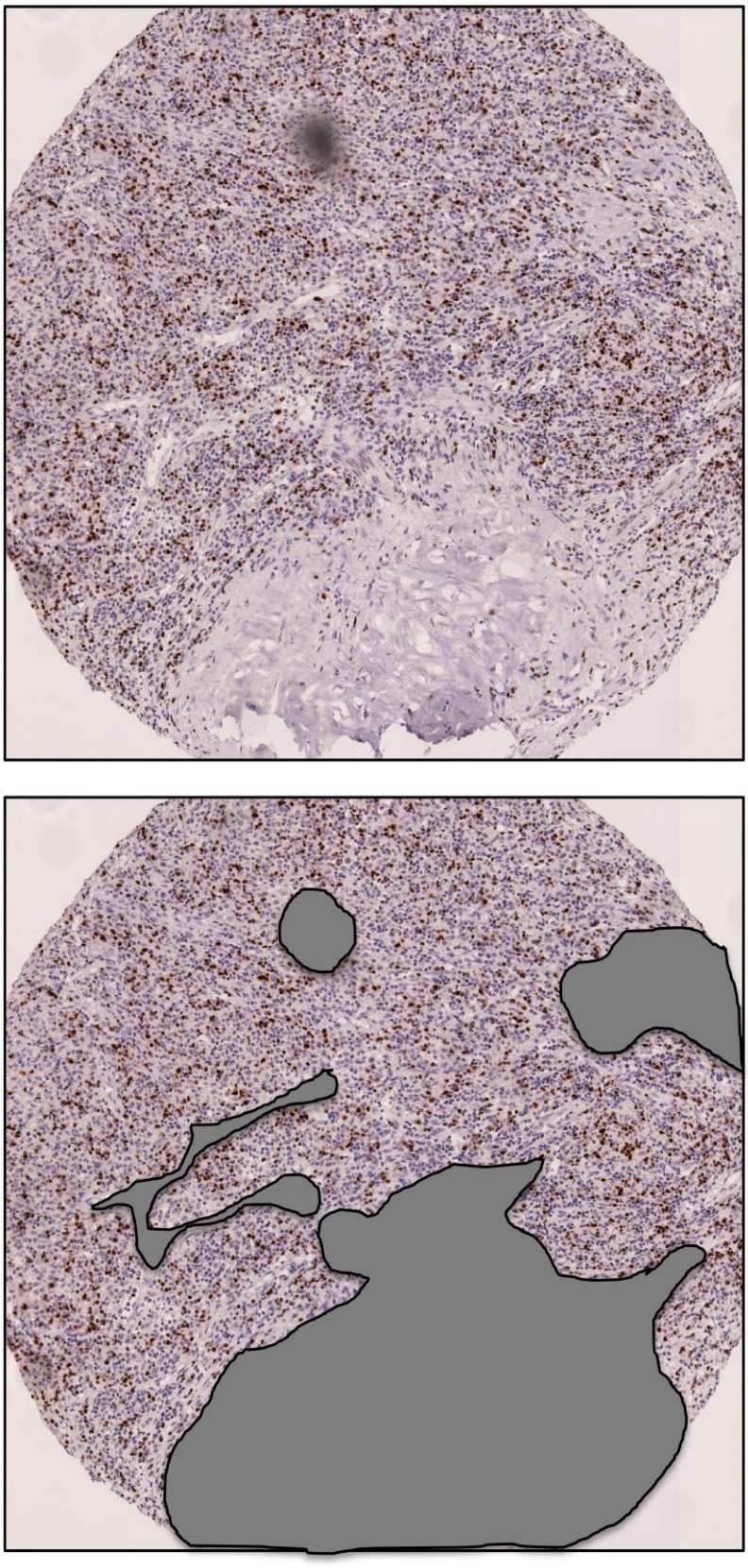

Figure 4.1 Selection of tumour representative regions.
Fibrotic and necrotic areas were excluded using the tools available on the Ariol system.

4.4 Results

4.4.1 Heterogeneous density of immune cells in the microenvironment of DLBCL

With minor exceptions, cells staining for the biomarkers investigated were homogeneously distributed across tumour sections. This was anticipated given the diffuse histological pattern of DLBCL and validates interpretation in partially represented cores where viable tissue is properly represented.

As expected, patients with DLBCL had heterogeneous infiltration of total T-cells, their subsets and macrophages at diagnosis (Table 4.2 and Figure 4.2).

CD3+ total T-cells were the most abundant cell type studied. We identified a balanced proportion of CD4 and CD8 towards a predominance of CD8+ T-cells when all cases were analysed simultaneously (Figure 4.2). The median ratio of CD4/CD8 for all measures taken was between 0.36 - 0.68. Using the antibody stated in Table 4.1, CD4 expression was typically dim and more heterogeneous. Moreover the expression of CD3 and CD8 correlated better than the expression of CD3 and CD4 (Pearson's-r for +cells/area 0.73 vs. 0.50, respectively). Although CD4 heterogeneity might represent biological inter-patient differences, a failure of IHC methods to properly identify CD4+ cells has to be considered as a possibility, specially when using a single antibody.

FoxP3 is a nuclear CD4+ T-cell transcription factor expressed by regulatory and activated T-cells. Its expression is clear and discrete and therefore training and analysis was facilitated. As can be appreciated in Figure 4.2, FoxP3 expression shows a skewed distribution. The number of cells expressing FoxP3 varied between samples (median 228.6, range 0 - 3197), but less so than other markers (e.g. TIA1, median 1648, range 31.7 - 5859).

TIA1 is a cytoplasmic marker expressed by NK-cells and by cytotoxic T-cells independently of their activation status. CD56 staining was done to explore the extent of infiltration by NK-cells and was almost absent in our series (data not

shown). The absence of NK-cells suggests that TIA1 expression is specific to T-cells. However, the mean and standard deviation distribution suggests it is more abundant than CD8. This might be related to the punctate cytoplasmic expression of TIA1.

CD68 is a cytoplasmic lysosomal protein expressed in all tissue macrophages.[329] Large and interdigitant cells such as macrophages are difficult to quantify. Image analysis training was challenging and focused on capturing the areas of dark brown staining present in the cell bodies. CD68 expression was particularly heterogeneous in this cohort.

Table 4.2 Descriptive statistics showing heterogeneity of expression of biomarkers

	Ariol			**Pannoramic Viewer**	
	Patients assessable, Number (%)	Cells/mm^2, Median (Range)	% Stained area, Median (Range)	Patients assessable, Number (%)	Nr pixels/mm^2, Median (Range)
CD3	258 (84)	2603 (50-7480)	11.8 (0-56.1)	243 (79)	$1.2x10^6$ ($1.5x10^4$-$2.4x10^{12}$)
CD4	206 (67)	438.4 (0-6730)	5.23 (0-61.8)	231 (75)	$2.3x10^5$ (147.4-$4.5x10^6$)
CD8	251 (81)	857.5 (25-3732)	6.44 (0-37.6)	252 (82)	$5x10^5$ (6566-6.6 $x10^7$)
FOXP3	253 (82)	228.6 (0-3197)	0.68 (0-13.6)	246 (80)	3.7 $x10^4$ (25.2-1.3 $x10^6$)
TIA1	249 (81)	1648 (32-5859)	9.6 (0-80.5)	250 (81)	$1.6x10^5$ (1079-2.5 $x10^6$)
CD68	252 (82)	1380 (85-6320)	7.1 (0-44.4)	249 (81)	$6.5x10^5$ (1.5x104-$6.1x10^6$)

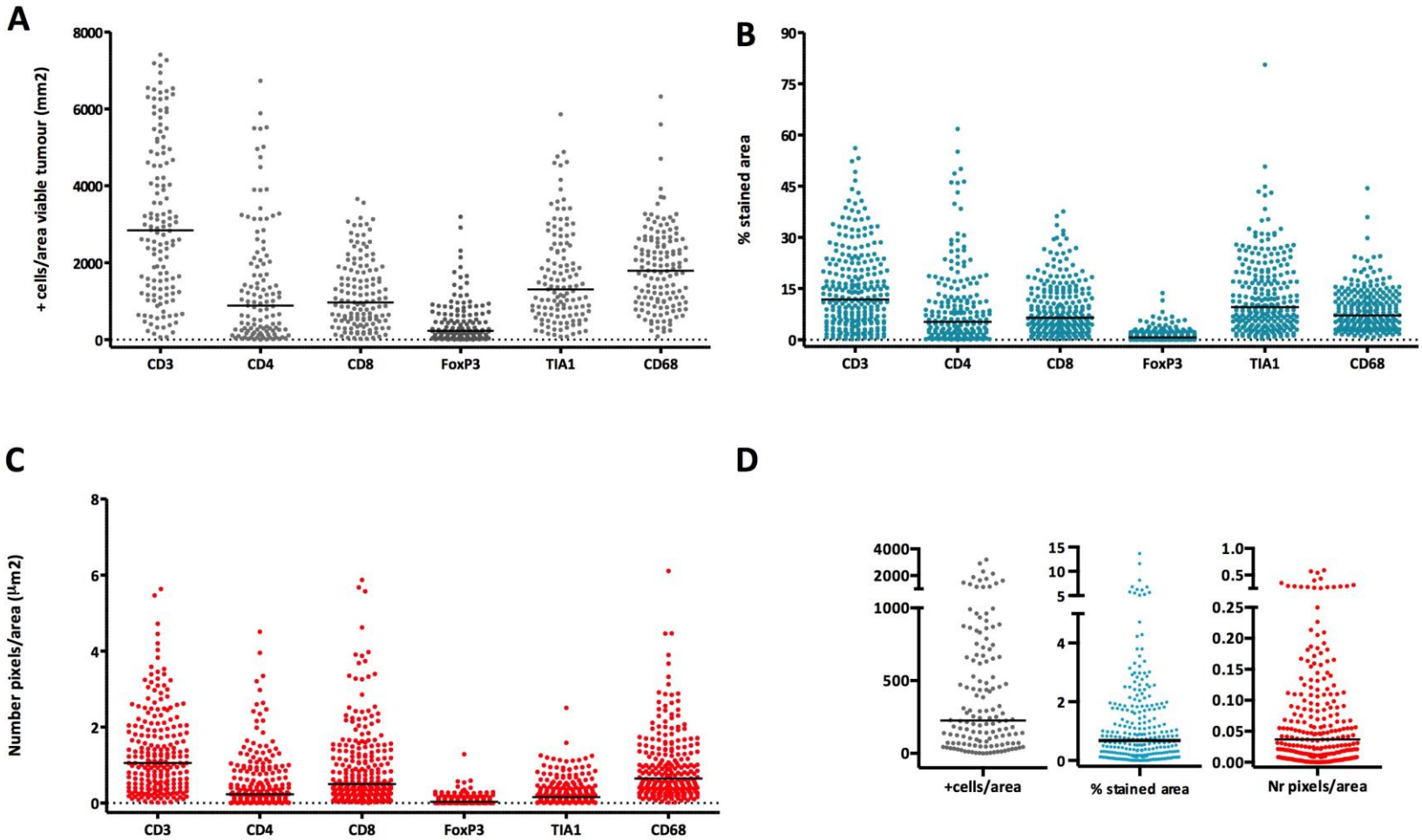

Figure 4.2 Biomarker distribution.

A: Distribution of cases according to number of positive cells per area of viable tumour (mm^2) as measured by the Ariol system. B: Distribution of cases according to % stained area as measured by the Ariol system. C: Distribution of cases according to number of stained pixels per area (μm^2) as measured by the Pannoramic Viewer system. D: Distribution of cases for FoxP3 according to all parameters analysed. Bars represent the median values.

4.4.2 Excellent concordance between two systems of semi-automated image analysis for the characterization of the microenvironment

Multiple studies have shown that manual quantification of abundantly expressed IHC markers is inconsistent. To this end a number of semi-automated systems for IHC analysis have been launched that allow analysing large stained areas and increase results consistency. However each one has its own features and scripts for analysis and it is unknown how robust results are when obtained from different computerized systems.

We therefore expected to detect considerable inconsistency between two semi-automated systems that use different methods of counting. This would have implications to the future of IHC studies utilising automated analysis. To test this hypothesis we examined the immune microenvironment in DLBCL by the Ariol and the Pannoramic Viewer systems.

Firstly we compared two different measures obtained from the Ariol system: a) number of positive stained cells/ area of viable tumour and b) % stained area for each marker. Against our own expectations, we detected an excellent correlation between the two Ariol measures for each of the markers studied (Pearson's r-values between 0.92 and 0.98, $p < 0.0001$) (Figures 4.3). Importantly the correlation was high even for the cytoplasmic markers TIA1 and CD68, for which training was more challenging and low interobserver agreements have been reported. These results suggest that, by applying our methodology, any of the two measures retrieved by the Ariol system can be used for future IHC studies.

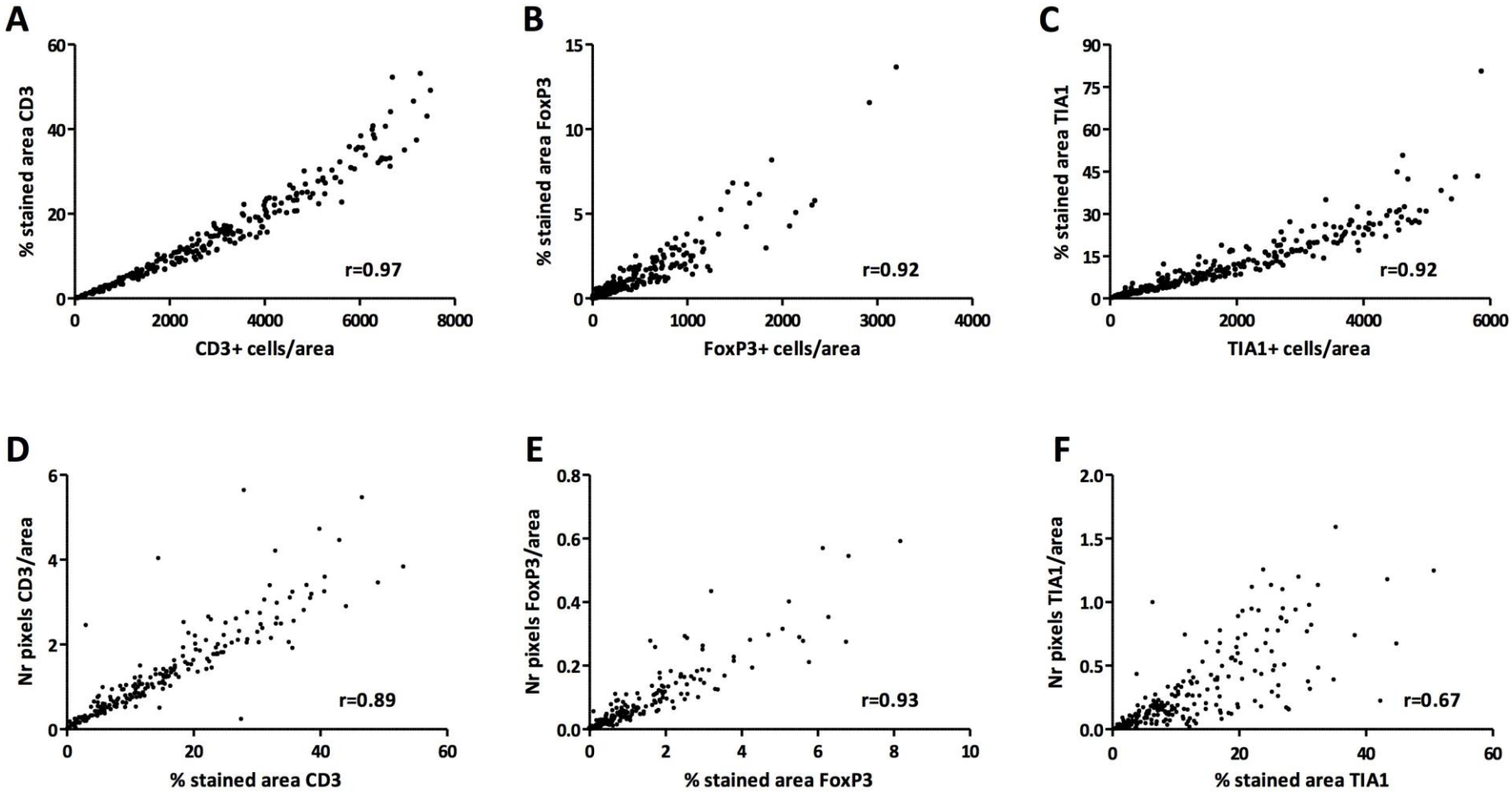

Figure 4.3 Expression correlation for CD3, FoxP3 and TIA1.

A-C: Correlation plots between the number of positive cells/area of viable tumour and % stained area for each case according to the Ariol System. D-F: Correlation plots between the % stained area as quantified by the Ariol System and the number of positive pixels/area according to the Pannoramic Viewer.

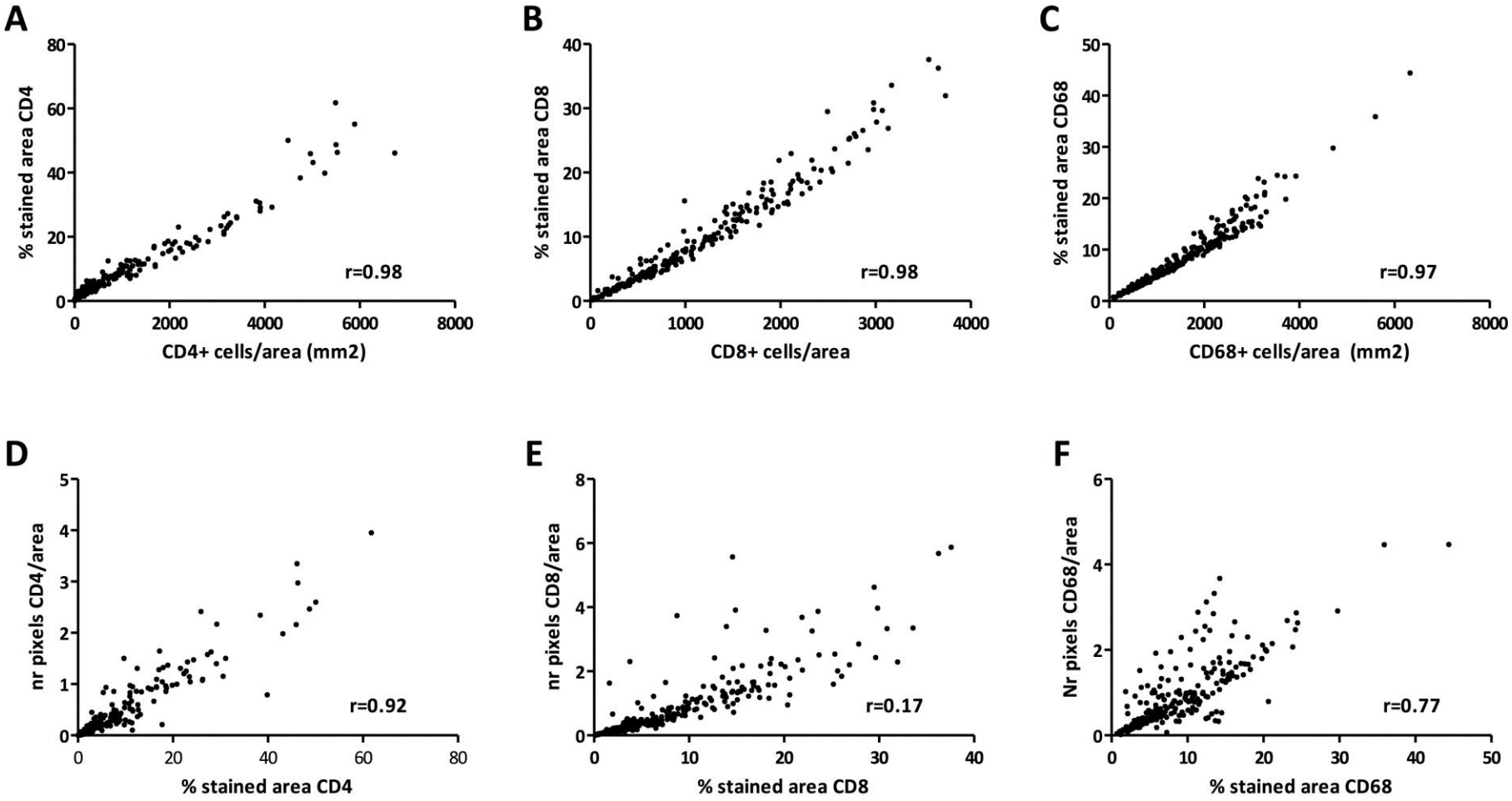

Figure 4.4 Expression correlation for CD4, CD8 and CD68.

A-C: Correlation plots between the number of positive cells/area of viable tumour and % stained area for each case according to the Ariol System. D-F: Correlation plots between the % stained area as quantified by the Ariol System and the number of positive pixels/area according to the Pannoramic Viewer.

We then compared the Ariol results with those obtained from the Pannoramic Viewer System, where quantification of the DAB stained pixels/area was performed. However, with the Pannoramic Viewer the tumour density is not taken into account and a precise selection of representative tumour areas is more demanding. Even so, we surprisingly detected significant correlations between the data retrieved from the two systems (Figure 4.3 and 4.4, D-F). Correlations were excellent for T-cells markers (Pearson's r-values between 0.89 and 0.92) and more modest for cytoplasmic proteins (CD68, r - 0.77; TIA1, r - 0.67). This indicates that, contrary to manual analysis of IHC, semi-automated systems are robust and preferable for future IHC studies.

4.4.3 CD3 and FoxP3 are potential predictors of response to R-CHOP

Our secondary aim was to investigate the prognostic impact of the immune biomarkers in DLBCL. For this purpose we specifically selected the 161 R-CHOP treated dataset. Table 4.3 details survival, hazard ratios, confidence intervals and p-values for all significant variables on univariate and multivariate analysis.

We did not detect any interactions between the clinical variables detailed in Table 3.1 and the biomarkers studied (data not shown). We applied strict quality criteria to image analysis to ensure robustness and reproducibility of the techniques. This meant that images were rejected for a number of markers that did not meet these criteria. The most common reason for exclusion was insufficient core size or degradation.

The following clinical variables were predictive of worse OS by univariate analysis (Table 4.3): age >60 years (P=0.02), stage III-IV (P=0.001) and ECOG performance status ≥2 (P=0.0001). Patients with high IPI had a lower probability of survival (P=0.001). Achieving a complete response to R-CHOP was the strongest prognostic variable on univariate analysis of this cohort (P<0.0001). Using the rpart package within the R software and a recursive splitting algorithm according to the criteria

described previously, no cutpoint discrimination could be established for the studied biomarkers that helped discriminating patients with different OS.

Regarding PFS, patients with stage III-IV (P=0.001), ECOG performance status ≥2 (P<0.0001), ≥2 extranodal areas (P=0.001) and IPI ≥3 (P= 0.001) had a lower probability of survival (Table 4.3). Patients treated at Bart's also had a lower PFS (three-year PFS of 52.4% vs 68.3%, P=0.03) compared to the ones treated at IPO. Using the rpart package we defined single cutpoints for CD3 and FoxP3 that segregated patients with different cell density and PFS. Although the number of pixels/area for CD3 presents with a Log-rank p-value slightly above the established 0.01 limit for significance, the robustness of the split is confirmed by the corresponding 2x2 table in Figure 4.4. For both biomarkers, patients with a higher cell density had a higher PFS after R-CHOP (Figures 4.5 and 4.6).

Using $CD3^+$ cells/area, the median PFS was 78.7 months for patients with high $CD3^+$ cell density (>1897 cells/mm^2, 86 patients) compared to 28 months for the remaining (p=0.005) (Figure 4.5C). When % stained area of CD3 was considered, 81 and 54 patients had high and low density. For these, median PFS was 82.1 vs 29.1 months, p=0.001 (Figure 4.5D). Finally, 64 and 68 patients were included in the high and low cohorts according to the Pannoramic Viewer analysis. Again, patients with a high CD3 infiltration exhibited a better median PFS (78.7 vs 30.3, p= 0.01) (Figure 4.5E).

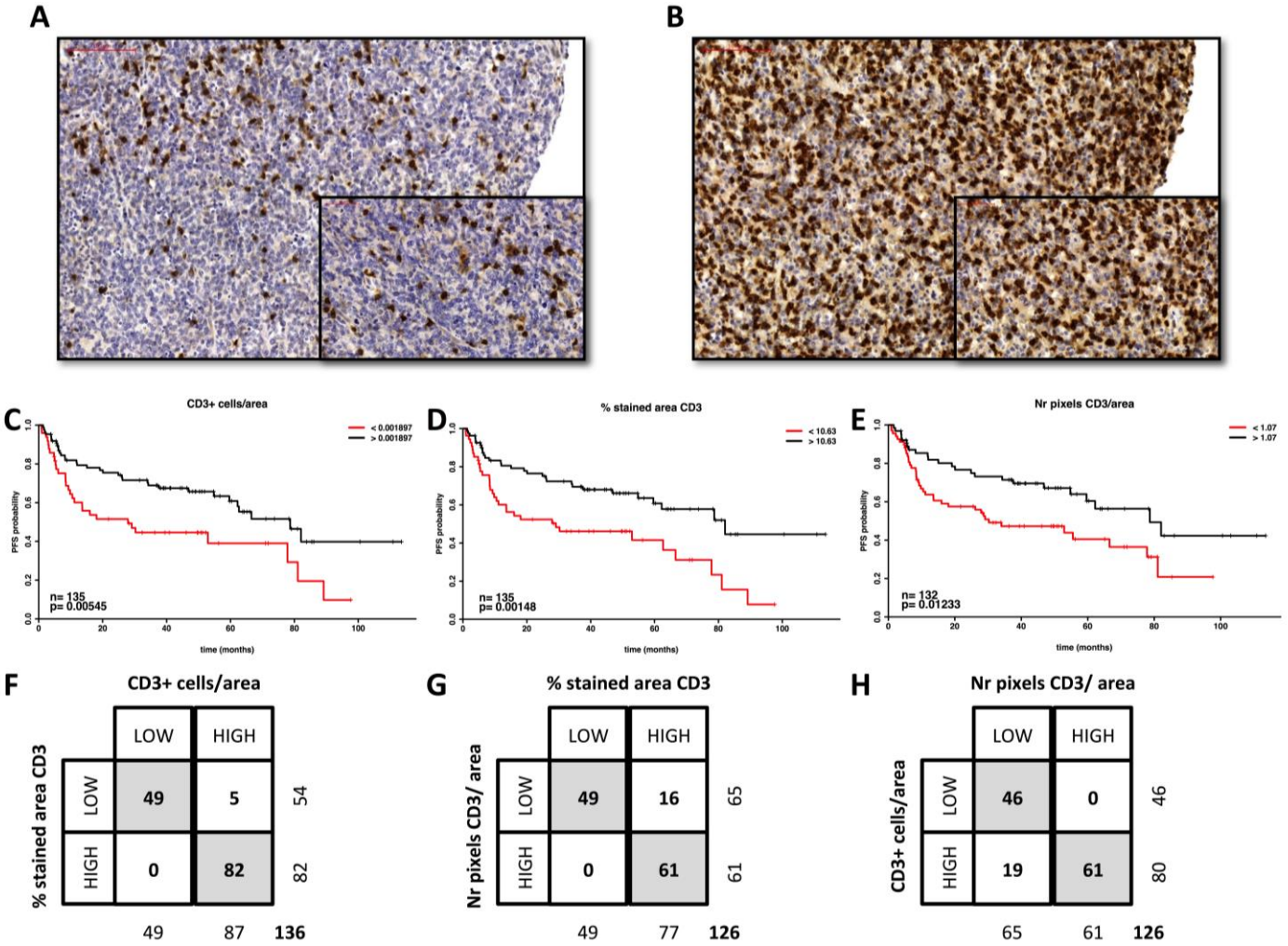

Figure 4.5 CD3 Expression and Outcome.

A&B: Examples of low (A) and high (B) expression of CD3. Magnification x20 and x40. CD3 expression is shown by HRP-DAB immunostaining. C-E: Kaplan Meier PFS analysis of patients based on low or high CD3 expressing cells/μm^2 (C), % stained area of CD3 (D) and number of stained pixels of CD3/area (E). F-H: Contingency tables representing the distribution of patients into low and high subgroups for each of the analyses.

We were also able to establish a cutpoint for each of the FoxP3 analyses performed using rpart (Figure 4.6). Patients allocated to the high $FoxP3^+$ cohorts had significantly higher CD3 expression compared to the remainder (p=0.0001). 78/137 patients were allocated to the high $FoxP3^+$ cell/area of viable tumour. High FoxP3 nuclei density was associated with a better median PFS (89.1 vs 26.2 months, p=0.0002) (Figure 4.6C). Similarly, patients in both the high FoxP3 % stained area (77/137 patients) and the high pixel number/area (72/132 patients) cohorts had a prolonged median PFS compared to patients in the low expression cohorts (undefined vs 27.9, and 82.1 vs 27.1 months, respectively) (Figures 4.6D and 4.6E).

Finally a Cox regression analysis was used to model PFS using the R-CHOP dataset (Table 4.3). All clinical variables significantly associated with PFS on univariate analysis together with the categorical data for CD3 and FoxP3 were included in the model. Using both backward and forward stepwise methods, the variables retaining independence for PFS prediction were stage III-IV and CD3 density quantified as % stained area of viable tumour according to Ariol. FoxP3 did not remain as prognostically significant on multivariate analysis.

Since we have used different parameters (+ cells/area of viable tumour, % stained area and number of brown pixels/area) to quantify each of the biomarkers, rpart retrieved different discriminatory cutpoints for each. Presuming that cohort discrimination is truly based on a biological impact of the immune cell infiltration and not a reflection of a methodology induced bias, we hypothesized that the agreement in allocation of each individual patient to high and low subgroups for each CD3, FoxP3 or TIA1 analyses would be high. To test this hypothesis, we plotted consensus matrices and explored the degree of allocation agreement for each subgroup.

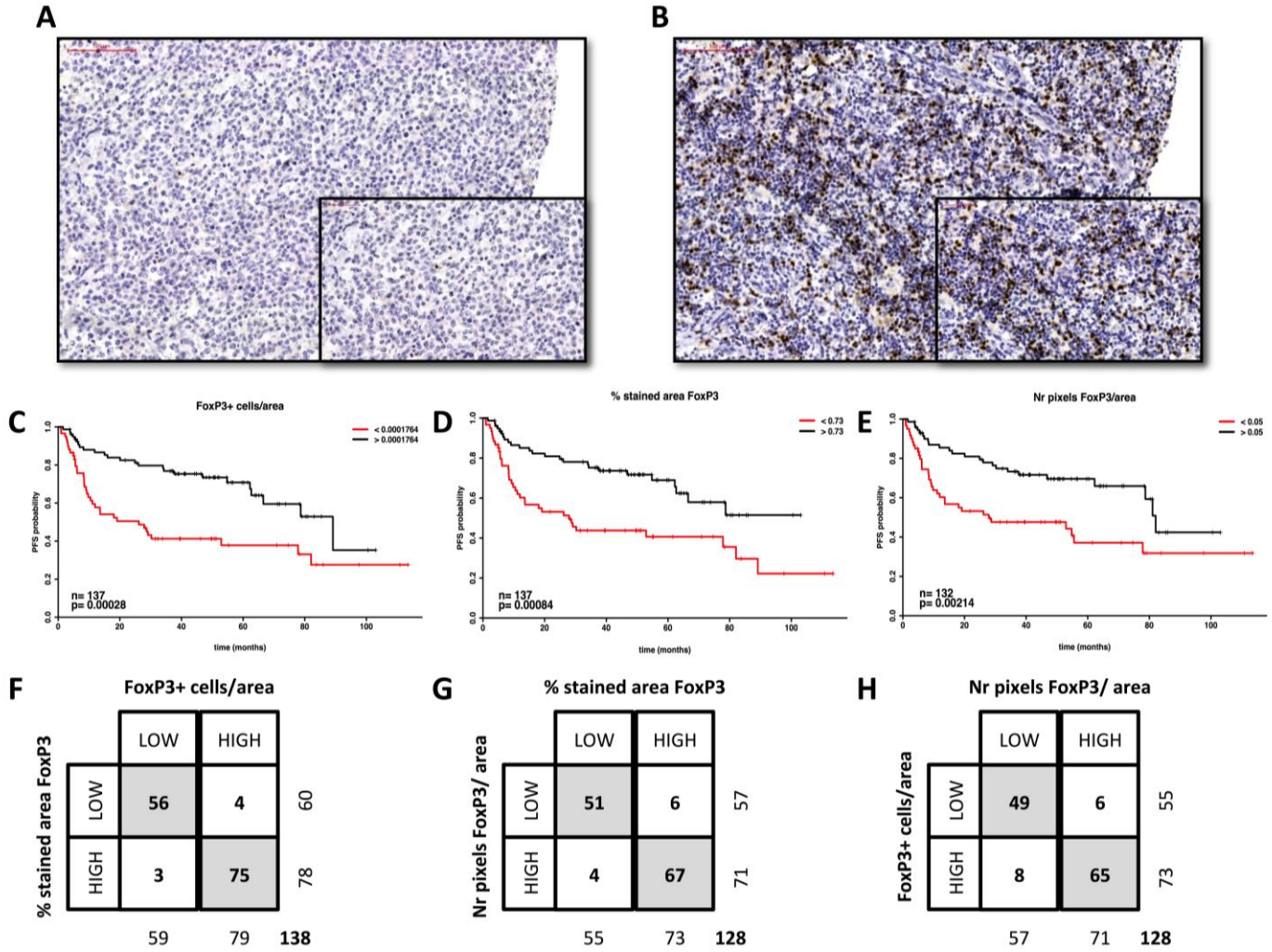

Figure 4.6 FoxP3 Expression and Outcome.

A&B: Examples of low (A) and high (B) expression of FoxP3. Magnification x20 and x40. C-E: Kaplan Meier PFS analysis of patients based on low or high FoxP3 expressing cells/μm^2 (C), % stained area of FoxP3 (D) and number of stained pixels of FoxP3/area (E). F-H: Contingency tables representing the distribution of patients into low and high subgroups for each of the analyses.

As can be appreciated in Figures 4.4 F-H and 4.5 F-H, allocation consistency is extremely good for CD3 and FoxP3 (Fisher's exact test p-value <0.0001 for each presented table), even when the Ariol and Pannoramic Viewer results were compared. This suggests these biomarkers as predictors of outcome after R-CHOP and validates recursive partitioning as an adequate method for cutpoint selection in continuous data with unknown demographic distribution.

Table 4.3 Univariate and multivariate survival analysis

UNIVARIATE ANALYSIS			
Overall Survival			
Variable	Median survival (m)	Hazard ratio (95% CI)	*P*
Age >60 years	92.1 vs undefined	1.96 (1.13, 3.34)	0.02
Stage III-IV	82.9 vs undefined	3.33 (1.51, 4.54)	0.0005
ECOG PS ≥2	30 vs undefined	3.84 (1.96, 9.09)	0.001
IPI ≥3	81 vs undefined	2.63 (1.44, 4.76)	0.001
Not achieving CR	17.2 vs undefined	6.66 (3.44, 14.3)	<0.0001
Progression-free survival			
Stage III-IV	34.2 vs 89.1	2.27 (1.42, 3.57)	0.0005
ECOG PS ≥2	11.1 vs 78.7	4.16 (1.85, 9.09)	0.0005
≥2 extranodal areas	10.5 vs 78.7	2.78 (1.43, 5.55)	0.003
Centre (Bart's vs IPO)	81 vs 52.9	1.6 (1.03, 2.59)	0.03
IPI ≥3	24.9 vs 82.1	2.43 (1.47, 4.0)	0.0004
CD3 +cells/area	78.7 vs 27.9	0.50 (0.31, 0.83)	0.005
CD3 %stained area	82.1 vs 29.1	0.46 (0.28, 0.75)	0.001
CD3 Nr pixels/area	78.7 vs 30.3	0.52 (0.31, 0.88)	0.01
FoxP3 +cells/area	89.1 vs 26.2	0.40 (0.24, 0.67)	0.0003
FoxP3 %stained area	undefined vs 27.9	0.43 (0.26, 0.71)	0.0008
FoxP3 Nr pixels/area	82.1 vs 27.9	0.45 (0.27, 0.76)	0.002
MULTIVARIATE ANALYSIS FOR PFS			
Variable		Hazard ratio (95% CI)	P
Stage III-IV		3.4 (1.72, 6.66)	< 0.0001
High % stained area CD3		0.42 (0.24, 0.73)	0.003

For IHC data, all hazard ratios with corresponding 95% confidence interval were calculated considering the HIGH versus LOW density for each variable. Results from backward stepwise (likelihood ratio) based on 120 cases. m: months; CI: confidence interval; Nr: number

4.5 Discussion

Developing robust prognostic biomarkers that are able to discriminate R-CHOP refractory patients is a research priority in DLBCL. As can be appreciated in our data and as expected, response to R-CHOP appears as a robust prognostic indicator of OS or PFS.

GEP studies suggested that the microenvironment has the potential for providing some of such prognostic biomarkers. However, a number of obstacles need to be acknowledged: 1) biomarkers need to be easily studied in any laboratory; 2) validation processes need to be robust and done under multi-institutional projects and clinical trials; 3) microarray technology is still reserved to the research setting and lacks well standardized procedural protocols and analysis; 4) only a few markers selected for the GEP studies mentioned have been evaluated by independent groups using different methodologies; 5) functional studies, the only ones able to provide definitive biological explanations towards the prognostic impact of specific molecules, are difficult to perform in the context of the lymphoma microenvironment.

IHC can be extended to the clinical practice, which makes it highly attractive as a diagnostic and prognostic tool. Nevertheless the results published regarding IHC analysis of the immune microenvironment in DLBCL and other lymphomas are contradictory. This can be explained by several reasons, the most important being the study of patient cohorts which are not representative of the disease the way is currently managed and, most importantly, by the use of inconsistent methodology. It is well recognized by expert pathologists that manual scoring of IHC staining is difficult to standardize, particularly when large tissue areas are analysed. Semi-automated image analysis is available, is ideal for scoring vast areas more representative of the microenvironment, and could help in improving scoring reproducibility. However each one has its own features and scripts for analysis and it is unknown how robust results are when obtained from different systems.

This study compares two different methods of semi-automated analysis for IHC staining using as a model the microenvironment of DLBCL and examining a large number of patients and area of diagnostic FFPE tissue per patient than any published work to date. Although our hypothesis was that using different systems, and hence different methodologies of analysis would decrease reproducibility and demonstrate that such approach is unreliable, if we could prove the opposite it would support semi-automated IHC as the way forward for assessing cell infiltration in large tissue areas using TMAs.

We competed two image analysis systems developed by independent companies (Ariol and Pannoramic Viewer), three image analysis methods (absolute numbers, percentage area and numbers of stained pixels), and as a secondary aim examined the prognostic impact of individual elements of the immune microenvironment in a cohort representative of the current clinical scenario in DLBCL. To our knowledge this approach has not been previously undertaken for IHC studies in lymphoma.

We demonstrate that the computerized results are highly reproducible, even when comparing different variables examined for each biomarker, such as cell density and % of area stained. Correlation data for the two Ariol measurements was extremely consistent, even for cytoplasmic proteins such as TIA1 or CD68. This was a surprising result given the difficulties in training the Ariol system to enumerate single cytotoxic T-cells or macrophages. Moreover, inter-observer variability of manual counting for macrophages is known to be very high. Even more surprising was the finding of highly acceptable reproducibility between the data retrieved from the Ariol and the Pannoramic Viewer. Correlations were excellent for T-cell surface markers, but more modest for cytoplasmic proteins.

Our data is in agreement with a recent validation study promoted by the LLBC.[329] This study, conducted by highly experienced haematopathologists, reports only low to moderate agreement in manual scoring of T-cell and macrophage markers when 4-5 scoring categories are used. However, comparison of semi-automated analyses set up

by two different operators in two different instruments using the same methodology (Ariol) was highly reproducible for T-cells markers. This is an important finding suggesting that operator-induced bias is not as relevant as has been thought previously and should not prevent researchers from using this methodology. The current study adds upon this by examining a large population of 309 patients and by comparing the computerized quantifications for a larger panel of markers, particularly including cytoplasmic proteins such as CD68. Moreover, we performed comparisons across the results retrieved from two semi-automated systems developed by different companies and three different assays for each staining.

Our own and the LLBC data indicates that automated systems of IHC analysis add the required robustness to IHC prognostic studies in an operator-independent manner and should be used in the future instead of manual analysis. Comparing manual and automated results was not the scope of this project and studies such as those of the LLBC support that manual scoring is highly variable and hence inadequate for outcome prediction studies. This is being explored further in Follicular Lymphoma by the LLBC.

Finally we explored the outcome potential of microenvironment biomarkers as assessed using the semi-automated systems in a representative dataset of 161 uniformly R-CHOP treated DLBCL patients. Whereas the clinical variables included in the IPI and achieving a complete remission after R-CHOP were predictive of OS, none of the biomarkers studied were, potentially reflecting the impact of salvage therapy. However we were able to validate CD3 and FoxP3 as predictors of PFS. For both biomarkers, patients with higher biomarker density had a lower risk of relapse after R-CHOP. Although studies highlighting CD3 as a potential marker in lymphoma are limited,[330] previous IHC analyses predominantly showed that a high infiltration of FoxP3 improves patient survival.[331] While functional studies point towards a negative prognostic impact for regulatory T-cells in lymphoma,[332-334] it is possible that regulatory T-cells are directly suppressing the malignant B-cells or are counteracting tumour supporting T-cells. Notably, we saw no evidence that CD68 expression by any

method is prognostic in our R-CHOP treated DLBCL patients. While the role of macrophages in tumour promotion has been robustly demonstrated in biological models of solid cancers, evidence in lymphoma is limited. GEP studies done in whole LN infiltrated with DLBCL suggested that the expression of macrophage and ECM related genes confer an improved outcome in the R-CHOP era. However, IHC studies are inconsistent, reflecting the difficulties in scoring CD68 but also likely relate to macrophage functional diversity in tissues that hardly can be mirrored by a single marker. While our results suggest that the problem of scoring inconsistency can be overcome by semi-automated analysis, macrophage functional complexity in lymphoma hasn't yet been address and is the aim of our next chapters.

While this study suggests that a high infiltration of T-cells is contributing, together with clinical parameters, for achieving and retaining a CR after R-CHOP in DLBCL, some limitations need to be acknowledged.

One important limitation is the retrospective nature of the study and the relatively small size of the R-CHOP treated cohort. Given the myriad of prognostic biomarkers suggested for this disease, any study of this kind should explore the most well established factors, from the IPI to genetic biomarkers. This should be done under a clinical trial gathering enough patients so as to reach sufficient statistical power for outcome prediction. Only then a proper multivariate analysis incorporating the strongest biomarkers can clarify which of them should be included in the diagnostic workup for DLBCL or change the clinical practice.

What this study clearly provides is evidence for considering semi-automated analysis of IHC for future studies exploring biomarkers in DLBCL or other diseases. This technology eliminates the underlying variability of manual analysis of IHC studies. However, we now face a challenge of what to suggest as methodology for future studies. It could be argued that discriminating prognostic cohorts based on analysis of immunohistochemical markers reflects only bias arising from this same methodology. Nevertheless, our use of multiple methodologies, technologies and markers, all suggesting a positive impact of increased immune cell infiltrate, supports the

hypothesis that this is representative of a real biological effect: that increased immune infiltrate leads to improved outcome. Moreover, is another finding supporting that the computerized analysis is a robust method for IHC analysis. Similar studies should now be done approaching other protein markers such as the COO or MYC.

Clearly, the results reported here are exploratory. Now that we have established that semi-automated systems are the tool of choice for analysis of IHC biomarkers, larger validation studies are required. First, it is essential to conduct an intergroup analysis of all semi-automated systems available, devise a consensus methodological approach and select laboratories that would be responsible for similar analyses under clinical trials where other established molecular and cytogenetic prognostic factors are investigated. This, in our opinion, would definitely answer whether the lymphoma microenvironment plays a role in outcome prediction in DLBCL and other lymphomas and bring forward methods to incorporate such biomarkers into clinical practice.

Chapter 5 Defining the diffuse large B-cell lymphoma associated-macrophage transcriptome

5.1 Introduction

As discussed in Chapter 1, GEP of DLBCL suggested that the stromal microenvironment might have an impact in outcome prediction in this disease. In contrast to what is generally found in solid tumours and other lymphoid malignancies, in DLBCL the expression of macrophage-related genes and ECM components was associated with an improved outcome. This data opens several hypotheses. The first is that in poor-risk DLBCL the malignant B-cells use strategies to change macrophages towards a tumour-permissive behaviour and that those changes are pictured in their transcriptome. The second is that a stromal response is associated with an improved response to R-CHOP or to an ability to control minimal residual disease after chemoimmunotherapy. Whereas the second possibility has been partially demonstrated by studies on the impact of macrophages in mediating efficacy of monoclonal antibody therapies, the first hypothesis has not yet been addressed.

Macrophage functions can be subverted by tumour cells to facilitate disease progression and immune evasion. Whereas a number of mechanisms by which cancer cells influence TAM function have been described, currently there is very limited understanding of the TAM polarisation status and effector function in human lymphoma, including DLBCL.

Mechanisms used by B-cells to influence macrophage activation have only recently been defined.

Crosslinking of the macrophage FcγR with Igs leads to a shift towards an M2-like immunossupressive phenotype. While this process might be adaptive to establish tolerance to autoantigens, it can be used by cancer cells to potentiate progression.[261]

In mice transplanted with melanoma, B-cell depletion shifted TAM polarisation from a tumour-facilitating M2 phenotype towards a classical M1 phenotype. In this model B-cell-derived IL-10 has been implicated in TAM activation.[260] However, using a mouse model of B-cell lymphoma, Leidi et al.[272] proposed that human M2-polarised macrophages have improved cytotoxicity against lymphoma cells after treatment with rituximab, which was further up-regulated by IL-10.

It is likely that B-cell-derived cytokines other than IL-10 can also influence macrophage activation. As has been discussed in Chapter 1, T-cell primed B-cells have been shown to produce Th1 and Th2 cytokines,[267] including IFN-γ, IL-12, IL-4 and IL-13, that were able to modify T-cell responses. It is likely that these impact on macrophage activation. However no data is available addressing this hypothesis, including in the context of B-cell malignancies.

In this study we addressed the hypothesis that the malignant DLBCL-associated macrophages are functionally distinct from macrophages encountered in non-malignant lymph nodes and that those differences are induced by the lymphoma cells to promote disease aggressiveness. Using methods other than IHC would overcome the well accepted phenomenon that the macrophage functional repertoire cannot be dissected using a limited number of markers.

5.2 Aims

We aimed to scrutinise the functional interactions between macrophages and malignant cells by exploring the transcriptome profile of highly pure populations of macrophages selected from human DLBCL samples compared to that from macrophages from reactive LN. Recognizing which genes are differentially expressed in DLBCL TAM might lead to a better understanding of their functions within the tumour microenvironment.

5.3 Materials and Methods

5.3.1 Samples

LN SCSs from patients with DLBCL were chosen based upon the availability of sufficient vials to perform these experiments. Table 5.1 provides clinical details and partial results from individual samples. Details on the methods used for cell suspension preparation, staining, flow-assisted cell sorting, RNA extraction, cDNA amplification, and final sample preparation for microarrays are provided in Chapter 2.

5.3.2 Microarray quality control assessment

Samples were run in two batches on separate days. In order to detect technical variability that could be reflected in the results, two samples were run in both experimental batches and the Pearson's correlation assessed. A series of quality control metrics provided by Affymetrix were utilized with the purpose of identifying failing samples and recognizing where along the processing chain potential problems occurred.

Hybridisation efficiency was evaluated using results from purposely added transcripts derived from Bacillus subtiliis. Due to the lack of other transcripts competing for their probe sets, the binding intensity is directly related to the quality of the hybridisation. Chip intensity can be further assessed by the mean of perfect match (PM) probes raw intensities. Also by assessing how different the ratio of positive exon-level versus negative intron-level probes is from the expected value, one can predict whether any sample has potential hybridisation problems. Analysis of polyA control RNAs spiked to each sample prior to amplification helps monitoring the quality of the labelling reaction.

Regarding sample quality controls, a number of parameters were assessed: the mean of the signal of all probe sets per sample; the mean relative log expression (RLE) and the expression signal.

Table 5.1 Samples used for transcriptomic analysis: patients characteristics and partial results

ID	Age	Gender	Date	AA stage	IPI	Response	Viability (%)	Cell # (x10^7)	CD36 yield	Purity (%)	RNA (ng/uL)	RIN	cDNA (ng/uL)	260/280
T2628	58	M	2009	I	0	CR	76	5.4	42357	90.5 [c]	0.62	7.5	385.7	1.94
T3531	60	M	2009	III	1	CR	40	3.0	65721	96.4	2.1	9.8	308.4	1.95
R3468	61	F	2004	IV	4	PR	45	3.0	47500	92.0	0.13	9.2	256.3	1.97
R8639	34	M	2007	I	0	CR	74	4.2	3780	93.9 [c]	0.03	7.4	288.4	1.92
T2114	42	F	2009	I	0	CR	50	0.7	20000	-	0.10	9.1	272.9	1.95
T4570	50	M	2010	III	2	PD	74	3.0	159268	98.6	4	9.7	389.9	1.92
R0433	86	M	2002 [b]	III	2	CR	35	2.1	22500	-	0.37	9.5	235.6	2.02
F7615	51	M	2000 [b]	III	1	CR	69	3.0	244432	95	8.3	9.7	366.2	1.94
T6932	48	M	2012	III	2	CR	62	2.0	46071	94.0	0.42	10	368.6	1.92
R9516	79	M	2008	IV	4	SD	78	1.3	60000	90	0.38	8.2	372.62	1.97
R6137	81	M	2005	IV	3	PR	77	2.0	22209	90	2.3	9.0	375.32	1.96
F8146	61	F	2000 [b]	II	2	CR	61	3.0	152324	96.4	8	10	368.36	1.95
R8756	59	M	2007	IV	2	CR	67	1.1	12310	92.1 [c]	0.3	7.6	295.2	1.95
T5900 [a]	21	M	2011				85	5.0	200000	92.0	5.7	7.0	359.2	1.96
T5353 [a]	19	M	2011				89	4.3	190000	97.0	20	8.0	357.95	1.96
T5996 [a]	18	F	2011				95	5.1	78388	92	1.6	8.5	372.8	1.95
T5848 [a]	68	F	2011				70	2.7	90380	97.1	0.45	9.3	354.8	1.96
T5424 [a]	73	M	2011				75	7.0	510852	96	10	9.7	305.8	1.96
T5175 [a]	37	F	2011				90	2.9	117416	96.4	6	10	431.38	1.98

[a] reactive LNs; [b] CHOP treated; [c] Double sorting; AA: Ann Arbor; IPI: International Prognostic Index; RIN: RNA integrity number; M: male; F: female; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease

5.3.3 Data normalisation

This analysis was performed using Bioconductor packages (www.bioconductor.org) within the R statistical environment. A robust multi-array average (RMA)[335] algorithm was applied for background signal adjustment and data normalisation.

Background signal is an inherent problem to microarray technology imputed to optical noise and non-specific binding (NSB) that can lead to falsely overestimated gene expressions. The RMA algorithm estimates and adjusts background intensity using a set of approximately 17000 generic background probes.

The background adjusted data is then normalised at the probe-level using a quantile method.[336] Quantile normalisation orders the data from each array from highest to lowest expression and identifies the average value at each quantile. The observed expression is then transformed to the quantile average in order to give each data set an equivalent distribution. This data manipulation is required to extract experimental variability associated with sample batching, inconsistent use of reagents between arrays or as a systemic experimental bias present in the study design.

Finally, to minimize statistical testing while retaining the highest probability of detecting differences in expression between samples, the data was filtered to obtain the 20% probes with higher binding variability across all samples using standard deviation.

5.3.4 Data analysis

To determine differentially expressed genes between the investigated groups, an empirical Bayes approach (limma package) was used. This method takes into account some characteristics of the whole data, including the variance of expression of all genes, the distribution of all variances and the number of samples under analysis to estimate a pooled pattern of expression for the study. Using this pattern it is possible to recognize data with unusual distribution that should be considered fortuitous but

would otherwise be taken as significant. This method is particularly important for studies with a low number of samples per group, where findings are not replicated enough and can merely be attributed to chance.

The Benjamini and Hochberg (BH) method of multiple hypothesis testing correction was applied in order to rectify the p-values for error induced by multiple measurements, where the null hypothesis is incorrectly rejected (Type 1 error). A double cut-off of adjusted p-value <0.05 and Log^2 FC >1 were set to specify a list of differential expressed genes.

Genes that passed these criteria were selected for hierarchical clustering. This method allows estimating correlations between two samples by evaluating the distance between measurements inputted for the samples. The closer the measurements (in this case gene expression), the shorter the distance and greater the correlation estimated. Samples that are highly correlated can be clustered in a group. The distance metric used for this study was Euclidean distance. The clustered data was represented in a dendogram using the Ward linkage criterion. The height measured from the bottom of the dendogram to the point where samples diverge represents the distance/degree of similarity between samples. Jacek Marzec performed all data normalisation and analysis.

5.3.5 Targeted gene expression validation by qRT-PCR

Amplified cDNA generated for the microarray experiments was further used for validation by qRT-PCR. The transcripts were chosen amongst the most differentially expressed between DLBCL and reactive samples. The TaqMan® Gene Expression Assays tested were: ALOX15, Hs00609608_m1; AQP9, Hs01035888_m1; CDH1, Hs01023894_m1; IDO1, Hs00984148_m1; C3AR1, Hs00377780_m1, CD1E, Hs00954575_m1; CTSL1, Hs00964650_m1 and ANKRD22, Hs00944018_m1. All assays were bought from Life Technologies.

The average of triplicate CT values was normalisedd to the house-keeping gene B2M due to it being the only one passing Levene's F-test of equality of variance among the

three explored (*B2M*, *GAPDH*, *RPL34*) across all samples. Results are expressed in RQ values with errors bars representing the standard error of the mean (SEM).

5.3.6 Hierarchical clustering of RQ values

RQ values were clustered in Cluster V3.0,[337] using Euclidian distance. The clustered data was represented in a dendogram built in Tree View[337] using the Average linkage criterion.

5.3.7 Gene enrichment analysis

The differentially expressed gene symbols were computed in Toppfun (http://toppgene.cchmc.org), an online bioinformatics tool that performs gene list functional enrichment. This data-mining tool compares the data inputted with annotated data from different categories: GO (gene ontology) terms, pathways, protein–protein interactions, transcription factor binding sites, gene tissue expressions and literature co-citation. Toppfun retrieves annotated information where the gene(s) of interest is/are over-represented, the significance of which is corrected with the BH test.

5.3.8 Generation of resting and polarised macrophage gene signatures

In order to perform a comparative analysis, we used bioinformatics approaches to develop macrophage gene signatures (Figure 5.1). Firstly, representative GEP datasets of human macrophages were chosen. The GEP dataset GSE22886[338] was used to construct a resting macrophage signature. This study comprehensively profiled six immune cell types (B, T, NK-cells, plasma cells, monocytes and neutrophils) and their activated and differentiated states. A “resting” macrophage-specific gene signature was depicted by comparing the GEP of monocyte-derived macrophages to all other experimental conditions. Secondly, the GEP dataset GSE5099[194] was chosen to delineate activated macrophage signatures. M1 and M2 signatures were defined by comparing IFN-γ/LPS and IL-4-treated macrophages, respectively, with the remaining

conditions. With the intention to exclude any gene that has been previously demonstrated to be expressed by DLBCL B-cells, Lenz et al.[34] and Basso et al.[339] datasets were used to inquire for the expression levels of genes included in the signatures and filter out the ones that passed the expression criteria. The design of this approach was performed by me. The bioinformatics analysis was performed by Aaron Newman and Ash Alizadeh, from Stanford School of Medicine, Palo Alto, CA.

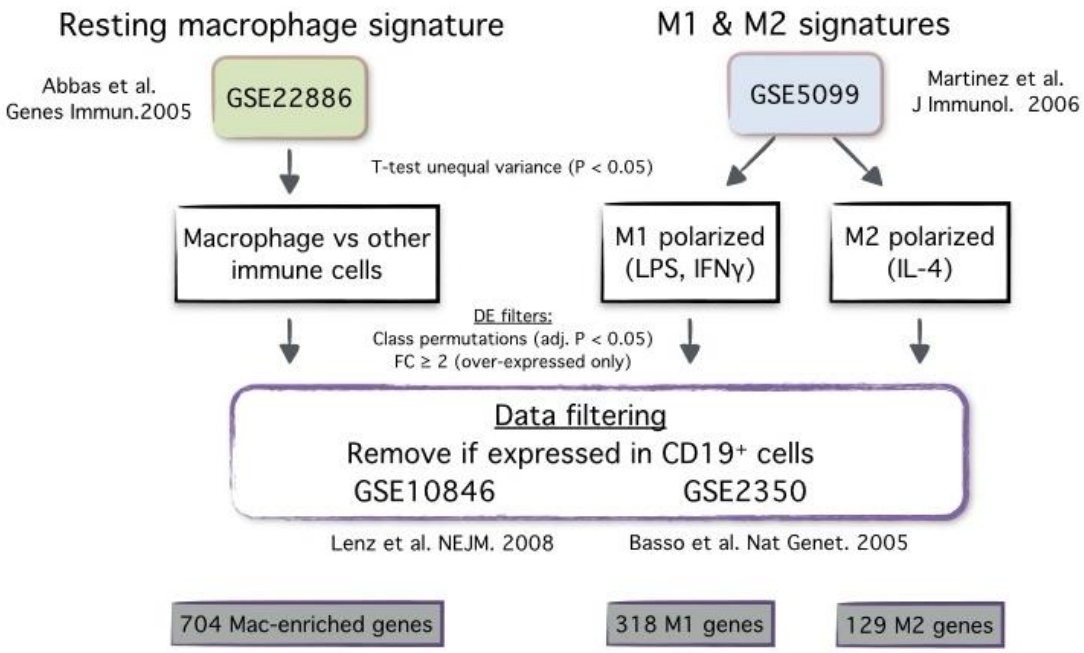

Figure 5.1 Development of macrophage gene signatures.
DE: differential expression; FC: fold change

5.4 Results

5.4.1 CD36 is expressed by the majority of tonsil CD68+ cells, representing a good single marker for macrophage cell sorting

Due to the limited number of LNs SCSs of DLBCL available in our tissue bank with enough vials to conduct this study, preliminary data was gathered using samples of other lymphoproliferative diseases and tonsils. Firstly we demonstrated, and posteriorly confirmed in our original data, that macrophages constitute a very rare population in SCSs as assessed by flow cytometry using CD68 intracellular staining (data not shown). Under this premise, we aimed at defining the simplest strategy to cell sort macrophages using only surface markers. This would enable us to minimize the staining steps to a minimum and avoid cell membrane permeabilisation that would likely affect RNA quality. We showed that the membrane scavenger receptor CD36 was co-expressed by the majority of CD68+ cells in tonsil SCSs (Figure 5.2).

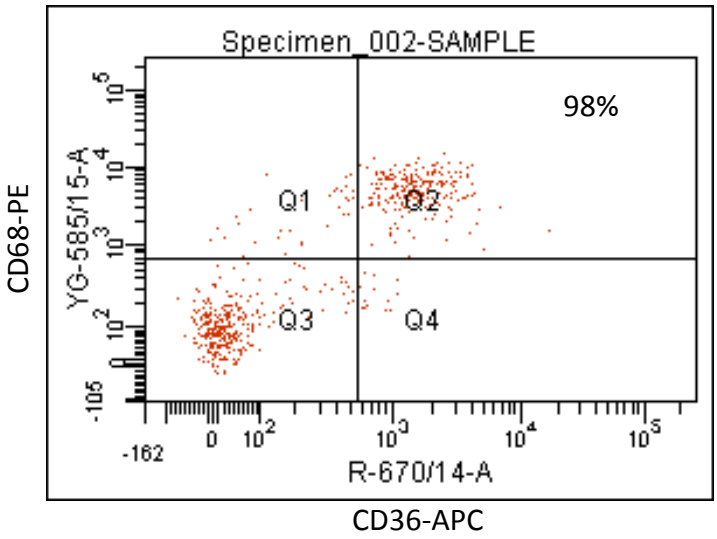

Figure 5.2 CD68+ cells co-express the membrane scavenger receptor CD36.
$5.0x10^5$ cells were fixed and permeabilised for intracellular staining with CD68. Cells are gated on singlets and live cells. Flow plot representative of a single biological replicate.

5.4.2 LN single cell suspensions of DLBCL are not representative of the tumour content when assessed by flow cytometry

For this study aiming at comparing the GEP of DLBCL-associated macrophages with that of reactive LN macrophages we were able to select 13 DLBCL and six control LN samples (Table 5.1). The cell sorting methodology and representative example of the sorting strategy were described in Chapter 2. Cell number and viability were assessed using an automated haematocytometer prior to antibody staining. From each sample, total CD20+ B-cells, CD3+ T-cells and CD36+ macrophages were isolated after gating on singlet, live leukocytes (CD45+). Cell yields were retrieved by the cell sorter.

Sample viability on thawing was variable but generally poorer for disease cases (35-78% for DLBCL compared to 75-90% for controls). Figure 5.3 represents an estimate of the percentage of macrophages, B and T-cells in the stored samples taking as denominator the total number of viable cells. As can be appreciated, B-cells were not the most abundant cell population in DLBCL SCSs, T-cells being more numerous in all samples.

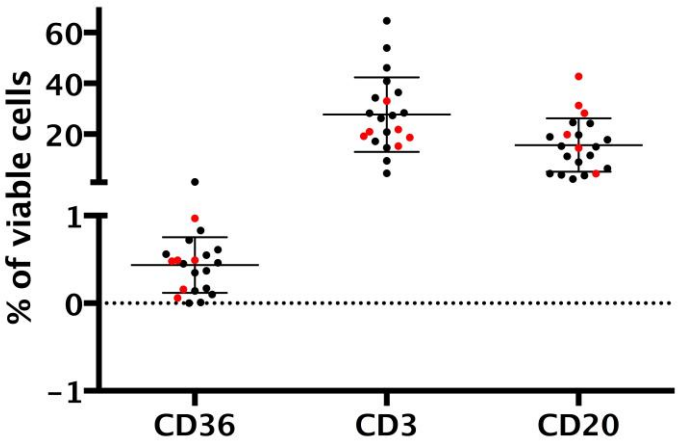

Figure 5.3 Proportion of macrophages, T-cells and B-cells in SCSs of DLBCL and reactive conditions.

Sample cell count and viability were taken using an automated haematocytometer before starting the staining protocol. Individual population cell counts were taken at the end of the cell sorting. Control samples are identified in red.

Cell purity after sorting was excellent for B and T-cells (>98% in all except one sample, data not shown). As can be appreciated in Table 5.1, macrophage purity after sorting was acceptable (median 94%, range 90-99%). However, three samples required double sorting in order to guarantee purity >90%. This strategy highly compromised cell yield but enabled the use of those samples for transcriptomic analysis.

CD36+ macrophages were a rare population both in malignant and in control samples. Macrophage cell yield was significantly higher for control samples (median $1.5x10^5$, range $8x10^4$ to $5x10^5$) compared to DLBCL samples (median $5x10^4$, range $4x10^3$ to $2.4x10^5$, p= 0.01).

5.4.3 Quality assessment of the experimental workflow employed in this study

As detailed in Chapter 2, macrophages were immediately pelleted after sorting by centrifugation at 8000rpm for 10min without any washing step, follow by RNA extraction. RNA concentrations and quality were assessed using the Agilent bioanalyser pico kit, designed for low RNA amounts. Although macrophage numbers were significantly higher in control samples, RIN numbers were excellent and comparable across disease and control samples, varying between 7.4 and 10. Figure 5.4 shows the electropherograms of a number of the samples used in this study and an internal control.

However, and as predicted, the RNA concentrations obtained from sorted DLBCL-associated macrophages we significantly lower (median 0.42, range 0.03 - 8 ng/uL) compared to reactive LN macrophages (median 5.9, range 0.4 - 20 ng/uL, p= 0.03).

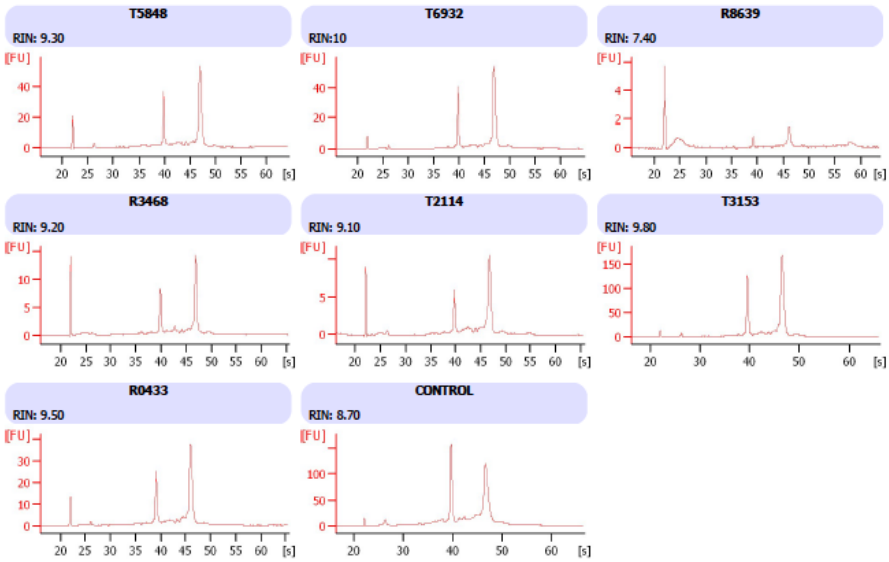

Figure 5.4 Excellent RNA quality obtained from macrophage populations.
Representative electropherograms for assessment of RNA quality of samples used in this study. With the exception of R8639, all samples had a RIN >9, suggestive of high quality RNA.

Given these results, nucleic acid amplification was required. RNA quantities taken forward for cDNA synthesis and amplification were equivalent across samples. A negative control was included with the study samples in two experimental batches. As can be appreciated in Table 5.1, cDNA yields and quality were similar across samples. Electrophoretic traces for amplified and fragmented cDNA were inspected for some samples to confirm integrity (Figure 5.5).

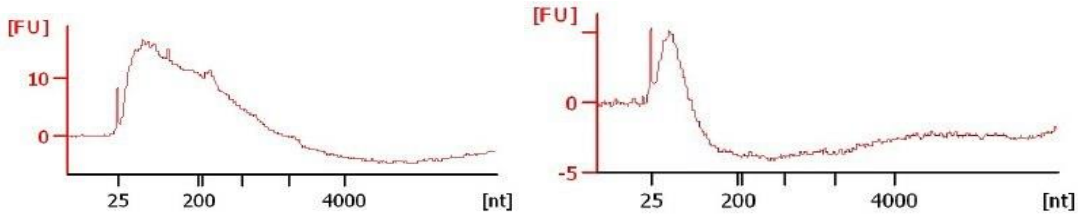

Figure 5.5 cDNA fragment size distribution analysis using the Agilent RNA 6000 Nano kit. The left side trace represents amplified cDNA; the right side trace represents fragmented cDNA.

After fragmentation, cDNA was biotin-labelled and hybridised to Affymetrix GeneChip® Human Gene 1.0 ST Arrays. These microarrays evaluate the expression of 28869 well-annotated genes with 764885 distinct probes that are distributed across the full length of the gene, guaranteeing whole transcript coverage. This is particularly useful for studies using potentially degraded samples.

None of the samples was flagged as an outlier using the quality control checks. Hybridisation was assured by evaluating the distribution of control probes (Figure 1, Appendix). Also, the ratio of positive versus negative probes was >0.7 (0.82-0.89, Figure 2, Appendix), suggesting that hybridisation occurred as expected. Mean intensities for all arrays were very consistent, with a variation of 0.05 (Figure 3, Appendix), as were mean intensities for PM and background probes (Figure 4, Appendix). Regarding the RLE signal (Figure 5, Appendix), sample R8639 stood out has having a higher mean compared to the remaining.

Finally, Pearson correlation of intensity signals for the duplicate samples shows a high correlation between technical replicates with individual R values 0.99 (Shown in Figure 5.6).

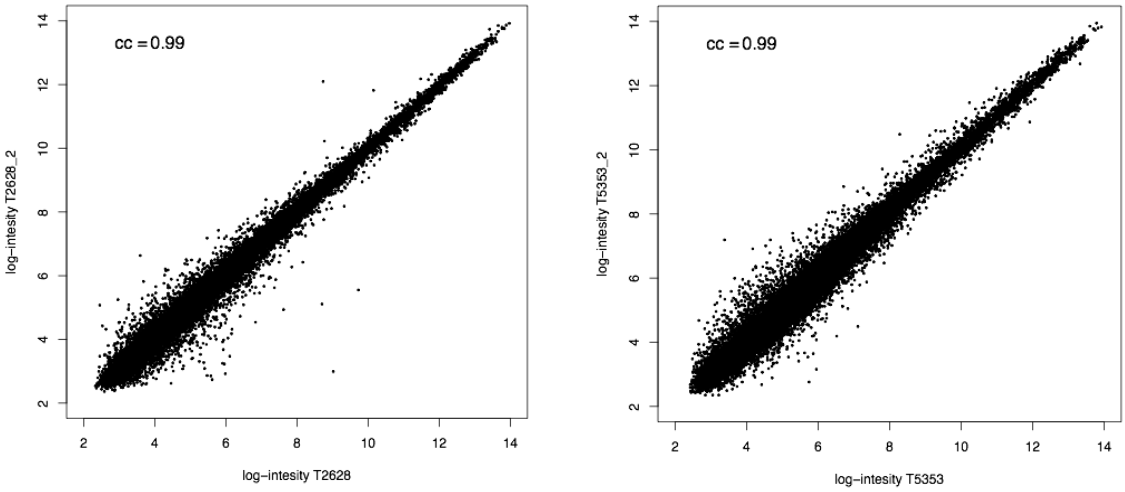

Figure 5.6 Correlation of intensity signals for the samples hybridised in the two experimental batches
cc, correlation coefficient.

5.4.4 Unsupervised hierarchical clustering

As already mentioned, hierarchical clustering is useful to detect similarities between samples. Undertaken in an unsupervised fashion means the whole filtered data (~12000 probes with highest expression variability) is used.

As can be appreciated in Figure 5.7, this analysis failed to fully resolve DLBCL TAM and control macrophages.

Whereas reactive LN-associated macrophages cluster together, showing a high degree of similarity, DLBCL TAM have higher GEP variability. Within the DLBCL samples R8756, R6137, F8146, R9516 and T6932 diverged earlier from the remaining samples. Samples R8639, R0433 and R3468 also segregated into a separate cluster. The remaining DLBCL TAM samples clustered closer to the control macrophages. We excluded this variability was due to a batch effect by observing close clustering of samples that were hybridised in different days.

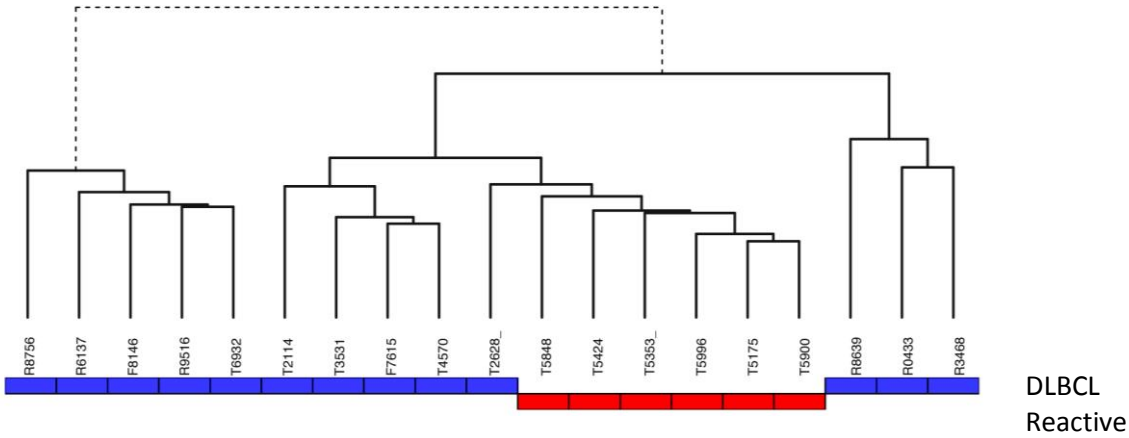

Figure 5.7 Unsupervised hierarchical clustering of DLBCL and reactive LN-associated macrophages.

Clustering based on RMA normalised expression of the top 12000 probes with the highest variability. Data is correlated by Euclidean distance using Ward linkage to define the dendogram.

5.4.5 Statistical analysis of differentially expressed genes between DLBCL TAM and reactive LN macrophages

A supervised hierarchical clustering analysis was performed to recognize the differentially expressed genes between DLBCL TAM and control macrophages. This was calculated by dividing the average logarithmic expression data for each group, generating a FC ratio. The FC cut-off selected for this experiment was ≥2 (Log^2 ≥1). P-values were corrected with BH False Discovery Rate (FDR) and a value of ≤0.05 accepted. Using these criteria, 208 probe sets for 202 well annotated genes were differentially expressed between the two groups. Of those, 63 were downregulated and the remaining were upregulated in DLBCL-associated macrophages versus reactive LN macrophages (Figure 5.8). The full data set for all significant differentially expressed genes is shown in detail in Table 1, Appendix.

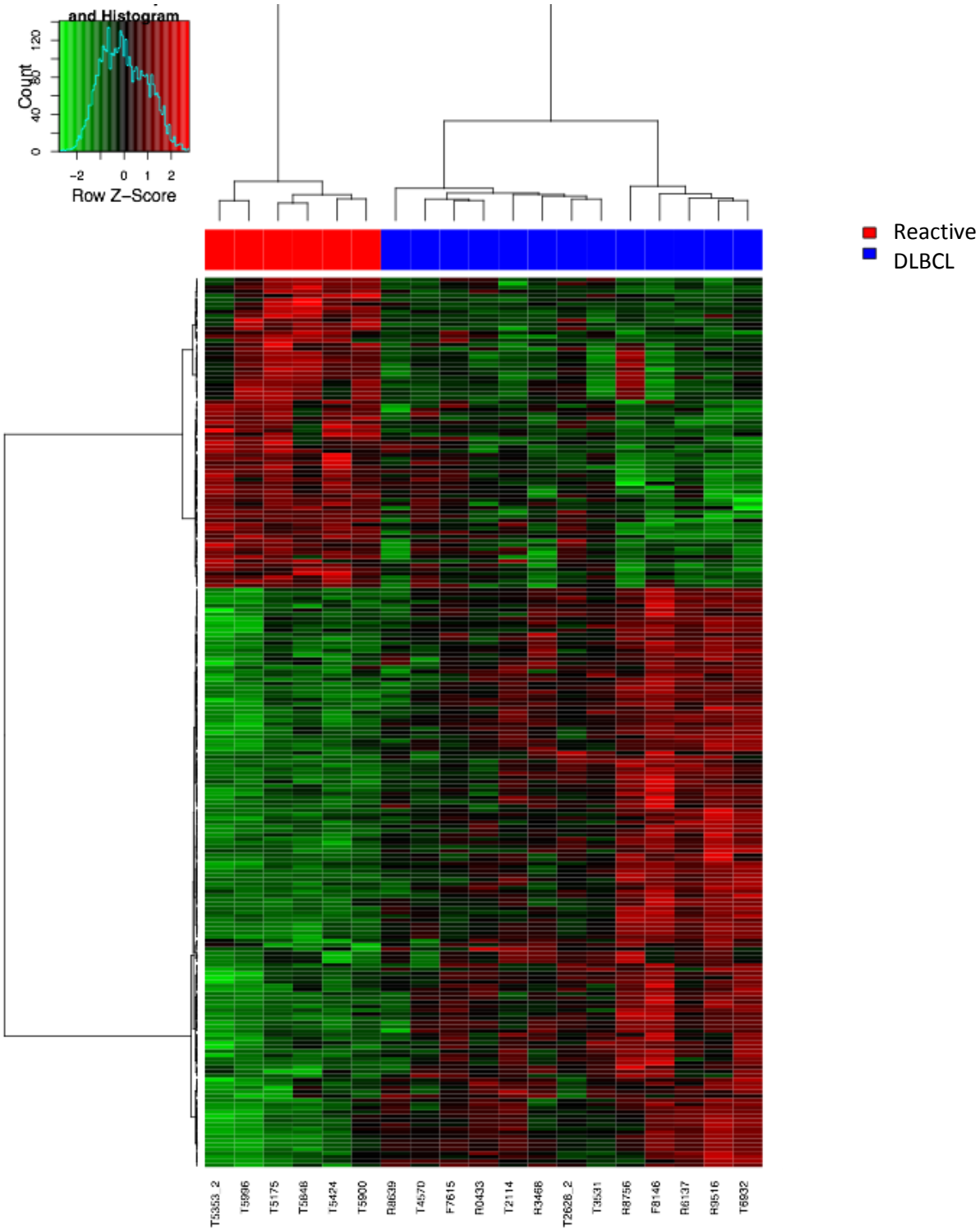

Figure 5.8 DLBCL TAM and reactive LN macrophages differ in expression of 202 genes.

Supervised hierarchical clustering analysis using a Euclidean distance measure and Ward linkage. Each column represents one sample, each row a transcript. Gene expression levels are represented on a scale of green to red colour indicative of low to high expression.

As can be visually appreciated in the heatmap, control samples have a more coordinated expression of the differentially expressed genes. To substantiate this, the six most highly differentially expressed transcripts are represented in Figure 5.9 plotted by sample. Representing the data in this manner suggests a higher level of heterogeneity in the GEP of TAM compared to controls.

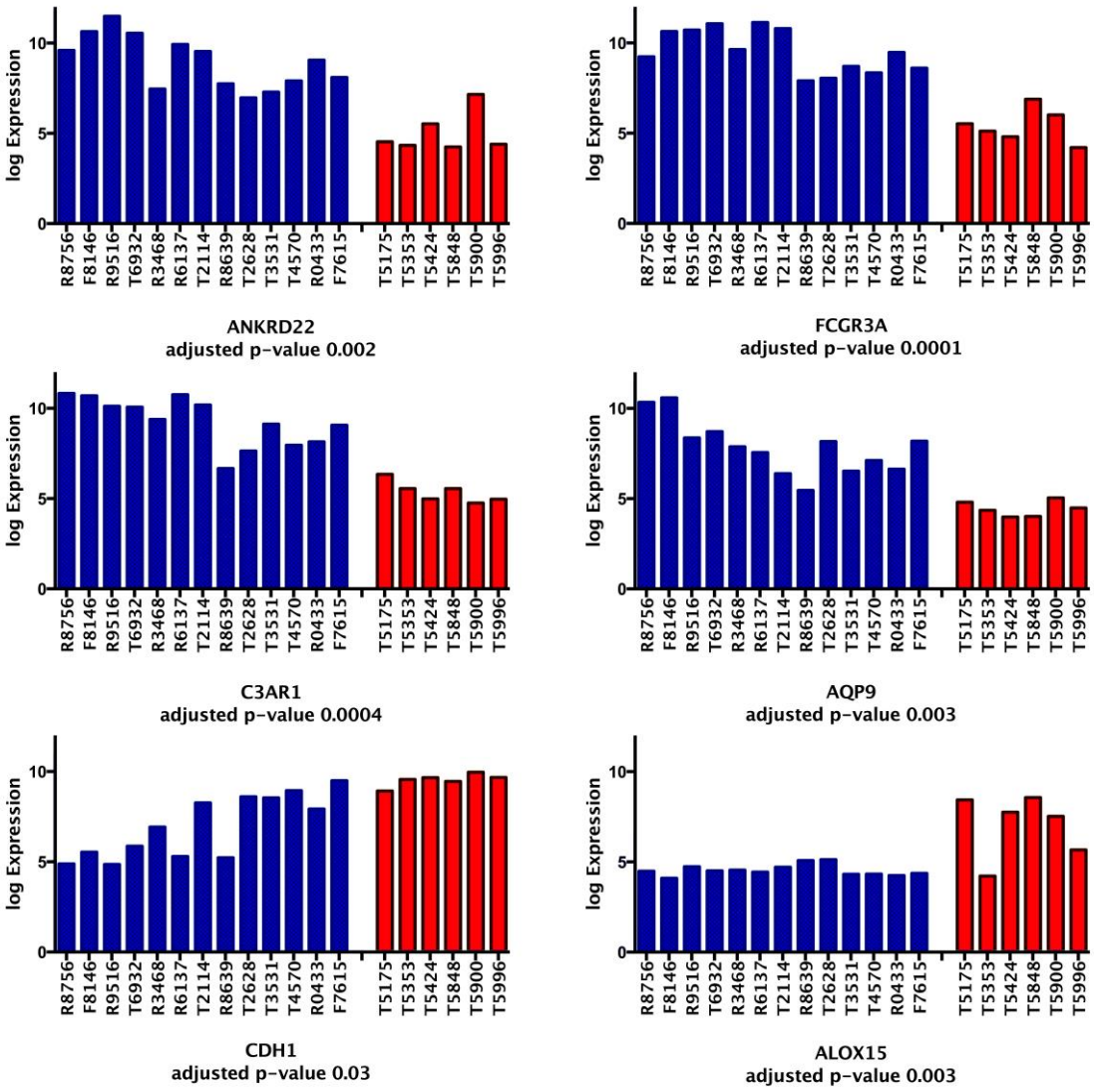

Figure 5.9 Expression heterogeneity of most differentially expressed genes in DLBCL and control macrophages.

Graphs represent each transcript and columns each sample. Log expression is represented for each gene and sample. Control samples are represented in purple and DLBCL samples in green. p-values were adjusted according to BH test.

5.4.6 Confirmation of differentially expressed genes by qRT-PCR

To validate expression of selected genes from our list, qRT-PCR was used. The transcripts chosen were among the highest differentially expressed.

In addition to the seven genes represented (Figure 5.10), ALOX15 was also tested. However, ALOX15 genes levels could be measured in control samples but were undetected in DLBCL TAM samples by qRT-PCR, which renders statistical analysis impossible. This corroborates the microarray data (Figure 5.9, right bottom graph) showing a significantly lower expression of ALOX15 in DLBCL TAM compared to controls.

As shown in Figure 5.10, GEP results could be validated by qRT-PCR, suggesting that the transcriptomic data indeed represents biological differences between the two groups. Although RQ values were not significantly different for CD1E, the FC direction was confirmed (Figure 5.10, B). FC results were higher for upregulated genes and lower for downregulated genes when tested by qRT-PCR compared to microarrays

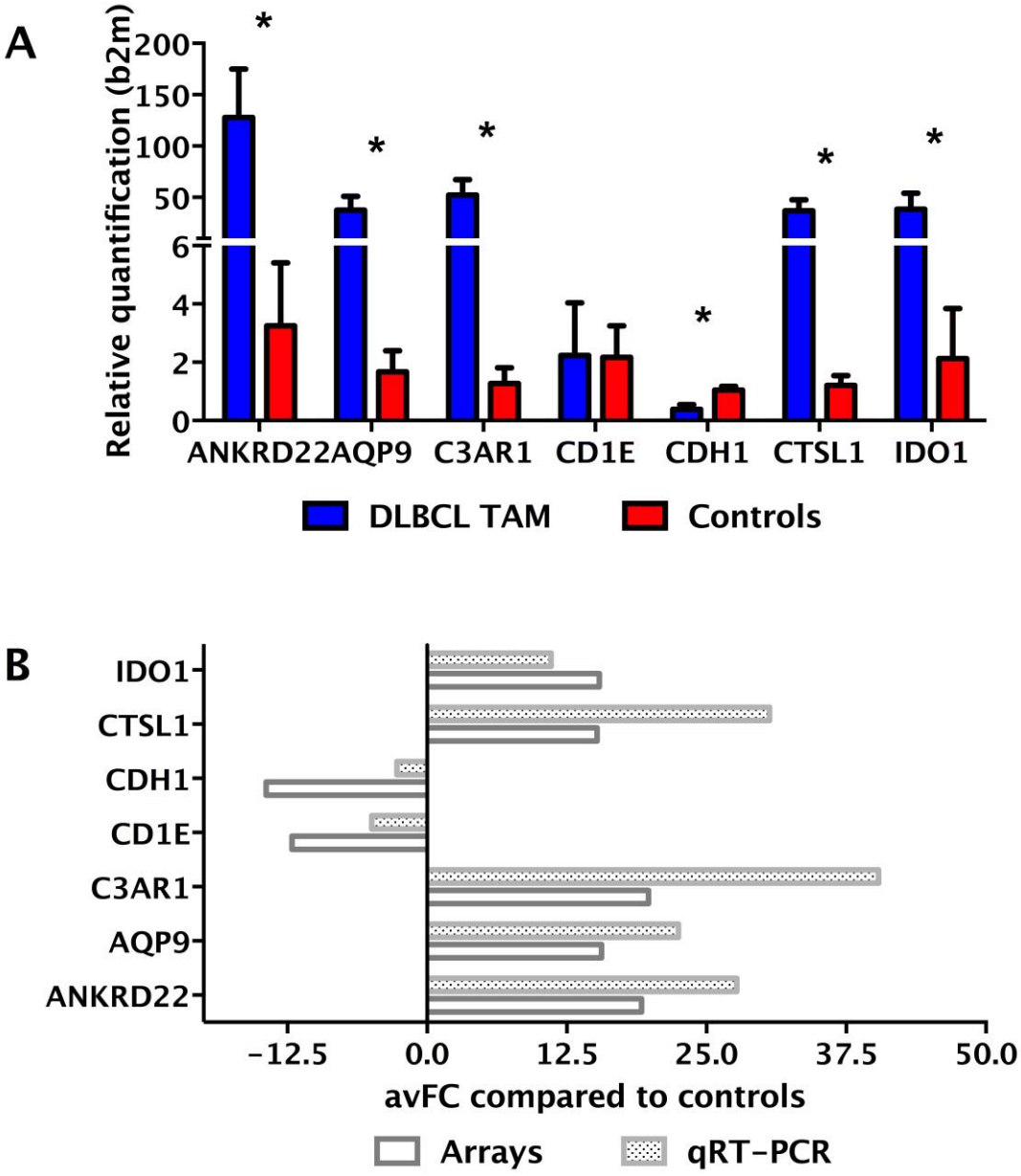

Figure 5.10 Validation of targeted differentially expressed genes by qRT-PCR.

50 ng of amplified cDNA was used in triplicate for each sample and analysed by qRT-PCR. CT values were normalised to B2M. A. Average RQ values were compared using T-test and p-values <0.05 deemed significant. Results plotted using mean +- SEM. B. Average fold change of gene expression in DLBCL TAM compared to controls using microarray and qRT-PCR methodologies.

5.4.7 Hierarchical clustering of targeted gene expression by qRT-PCR

An unsupervised hierarchical cluster analysis was performed on the RQ values for the eight genes tested by qRT-PCR. This simplified methodology divided DLBCL TAM in a separate cluster from control samples (Figure 5.11). Regarding control samples, C5 (T5424) clustered closer to lymphoma samples, whereas C1 (T5175) clustered separately from all other samples. The remaining control samples collected in a separate group.

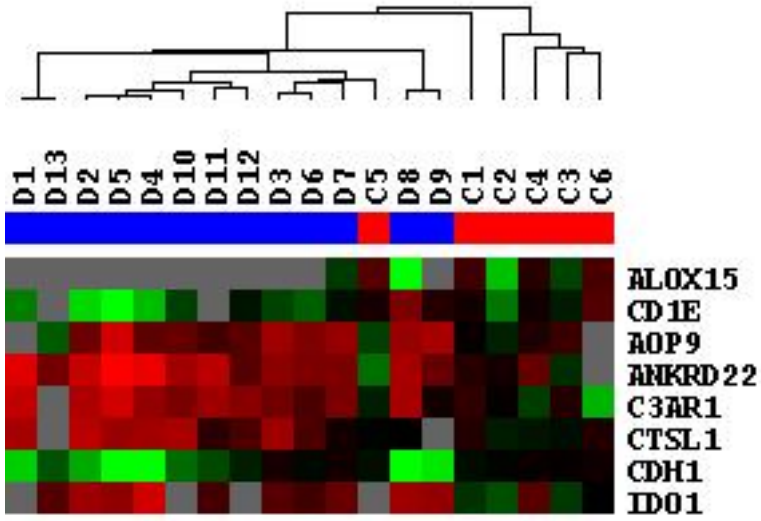

Figure 5.11 Hierarchical cluster analysis of RQ values of seven transcripts resolves DLBCL TAM from controls.

Data is correlated using Euclidean distance and an average agglomeration to define the linkage tree. Each column represents one sample, each row a gene. Expression is represented in a continuum from red (for high expression) to green (for low expression). Absent data is coloured in gray. DLBCL samples are labelled from D1 – D12 and reactive samples are labelled from C1 to C6.

5.4.8 Gene enrichment analysis establishes links between our gene set, the "LN signature" of DLBCL and M1 and M2 polarised macrophage functions

The most relevant results retrieved by Toppfun are presented here according to category. More comprehensive information for each of the sub-sections will be provided in Appendix, Table 2. All p-values stated are corrected for the FDR with the BH test.

5.4.8.1 Gene ontology (GO)

The GO consortium gathers several genome databases that use a standardized nomenclature for gene products according to biological processes and molecular functions where they have been implicated.

An important finding that highly supports our methodological approach and results was that our differentially expressed genes are significantly implicated in cellular functions where macrophages are known to play a role: immune response ($p=1.5x10^{-35}$), defence response ($p=1.1x10^{-27}$), innate immune response ($p=2.5x10^{-18}$) or inflammatory response ($p=5.6x10^{-21}$). This subanalysis suggests that both M1 and M2-related genes are present in our dataset as both "Response to IFN-γ" and "Response to LPS", prototypic M1 stimuli; or "Response to wounding", a prototypic M2 functional activation, were on our top 20 GO cellular functions (Figure 5.12, A).

Receptor activation and signalling were molecular functions implicated in our dataset (Figure 5.11, B). Of note, Ig binding, known to play a role in M2 activation, and LPS receptor signalling, involved in M1 activation, were both enriched.

5.4.8.2 Pathways

Toppfun inquires for overlaps of the data inputed with data sourced in pathway databases, such as the Kyoto Encyclopedia of Genes and Genomes or Panther database. Again, significant terms point towards conditions in which macrophages have a central role. Prototypical M2 diseases were listed: tuberculosis ($p=2.034x10^{-4}$), leishmaniasis ($p=9.3x10^{-4}$). However, M1-related IFN-γ signalling pathway appeared

more significantly enriched ($p=8.9x10^{-6}$).

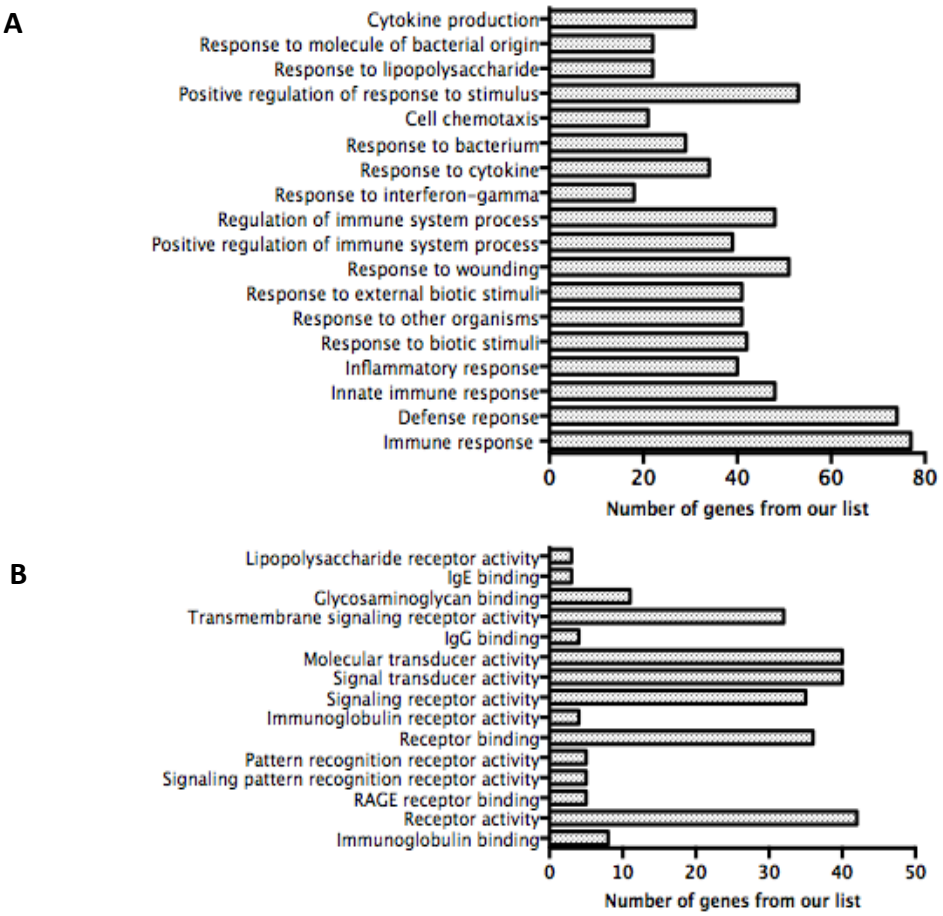

Figure 5.12 GO enrichment analysis.

A. Most highly represented cellular functions, p-values between $1.5x10^{-35}$ and $3.1x10^{-11}$. B. Most highly enriched molecular functions, p-values between $3.9x10^{-8}$ and $5.0x10^{-4}$. Bars correspond to the number of genes from our list implicated in each function, decreasing in significance from the bottom of the graph.

5.4.8.3 Co-expression

Toppfun also overlaps gene lists provided with annotated gene signatures indexed in Pubmed and gathered in the Gene Signature DataBase.

The most significant hit in the co-expression category ($p=4.6 \times 10^{-60}$), and in fact the top hit from this bioinformatics analysis, was a GEP study by Van Loo et al.[340] on full tissue sections of THRLBCL, a pathological DLBCL sub-entity where the bulk of the tumour is composed of T-cells and macrophages. Sixty-three of our upregulated genes in DLBCL were amongst a list of 373 the authors described as being upregulated in THRLBCL compared to nodular lymphocyte predominant Hodgkin Lymphoma (NLPHL) and reactive LNs. Whereas NLPHL and reactive LNs overexpressed established B-cell genes, THRLBCL genes were related to the microenvironment. Importantly, 11/20 of our top differentially expressed genes were overlapping with Van Loo's data, including: C3AR1 (#2), ANKRD22 (#3), FCGR1A (#4), IDO1 (#8) or CXCL10 (#9).

Another interesting finding of this sub-analysis was an over-representation of gene sets that represent cell states and perturbations within the immune system generated as part of the Human Immunology Project Consortium and the Mouse Immunological Genome Project. Significant overlap was detected between our dataset and 18 gene sets of the human project (p-values between 2.7×10^{-46} and 4.0×10^{-18}), the vast majority of them referring to differentially expressed genes between monocytes and other cells of the immune system, including B-cells.

In the same line, we found a highly significant overlap between our differentially expressed genes and the ones defining several mouse immune cell populations. The 50 most significant hits (p-values between 4.4×10^{-55} and 1.3×10^{-22}) correspond to myeloid cell populations.

Neither of the two studies defining the LN and "stromal-1" gene signatures figured in the Gene Signature DataBase and hence was retrieved in our analysis. Those

signatures were enriched for transcripts likely derived from macrophages, but a formal demonstration of that was deemed impossible due to the way the original studies were designed. However, the LN-signature was introduced by a secondary study,[341] which permitted it to result in our analysis. We found a significant overlap of our gene list with the LN signature (25 out of 161 genes, $p=3.5x10^{-22}$). Among some of those genes were: FCGR1A, CXCL10, CXCL9, CCL2, C3AR1, CD14, or CTSL. On the contrary, only six of our transcripts were represented in the "stromal-1" signature.

5.4.8.4 Cancer Modules

Through an integrated analysis of 1975 microarray studies spanning 22 tumour types, Segal et al. defined cancer modules, sets of genes that act in concert to carry out a specific function in cancer tissues.[342] Toppfun analysis showed that the most enriched (53/531 genes, $p=1.7x10^{-27}$) was the "immune (humoral) and inflammatory response" module. Importantly, this module has been shown to be over-represented in DLBCL, independently of the molecular subtype.

5.4.8.5 Gene family

Five metallothionein (MT) genes appeared overexpressed in DLBCL TAM compared to controls ($p=7.1x10^{-7}$). MTs are metal-binding proteins that respond to oxidative stress and acute phase cytokines such as IL-1, IL-6, and TNF-α and have been implicated in immune regulatory effects. MT knockout mice show impaired wound healing and higher susceptibility to inflammation. It was suggested by a single study that MT expression is associated with poor outcome in DLBCL.[343]

5.4.8.6 Upstream regulators

Finally we used another bioinformatics tool, Ingenuity pathway analysis (http://www.ingenuity.com/products/ipa), to inquire for upstream regulators of expression of our gene set. Both prototypical M1 (IFN-γ, $p=2.5x10^{-30}$) and M2 (IL-10, $p=2.4x10^{-17}$; Ig, $p=1.5x10^{-16}$; IL-13, $p=5.0x10^{-15}$; IL-1β, $p=9.5x10^{-14}$; IL-4, $p=3.0\ x10^{-6}$) cytokines were significantly reported. The most significant transcription factors identified were the M2, IL-10-induced STAT3 ($p=5.0\ x10^{-11}$) and the M1-related STAT1 ($1.9\ x10^{-6}$).

5.4.9 Comparative analysis with macrophage gene signatures

As can be appreciated in Figure 5.1, "resting" and "activated" macrophage gene signatures were delineating for comparative analysis. Data overlap was tested both before and after the filtering step described.

5.4.9.1 Comparison with a "resting" macrophage-specific signature

To confirm the specificity of the "resting" macrophage signature we plotted the expression of the 839 genes across all immune cell populations from the study of Abbas et al.[338] As can be appreciated in the heatmap (Figure 5.13), monocyte-derived macrophage replicates (gray square) showed concerted expression of the genes selected. Additionally, a higher degree of similarity was detected between the myeloid cells included in the study, particularly monocytes and dendritic cells compared to other immune cells. B-cells have a different expression pattern compared to macrophages. This data validates the bioinformatics approach to define macrophage signatures.

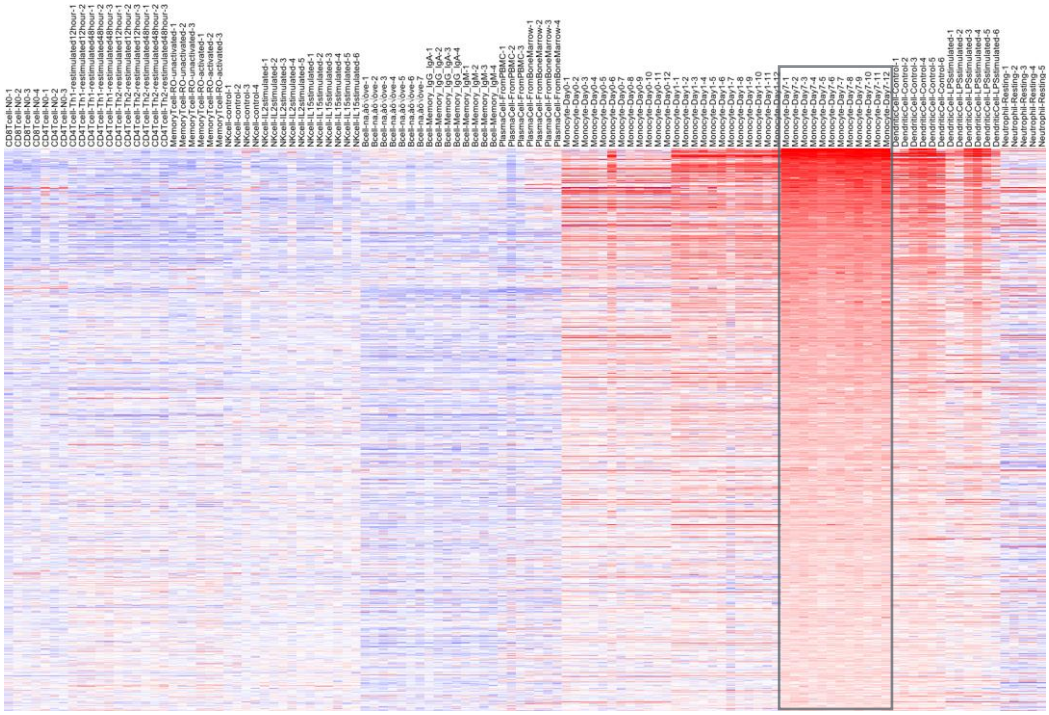

Figure 5.13 Heatmap of "resting" macrophage-enriched genes.
As outlined in Figure 5.1, 704 genes were chosen to represent a macrophage-specific signature. Gene expression is represented on a scale with a Log^2 range of +6 (red) to -6 (blue).

Among the 202 differentially expressed genes between DLBCL TAM and controls, 51 were also present in the "resting" macrophage-enriched gene list, representing a highly significant overlap as per the probability hypergeometric test ($p=1.2 \times 10^{-39}$). All but six of these were upregulated in DLBCL TAM versus control macrophages.

Finally we investigated the survival impact of each of the genes in the "resting" macrophage signature in an R-CHOP dataset.[34] The 51 overlapping genes from our dataset were variably associated with OS: 22 with improved OS, 18 with worse OS and 11 had no impact. Overlapping genes and survival impact are depicted in Figure 5.14.

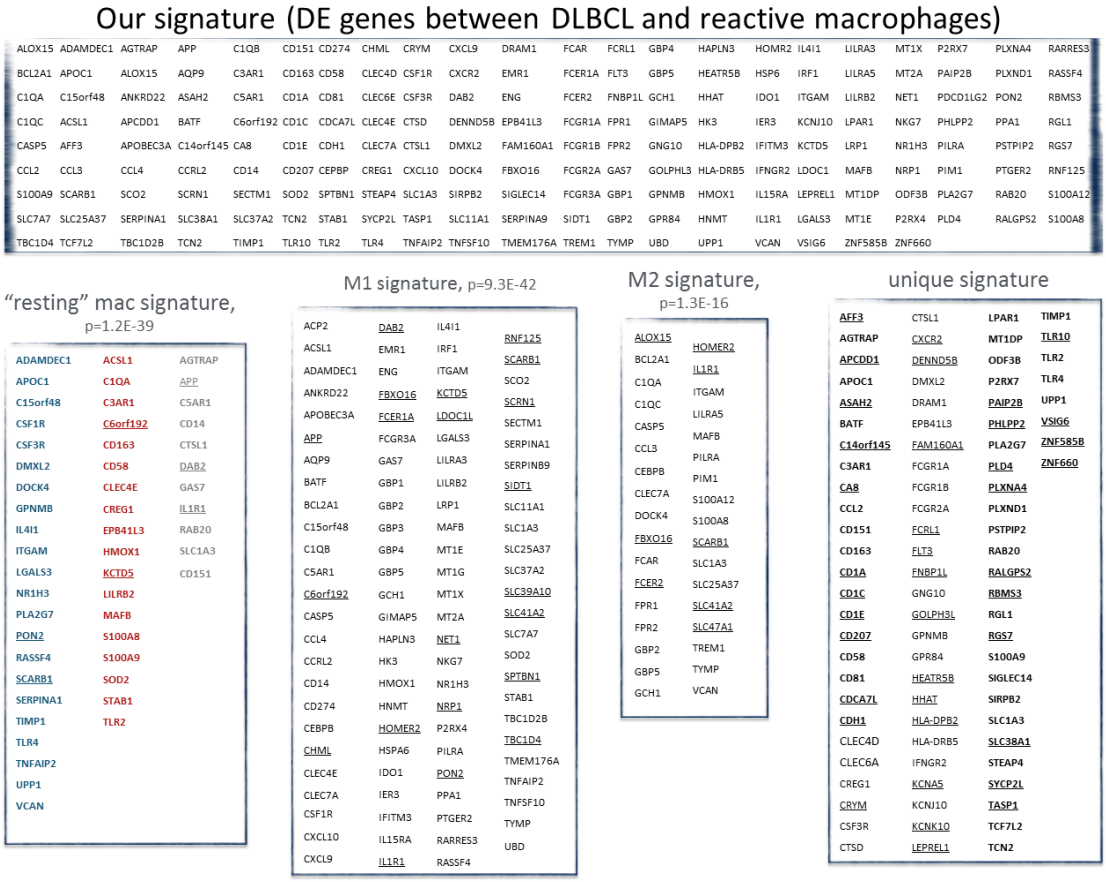

Figure 5.14 Comparative analyses of our DLBCL macrophage associated gene signature with other macrophage gene signatures.

Our gene expression signature is depicted in the top box. Overlapping genes between our signature and the ones developed by bioinformatics analysis are given in each box. The degree of significance was tested using hypergeometric distribution. R-CHOP OS impact of "resting" macrophage genes is represented in a colour code of red for worse outcome, blue for improved outcome and gray for negligible OS effect. Underlined genes are downregulated in DLBCL TAM versus controls.

5.4.9.2 Comparison with "activated" macrophage signature

Martinez et al.[194] stated in their original publication that the transcriptional profile variance from "resting" to M1 polarised is significantly greater than to M2 polarised (90% versus 8%), which can be explained by two reasons: M1 activators are more potent than M2 activators; and the addition of CSF-1 during the in vitro maturation process already switches cells towards an M2 phenotype, hence decreasing the impact of M2 activation. This justifies the identification of smaller M2-specific datasets compared to M1-specific datasets (see Figure 5.1).

Given that it is conceivable that macrophage activation triggers expression of other than macrophage-restricted genes (for example expression of a chemokine), enrichment analysis was performed before the B-cell expression-filtering step (Figure 5.1).

The M1-polarised signature comprised 3705 genes, whereas the M2-polarised signature contained 959 genes (data not shown). As can be appreciated in Figure 5.1, these figures reduced to 318 M1 and 129 M2 genes after the expression-filtering step. Some transcripts were overlapping in M1 and M2 signatures, which is expected as selection was based on both up and downregulation compared to the remaining experimental conditions. Typically, overlapping genes between the two activation states were upregulated in LPS/IFN-γ and downregulated in IL4-treated macrophages.

A comparative analysis between macrophage-polarised signatures and our own signature of 202 genes demonstrated that 99 genes were in common with an M1 transcriptome ($p=9.3 \times 10^{-42}$) and 34 ($p=1.3 \times 10^{-16}$) were shared with an M2 transcriptome. The majority of the overlapping genes were upregulated in DLBCL TAM compared to reactive LN macrophages, irrespective of the polarisation status (Figure 5.13).

An additional 83 genes were exclusive to the DLBCL TAM transcriptome. Interestingly, within this gene set, a significantly higher number of transcripts were downregulated in DLBCL TAM compared to controls: 29 of 133 genes in polarised signatures compared to 39 of 83 genes in unique signature, p=0.0001. Fifteen of these genes

were among the ones overlapping with the "resting"-macrophage signature. Whereas this kind of analysis supports the quality of our methodology and suggests that a bidirectional macrophage polarisation is present in DLBCL, among our DLBCL TAM signature are relevant genes that have not been retrieved in this comparative analysis, including some FcγR genes (FCGR1A, FCGR1B, FCGR2A), the macrophage attractant CCL2,[221] or the prototypic human M2 gene CTSD.[195] Additionally, specific genes have been functionally validated in murine models studying TAM, such as GPNMB[235] or C1QC.

5.5 Discussion

Our aim for this chapter was to recognize changes in gene expression within macrophages selected from LN infiltrated with DLBCL compared to macrophages selected from reactive LN. The underlying premise for using a high throughput approach was that, contrary to targeted protein studies, GEP could improve our understanding of the macrophage functional variability in DLBCL. To date there is a lack of global GEP from TAM selected from human cancer, thereby most assumptions regarding TAM functional features are derived from inbred mouse tumour models.

The choice of DLBCL samples was totally dependent on availability of enough vials to perform cell sorting. Under this limitation, cases included in this study are not as representative of DLBCL heterogeneity as we wanted, with a bias towards a better biology group (low median age at diagnosis - 59 years - and 77% of cases with low/low-intermediate IPI). Age-matching control samples would be ideal in any study of immune cell function, but was impossible, since most LN reactive pathologies are diagnosed in young adults. These limitations raise even more the requirement of future validation of these findings in larger datasets.

CD36 is a membrane scavenger receptor that participates in the recognition of apoptotic cells,[344] thereby assuming an important role in maintaining homeostasis in

tissues populated by B-cells with a high proliferation index. The expression of CD36 is enhanced in macrophages by IL-4[345] and CSF-1, traditional M2 cytokines. Using mouse and human CD36-deficient macrophages, Huang et al. proposed that this receptor is required for the acquisition of phenotypic and functional attributes of M2 activation.[346] This suggests that macrophage isolation based on CD36 expression is biased towards an alternatively activated population. Using flow cytometry, we saw that CD36 is homogeneously expressed in the whole population of CD68+ macrophages in tonsils, therefore being an excellent marker for cell sorting. Moreover, expression of well-recognized M1 markers was also detected in the transcriptomic analysis, suggesting that CD36-expressing macrophages might not be on the extreme of cell activation spectrum but have a mixed activation pattern adjusted to the requirements of their vicinity.

Even using a single marker, therefore minimizing the risk of cell loss during sample preparation, CD36+ cells are rare in stored samples of both lymphoma and reactive samples. It is possible that macrophages are more prone to cell death during storage and preparation. That should also be the case of malignant B-cells, since we found that, although these should be the most abundant cells in DLBCL samples, T-cells outnumbered them. To examine this we would have to test fresh samples in parallel with frozen samples and this is planned for future studies, but is not within the scope of this thesis.

Given that we were working with stored samples with poor viability and the low macrophage cell numbers obtained by flow sorting, it was of paramount priority to perform quality controls in all experimental steps of this study. Despite these features, except for sample R8639, which had a low but still acceptable RIN for GEP studies, RNA quality was excellent for all the remaining samples, supporting our protocol for cell isolation and RNA extraction when working with rare cell populations.

Even though samples with as low as 50 ng of cDNA are feasibly studied by GEP, not all samples in our study reached that amount of RNA. Therefore, cDNA amplification had to be performed. The Ovation Pico WTA System used here for cDNA synthesis and amplification has some unique features, including: (1) the use of random primers scattered along the transcriptome, which circumvents the amplification bias towards 3' regions of the gene when only oligo-DT primers are used; and (2) the use of a linear amplification method that retrieves multiple copies of single stranded DNA that is complementary to the mRNA and can be directly hybridised to probes immobilized on various microarray platforms. In a comparative analysis of amplifications methods for low RNA yields, it has been shown that the WT-Ovation pico system was the most suitable, providing reproducible results and good quality transcriptome analysis.[347] Indeed we obtained adequate electrophoretic traces after amplification. Moreover, none of the samples was flagged as defective in the quality control performed after array hybridisation, suggesting that the experimental work-flow undertaken was well thought.

Sample R8639 had the lowest RIN and the most variable relative log expression. Since no formal recommendation was given by the quality control checks to consider this sample an outlier, we did not exclude it in the current study. Yet, given the known impact of potential outliers in array data normalisation, leading to decreased statistical power and biological significance of results, it would be worth performing a new analysis in the future excluding R8639.

Unsupervised hierarchical clustering did not resolve DLBCL TAM samples from reactive macrophage samples. This, in itself is an interesting finding indicating that macrophage heterogeneity in DLBCL should be considered and bringing into question what would be suitable controls for this study. In any case, a perfect clustering of DLBCL and control macrophages would be surprising for two main reasons: (1) the cells from the microenvironment are being influenced by the tumour cells, but unlikely in a homogeneous fashion; and (2) as discussed in the introduction, macrophage functional heterogeneity is to be expected, even in tumour samples.

The supervised analysis identified 202 genes that were differentially expressed in DLBCL TAM relative to reactive controls. The identification of substantial gene expression variability in DLBCL TAM by unsupervised analysis suggested that corrected p-values would be modest and therefore the gene set would be small. Even among the genes that passed inclusion criteria, expression heterogeneity in DLBCL TAM was evident (Figure 5.9) and greater than that of control macrophages. Although these findings could potentially reflect technical problems with our data, we do not have clear evidence for that and instead interpret them as biological variability that is inherent of human samples, particularly in bystander cells of the tumour microenvironment. Going back to our original hypothesis of DLBCL cells influencing macrophage behaviour, it will be interesting to test whether such GEP variability of macrophages is related to unknown underlying differences of the tumour cells.

The accuracy of microarray platforms has improved substantially since the introduction of these technologies. Whole transcriptome profiling represents an opportunity to functionally model a disease and bioinformatics approaches can help in giving biological meaning to high throughput data. However, when it comes to teasing out the significance at the single gene level, confirmation of expression differences by other methods, such as qRT-PCR is highly advisable. With microarray technologies gene expression values can be widely affected by sampling, normalisation techniques and the analysis strategies used. As can be appreciated in Figure 5.10, GEP results could be validated by qRT-PCR, suggesting that the transcriptomic data is trustfully representing biological differences between the two groups. However we should bear in mind that amplified cDNA was used in the validation step. Ideally results should have been checked with non-amplified material. A change in the magnitude of fold change values when array and qPCR experiments have been compared has been noted in our data (Figure 5.10) and previously in the literature.[348]

Although our primary aim when performing validation by qRT-PCR was to ascertain

the validity of the transcriptomic data, we thought of testing whether the expression patterns of the eight genes studied could resolve DLBCL TAM from reactive controls. Surprisingly, qRT-PCR results for only eight genes that were among the highest differentially expressed of our dataset helped clustering DLBCL TAM samples apart from controls. This suggests that simpler and more feasible strategies can be used to explore differences between these two groups. Therefore, qRT-PCR will be taken forward for validation studies explained in Chapter 6.

A common approach to functionally contextualize transcriptomic data is to perform gene enrichment analysis. We used a free, online tool that performs a thorough comparative analysis using a large number of data sources. This, in turn, increases the depth of functional annotation of the data inputted.

The functional enrichment analysis undertaken here exposed critical elements of our data that sustain their quality and highlight their novelty:

(1) Functional annotations are macrophage-related

As can be appreciated, macrophages have instrumental roles in the GO cellular functions suggested, including innate immune and defence responses. Moreover, our gene set overlaps with gene signatures of myeloid cell populations, in particular monocyte and macrophage populations developed by the Human Immunology Project Consortium and the Mouse Immunological Genome Project.

Additionally, a quarter of our genes were represented in the “immune and inflammatory response” gene cluster developed using cancer transcriptomes. Importantly, the expression of genes of this group was detected in both GCB and ABC DLBCL datasets, which might suggest that relevant features of the microenvironment are independent of the molecular features of the malignant B-cells.

Furthermore, the top hit of this enrichment analysis was a GEP study on THRLBCL, a pathological DLBCL sub-entity where the bulk of the tumour is composed of macrophages.

We believe these findings provide additional support that our data not only reflects the transcriptomic features of macrophages, but also unique GEP changes characteristic of the DLBCL TAM. In consequence, our study has potential to generate hypotheses for functional validation of macrophage biology in human DLBCL.

(2) The LN signature of DLBCL is derived from transcriptomic features of TAM in the microenvironment

One of the most significant hits in the bioinformatics analysis was the overlap between our and the LN signature described by Rosenwald and collaborators,[76] providing compelling evidence that this signature genes result from transcriptional activation of the macrophages in the stromal microenvironment. This signature was linked to a better OS, contradicting what is found in other tumour models where TAM have a pro-tumoural behaviour.

(3) DLBCL TAM have a bidirectional M1 and M2 functional activation

A finding of utmost relevance is the differential expression of both M1 and M2 genes by DLBCL TAM. Evidence for this comes from different layers of our analysis. Over-representation analysis retrieved functional attributes of M1 and M2 activated macrophages.

Pro-inflammatory M1 cytokines, including IFN-γ, are likely more abundant in DLBCL than in reactive pathologies, as “Response to IFN-γ”, “IFN-γ signalling pathway” or “Inflammatory response” were enriched terms. Indeed, a significant upstream regulator of expression of our genes was the prototypical M1 cytokine IFN-γ. STAT1, an IFN-γ-induced transcription factor that induces expression of M1 genes was also detected.

However, all traditional M2 cytokines (IL-4, IL-13, IL-10) were significantly reported, suggesting that a number of our differentially expressed genes are M2-skewed. Most importantly, our data allows hypothesizing that B-cells are interacting with macrophages, driving them towards an alternative phenotype. Fcγ receptor activation by Igs is known to shift towards an M2 immunossupressive phenotype. This phenomenon might be in place in DLBCL, since “Ig binding” was the most enriched

GO molecular function identified and FcγR genes were highly overexpressed in DLBCL compared to controls.

The development of macrophage signatures served two purposes in this project. We wanted to reinforce two statements that could be withdrawn from the analysis already discussed:

(1) Firstly, that the methodology used is appropriate to explore the GEP of rare populations in human complex tissues. To support it, we compared our genes of interest with a list of macrophage-specific genes. Finding a significant enrichment would be an indication of accuracy of our results. In fact, a comparative analysis between a macrophage-enriched signature and our own signature revealed a highly significant overlap. To tease out the prognostic relevance of the expression of those genes, we used GEP data from R-CHOP treated patients and demonstrated that the 51 overlapping genes were variably associated with OS, with 22 being associated with improved OS. It would be interesting to explore the outcome impact of the remaining genes from our signature and develop multivariate prognostic models together with the IPI.

(2) Secondly, we would like to provide additional evidence for the functional heterogeneity of DLBCL TAM. Again reiterating this phenomenon, a comparative analysis between macrophage-polarised signatures and our own signature demonstrated that 99 genes were in common with an M1 transcriptome and 34 were shared with an M2 transcriptome, the larger part of them being upregulated in DLBCL TAM compared to control macrophages, irrespective of the polarisation status.

A large number of genes from our DLBCL TAM transcriptomic signature were, however, not represented in the macrophage signatures, reflecting the limitations of this kind of bioinformatics approaches. Some of those have an established role in TAM biology. However, others are quite novel and may open a window for the recognition of TAM specificities in DLBCL. Although we do not have a mechanistic

explanation for this, we found that within this unique gene set, a significantly higher number of transcripts were downregulated in DLBCL TAM compared to controls: 29 of 133 genes in polarised signatures compared to 39 of 83 genes in unique signature, p=0.0001.

Some relevant data could only be identified by performing literature searches, which highlights the problem of relying solely on annotated data for microarray interpretation. It is known that while two-thirds of all the genes are annotated by at least one functional annotation, the remaining one-third is yet to be annotated. We used individual gene names to complete the knowledge of their functional implications in the context of macrophage biology and searched for studies implicating pertinent genes and lymphoma biology. During this literature search two papers were found of most relevance for our work.

Hartmann et al.[208] performed GEP of minute amounts of macrophages selected by laser-microdissection from the DLBCL subtype THRLBCL. This study conveys key findings that overlap with our own and therefore helps to support some of our conclusions for this chapter:

(1) GEP changes in macrophages might be independent of the pathological subtypes of DLBCL and are not a mere reflection of the extent of infiltration of this cell type within the tumour microenvironment;
(2) LN macrophages have a high degree of transcriptomic resemblance, independently of pathologic alterations, impeding a clear-cut clustering by unsupervised hierarchical analysis;
(3) DLBCL macrophages overexpressed genes codifying for the metal-binding proteins MTs (overlapping with our signature MT2A, MT1G, MT1X), suggesting a role for these molecules in macrophage biology in DLBCL that could be functionally teased out;
(4) DLBCL TAM have a unique bi-directional M1 and M2 transcriptome.

Finally, by acknowledging fundamental differences that exist between ours and their study design, method of cell selection, array normalisation and statistical testing of differentially expressed genes, the coincident findings are more surprising and should be taken as biologically genuine.

To overcome the need for cell isolation, with inherent potential for inducing changes in the GEP of selected cells, bioinformatics analysis can be used to determine cell-specific gene signatures in the tumour microenvironment. Doig et al.[349] described GEP clusters reflecting stromal elements conserved across a variety of tumour datasets, including DLBCL. Using Lenz et al.[34] dataset, the authors discovered that well known macrophage-specific transcripts (e.g. CD14, CSF1R, CD163) and previously unrecognised macrophage-related genes showed very similar, potentially concerted expression profiles across all patients and other cancer datasets.

We found a very significant enrichment of our signature with transcripts belonging to this macrophage signature (39 of 161 genes, $p=7.8 \times 10^{-48}$), which again suggests:
(1) that our approach to define the DLBCL TAM transcriptome is valid and is likely providing authentic data despite all potential biases introduced by technically manipulation;
(2) whereas the expression of the common genes between ours and the macrophage signature might be, as the authors state, an expected finding resulting from the relatively higher abundance of macrophages in the microenvironment of DLBCL compared to reactive LNs, other differentially expressed genes can potentially reflect the uniqueness of the DLBCL TAM;
(3) these data similarities also validate in silico approaches to explore the function of cell subsets that are hard to study due to limited amounts or sensitiveness to laboratory manipulation, including macrophages.

The impact of stromal cells in transcriptomic features of whole tumours is indisputable. The microenvironment contribution on GEP reflects not only the extent of infiltration of a given cell and the relative abundance of all cells on the

microenvironment, but also likely the functional attributes of those same cell populations within the tumours. All transcriptomic analysis of DLBCL tumours described the impact of stromal and macrophage transcripts for prediction outcome modelling, before and after the introduction of R-CHOP. Our data provides novel insights on the functional heterogeneity of macrophages in DLBCL.

Chapter 6 Attempts for functional validation of transcriptomic analysis

6.1 Introduction

As already discussed in Chapter 1, independent transcriptomic studies demonstrated that whole DLBCL tumours display a distinctive macrophage and stromal remodelling gene signature that distinguishes these patients from patients with other histological subtypes of NHL. In our previous chapter we showed that the GEP of macrophages selected from DLBCL LN exhibits significant similarities with other transcriptomic analyses undertaken using DLBCL human samples. Therefore, it is conceivable that the malignant DLBCL cells are influencing macrophages towards this specific transcriptome, either directly or indirectly by inducing changes in other cells of the microenvironment.

A limited number of publications tried to explore this hypothesis in functional studies. Mueller et al. used co-culture experiments to provide a mechanism by which human DLBCL B-cells can directly influence monocyte function.[350] The authors suggested that B-cell-derived CCL5 was the culprit monocyte attractant to tumours. The later in turn sustained normal and malignant B-cell survival and proliferation through production of B-cell activating factor (BAFF) and IL-2.

Lin et al. explored functional changes of CD14+ peripheral blood monocytes in patients with DLBCL.[351] The authors demonstrated that immunossupressive HLA-$DR^{-/low}$ monocytes are expanded in DLBCL and are able to decrease T-cell proliferation and Th1 responses to foreign antigens. Although likely, this study does not formally demonstrate that DLBCL B-cells drive monocyte immunossupressive functions.

Using follicular lymphoma as a model, Guilloton et al. showed that malignant cells can induce the production of CCL2 by autologous mesenchymal stromal cells, which in turn attracts and skews monocytes towards an IL-10 secreting immunossupressive phenotype.[221] This effect could be abolished by CCL2 inhibition. The same group was

able to establish another mechanism by which macrophages favour B-cell proliferation.[284] Monocytes and monocyte-derived macrophages could trans-present IL-15 to B-cells, and in cooperation with T-cell derived CD40L, favour B-cell proliferation through STAT5 activation.

In this chapter we worked under the hypothesis that DLBCL TAM transcriptome and proteome is shaped by the malignant B-cells.

6.2 Aims

For this chapter we aimed to perform a functional validation towards recognizing the effect of malignant B-cells in macrophage GEP and protein phenotype. The experimental design was based on the findings from Chapter 5.

6.3 Materials and methods

6.3.1 Samples

PBMCs from healthy donors were isolated from buffy cones as described in section 2.1.3 and used to select monocytes. Tonsil SCSs were obtained from our tissue bank storage, prepared as described in section 2.1.2 and used to select control B-cells. Two DLBCL cell lines were selected for co-culture experiments based on their molecular profile. The GCB cell line Su-DHL4 was gifted from Dr. A Letai and the ABC cell line Ri1 was gifted from Dr. M Capasso. Both cell lines were maintained as described in section 2.1.4.

6.3.2 Co-culture experiments

Healthy CD14+ monocytes were positively selected from PBMCs using microbeads as described in section 2.8.1. Monocyte purity checks were performed by flow cytometry using CD14 antibody (Pacific Blue™ anti-human CD14 antibody, clone

M5E2, eBiosciences). Monocytes were matured for 7 days in petri dishes at $4.5x10^6$/ml concentration in DMEM medium with 10% human AB serum. The adherent cell fraction was then harvested using a cell scraper, washed and re-adhered at a concentration of 2 - 5 $X10^6$/ml at 37^oC into 24-well plates for 24h prior to co-culture set-up (see section 2.8.3).

Reactive B-cells were positively selected from tonsil SCSs as detailed in section 2.8.2 and resuspended in DMEM medium with 10% human AB serum. These and DLBCL cell lines with acceptable viability (>85%) were layered in triplicate over the adherent macrophages at a 1:1 concentration or placed in transwell inserts as detailed in section 2.8.4. Macrophages treated with 100ng/ml of LPS were used as positive control. Macrophages cultured alone constituted our negative internal controls for each biological experiment.

After 24h, co-culture supernatant was centrifuged and stored at -20^oC for cytokine quantification (see section 2.9). Adherent macrophages were harvested and used for surface marker analysis using flow cytometry (see section 2.4.2). Antibodies used are described in Table 6.1. A minimum of 10000 events gated on compensated viable-singlet cells were acquired using the four laser BD Fortessa flow cytometer. Results are expressed as the difference of MFI compared to macrophages cultured alone in each individual experiment.

Additionally, macrophages were sorted using size and intracellular complexity allowing clear separation from contaminant B-cells. After sorting, RNA was extracted from each cell pellet, assessed for quality, converted to cDNA and analysed by qRT-PCR. NanoDrop spectrophotometer 260/280 values were comparable across all samples and were within the accepted ranges for good quality RNA (data not provided). CT values generated by the qRT-PCR reaction were normalised to GAPDH and converted to RQ values as described in section 2.7.2.2. The TaqMan® Gene Expression Assays tested were described in Chapter 5.

6.3.3 Statistical analysis

Statistical analyses were performed with GraphPad Prism software using the Students T-test or Mann Whitney U-test when appropriate.

Table 6.1 Antibodies used in this study

Marker	Clone	Isotype	Fluorochrome	Source
CD80	L307.4	Mouse IgG1, κ	PE	BD
CD86	2331 (FUN-1)	Mouse IgG1, κ	FITC	BD
PD-L1	MIH1	Mouse IgG1, κ	APC	BD
HLA-DR	G46-6	Mouse IgG_{2a}, κ	APC-H7	BD
CD36	CB38	Mouse IgM, κ	APC	BD
Fixable viability dye			eFluor® 780	eBioscience

6.4 Results

6.4.1 Investigating changes in the expression of macrophage activation markers after co-culture

Given we had no previous experience with this experimental model, the first approach was to investigate the dynamics of expression of macrophage activation markers after co-culture. None of the markers explored was found to be differentially expressed in our GEP studies. However we thought they could be informative of the utility of the model.

As can be appreciated in Figure 6.1, LPS significantly enhanced the expression of CD80, CD86 and HLA-DR, suggesting macrophage responsiveness and that this model could be valuable to test the influence of B-cells.

LPS significantly increased CD86 expression compared to untreated macrophages and all remaining co-culture conditions. Whereas reactive B-cells cultured in contact with macrophages had no significant impact on CD86 expression, all other experimental conditions significantly decreased CD86 expression. This is visually clearer in the right side graph of Figure 6.1 A, representing CD86 expression dynamics for each replicate. Co-culture with the ABC cell line Ri1 produced the most significant reduction compared to untreated macrophages and macrophages cultured with reactive B-cells.

Likewise, CD80 expression was upregulated by LPS. However the magnitude of change was more heterogeneous across biological replicates. In two of six of the replicates, LPS decreased CD80 expression. In fact, these two biological replicates display a different expression dynamics suggestive of being less responsive in co-culture. Otherwise, experimental conditions were associated with enhanced expression of CD80, particularly in contact with B-cells and the Ri1 cell line.

HLA-DR expression varied in a more heterogeneous fashion in all experimental conditions, and contrary to the results of the other markers, HLA-DR did not exhibit a normal distribution (data not shown). Even applying the Wilcoxon matched-pairs ranked test we did not find any significant results, supporting experimental heterogeneity.

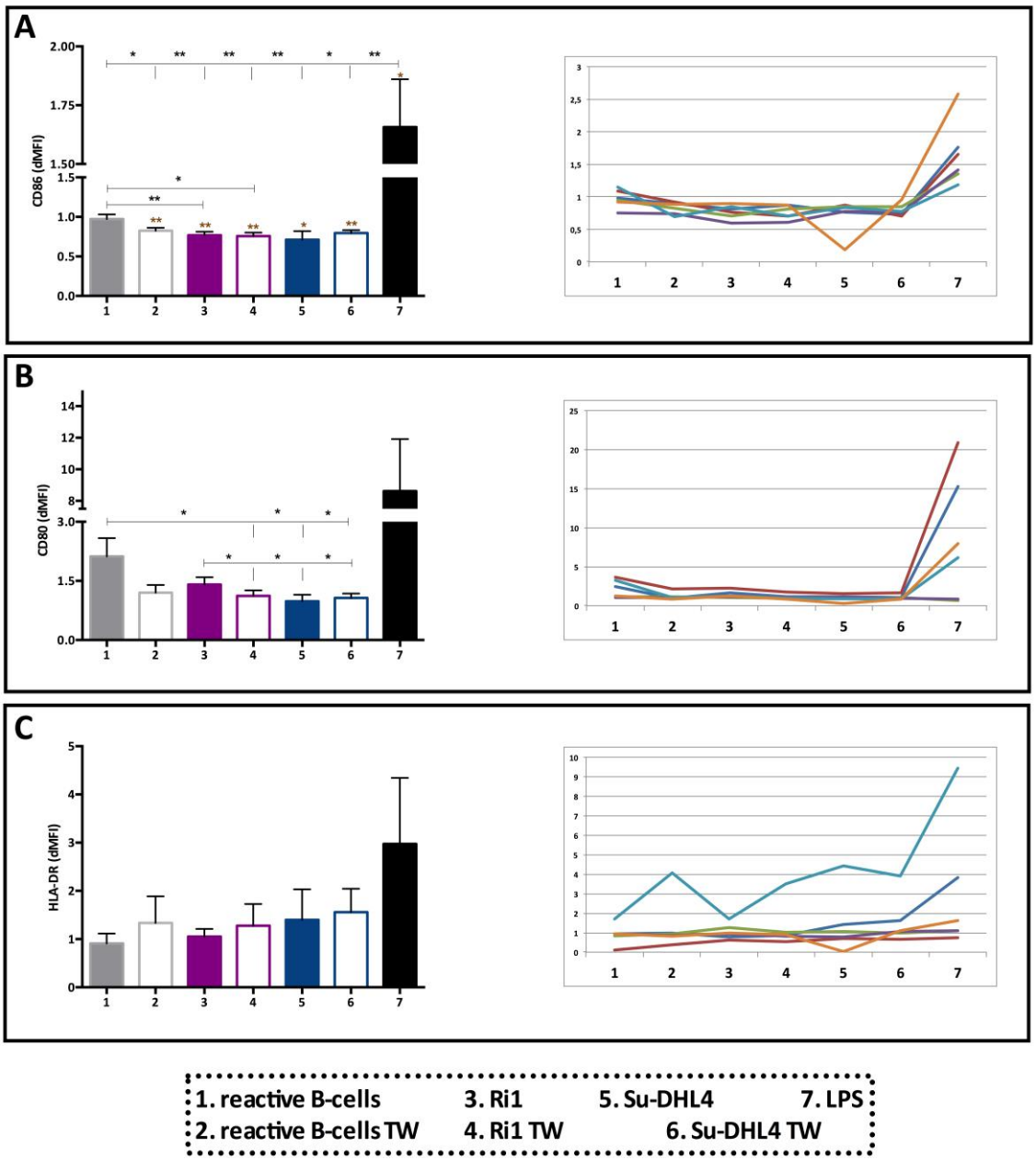

Figure 6.1 Expression dynamics of activation markers CD80 (A), CD86 (B) and HLA-DR (C) in macrophages harvested from co-culture.

Samples were incubated with conjugated mAbs for 30mins on ice in the dark and analysed by flow cytometry. The left side graphs represent the change of MFI compared to macrophages cultured alone (median +- SEM of six independent experiments). The right side graphs portray the expression dynamics according to each biological replicate. Paired Students T-test with *p <0.05, ** p <0.005. Brown symbols represent comparisons with control macrophages.

6.4.2 Exploring whether changes in gene expression found by GEP can be mimicked using the co-culture system

In order to comprehend the expression dynamics of selected transcripts in macrophages in this co-culture model we had to sort them after co-culture. Although several methods were used to eliminate B-cells from the harvested cell pellet (from extensive washing to treatment with trypsin or less aggressive digesting solutions), contamination with B-cells could not be avoided (confirmed by flow cytometry, data not shown). Macrophages were selected according to size and intracellular complexity using the forward and side scatter pattern.

The transcripts studied in three independent biological replicates were chosen from the top differentially expressed genes between DLBCL TAM compared to reactive macrophages in GEP studies (Chapter 5). The transcripts studied were *AQP9*, which was among the most upregulated; and *IDO1* and *CDH1*, both among the most downregulated in DLBCL TAM compared to reactive macrophages.

Given the necessity to select macrophages after co-culture, we firstly explored whether the sorting affected gene expression results.
Indeed, relevant changes occurred after flow sorting in expression of the transcripts investigated. However, this effect did not seem to be similar neither across genes nor across samples. mRNA levels for the three assays changed in an heterogeneous fashion, as is illustrated in Figure 6.2.

Next we are going to describe the findings for each of the three genes studied using qRT-PCR.

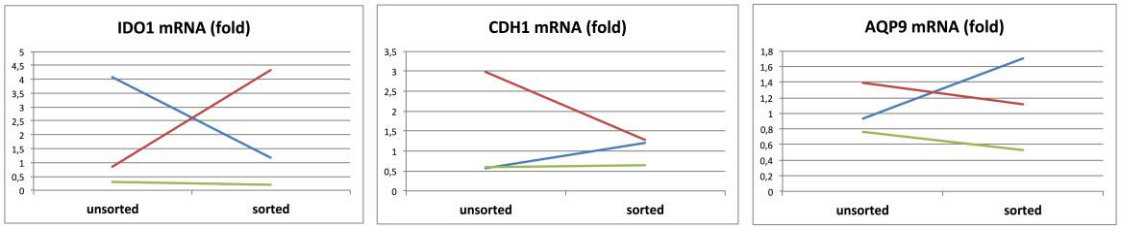

Figure 6.2 Targeted mRNA expression is affected by sorting after co-culture.
50ng of amplified cDNA was used in triplicate for each sample and analysed by qRT-PCR. CT values were normalised to GAPDH. RQ values, calculated using as calibrator the average ΔCT of untreated macrophages, are represented for unsorted and sorted macrophages.

AQP9 belongs to the family of aquaporins (AQP) which are water-selective membrane channels critical for the regulation of cell volume and migration.[352] AQP9 is the most prominent AQP in human inflammatory cells,[353] including macrophages, and was the sixth most upregulated gene in DLBCL TAM compared to reactive LN macrophages in our array studies (Log^2 FC 3,35, adjusted p-value 0.003).

AQP9 is among the transcripts defining an M1 signature (Figure 5.13) and hence should increase when macrophages are triggered by LPS. Accordingly, we were able to document a rise in this membrane channel in macrophages under LPS treatment (Figure 6.3 A). Compared to untreated macrophages, all conditions induced a decrease in *AQP9*, with the exception of two of the three samples when in contact with B-cells. As can be appreciated in Figure 6.3 A, right side panel, reactive B-cells did not induce a significant change in *AQP9* expression when in contact due to sample heterogeneity, but decreased it when in transwell in all samples.

The second molecule investigated, e-cadherin (*CDH1*), was among the most downregulated genes in our TAM signature (Log^2 FC -2,6, adjusted p-value 0,03). Van den Bossche et al. described that *CDH1* expression is enhanced in macrophages exposed to the prototypic M2 cytokine IL-4 and IL-10.[354] Mechanistically, e-cadherin was shown to be involved in macrophage interactions with T-cells in a Th2 immune context. In addition, M1 stimuli repress *CDH1* in macrophages.[354]
CDH1 expression in untreated macrophages was variable, as can be appreciated in

Figure 6.3 B. In all experimental conditions a decrease of *CDH1* mRNA levels was detected compared to controls particularly in contact conditions with both reactive and B-cell lines. As predicted, the M1 stimuli LPS induced a decrease in *CDH1* expression.

Finally, we investigated the expression dynamics of indoleamine 2,3-dioxygenase 1 (*IDO1*), a catalyser of the rate-limiting step in the immunoregulatory pathway of tryptophan catabolism. IDO1 activity in human macrophages is triggered by direct interaction with CD4+ T-cells and controlled by the M1 cytokine IFN-γ.[355]

Contact with reactive B-cells led to an important increase in *IDO1* mRNA levels in macrophages, and even more so with LPS treatment (Figure 6.3 C), corroborating published data. All other conditions induced a more depreciable and heterogeneous increase in *IDO1* transcript levels compared to untreated macrophages. We also inspected IDO1 protein expression and results will be discussed in our next chapter.

Moreover, we addressed whether we could mirror, using this co-culture model, the expression changes documented in our transcriptomic analysis in two other genes, *CTSL1* and *C3AR1*. Both genes increased in all conditions compared to untreated macrophages, with no significant differences detected (data not shown).

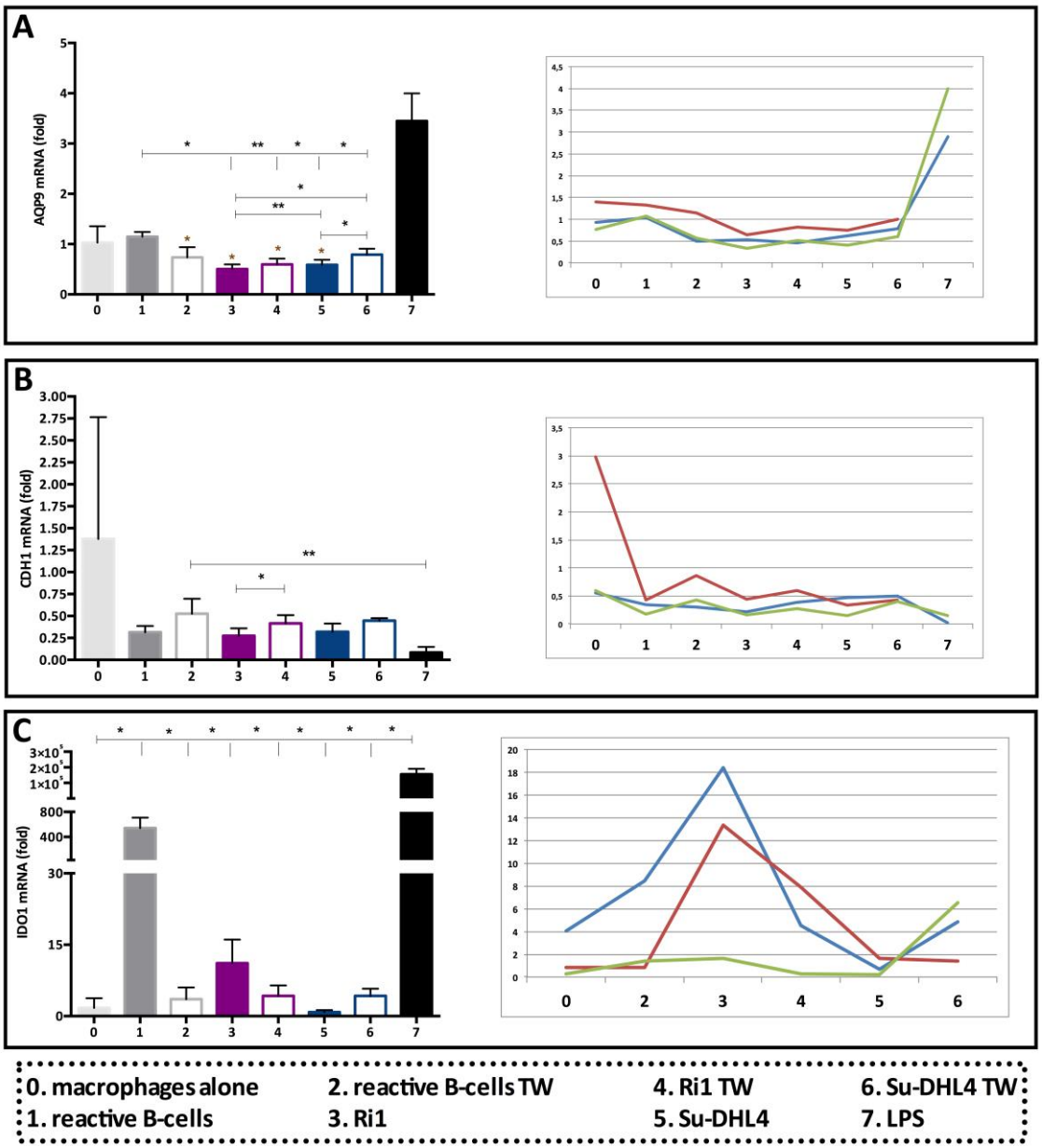

Figure 6.3 Gene expression changes of targeted genes in macrophages harvested from co-culture.

50ng of amplified cDNA was used in triplicate for each sample and analysed by qRT-PCR. CT values were normalised to GAPDH. The left side graphs represent RQ values (median +- SEM of three independent experiments). The right side graphs depict the expression dynamics according to each biological replicate. Paired Students T-test with *p <0.05, **p <0.005. Brown symbols represent comparisons with control macrophages.

6.4.3 Expression dynamics of the immunoregulatory molecule PD-L1

TAM can exhibit immunossupressive functions (see Figure 1.2) through diverse mechanisms, from the production of soluble molecules, including IDO-1; to the direct crosstalk with T-cells through, for instance, e-cadherin or PD-L1. Macrophages can induce profound suppression of T-cell proliferation and cytokine secretion through crosstalk between PD-L1 (CD274) and the T-cell surface receptor PD-1.[228,237,356] According to published data, PD-L1 can be induced in macrophages by both M1 (LPS and IFN-γ)[357] and M2 (IL-4 and IL-10) cytokines.[358]

In our transcriptomic studies, PD-L1 was upregulated in DLBCL TAM compared to controls. We hence decided to investigate PD-L1 protein expression dynamics in macrophages after co-culture (Figure 6.4).

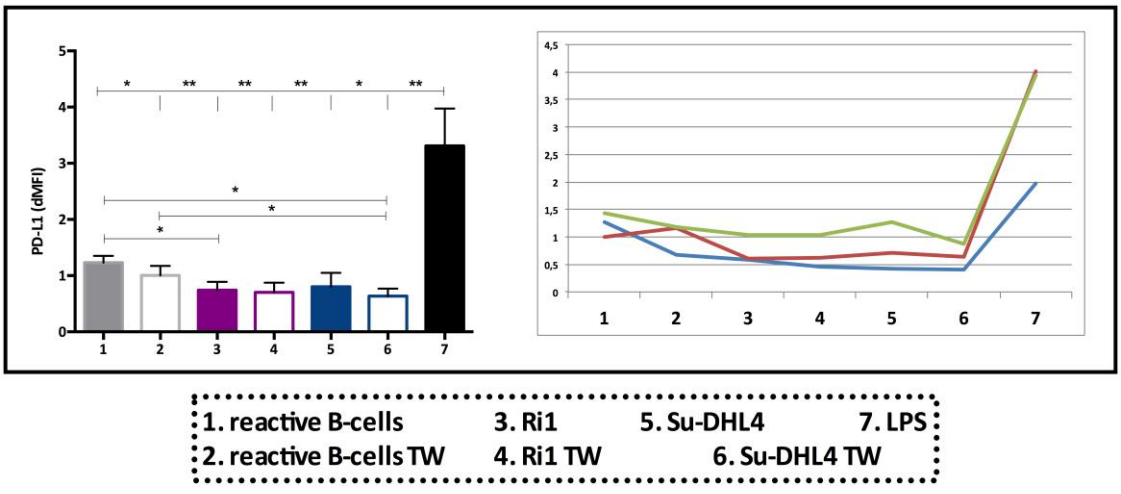

Figure 6.4 Expression of PD-L1 protein in macrophages after co-culture.
PD-L1 protein expression was tested using flow cytometry, results being the ratio of MFI compared to untreated macrophages. Paired Students T-test with *p <0.05, **p <0.005.

Using flow cytometry we confirmed that LPS induces an increase in PD-L1 basal levels in macrophages. Reactive B-cells enhanced, whereas DLBCL cell lines decreased PD-L1 expression. However responses were variable across the three replicates studied.

6.4.4 Investigating changes in CD36 expression in macrophages after co-culture

Given that CD36 was used for positive selection of macrophages in our GEP studies, we investigated the dynamics of expression of this scavenger receptor in macrophages after co-culture (Figure 6.5).

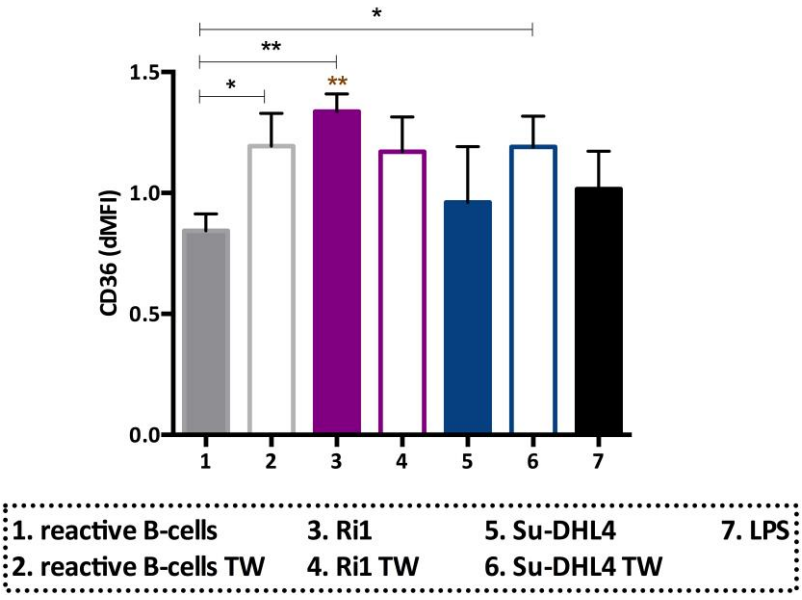

Figure 6.5 CD36 expression dynamics in macrophages harvested from co-culture.
Sample were incubated with an anti-CD36 APC-conjugated mAb for 30mins on ice in the dark, washed and analysed by flow cytometry. Bars represent the change of MFI compared to macrophages cultured alone. Mean and SEM of six independent biological experiments are displayed. Paired Students T-test with *p <0.05, **p <0.005.

CD36 expression was detected in 100% of macrophages and differences in expression intensity according to co-culture conditions were modest, as can be appreciated in Figure 6.5.

The Ri1 ABC cell line significantly increased CD36 expression in macrophages compared to control macrophages. Co-culture in contact with reactive B-cells was the only condition inducing a decrease of CD36 expression in macrophages compared to control. The magnitude of reduction in expression intensity induced by reactive B-cells was significantly different from the ones detected when reactive B-cells were cultured in transwell, when macrophages were cultured in contact with Ri1 cell lines or when the GCB Su-DHL4 cell line was placed in transwell.

6.5 Discussion

Over the last decade an abundance of transcriptomic studies have been reported in DLBCL. This high throughtput technique led to major findings in DLBCL biology. The COO model illuminated on the heterogeneity of this disease and served as starting point for a multiplicity of functional studies of genetic aberrations shown to be specific to each molecular subgroup.

On the contrary, whereas GEP of whole DLBCL tumours hinted towards the importance of a stromal response in disease biology and outcome, the lack of further functional validation of this hypothesis is remarkable.

GEP consensus clustering analysis performed by Margaret Shipp's group in 2005 delineated three groups of DLBCL independently of the molecular COO. Following this discovery, the group elegantly provided validation data and novel insights into disease pathogenesis. For example, Caro et al. demonstrated that the DLBCL cluster harbouring an oxidative phosphorylation GEP signature displayed enhanced mitochondrial energy transduction and glutathione levels and was, notably, selectively sensitive to manipulation of glutathione synthesis compared to the remaining patient's subsets.[359] We believe that with additional, more comprehensive validation, GEP signatures of the microenvironment in DLBCL can provide similar

insights into the biology of DLBCL.

Towards that end, we defined the transcriptome of the DLBCL TAM and suggested that macrophages are indeed implicated in the expression of prognostic immune signatures previously reported in this disease and hence constitute an excellent cell population to study at the functional level.

Under the hypothesis that differences in GEP detailed in Chapter 5 result from the influence of malignant B-cells on macrophages we used a co-culture system involving these two cell types.

Work presented in this chapter using the current co-culture model does not fully support this hypothesis.

In order to determine whether this model could be used to study the dynamics of macrophage phenotype, we firstly investigated the expression pattern of the macrophage activation markers CD80, CD86 and HLA-DR after 24h in co-culture.

The co-stimulatory molecules CD80 and CD86 are important deliverers of T-cell activation signals during immune responses. It is established that CD80 and CD86 help determining to what extent macrophages are skewed towards a pro-inflammatory phenotype, since their expression increases when macrophages are stimulated with LPS.[360] However, it has been suggested that the expression dynamics of both molecules is differential under IFN-γ stimulation, with CD86 expression increasing and CD80 expression decreasing.[360] Moreover, the kinetics of modulation is also different, with changes occurring between 18 and 42h after activation and with CD86 preceding CD80 expression after LPS treatment for approximately 12h.[361]

In our snapshot analysis at 24h, we indeed confirmed that, in most cases, LPS increased the expression of CD80, CD86 and HLA-DR. The lack of response to such potent activation in some replicates regarding CD80 and HLA-DR most likely represents inter-sample heterogeneity, given that the same samples responded to

LPS by increasing CD86 expression. These dampened responses could be attributed to a pre-existent activation state towards a pro-inflammatory phenotype in basal conditions or to a different kinetics with later responses in some cases.
After 24h, CD86 expression decreased in all experimental conditions compared to control macrophages, with only two biological replicates being exceptions when cultured in contact with reactive B-cells. On the other hand CD80 expression increased throughout, particularly when macrophages were cultured in contact with reactive B-cells and the ABC cell line Ri1. Overall no clear differences were induced by reactive B-cells compared to cell lines at 24h of co-culture and the relevance of this expression pattern is unclear. Since only one time point was applied, we are unaware of the kinetics of expression of these molecules, which could help clarifying the significance of these findings.

We then intended to demonstrate that macrophage gene expression changes revealed in our microarray studies could be mirrored using the co-culture model. We recognize that a more comprehensive approach using whole transcriptomic analysis or customized qRT-PCR arrays would probably be more useful to demonstrate this assumption.

Whereas B-cell contamination did not constitute a problem for protein studies using flow cytometry due to differences in cell size and internal complexity, it made macrophage sorting necessary before qRT-PCR. We demonstrated that targeted gene expression is affected by the sorting, a phenomenon that is remarkably under-reported in the literature. Changes were heterogeneous and suggested that the effect of cell sorting could not be predicted, nor was dependent on the biological sample. Our approach was to sort macrophages from all co-culture conditions. This was taken forward in three independent biological replicates.

We chose to study transcripts that: (1) were maximally differentially regulated between reactive and DLBCL macrophages; (2) have been previously implicated in T-cell regulatory functions; (3) have not been properly studied in DLBCL. If we could

prove a direct influence of B-cells on expression of these genes in macrophages, it would constitute an interesting mechanism of immunoregulation in DLBCL that should be experimentally tested.

In agreement with previously published data, *AQP9* and *IDO1* were upregulated and *CDH1* repressed by LPS. *CDH1* basal levels were extremely variable in the three samples tested. This biological heterogeneity can be related to differences in cytokine levels or oxygen tension in healthy donors.

The expression dynamics of *AQP9* and *CDH1* was similar, with a decrease in expression being induced in all experimental conditions compared to untreated macrophages. Contrary to our transcriptomic data, no clear differences arose from culture with reactive B-cells and DLBCL cell lines. Moreover, *IDO1* was clearly upregulated by reactive B-cells compared to cell lines, again not substantiating GEP findings. This discrepancy in the kinetics of expression of the M1 molecules *AQP9* and *IDO1* suggests that related cytokines such as IFN-y might not play an import role in this model.

We also attempted to demonstrate a differential effect of co-culture conditions in the expression of macrophage PD-L1 protein. However responses were variable across the three samples studied.

A number of aspects related to the experimental design can justify the heterogeneity of our results and, as a consequence, the inability to support our hypothesis.

For this investigation we used healthy human monocyte-derived macrophages that were matured in vitro in DMEM media and human serum. After our experiments were finalized an important consensus publication was published proposing a standardized methodology for macrophage studies that will allow achieving more reproducibility across experiments and laboratories.[213] The panel of experts suggested that monocytes should be matured in media with purified endotoxin-free recombinant CSF-1. We recognize that using different batches of human serum as the source of CSF-1 increases the potential of obtaining experimental variability.

Additionally it is likely that commercialized serum will contain variable amounts of IFN that can cause confounding effects. We tested this hypothesis and did not detect IFN-γ in supernatants (see following chapter). However, it has to be hypothesized that soluble factors present in human serum (including CSF-1) influence macrophage activation. Our choice to use human serum was based on data demonstrating that CSF-1 directs macrophages towards an M2 phenotype.

Additionally, peripheral blood monocytes from different donors (from which no clinical or demographical information is available) are likely to be differently activated.
We believe to have overcome these two limitations by establishing as controls unmanipulated macrophages that were cultured in the same circumstances in each biological experiment.

The main drawback of our co-culture model is the use of DLBCL cell lines. Although malignant cell lines provide valuable preliminary data, they have limitations. Their infinite growth in culture, independent of microenvironment signals, likely make them unsuitable cells to tease out the potential effect of cell interactions in vivo in the original tumour. Moreover they are unable to reflect patient tumour heterogeneity.

Finally, due to a complex experimental design and cumbersome procedural manipulations, the analysis had to be limited to a single time-point and to reduced read-out analysis. In our opinion this is an extremely important aspect that limits interpretation of the co-culture system.

It was entirely expected that such simplified model was limited to replicate the complex crosstalk that can be established between DLBCL malignant cells and the non-malignant cells in the microenvironment. An in vitro culture system using patient samples, which much more closely resemble the original tumour, would be a valuable tool to test our hypothesis. This was intended but deemed impossible in the current

studies. The extremely low viability on thawing of primary B-cells (data not shown, but indirectly illustrated by the modest RIN numbers) was not surprising, given that SCSs were previously stored, submitted to cell sorting and then re-stored.

A three cell (macrophages, B and T-cells in an autologous setting) co-culture model using fresh samples, and testing a larger panel of markers would help us to finally address our hypothesis.

Chapter 7 Exploring macrophage heterogeneity in diffuse large B-cell lymphoma

7.1 Introduction

As already mentioned, there are limited studies towards understanding the polarisation status and effector function heterogeneity of macrophages in human cancer. Our data presented in Chapter 4 highlights important differences between the DLBCL TAM and reactive LN macrophages. The GEP analysis undertaken also hints towards genuine macrophage heterogeneity within the DLBCL samples, not related to patient characteristics. This is an interesting finding, likely related to a distinct impact of the tumour cells on macrophage function that has potential prognostic value.

Remarkably, most GEP findings were hardly validated at the protein level, likely due to problems inherent to IHC analysis of macrophages. CD68 is a non-specific pan-macrophage marker highly difficult to analyse.[329] Diagnostic DLBCL samples have heterogeneous infiltrations of CD68+ macrophages. In chapter 4, we suggested that the extent of total macrophages has no impact on patient's outcome, contradicting the prognostic models developed using transcriptomic analysis of DLBCL LNs. This data exposes a major limitation of IHC studies to functionally characterise TAM in the microenvironment. CD163 has been claimed to specifically identify M2 macrophages. A subset of STAT1+ M1 macrophages has been identified using double IF staining in follicular lymphoma and linked to a worse outcome.[362] However recognizing global M1 and M2 populations using broad prototypic markers brings limited information regarding the detailed functional activation of macrophage subsets in the microenvironment. Identifying novel proteins, as SPARC has been identified, might be the way forward to recognize subsets of macrophages with more restricted functions, with an outcome impact potential and amenable to be targeted.

In this study we aimed to consolidate our previous finding that macrophages are functionally heterogeneous in the DLBCL microenvironment. And with the assumption

that the malignant B-cells are partially contributing to macrophage polarisation in DLBCL, we hypothesized that a deeper analysis of GEP of TAM could help to segregate different groups of DLBCL establishing different patterns of crosstalk with macrophages in the microenvironment.

Additionally, we predicted that novel markers discovered using GEP could help to better characterise the functional polarisation of TAM in the DLBCL microenvironment.

7.2 Aims

In this study we aimed at exploring the DLBCL TAM trancriptomic heterogeneity. After depicting subgroups within the TAM samples studied in Chapter 5, we intended to explore the cytokine and chemokine profile of corresponding malignant B-cells, hoping to illuminate on potential molecules partially responsible for the different macrophage polarisation statuses identified in the transcriptomic analysis.

In the same line, we wanted to reinforce the underlying macrophage heterogeneity by demonstrating that DLBCL TAM express previously unexplored M1 and M2 proteins.

7.3 Materials and methods

7.3.1 Samples

The transcriptomic analysis was performed using samples described in Chapter 5.

IHC and IF studies were performed using FFPE tissue from DLBCL patients treated with R-CHOP at St. Bartholomew's Hospital. Patient's characteristics have been described in Chapter 3.

7.3.2 Cytokine studies

CBAs were used to test for the presence of cytokines and chemokines in culture media. These studies were undertaken using samples from the co-culture experiments explained in Chapter 6 and from cell line supernatant. A detailed methodology is given in Chapter 2.

7.3.3 IHC Staining and IF analysis

Single protein immunostainings were performed as previously described. IF was used in this study to detect co-expression of two proteins of interest in macrophages in FFPE tissue of DLBCL and reactive LNs. A detailed staining protocol is provided in section 2.3. Primary antibodies and conditions of use are described in Table 7.1.
Slides were scanned using an Olympus BX61 microscope and analyzed using the Ariol SL-50 visual analysis software. Scanned slides were observed on a computer screen. Single channel and overlapping images were used for final analysis quantifying single and double positive cell counts.

7.3.4 Statistical analysis

Statistical analyses were performed with GraphPad Prism software using the Students T-test or Mann Whitney U-test when appropriate.

Table 7.1 Antibodies used in this study

Antibody	Clone	Species	Dilution	Source
CD68	PGM1	Mouse	1/400	Dako
CD163	10D6	Mouse	1/2800	Leica Biosystems
ALOX15	3D8	Mouse	1/250	AbD Serotec
IDO1	polyclonal	Rabbit	1/50	Sigma-Aldrich

7.4 Results

7.4.1 Transcriptome heterogeneity in DLBCL TAM

The gene and protein expression studies just described suggest that the co-culture model developed does not mirror the GEP data. This was entirely expected, and limitations of the model were already discussed.

In this context, we considered addressing the previous finding of macrophage GEP heterogeneity within DLBCL TAM.

Microarray unsupervised hierarchical clustering based on the 12000 most variables probesets of the 13 samples studied showed that DLBCL TAM agglomerated in two main clusters (Figure 7.1 A). No differences in clinical features were detected between these two groups (data not shown).

We hence thought of interrogating for differential expressed genes between the two groups (identified in Figure 7.1 A and B in gray and blue). Using a FC ≥3 and adjusted p-value ≥0.05 we identified 150 well-annotated genes that were differentially expressed between the DLBCL TAM groups. The genes we investigated in our co-culture model were among the differentially expressed genes between the two groups, as can be visualized in Figure 7.1 C. This could be expected by simply inspecting the two-dimensional scaling image, showing that the DLBCL TAM population 1 (in gray) clusters closer to reactive LN macrophages. Indeed, both *CD1E* and *CDH1*, found to be upregulated in reactive controls versus the whole DLBCL TAM group, were both upregulated in the DLBCL TAM population 1. To test the degree of similarity between the reactive macrophages DLBCL TAM population 1, we looked at differentially expressed genes between the two groups. Top transcripts from our original analysis were still differentially expressed, including FCγR genes, MT genes, *ANKRD22, AQP9, IDO1, C3AR1* or *CTSL1*.

7.4.2 Investigating soluble factors differentially expressed between DLBCL TAM groups

As has already been described, it is likely that malignant B-cells influence macrophage polarisation and function through the production of soluble factors. Cytokines and chemokines are deregulated in tumours. These can contribute in an autocrine fashion to tumour growth, but can also affect the microenvironment.

We hypothesized that the differences documented in DLBCL TAM GEP profiles relate to differences in the cytokine and chemokine profile of corresponding malignant B-cells.

To this end we selected, among the pure (>95%) B-cell samples stored from DLBCL and reactive LN, three samples from each group (group 1: R0433, T2628, T4570; group 2: T6932, R8756, R9516) and two from reactive B-cell samples (T5900, T5996) and performed whole trancriptome analysis using Affymetrix U133 Plus 2.0 arrays. A targeted analysis of cytokines, chemokines and adhesion molecules was undertaken.

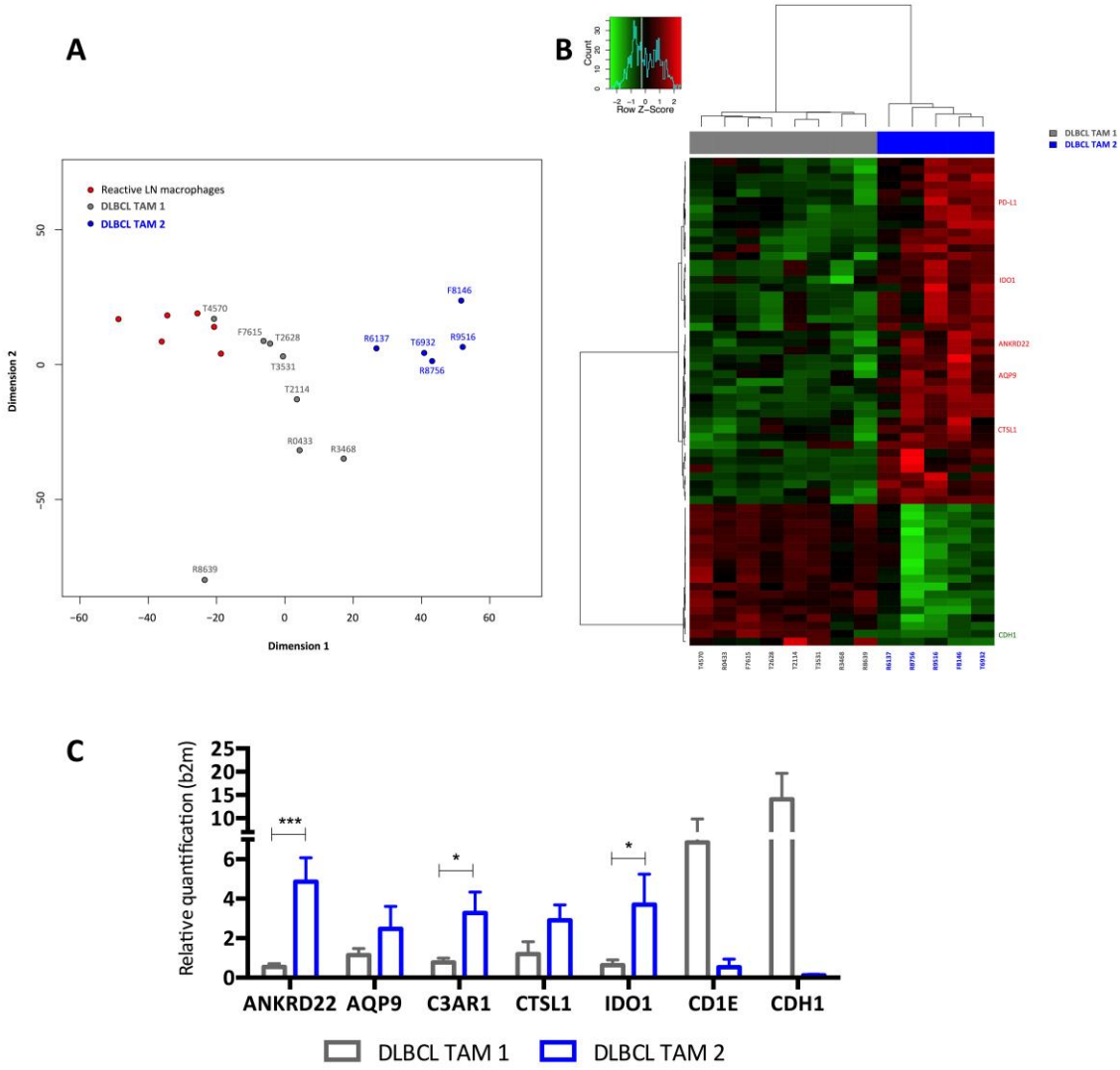

Figure 7.1 GEP analysis suggesting two groups of DLBCL TAM.

A. Two-dimensional scaling based on expression of 12,000 genes, distances between samples calculated using Euclidean distance metrics. B. Heatmap representing differentially expressed genes between the two patient groups with a FC≥4 and adjusted p-value ≤0.05. Euclidean distance and Ward linkage applied. C. qRT-PCR results for targeted genes comparing the two DLBCL TAM groups. Bars represent mean + SEM, unpaired T-test, * p-value ≤0.05, ** p-value ≤0.005.

Within the archetypal molecules involved in macrophage activation, the expression of M2 cytokines IL-4, IL-13 and IL-10 or IL-6 was not significantly different between B-cells of both DLBCL groups and controls. On the contrary, IFN-γ, the prototypical M1 cytokine was the most differentially expressed soluble factor between DLBCL B-cells, with an absolute FC of group 1 versus group 2 of -9.4 and an adjusted p-value of $6.4X10^{-5}$ (Figure 6.7).

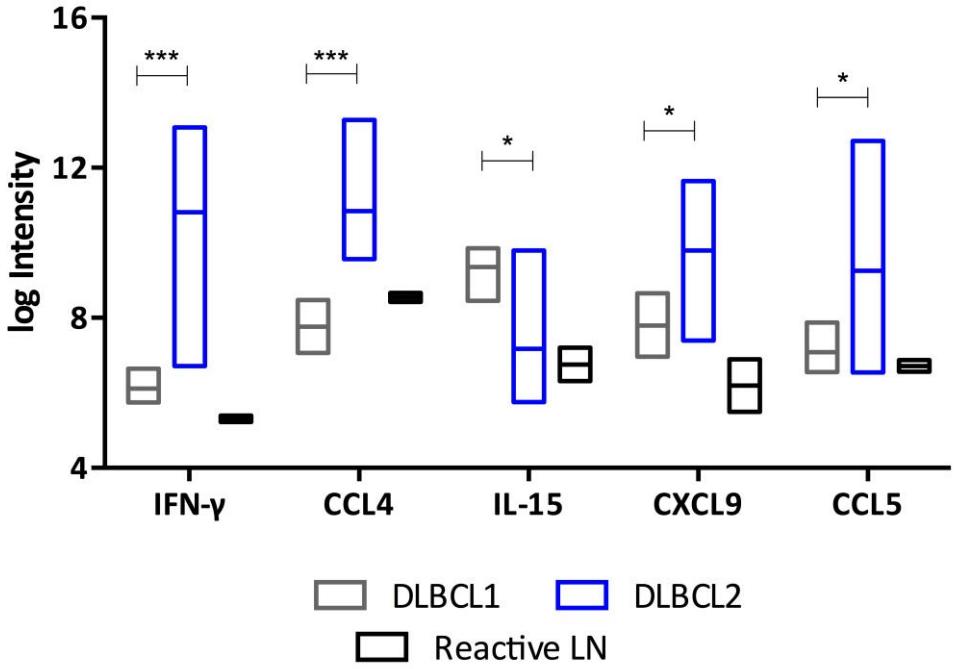

Figure 7.2 Differentially expressed soluble factors between B-cells.

Initial RNA integrity was comparable across samples (RIN between 7.4 and 8.9). 12.5 μg of fragmented and labelled aRNA were hybridised to Affymetrix U133 Plus 2.0 arrays according to manufacturer's recommendations and analysed in a single batch. Data was normalised using RMA and filtered using 20% standard deviation of signal intensity as threshold on O-miner (http://o-miner.org/onlinetool/index.html). Differentially expressed genes were tested using the limma method. Boxplots represent the log intensity for each group, and significance was tested using BH FDR.

We hence explored whether there was a predominance of M1-skewed genes overexpressed in the DLBCL TAM group 2 compared to group 1. We confirmed that DLBCL TAM group 2 preferentially upregulated M1 genes compared to group 2 (56 of 62 M1 genes upregulated in group 2 versus 5 of 62 in group 1).

As can be appreciated in Figure 7.2, the two reactive B-cell samples profiled showed a much lower and tighter expression of IFN-γ, CCL4 or CCL5 compared to DLBCL samples. CCL5, previously implicated in interactions between DLBCL cells and monocytes was also significantly upregulated in group 2 compared to group 1. CCL4 expression between DLBCL groups was also strikingly different. Although CCL4 is predominantly produced by macrophages, it can be secreted by normal and malignant B-cells upon BCR crosslinking, suggesting that differences in serum levels of this chemokine can be expected in DLBCL according to oncogenic BCR signalling activation.

We then investigated whether DLBCL cell lines exhibited constitutional secretion of IFN-γ or CCL5 using CBAs. This experiment had as objective identifying cell lines with a discrepant cytokine profile that could be used in co-culture models instead of primary cells.

Using CBA as described in Chapter 2, DLBCL cell lines did not seem to secrete either IFN-γ or CCL5 in the conditions tested.
IFN-γ was undetectable in all conditions (data not shown). The two culture media where DLBCL cell lines are generally maintained, RPMI or IMDM with FCS, had undetectable CCL5 (Figure 7.3). However, human serum utilised to mature macrophages in vitro and in our co-culture system, contains traceable amounts of CCL5. DMEM with human serum had similar levels of CCL5 as the same media extracted from wells where cell lines where grown for 24h. DMEM media alone had undetectable CCL5.

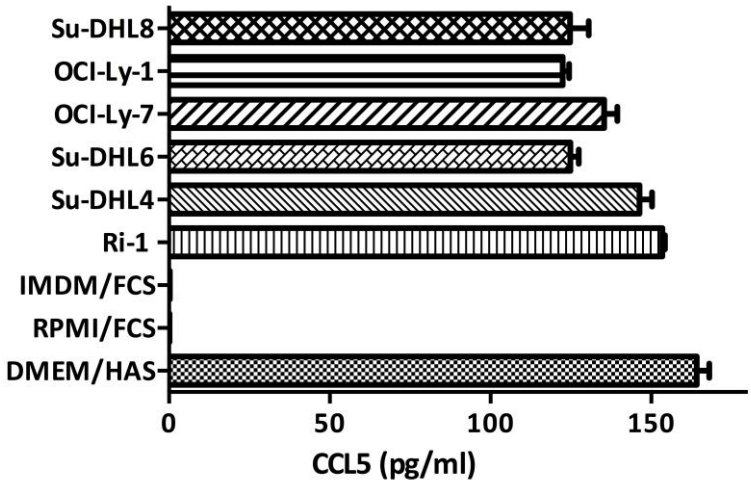

Figure 7.3 CCL5 is detected in the cell culture media used in co-culture experiments.
DLBCL cell lines were cultured in triplicate in 24 well plates in DMEM and human AB serum (HAS) for 24h, in an attempt to mimic co-culture conditions. For each condition, 50 ml of culture media was tested in triplicate using CBAs. Media only was tested in parallel. Bars represent the mean and SEM.

In this context, we treated macrophages for 24h with the five soluble factors described in Figure 7.2 and analysed the expression of CD80 and PD-L1. The objective of this experiment was to investigate whether such factors could induce expression changes similar to the ones obtained in our co-culture models.

In the absence of standardized protocols for macrophage stimulation with some of the molecules, different concentrations of CCL4 (10 and 20 ng/ml), CCL5 (1 and 5 ng/ml) and IL-15 (50 and 100 ng/ml) were used. The data presented refers to the dose inducing relevant expression changes compared to untreated macrophages. In this pilot experiment (Figure 7.4) we confirmed our previous and others data describing LPS and IFN-γ as potent activators and CD80 and PD-L1 as M1-induced proteins in macrophages. With the exception of LPS and IFN-γ, the added molecules produced modest changes compared to untreated macrophages. However, those changes were within the range of the ones induced by reactive B-cells or DLBCL cell lines (see Figures 6.1 and 6.3).

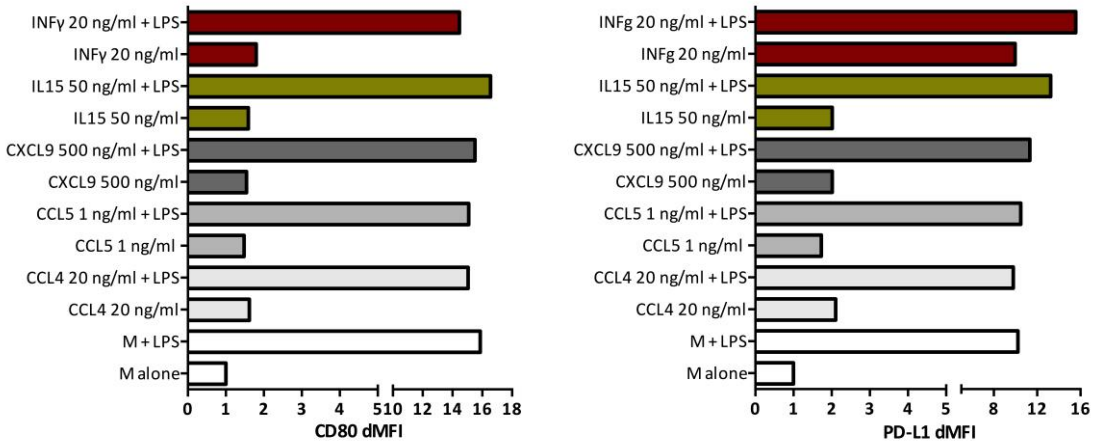

Figure 7.4 Expression of CD80 and PD-L1 in macrophages treated with soluble factors.
Cells were harvested from 24 well plates 24h after treatment, stained for CD80 and PD-L1 by flow cytometry. Bars represent the difference in MFI compared to untreated macrophages from one biological replicate.

All treatment conditions induced an increase of CD80 and PD-L1 MFI compared to controls. We found the same effect on CD80 in all co-culture settings, but a different effect on PD-L1, which expression decreased when DLBCL cell lines were added in contact or in transwell.

7.4.3 Investigating Th1 and Th2 cytokines in co-culture supernatant

Finally we investigated whether we could identify Th1 and Th2 cytokines in co-culture media. IFN-γ was undetectable in all five experiments. The Th1 cytokine TNF-α and the Th2 cytokines IL-6 and IL-10 could be identified in our model (Figure 7.5). Stimulation with LPS led to a considerable secretion of these cytokines by macrophages. Although in a significantly lower concentration IL-6, IL-10 and TNF-α were traced in supernatants from macrophages and reactive B-cells in contact conditions. When reactive B-cells were cultured alone, none of the three cytokines could be identified. All remaining conditions were negative.

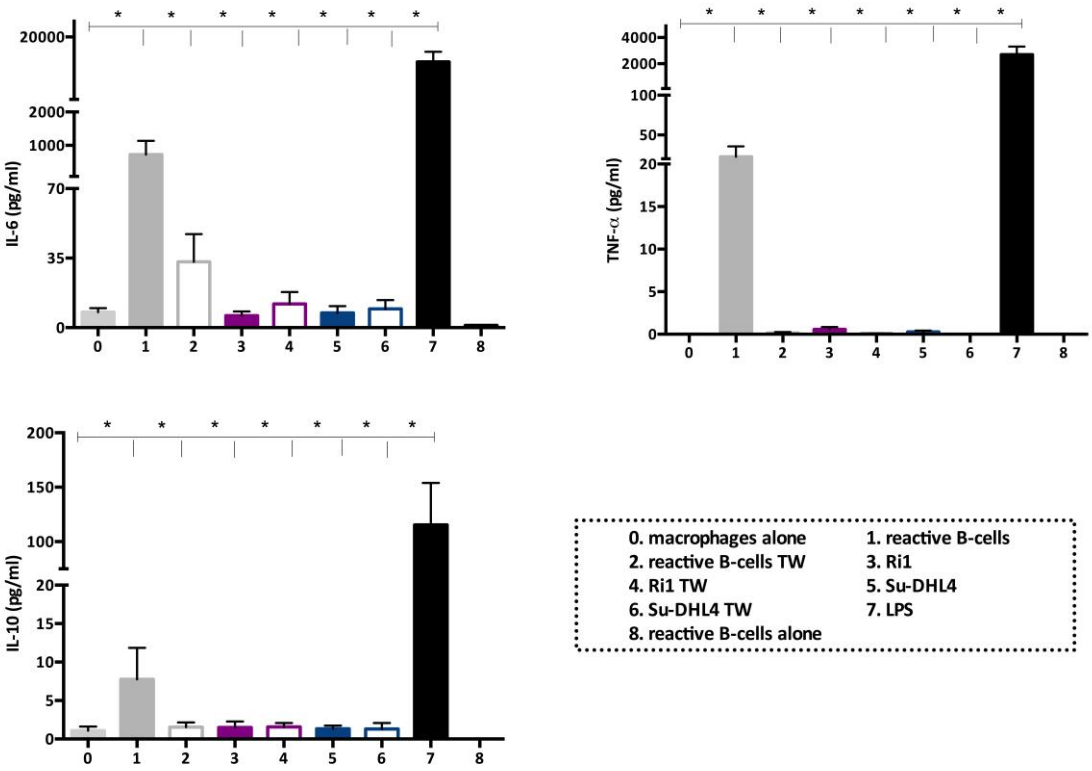

Figure 7.5 Detection of Th1 and Th2 cytokines in co-culture supernatants.
50 ml of co-culture supernatant was tested using CBA as described. Results represent the mean and SEM of five independent experiments. Bars represent mean + SEM, unpaired T-test, * p-value ≤0.05.

7.4.4 Using immunohistochemistry and immunofluorescence to discover specific macrophage subsets in the microenvironment

We hypothesized that the expression of proteins codified by M1 and M2 genes previously identified by GEP could help identifying subsets of macrophages in the microenvironment, as well as corroborating the existence of macrophage heterogeneity in human DLBCL. Using novel markers might aid defining cell subsets with potentially different functions.

Most studies trying to define M2 macrophage density in lymphoma used IHC for CD163. We claim that, in some circumstances, this is an inadequate approach for the

purpose.

Manual counting of both markers would likely produce ambiguous data, with CD163 outnumbering CD68, particularly in cases with heavy infiltration (Figure 7.6 A, case 1). This can be better appreciated in a correlation plot (Figure 7.6 C) showing a significant number of cases (~ 25%) scored higher for CD163 than for CD68.

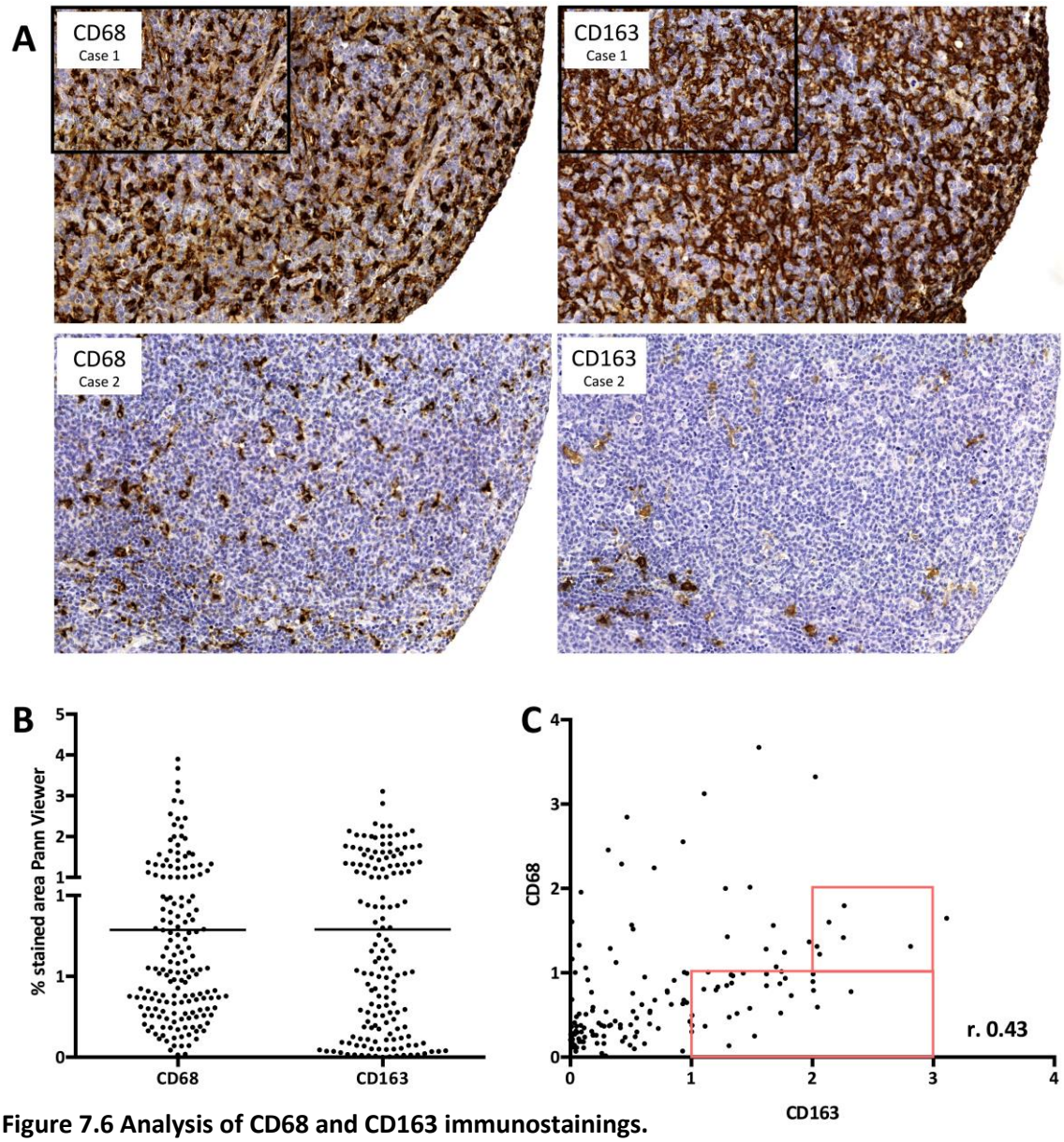

Figure 7.6 Analysis of CD68 and CD163 immunostainings.

A. Consecutive TMA sections were stained with CD68-KP1 and CD163. Images taken from the same core areas for two illustrative cases of heavy and low cell density. Amplification x20 and x40. B. Distribution of cases according to % stained area for CD68 and CD163 using Pannoramic Viewer. C. Correlation between results plotted in Figure B.

7.4.4.1 ALOX15 helps identifying a macrophages M2 subset in DLBCL

Arachidonate 15-Lipoxygenase (ALOX15) is a lipid-peroxiding enzyme involved in the clearance of apoptotic cells by macrophages. It has been demonstrated that ALOX15 belongs to an IL-4-induced, M2 transcriptional programme in human macrophages controlled by the transcription factor MYC. Coordinate induction of PPAR-γ and ALOX15 mediates interleukin-4-dependent transcription of the CD36 in macrophages. Immunostaining for ALOX15 in reactive LN TMAs suggested its expression was restricted to the stromal compartment (illustrative example pictured in Figure 7.7 A).

Morphologically, a substantial number of ALOX15+ cells in DLBCL were smaller than macrophages, likely representing monocytes. By performing IF in TMAs we determined that ALOX15 expression is circumscribed to a small subset of CD68+ cells in DLBCL (Figure 7.7 A). Hence, a comparative analysis could be done using single ALOX15 immunostaining.

Quantification of ALOX15+ cells in TMA corroborated GEP data (logFC -2,513 of DLBCL TAM versus controls, p=0.003). R-CHOP treated DLBCL showed a significantly lower expression of this enzyme compared to reactive LNs (Figure 7.7 B). Although a relatively small number of DLBCL cases were investigated (results available for 64 of the 77 patients treated with R-CHOP at Bart's), a training validation method suggested that having ALOX15+ cells in the stromal compartment was associated with an improved PFS (Figure 7.7 C). It will now important to validate these results in an independent cohort.

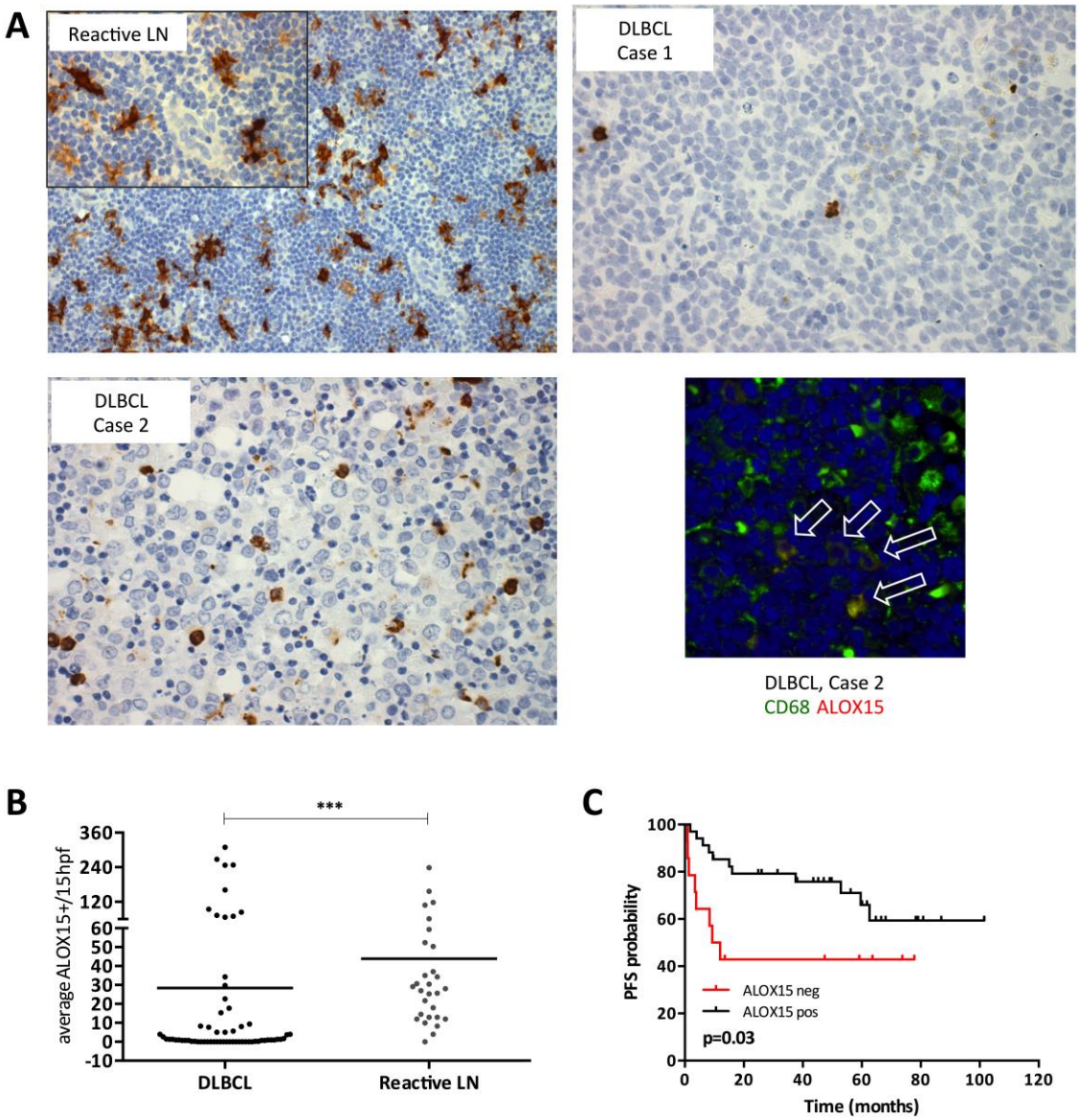

Figure 7.7 Expression of ALOX15 in reactive LNs and DLBCL.

A. HRP-DAB immunostaining for ALOX15 in representative cases of reactive LN (x20 and x40 magnification) and DLBCL (x40 magnification). IF studies performed in DLBCL TMAs showed that ALOX15 is restricted to CD68+ cells (white arrows, 20x magnification). B. Manual quantification of ALOX15 in DLBCL (64 cases) and reactive LN (29 cases) TMAs. Mann-Whitney U-test p-value <0.001. C. PFS according to absence or presence of ALOX15 expression (HR for negative cases 2.6, 95% CI 1.120-9.891).

7.4.4.2 IDO1 helps identifying a macrophages M1 subset in DLBCL

IDO1 transcript levels were upregulated in DLBCL macrophages compared to reactive controls, particularly in a subset of patients (group 2). Its promoter includes transcription factor sites that confer responsiveness to IFN-γ and it has been shown that macrophages, DCs and endothelial cells increase IDO1 expression when exposed to this cytokine and to LPS.

IDO1 protein expression was previously explored in DLBCL by a single group. The authors established a direct correlation of IDO1 expression by malignant B-cells with a worse survival in R-CHOP treated patients.

Corroborating previous findings, IDO1 expression was detected in stromal cells in reactive and DLBCL LN, likely in monocytes/macrophages, DCs and endothelial cells. In some cases (on example is illustrated in Figure 7.8, case 1), IDO1 appears to be expressed by malignant B-cells, but with a much lower intensity compared to stromal cells. Even within the stromal compartment, expression was variable, as is illustrated in case 2.

IF staining in DLBCL and reactive LN TMA showed co-localization of IDO1 and CD68 in the cytoplasm of a subset of small and large, interdigitating cells (identified in Figure 7.8 A). Both CD68+/IDO- macrophages and CD68-/IDO1+ cells were also identified.

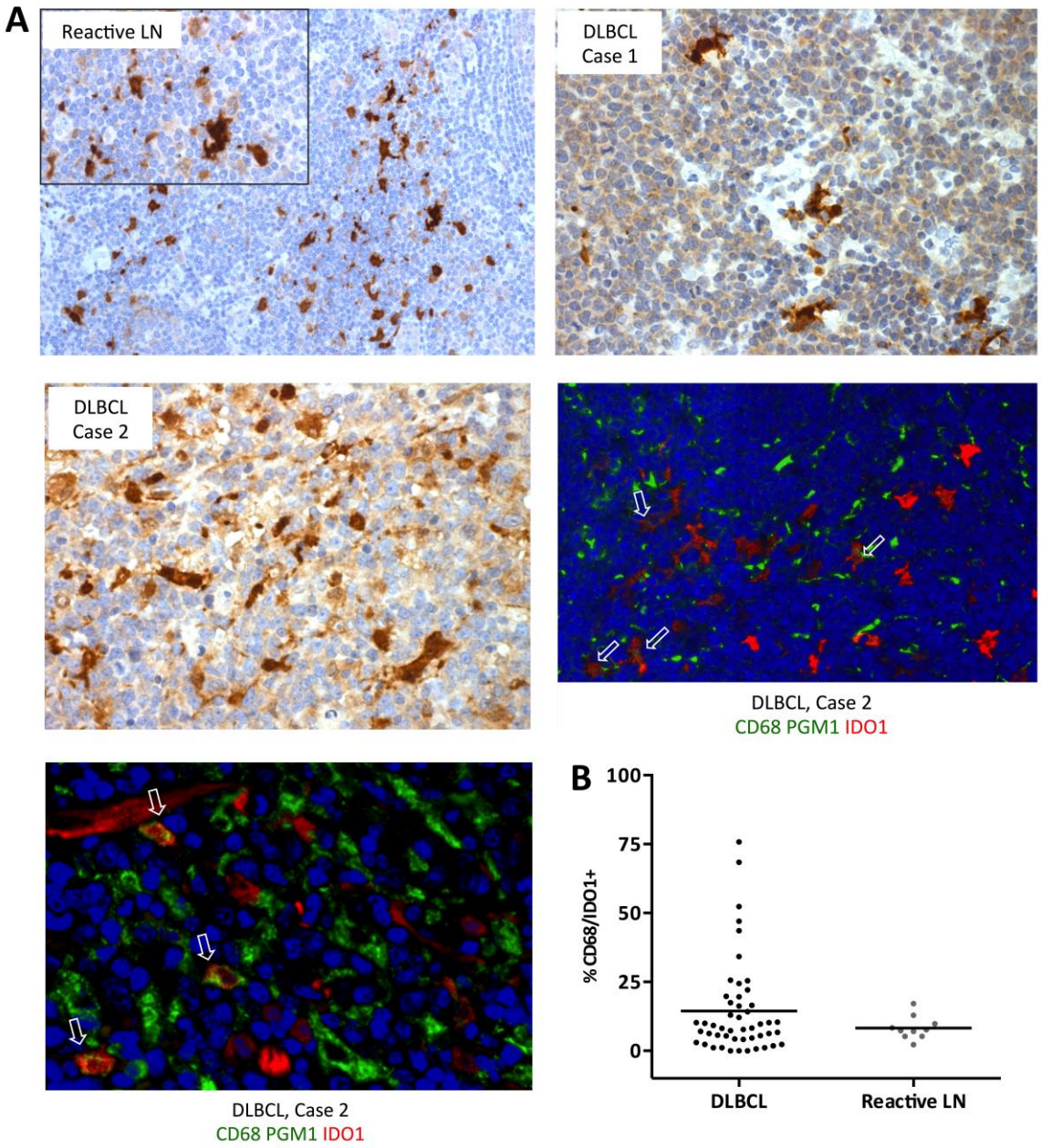

Figure 7.8 Expression of IDO1 in reactive LNs and DLBCL.

A. HRP-DAB immunostaining for IDO1 in representative cases of reactive LN (x20 and x40 magnification) and DLBCL (x40 magnification). Representative IF results for case 2 are illustrated, with white arrows identifying CD68+/IDO+ cells (40x and 63x magnification). B. The %CD68+/IDO1+ within the total macrophage population in DLBCL and control cases was manually quantified using IF.

IF analysis was performed manually on a computer screen. Firstly, CD68+ cells were quantified using the green channel. By switching the red channel on and off, CD68+/IDO1+ cells were then estimated for each case. Results for 59 R-CHOP treated patients and ten reactive LN samples are presented (Figure 7.8 B).

According to our data, the proportion of IDO1+ macrophages in DLBCL is variable. Two-thirds of DLBCL patients had <12.5% infiltration, whereas the remaining had a denser infiltration of dual positive cells. This is in accordance to our transcriptomic studies, in which IDO1 transcript was upregulated in a smaller DLBCL subset (group 2, five patients) compared to the remaining (group 1, eight patients). Additionally, although control cases studied are scarce, the proportion of IDO1+ macrophages seems to be lower than in DLBCL, again supporting our GEP results. These findings, however, need to be validated in a larger dataset.

7.5 Discussion

Given the impossibility of designing an optimal co-culture model, we thought of addressing DLBCL TAM heterogeneity that was hinted by the initial unsupervised hierarchical clustering of GEP clusters.

Supervised analysis showed that transcripts investigated in our co-culture model were among the differentially expressed genes between the two groups. However we believe these results do not devaluate the initial GEP results, which put into relevance the homogeneity of reactive LN macrophage GEP compared to DLBCL TAM GEP. To corroborate this, when results from reactive macrophages and samples belonging to the DLBCL TAM population 1 were compared, top transcripts were still differentially expressed. This is visually illustrated in the multidimensional scaling results placing DLBCL TAM population 1 in an intermediate position between controls and DLBCL TAM population 2.

We then postulated that TAM GEP relate to differences in the cytokine and chemokine profile of corresponding malignant B-cells. Recognizing differentially

expressed soluble factors could help to design simpler models to test their influence in macrophage transcriptome and proteome and to illuminate on potential cell interactions.

Emulating T helper-cell functional classification, it has been proposed that effector B-cell subsets can be identified by the pattern of cytokines secreted when stimulated with antigen and T helper cells.[263] When triggered by Th1 cells, B-cells are able to produce Th1 cytokines including IFN-γ, which in turn are able to amplify a Th1 immune response by affecting T-cell polarisation. This effector classification has not been addressed in the context of B-cell malignancies.

ABC DLBCL cell lines harbouring MYD88 mutations can secrete IL-6 and IL-10.[125] Studies using human DLBCL samples confirmed an increase in these JAK/STAT pathway-related serum cytokines and established a link between higher levels and a worse outcome.[363-365] It can be envisaged that DLBCL-derived IL-6 and IL-10, similarly to what has been described for IL-10 in BL and in other tumour contexts, polarise macrophages towards an M2, immunossupressive phenotype and sustain their survival in the microenvironment. Data developed in our laboratory (Hallam, SL, unpublished) confirmed a rise in circulating concentrations of IL-10 and IL-6 in mice transplanted with an aggressive B-cell NHL. In this model IL-6 decreased after macrophage ablation.

The expression of genes codifying the M2 cytokines IL-4, IL-13 and IL-10 or IL-6 was not significantly different between B-cells of both DLBCL groups and controls. However, the prototypical M1 cytokine IFN-γ was the most differentially expressed soluble factor between DLBCL B-cell groups. It is possible that B-cell-derived IFN-γ can intervene in macrophage M1 polarisation. We confirmed that DLBCL TAM group 2 preferentially upregulated M1 genes compared to group 2, which could corroborate the existence of different B-cell polarised effector population and an IFN-γ enriched milieu.

However, correlations between cytokine mRNA levels and secretion are generally poor. In an attempt to identify cell lines with a discrepant cytokine profile that could be used in co-culture models instead of primary cells, we measured cytokines in supernatants. None of the six DLBCL cell lines exhibited constitutional secretion of IFN-γ or other molecules tested, which could have been predicted given the lack of polarising triggers.

Measuring cytokines in DLBCL LN immediately after collection, yet not revealing their source, would be the only way to support the relevance of our findings. This could not be performed in our studies and is hard to pursue given the obstacles in obtaining fresh material. Although FFPE immunostaining could be used, we suggest that problems of quantification would be encountered.

IFN-γ was also undetected in all co-culture conditions, suggesting: (1) that this model is not adequate to demonstrate B-cell polarisation; (2) that IFN-γ, if produced, is not release to the media. On the other hand, TNF-α and the Th2 cytokines IL-6 and IL-10 were identified in supernatants and likely are macrophage-derived. Stimulation with LPS led to a considerable secretion of these cytokines, suggesting that macrophages are able to establish a secretory response to pro-inflammatory stimuli. Moreover, reactive B-cells cultured alone did not secrete any of the three molecules. Cell lines had no influence on cytokine secretion. Instead, IL-6, IL-10 and TNF-α were detected in supernatants from macrophages and reactive B-cells in contact conditions, suggesting macrophage activation by B-cells. However, this was not formally demonstrated in our experiments. It would be interesting to perform intracellular cytokine staining by flow cytometry of both B-cells and macrophages, since there is a possibility that both cell subsets trigger each other to produce these cytokines. These findings should be pursued in future experiments using primary DLBCL B-cells.

Other approach to validate GEP data is to perform protein studies using IHC. The availability of TMAs of DLBCL and reactive LN in our laboratory permitted analysing

protein expression patterns in DLBCL TAM and reactive macrophages in parallel, emulating the transcriptomic studies.

In this part of our investigations we wanted to clarify whether both M1 and M2 macrophages, expressing novel markers identified by GEP, are present in DLBCL tumours.

As already mentioned, it has been postulated that CD163 identifies M2 macrophages in the microenvironment of lymphoma. We think defining smaller, functionally skewed subsets would be more useful. Moreover, we claim that CD163, yet being a macrophage specific marker, is an inadequate M2 marker and hope to convey it by showing that one quarter of cases analysed for % stained area of CD68 and CD163 in parallel score higher for CD163 than for CD68.

Within the list of differentially expressed genes resulting from our transcriptomic studies we investigated several markers, but chose to show only results for ALOX15 and IDO-1 due to holding more comprehensive results. The remaining proteins (ANKRD22, CTSL1, PD-L1, AQP9, CDH1) are still under investigation.

The expression pattern of ALOX15 in LNs is underreported in the literature. According to our data, ALOX15 expression is restricted to cells of the microenvironment in DLBCL and reactive LNs and co-localizes with CD68 in all cases studied. Some double positive cells have morphology not suggestive of macrophages. It is possible that these are monocytes instead. It is known that CD68 staining can be identified in cells other than macrophages, such as mast cells. In our studies we used the monoclonal antibody PGM1, which is more specific to macrophages, giving less positivity to mast cells than the monoclonal antibody KP1. However, due to staining pattern of ALOX15 and its universal co-localization with CD68, it would be important to perform co-staining with triptase to exclude we are observing mast cells. Certainly, ALOX15 allows identifying only a subset of CD68+ cells, which is easily quantifiable using HRP-DAB immunostaining. Quantification of ALOX15+ cells in TMA corroborated GEP data with DLBCL cases showing a significantly lower expression of this enzyme compared

to reactive LNs. We suggest that ALOX15 presence in the microenvironment of DLBCL is associated with a longer PFS, but agree that an independent study performed in a larger dataset needs to be undertaken to validate this potential biomarker.

Similarly, we suggest that IF using IDO1 and CD68 permits identifying an M1 macrophage subset, likely displaying immunoregulatory functions in the microenvironment. By analysing a small dataset of patients and controls, we found that the proportion of IDO1+ macrophages in DLBCL is variable, being heavier in one third of patients. These cases can approximate to the ones belonging to DLBCL TAM group 2, which had a higher IDO1 gene expression. These findings, however, need to be validated in a larger dataset. Moreover, we need to gain more experience in analysing IF data before suggesting any methodology to replicate such results in independent studies. We recognize that interpretation of IDO1 is difficult due to a wide expression intensity range in LN cells. Automated analysis would certainly increase reproducibility and a method is being devised for that purpose using the Ariol system.

Despite some limitations, we demonstrate in these studies that M1 and M2 macrophage subsets co-inhabit in DLBCL tumours. Using novel markers that have a better functional translation might help to interpret the role of subpopulations in the microenvironment.

Although it is plausible that a single cell is able to express both M1 and M2 markers simultaneously, we currently do not have data demonstrating it. RNA *in situ* hybridisation and IF using M1 and M2 proteins will help to address this hypothesis. However, it is possible that studying targeted markers using these techniques still does not rule out this hypothesis, and we plan to perform macrophage single cell sorting and RNA sequencing to definitely clarify this matter.

Chapter 8 Final discussion and further work

Whereas only the IPI is established as a robust prognostic model, a number of biomarkers have the potential of better identifying poor-risk DLBCL patients. The underlying biological diversity of this disease explains the heterogeneous responses to first line chemoimmunotherapy and warrants the development of novel prognostic models for this disease incorporating such biomarkers. In this thesis we investigated biomarkers of the tumour and the microenvironment of DLBCL.

The hypotheses underlying our investigation were based on two critical findings of transcriptomic analysis of DLBCL:

- Two molecularly distinct forms based upon the cell-of-origin (COO) of DLBCL could be identified using GEP, which assume distinct genetic changes, oncogenic signalling pathways and response to R-CHOP. Thus, the COO molecular classification provides prospects for individualized treatment approaches that could improve patient's outcome. The obstacle to this enticing clinical scenario is the inexistence of robust strategies that enable us to recognize patients with GCB and ABC DLBCL.

- The differential expression of genes derived from the microenvironment can be used to identify patients with different outcome after R-CHOP. GEP-based prognostic signatures devised by independent research groups using independent patient cohorts incorporated genes encoding for tumour-infiltrating macrophages and matrix components. However, this data largely lacks independent validation. Moreover, a clear understanding of the macrophage functional heterogeneity in DLBCL is lacking.

The translational impact of transcriptomic studies has been demonstrated to be much limited compared to what is proposed in the original studies.

A number of obstacles to the acceptance of GEP-derived biomarkers can be acknowledged, including:

- Biomarkers need to be easily studied in any laboratory;
- Validation processes need to be robust;
- Microarray technology is still reserved to the research setting.

Undoubtedly, the methodology applied in independent studies exploring GEP-based biomarkers has an important impact on their validation. IHC has been used both as a surrogate to classify patients according to the COO classification and to enumerate and functionally characterize the microenvironment in DLBCL. Importantly, the Hans algorithm is being utilised to recognize ABC-DLBCLs in clinical trials offering NF-kB targeting agents to patients with this subtype.

Our first two studies encapsulate the problems and hopes of IHC methodologies for validation of prognostic biomarkers in DLBCL.

Our first study provides evidence that IHC algorithms are unreliable predictors of the molecular classification and hence should not be used for clinical decisions in DLBCL. Our second study establishes semi-automated methods as the best for IHC studies assessing cell infiltration in large tissue areas using TMAs.

While we suggest that the strategies used in previous investigations defining COO IHC algorithms are hardly replicated, we do not fully discard the notion that IHC can be used for the molecular classification of DLBCL.

We propose several approaches that could help demonstrating whether IHC can indeed be utilised for this purpose:

- Future investigations could be done in RCTs involving expert haematopathologists, haemato-oncologists and statisticians from different countries;

- An exhaustive document could arise from these trials that details guidelines for every single procedural step known to impact on IHC results, from tissue fixation to image and statistical analysis;

- In order to demonstrate that the methodologies can indeed be put in place in a robust manner, fresh tissue samples from a single patient could be distributed to different laboratories and final results could be compared;

- Once this is confirmed, the dissemination of this protocol could be done by experts in international meetings gathering other pathologists responsible for optimization of the techniques in their own countries;

- Different institutions could be invited to report their results in follow-up meetings, stimulating discussion and further optimization of the established guidelines.

- While we used the immune microenvironment in DLBCL as a model, the findings from our second study suggest that computerized systems eliminate the underlying variability of manual counting of IHC immunostainings and could also be explored for the analysis of proteins incorporated in the COO algorithms. A consensus methodological approach could be developed in the context of RCTs incorporating different systems of semi-automated analysis as delineated above.

Our subsequent studies address the unmet need for comprehensive analysis of macrophage activation status and effector function in human DLBCL. We conducted transcriptomic studies on purified macrophages from DLBCL and reactive LN. Using this data we explored in vitro cell systems to capture the functional dialogue of malignant B-cells and macrophages in the microenvironment. In the last study we explore macrophage heterogeneity in DLBCL, either by dissecting the findings of the transcriptome analysis and establishing parallels with paired B-cell analysis, or by exploring double IF staining to recognize macrophage populations in LN.

Previous publications using GEP of DLBCL correlated macrophage gene expression and outcome in DLBCL. Immunostaining for CD68 is taken as a surrogate of the extent of macrophage infiltration in tissues. However, difficulties posed by its analysis have to date hampered the ability to obtain consistent results. The computerized analysis of CD68 undertaken here using different machines and methods showed extremely good correlation, suggesting that this approach will be useful in clarifying the role of CD68 in outcome prediction in lymphoma. However, macrophages are highly plastic cells, assuming unique phenotypes and functions contextualized in their dynamic microenvironment. Consequently, IHC analysis of single markers is unable to represent subsets of polarised macrophages.

Our approach was to use GEP to clarify the functional repertoire of macrophages in DLBCL. A number of technical aspects could detract from our findings. However, we assured that all procedural steps were subjected to strict quality controls and demonstrate that it is feasible to undertake this kind of study in stored samples.

Unsupervised hierarchical clustering alone did not resolve disease from control samples in our study. This is an enticing discovery suggesting macrophage heterogeneity in DLBCL. In fact, it is doubtful that microenvironmental influences of the tumour cells work in a homogeneous fashion.

By performing a supervised hierarchical clustering analysis, we identified a 202-gene signature that distinguishes DLBCL TAM from reactive controls. The recognition of substantial gene expression variability in DLBCL TAM by unsupervised analysis suggested that corrected p-values would be modest and therefore the gene set would be small. Indeed expression variability was higher in DLBCL TAM. Although these findings could potentially reflect technical problems with our data, we do not have clear evidence for that and instead interpret them as biological variability that is inherent of human samples, particularly in bystander cells of the tumour microenvironment. In future studies we would like to test whether such GEP variability of macrophages is paralleled by underlying differences of the malignant B-cells.

A key approach taken in our GEP studies was the bioinformatics analysis. Whole transcriptome profiling represents an opportunity to functionally model a disease and bioinformatics approaches can help in giving biological meaning to high throughput data.

The functional enrichment analysis undertaken exposes that our data not only reflects the transcriptomic features of macrophages, but also unique GEP changes characteristic of the DLBCL TAM. Our gene expression signature significantly overlaps with other GEP-defined signatures of DLBCL, including the LN-signature and others developed in THRLBCL. Moreover, an enrichment of our macrophage-related genes was detected in both GCB and ABC cases, suggesting that relevant features of the microenvironment are independent of the molecular features of the malignant B-cells.

Our studies suggest that DLBCL TAM have a bidirectional M1 and M2 functional activation. Additionally, univariate survival analysis points towards a variable impact of the expression of M2 genes in R-CHOP treated cases, refuting the common assumption that M2 macrophage activation is necessarily associated with poor survival. In future studies we would like to explore the outcome impact of the

remaining genes from our signature and develop multivariate prognostic models together with the IPI.

Overall, these data commonalities validate *in silico* approaches to explore the function of cell subsets in complex tissues that are hard to study due to limited cell amounts or sensitiveness to laboratory manipulation. However, functional studies are the only ones able to provide definitive biological explanations towards the impact of specific molecules. In this context we conducted in vitro studies to functionally validate the trancriptomic studies.

Macrophage activation is influenced by a multitude of factors. Numerous cells, cytokines and growth factors interact *in vivo* to determine macrophage function. Ultimately there is no macrophage identical to another, as dynamic changes in the microenvironment have a great potential to affect the cell transcriptome and function. Such microenvironmental complexity is hardly mimicked in a culture plate, more so when using a two-cell co-culture system, as we did in our fourth results chapter.

In our last study we propose B-cell derived IFN-γ as the culprit molecule involved in a preferential M1 activation pattern in a subset of DLBCL TAM. However we did not formally test this experimentally in the current investigations and this will be subject of future experiments. To robustly demonstrate the hypothesis that the malignant DLBCL cells are influencing macrophages towards a specific transcriptome, we suggest in the future using three cell (macrophages, B and T-cells) co-culture models of primary samples, in an autologous setting and investigating a larger panel of markers. In this manner we shall better replicate the complex crosstalk that can be established between malignant and non-malignant cells in the microenvironment.

Using molecules that have a better functional translation might help to interpret the role of macrophage subpopulations in the microenvironment. By studying novel markers depicted from the transcriptomic analysis we clarify that macrophages

express both M1 and M2 markers in DLBCL tumours.

Although it is likely that a single macrophage is able to express both M1 and M2 molecules simultaneously, we do not yet have data demonstrating this. As our own and other genomic data become available underlining macrophage heterogeneity, novel techniques to label several proteins and RNA transcripts concurrently in situ will be required to validate potential prognostic subsets in large datasets of patients. Furthermore, transcriptomic studies of single macrophages acquired by cell sorting would help clarifying whether subsets of macrophages within DLBCL have different activation profiles or whether a macrophage can acquire a unique and hybrid activation pattern. We suspect both situations are in place in vivo and aim to demonstrate it in future studies.

In conclusion, the work developed in this thesis significantly deepened the understanding of the role of IHC methodologies for the validation of biomarkers for DLBCL. Although there are still large gaps to fulfil in order to comprehend the more complex phenomenon of macrophage activation in DLBCL, our investigations significantly moved the field forward by demonstrating cell functional heterogeneity and raising novel questions that will hardly be answered through simplified experimental designs.

Chapter 9 References

1 Swerdlow SH, C. E., Harris NL, et al. . *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues* 4th Edition edn, (IARC Press, 2008).

2 Ferlay, J. *et al.* Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *European journal of cancer* **49**, 1374-1403, doi:10.1016/j.ejca.2012.12.027 (2013).

3 Montoto, S. & Fitzgibbon, J. Transformation of indolent B-cell lymphomas. *J Clin Oncol* **29**, 1827-1834, doi:10.1200/JCO.2010.32.7577 (2011).

4 Coutinho, R. *et al.* HIV status does not impair the outcome of patients diagnosed with diffuse large B-cell lymphoma treated with R-CHOP in the cART era. *Aids* **28**, 689-697, doi:10.1097/QAD.0000000000000133 (2014).

5 Ok, C. Y. *et al.* Prevalence and clinical implications of epstein-barr virus infection in de novo diffuse large B-cell lymphoma in Western countries. *Clinical cancer research : an official journal of the American Association for Cancer Research* **20**, 2338-2349, doi:10.1158/1078-0432.CCR-13-3157 (2014).

6 Lopez-Guillermo, A. *et al.* Diffuse large B-cell lymphoma: clinical and biological characterization and outcome according to the nodal or extranodal primary origin. *J Clin Oncol* **23**, 2797-2804, doi:10.1200/JCO.2005.07.155 (2005).

7 Carbone, P. P., Kaplan, H. S., Musshoff, K., Smithers, D. W. & Tubiana, M. Report of the Committee on Hodgkin's Disease Staging Classification. *Cancer research* **31**, 1860-1861 (1971).

8 Sehn, L. H. *et al.* Impact of concordant and discordant bone marrow involvement on outcome in diffuse large B-cell lymphoma treated with R-CHOP. *J Clin Oncol* **29**, 1452-1457, doi:10.1200/JCO.2010.33.3419 (2011).

9 Cheson, B. D. *et al.* Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. *J Clin Oncol*, doi:10.1200/JCO.2013.54.8800 (2014).

10 Gurbuxani, S., Anastasi, J. & Hyjek, E. Diffuse large B-cell lymphoma--more than a diffuse collection of large B cells: an entity in search of a meaningful classification. *Arch Pathol Lab Med* **133**, 1121-1134, doi:10.1043/1543-2165-133.7.1121 (2009).

11 Achten, R., Verhoef, G., Vanuytsel, L. & De Wolf-Peeters, C. Histiocyte-rich, T-cell-rich B-cell lymphoma: a distinct diffuse large B-cell lymphoma subtype showing characteristic morphologic and immunophenotypic features. *Histopathology* **40**, 31-45 (2002).

12 Monti, S. *et al.* Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood* **105**, 1851-1861, doi:2004-07-2947 [pii]
10.1182/blood-2004-07-2947 (2005).

13 Hummel, M. *et al.* A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* **354**, 2419-2430, doi:10.1056/NEJMoa055351 (2006).

14 Aukema, S. M. *et al.* Double-hit B-cell lymphomas. *Blood* **117**, 2319-2331, doi:10.1182/blood-2010-09-297879 (2011).

15 Snuderl, M. *et al.* B-cell lymphomas with concurrent IGH-BCL2 and MYC rearrangements are aggressive neoplasms with clinical and pathologic features distinct from Burkitt lymphoma and diffuse large B-cell lymphoma. *Am J Surg Pathol* **34**, 327-340, doi:10.1097/PAS.0b013e3181cd3aeb (2010).

16 Salaverria, I. & Siebert, R. The gray zone between Burkitt's lymphoma and diffuse large B-cell lymphoma from a genetics perspective. *J Clin Oncol* **29**, 1835-1843, doi:10.1200/JCO.2010.32.8385 (2011).

17 Proctor, I. E., McNamara, C., Rodriguez-Justo, M., Isaacson, P. G. & Ramsay, A. Importance of expert central review in the diagnosis of lymphoid malignancies in a regional cancer network. *J Clin Oncol* **29**, 1431-1435, doi:10.1200/JCO.2010.31.2223 (2011).
18 McKelvey, E. M. *et al.* Hydroxyldaunomycin (Adriamycin) combination chemotherapy in malignant lymphoma. *Cancer* **38**, 1484-1493 (1976).
19 Fisher, R. I. *et al.* Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N Engl J Med* **328**, 1002-1006, doi:10.1056/NEJM199304083281404 (1993).
20 Coiffier, B. *et al.* CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* **346**, 235-242, doi:10.1056/NEJMoa011795 (2002).
21 Coiffier, B. *et al.* Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood* **116**, 2040-2045, doi:10.1182/blood-2010-03-276246 (2010).
22 Sehn, L. H. *et al.* Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. *J Clin Oncol* **23**, 5027-5033, doi:10.1200/JCO.2005.09.137 (2005).
23 Pfreundschuh, M. *et al.* CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol* **7**, 379-391, doi:10.1016/S1470-2045(06)70664-7 (2006).
24 Pfreundschuh, M. *et al.* CHOP-like chemotherapy with or without rituximab in young patients with good-prognosis diffuse large-B-cell lymphoma: 6-year results of an open-label randomised study of the MabThera International Trial (MInT) Group. *Lancet Oncol* **12**, 1013-1022, doi:10.1016/S1470-2045(11)70235-2 (2011).
25 Delarue, R. *et al.* Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. *Lancet Oncol* **14**, 525-533, doi:10.1016/S1470-2045(13)70122-0 (2013).
26 Cunningham, D. *et al.* Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet*, doi:10.1016/S0140-6736(13)60313-X (2013).
27 Gisselbrecht, C. *et al.* Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. *J Clin Oncol* **28**, 4184-4190, doi:10.1200/JCO.2010.28.1618 (2010).
28 Schumacher, M., Binder, H. & Gerds, T. Assessment of survival prediction models based on microarray data. *Bioinformatics* **23**, 1768-1774, doi:10.1093/bioinformatics/btm232 (2007).
29 Mehta, S. *et al.* Predictive and prognostic molecular markers for cancer medicine. *Therapeutic advances in medical oncology* **2**, 125-148, doi:10.1177/1758834009360519 (2010).
30 A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* **329**, 987-994, doi:10.1056/NEJM199309303291402 (1993).
31 Sehn, L. H. *et al.* The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood* **109**, 1857-1861, doi:10.1182/blood-2006-08-038257 (2007).
32 Ziepert, M. *et al.* Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. *J Clin*

Oncol **28**, 2373-2380, doi:10.1200/JCO.2009.26.2493 (2010).
33 Zhou, Z. *et al.* An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood* **123**, 837-842, doi:10.1182/blood-2013-09-524108 (2014).
34 Lenz, G. *et al.* Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* **359**, 2313-2323, doi:359/22/2313 [pii]
10.1056/NEJMoa0802885 (2008).
35 Savage, K. J. *et al.* MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood* **114**, 3533-3537, doi:blood-2009-05-220095 [pii]
10.1182/blood-2009-05-220095 (2009).
36 Barrans, S. L. *et al.* The t(14;18) is associated with germinal center-derived diffuse large B-cell lymphoma and is a strong predictor of outcome. *Clinical cancer research : an official journal of the American Association for Cancer Research* **9**, 2133-2139 (2003).
37 Barrans, S. L. *et al.* Rearrangement of the BCL6 locus at 3q27 is an independent poor prognostic factor in nodal diffuse large B-cell lymphoma. *Br J Haematol* **117**, 322-332 (2002).
38 Klapper, W. *et al.* Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. *Blood* **119**, 1882-1887, doi:10.1182/blood-2011-10-388470 (2012).
39 Riihijarvi, S., Taskinen, M., Jerkeman, M. & Leppa, S. Male gender is an adverse prognostic factor in B-cell lymphoma patients treated with immunochemotherapy. *European journal of haematology* **86**, 124-128, doi:10.1111/j.1600-0609.2010.01541.x (2011).
40 Michael Pfreundschuh , G. H., Samira Zeynalova , Viola Pöschel , Andreas Viardot , Matthias Hänel, Ulrich Keller , Marcel Reiser , Markus Löffler, Niels Murawski. in *53rd ASH Annual Meeting and Exposition* Vol. 118 (Blood, San Diego, 2011).
41 Carson, K. R. *et al.* Increased body mass index is associated with improved survival in United States veterans with diffuse large B-cell lymphoma. *J Clin Oncol* **30**, 3217-3222, doi:10.1200/JCO.2011.39.2100 (2012).
42 Larsson, S. C. & Wolk, A. Body mass index and risk of non-Hodgkin's and Hodgkin's lymphoma: a meta-analysis of prospective studies. *European journal of cancer* **47**, 2422-2430, doi:10.1016/j.ejca.2011.06.029 (2011).
43 Hong, F. *et al.* The role of body mass index in survival outcome for lymphoma patients: US intergroup experience. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **25**, 669-674, doi:10.1093/annonc/mdt594 (2014).
44 Pfreundschuh, M. *et al.* Prognostic significance of maximum tumour (bulk) diameter in young patients with good-prognosis diffuse large-B-cell lymphoma treated with CHOP-like chemotherapy with or without rituximab: an exploratory analysis of the MabThera International Trial Group (MInT) study. *Lancet Oncol* **9**, 435-444, doi:10.1016/S1470-2045(08)70078-0 (2008).
45 Recher, C. *et al.* Intensified chemotherapy with ACVBP plus rituximab versus standard CHOP plus rituximab for the treatment of diffuse large B-cell lymphoma (LNH03-2B): an open-label randomised phase 3 trial. *Lancet* **378**, 1858-1867, doi:10.1016/S0140-6736(11)61040-4 (2011).
46 Chung, R. *et al.* Concordant but not discordant bone marrow involvement in diffuse large B-cell lymphoma predicts a poor clinical outcome independent of the International Prognostic Index. *Blood* **110**, 1278-1282, doi:10.1182/blood-2007-01-070300 (2007).
47 Takahashi, H. *et al.* Prognostic impact of extranodal involvement in diffuse large B-cell lymphoma in the rituximab era. *Cancer* **118**, 4166-4172, doi:10.1002/cncr.27381 (2012).
48 Gutierrez-Garcia, G. *et al.* Clinico-biological characterization and outcome of primary nodal and extranodal diffuse large B-cell lymphoma in the rituximab era. *Leuk Lymphoma* **51**, 1225-1232, doi:10.3109/10428194.2010.483301 (2010).

49 Moskowitz, C. H. *et al.* Risk-adapted dose-dense immunochemotherapy determined by interim FDG-PET in Advanced-stage diffuse large B-Cell lymphoma. *J Clin Oncol* **28**, 1896-1903, doi:10.1200/JCO.2009.26.5942 (2010).
50 Casasnovas, R. O. *et al.* SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. *Blood* **118**, 37-43, doi:10.1182/blood-2010-12-327767 (2011).
51 Ott, G. *et al.* Immunoblastic morphology but not the immunohistochemical GCB/nonGCB classifier predicts outcome in diffuse large B-cell lymphoma in the RICOVER-60 trial of the DSHNHL. *Blood* **116**, 4916-4925, doi:10.1182/blood-2010-03-276766 (2010).
52 Montes-Moreno, S. *et al.* Aggressive large B-cell lymphoma with plasma cell differentiation: immunohistochemical characterization of plasmablastic lymphoma and diffuse large B-cell lymphoma with partial plasmablastic phenotype. *Haematologica* **95**, 1342-1349, doi:10.3324/haematol.2009.016113 (2010).
53 Kobayashi, T. *et al.* Microarray reveals differences in both tumors and vascular specific gene expression in de novo CD5+ and CD5- diffuse large B-cell lymphomas. *Cancer research* **63**, 60-66 (2003).
54 Tagawa, H. *et al.* Comparison of genome profiles for identification of distinct subgroups of diffuse large B-cell lymphoma. *Blood* **106**, 1770-1777, doi:10.1182/blood-2005-02-0542 (2005).
55 Ennishi, D. *et al.* CD5 expression is potentially predictive of poor outcome among biomarkers in patients with diffuse large B-cell lymphoma receiving rituximab plus CHOP therapy. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* **19**, 1921-1926, doi:10.1093/annonc/mdn392 (2008).
56 Johnson, N. A. *et al.* Diffuse large B-cell lymphoma: reduced CD20 expression is associated with an inferior survival. *Blood* **113**, 3773-3780, doi:10.1182/blood-2008-09-177469 (2009).
57 Kuppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nature reviews. Cancer* **5**, 251-262, doi:10.1038/nrc1589 (2005).
58 de Villartay, J. P., Fischer, A. & Durandy, A. The mechanisms of immune diversification and their disorders. *Nature reviews. Immunology* **3**, 962-972, doi:10.1038/nri1247 (2003).
59 Shaffer, A. L., Rosenwald, A. & Staudt, L. M. Lymphoid malignancies: the dark side of B-cell differentiation. *Nature reviews. Immunology* **2**, 920-932, doi:10.1038/nri953 nri953 [pii] (2002).
60 Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nature reviews. Immunology* **8**, 22-33, doi:nri2217 [pii] 10.1038/nri2217 (2008).
61 Polyak, K., Haviv, I. & Campbell, I. G. Co-evolution of tumor cells and their microenvironment. *Trends Genet* **25**, 30-38, doi:S0168-9525(08)00309-0 [pii] 10.1016/j.tig.2008.10.012 (2009).
62 Limpens, J. *et al.* Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. *Blood* **85**, 2528-2536 (1995).
63 Klein, U. *et al.* Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A* **100**, 2639-2644, doi:10.1073/pnas.0437996100 (2003).
64 Wagner, S. D. & Neuberger, M. S. Somatic hypermutation of immunoglobulin genes. *Annu Rev Immunol* **14**, 441-457, doi:10.1146/annurev.immunol.14.1.441 (1996).
65 Shen, H. M., Peters, A., Baron, B., Zhu, X. & Storb, U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* **280**, 1750-1752 (1998).
66 Peled, J. U. *et al.* The biochemistry of somatic hypermutation. *Annu Rev Immunol* **26**, 481-511, doi:10.1146/annurev.immunol.26.021607.090236 (2008).

67 Cattoretti, G. *et al.* Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell* **7**, 445-455, doi:10.1016/j.ccr.2005.03.037 (2005).
68 Saito, M. *et al.* BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **106**, 11294-11299, doi:10.1073/pnas.0903854106 (2009).
69 Ranuncolo, S. M., Polo, J. M. & Melnick, A. BCL6 represses CHEK1 and suppresses DNA damage pathways in normal and malignant B-cells. *Blood Cells Mol Dis* **41**, 95-99, doi:10.1016/j.bcmd.2008.02.003 (2008).
70 Ye, B. H. *et al.* The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet* **16**, 161-170, doi:10.1038/ng0697-161 (1997).
71 Tunyaplin, C. *et al.* Direct repression of prdm1 by Bcl-6 inhibits plasmacytic differentiation. *Journal of immunology* **173**, 1158-1165 (2004).
72 Basso, K. & Dalla-Favera, R. BCL6: master regulator of the germinal center reaction and key oncogene in B cell lymphomagenesis. *Adv Immunol* **105**, 193-210, doi:10.1016/S0065-2776(10)05007-8 (2010).
73 Bubendorf, L. *et al.* Hormone therapy failure in human prostate cancer: analysis by complementary DNA and tissue microarrays. *Journal of the National Cancer Institute* **91**, 1758-1764 (1999).
74 Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503-511, doi:10.1038/35000501 (2000).
75 Alizadeh, A. *et al.* The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harbor symposia on quantitative biology* **64**, 71-78 (1999).
76 Rosenwald, A. *et al.* The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* **346**, 1937-1947, doi:10.1056/NEJMoa012914
346/25/1937 [pii] (2002).
77 Williams, P. M. *et al.* A novel method of amplification of FFPET-derived RNA enables accurate disease classification with microarrays. *J Mol Diagn* **12**, 680-686, doi:10.2353/jmoldx.2010.090164 (2010).
78 Rimsza, L. M. *et al.* Gene expression predicts overall survival in paraffin-embedded tissues of diffuse large B-cell lymphoma treated with R-CHOP. *Blood* **112**, 3425-3433, doi:10.1182/blood-2008-02-137372 (2008).
79 Scott, D. W. *et al.* Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. *Blood* **123**, 1214-1217, doi:10.1182/blood-2013-11-536433 (2014).
80 Shipp, M. A. *et al.* Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* **8**, 68-74, doi:10.1038/nm0102-68
nm0102-68 [pii] (2002).
81 Lossos, I. S. *et al.* Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc Natl Acad Sci U S A* **97**, 10209-10213, doi:10.1073/pnas.180316097
180316097 [pii] (2000).
82 Lenz, G. *et al.* Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A* **105**, 13520-13525, doi:0804295105 [pii]
10.1073/pnas.0804295105 (2008).
83 Shaffer, A. L., 3rd, Young, R. M. & Staudt, L. M. Pathogenesis of human B cell lymphomas. *Annu Rev Immunol* **30**, 565-610, doi:10.1146/annurev-immunol-020711-075027 (2012).

84 Jardin, F. Next generation sequencing and the management of diffuse large B-cell lymphoma: from whole exome analysis to targeted therapy. *Discovery medicine* **18**, 51-65 (2014).
85 Zhang, J. *et al.* Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* **110**, 1398-1403, doi:10.1073/pnas.1205299110 (2013).
86 Xu-Monette, Z. Y. *et al.* Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood* **120**, 3986-3996, doi:10.1182/blood-2012-05-433334 (2012).
87 Ichikawa, A. *et al.* Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N Engl J Med* **337**, 529-534, doi:10.1056/NEJM199708213370804 (1997).
88 Wilson, W. H. *et al.* Relationship of p53, bcl-2, and tumor proliferation to clinical drug resistance in non-Hodgkin's lymphomas. *Blood* **89**, 601-609 (1997).
89 Young, K. H. *et al.* Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood* **112**, 3088-3098, doi:10.1182/blood-2008-01-129783 (2008).
90 Zainuddin, N. *et al.* TP53 mutations predict for poor survival in de novo diffuse large B-cell lymphoma of germinal center subtype. *Leuk Res* **33**, 60-66, doi:10.1016/j.leukres.2008.06.022 (2009).
91 Bea, S. *et al.* Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction. *Blood* **106**, 3183-3190, doi:10.1182/blood-2005-04-1399 (2005).
92 Wright, G. *et al.* A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A* **100**, 9991-9996, doi:10.1073/pnas.1732008100
1732008100 [pii] (2003).
93 Coutinho, R. *et al.* Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma: implications for therapeutic strategies. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 6686-6695, doi:10.1158/1078-0432.CCR-13-1482 (2013).
94 Berglund, M. *et al.* Evaluation of immunophenotype in diffuse large B-cell lymphoma and its impact on prognosis. *Mod Pathol* **18**, 1113-1120, doi:10.1038/modpathol.3800396 (2005).
95 De Paepe, P. & De Wolf-Peeters, C. Diffuse large B-cell lymphoma: a heterogeneous group of non-Hodgkin lymphomas comprising several distinct clinicopathological entities. *Leukemia* **21**, 37-43, doi:2404449 [pii]
10.1038/sj.leu.2404449 (2007).
96 Barrans, S. L. *et al.* Germinal center phenotype and bcl-2 expression combined with the International Prognostic Index improves patient risk stratification in diffuse large B-cell lymphoma. *Blood* **99**, 1136-1143 (2002).
97 Iqbal, J. *et al.* BCL2 predicts survival in germinal center B-cell-like diffuse large B-cell lymphoma treated with CHOP-like therapy and rituximab. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17**, 7785-7795, doi:10.1158/1078-0432.CCR-11-0267 (2011).
98 Klapper, W. *et al.* Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **22**, 2226-2229, doi:10.1038/leu.2008.230 (2008).
99 Johnson, N. A. *et al.* Concurrent Expression of MYC and BCL2 in Diffuse Large B-Cell Lymphoma Treated With Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and

Prednisone. *J Clin Oncol*, doi:10.1200/JCO.2011.41.0985 (2012).
100 Obermann, E. C., Csato, M., Dirnhofer, S. & Tzankov, A. Aberrations of the MYC gene in unselected cases of diffuse large B-cell lymphoma are rare and unpredictable by morphological or immunohistochemical assessment. *J Clin Pathol* **62**, 754-756, doi:10.1136/jcp.2009.065227 (2009).
101 Barrans, S. *et al.* Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J Clin Oncol* **28**, 3360-3365, doi:10.1200/JCO.2009.26.3947 (2010).
102 Green, T. M. *et al.* Immunohistochemical Double-Hit Score Is a Strong Predictor of Outcome in Patients With Diffuse Large B-Cell Lymphoma Treated With Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone. *J Clin Oncol*, doi:10.1200/JCO.2011.41.4342 (2012).
103 Niitsu, N., Okamoto, M., Miura, I. & Hirano, M. Clinical features and prognosis of de novo diffuse large B-cell lymphoma with t(14;18) and 8q24/c-MYC translocations. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **23**, 777-783, doi:10.1038/leu.2008.344 (2009).
104 Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nature reviews. Drug discovery* **10**, 261-275, doi:10.1038/nrd3428 (2011).
105 Pasqualucci, L. *et al.* Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood* **101**, 2914-2923, doi:10.1182/blood-2002-11-3387 (2003).
106 Cerchietti, L. C. *et al.* A small-molecule inhibitor of BCL6 kills DLBCL cells in vitro and in vivo. *Cancer Cell* **17**, 400-411, doi:10.1016/j.ccr.2009.12.050 (2010).
107 Salaverria, I. *et al.* Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. *Blood* **118**, 139-147, doi:10.1182/blood-2011-01-330795 (2011).
108 Davis, R. E., Brown, K. D., Siebenlist, U. & Staudt, L. M. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med* **194**, 1861-1874 (2001).
109 Ferch, U. *et al.* Inhibition of MALT1 protease activity is selectively toxic for activated B cell-like diffuse large B cell lymphoma cells. *The Journal of experimental medicine* **206**, 2313-2320, doi:10.1084/jem.20091167 (2009).
110 Hailfinger, S. *et al.* Essential role of MALT1 protease activity in activated B cell-like diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* **106**, 19946-19951, doi:10.1073/pnas.0907511106 (2009).
111 Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676-1679, doi:10.1126/science.1153629 (2008).
112 Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-721, doi:nature07968 [pii]
10.1038/nature07968 (2009).
113 Ruan, J. *et al.* Bortezomib plus CHOP-rituximab for previously untreated diffuse large B-cell lymphoma and mantle cell lymphoma. *J Clin Oncol* **29**, 690-697, doi:10.1200/JCO.2010.31.1142 (2011).
114 Ribrag, V. *et al.* Efficacy and toxicity of 2 schedules of frontline rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone plus bortezomib in patients with B-cell lymphoma: a randomized phase 2 trial from the French Adult Lymphoma Study Group (GELA). *Cancer* **115**, 4540-4546, doi:10.1002/cncr.24518 (2009).
115 Refaeli, Y. *et al.* The B cell antigen receptor and overexpression of MYC can cooperate in the genesis of B cell lymphomas. *PLoS biology* **6**, e152, doi:10.1371/journal.pbio.0060152 (2008).

116 Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88-92, doi:10.1038/nature08638 (2010).
117 Yang, Y. *et al.* Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. *Cancer Cell* **21**, 723-737, doi:10.1016/j.ccr.2012.05.024 (2012).
118 Advani, R. H. *et al.* Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765) has significant activity in patients with relapsed/refractory B-cell malignancies. *J Clin Oncol* **31**, 88-94, doi:10.1200/JCO.2012.42.7906 (2013).
119 Chen, L. *et al.* SYK-dependent tonic B-cell receptor signaling is a rational treatment target in diffuse large B-cell lymphoma. *Blood* **111**, 2230-2237, doi:10.1182/blood-2007-07-100115 (2008).
120 Cheng, S. *et al.* SYK inhibition and response prediction in diffuse large B-cell lymphoma. *Blood* **118**, 6342-6352, doi:10.1182/blood-2011-02-333773 (2011).
121 Friedberg, J. W. *et al.* Inhibition of Syk with fostamatinib disodium has significant clinical activity in non-Hodgkin lymphoma and chronic lymphocytic leukemia. *Blood* **115**, 2578-2585, doi:10.1182/blood-2009-08-236471 (2010).
122 Naylor, T. L. *et al.* Protein kinase C inhibitor sotrastaurin selectively inhibits the growth of CD79 mutant diffuse large B-cell lymphomas. *Cancer research* **71**, 2643-2653, doi:10.1158/0008-5472.CAN-10-2525 (2011).
123 Kloo, B. *et al.* Critical role of PI3K signaling for NF-kappaB-dependent survival in a subset of activated B-cell-like diffuse large B-cell lymphoma cells. *Proc Natl Acad Sci U S A* **108**, 272-277, doi:1008969108 [pii]
10.1073/pnas.1008969108 (2011).
124 Abubaker, J. *et al.* PIK3CA mutations are mutually exclusive with PTEN loss in diffuse large B-cell lymphoma. *Leukemia* **21**, 2368-2370, doi:10.1038/sj.leu.2404873 (2007).
125 Ngo, V. N. *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **470**, 115-119, doi:10.1038/nature09671 (2011).
126 Lam, L. T. *et al.* Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B-cell lymphoma. *Blood* **111**, 3701-3713, doi:blood-2007-09-111948 [pii]
10.1182/blood-2007-09-111948 (2008).
127 Dufner, A. & Schamel, W. W. B cell antigen receptor-induced activation of an IRAK4-dependent signaling pathway revealed by a MALT1-IRAK4 double knockout mouse model. *Cell communication and signaling : CCS* **9**, 6, doi:10.1186/1478-811X-9-6 (2011).
128 Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303, doi:10.1038/nature10351 (2011).
129 Beguelin, W. *et al.* EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* **23**, 677-692, doi:10.1016/j.ccr.2013.04.011 (2013).
130 Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature genetics* **42**, 181-185, doi:10.1038/ng.518 (2010).
131 McCabe, M. T. *et al.* EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108-112, doi:10.1038/nature11606 (2012).
132 Alencar, A. J. *et al.* MicroRNAs are Independent Predictors of Outcome in Diffuse Large B-cell Lymphoma Patients Treated with R-CHOP. *Clinical cancer research : an official journal of the American Association for Cancer Research*, doi:1078-0432.CCR-11-0224 [pii]
10.1158/1078-0432.CCR-11-0224 (2011).
133 Montes-Moreno, S. *et al.* MicroRNA expression in diffuse large B-cell lymphoma treated with chemoimmunotherapy. *Blood*, doi:10.1182/blood-2010-11-321554 (2011).
134 Staudt, L. M. & Dave, S. The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv Immunol* **87**, 163-208, doi:S0065-2776(05)87005-1 [pii]

10.1016/S0065-2776(05)87005-1 (2005).
135 Jais, J. P. *et al.* The expression of 16 genes related to the cell of origin and immune response predicts survival in elderly patients with diffuse large B-cell lymphoma treated with CHOP and rituximab. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **22**, 1917-1924, doi:10.1038/leu.2008.188 (2008).
136 Lossos, I. S. *et al.* Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med* **350**, 1828-1837, doi:10.1056/NEJMoa032520
350/18/1828 [pii] (2004).
137 Alizadeh, A. A. *et al.* Prediction of survival in diffuse large B-cell lymphoma based on the expression of 2 genes reflecting tumor and microenvironment. *Blood* **118**, 1350-1358, doi:10.1182/blood-2011-03-345272 (2011).
138 Efron, B. The bootstrap and Markov-chain Monte Carlo. *Journal of biopharmaceutical statistics* **21**, 1052-1062, doi:10.1080/10543406.2011.607736 (2011).
139 Efron, B. & Tibshirani, R. Empirical bayes methods and false discovery rates for microarrays. *Genetic epidemiology* **23**, 70-86, doi:10.1002/gepi.1124 (2002).
140 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
141 Pages, F. *et al.* Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene* **29**, 1093-1102, doi:10.1038/onc.2009.416 (2010).
142 Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* **3**, 991-998, doi:10.1038/ni1102-991 (2002).
143 Balkwill, F. & Mantovani, A. Inflammation and cancer: back to Virchow? *Lancet* **357**, 539-545, doi:10.1016/S0140-6736(00)04046-0 (2001).
144 DeNardo, D. G., Andreu, P. & Coussens, L. M. Interactions between lymphocytes and myeloid cells regulate pro- versus anti-tumor immunity. *Cancer Metastasis Rev* **29**, 309-316, doi:10.1007/s10555-010-9223-6 (2010).
145 Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883-899, doi:10.1016/j.cell.2010.01.025 (2010).
146 Qian, B. Z. & Pollard, J. W. Macrophage diversity enhances tumor progression and metastasis. *Cell* **141**, 39-51, doi:S0092-8674(10)00287-4 [pii]
10.1016/j.cell.2010.03.014 (2010).
147 Hsie, A. W. *et al.* Evidence for reactive oxygen species inducing mutations in mammalian cells. *Proc Natl Acad Sci U S A* **83**, 9616-9620 (1986).
148 Pollard, J. W. Trophic macrophages in development and disease. *Nature reviews. Immunology* **9**, 259-270, doi:nri2528 [pii]
10.1038/nri2528 (2009).
149 Biswas, S. K. & Mantovani, A. Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. *Nat Immunol* **11**, 889-896, doi:ni.1937 [pii]
10.1038/ni.1937 (2010).
150 Egeblad, M., Nakasone, E. S. & Werb, Z. Tumors as organs: complex tissues that interface with the entire organism. *Developmental cell* **18**, 884-901, doi:10.1016/j.devcel.2010.05.012 (2010).
151 MacDonald, K. P. *et al.* An antibody against the colony-stimulating factor 1 receptor depletes the resident subset of monocytes and tissue- and tumor-associated macrophages but does not inhibit inflammation. *Blood* **116**, 3955-3963, doi:10.1182/blood-2010-02-266296 (2010).
152 van Furth, R. *et al.* The mononuclear phagocyte system: a new classification of macrophages, monocytes, and their precursor cells. *Bulletin of the World Health Organization* **46**, 845-852 (1972).
153 Varol, C. *et al.* Monocytes give rise to mucosal, but not splenic, conventional dendritic

cells. *The Journal of experimental medicine* **204**, 171-180, doi:10.1084/jem.20061011 (2007).
154 Wiktor-Jedrzejczak, W. *et al.* Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse. *Proc Natl Acad Sci U S A* **87**, 4828-4832 (1990).
155 Dai, X. M. *et al.* Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects. *Blood* **99**, 111-120 (2002).
156 DeKoter, R. P., Walsh, J. C. & Singh, H. PU.1 regulates both cytokine-dependent proliferation and differentiation of granulocyte/macrophage progenitors. *The EMBO journal* **17**, 4456-4468, doi:10.1093/emboj/17.15.4456 (1998).
157 Geissmann, F., Jung, S. & Littman, D. R. Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity* **19**, 71-82, doi:S1074761303001742 [pii] (2003).
158 Auffray, C. *et al.* Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science* **317**, 666-670, doi:10.1126/science.1142883 (2007).
159 Nahrendorf, M. *et al.* The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. *J Exp Med* **204**, 3037-3047, doi:10.1084/jem.20070885 (2007).
160 Cros, J. *et al.* Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. *Immunity* **33**, 375-386, doi:10.1016/j.immuni.2010.08.012 (2010).
161 Mildner, A. *et al.* Microglia in the adult brain arise from Ly-6ChiCCR2+ monocytes only under defined host conditions. *Nat Neurosci* **10**, 1544-1553, doi:10.1038/nn2015 (2007).
162 Ajami, B., Bennett, J. L., Krieger, C., Tetzlaff, W. & Rossi, F. M. Local self-renewal can sustain CNS microglia maintenance and function throughout adult life. *Nat Neurosci* **10**, 1538-1543, doi:10.1038/nn2014 (2007).
163 Schulz, C. *et al.* A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336**, 86-90, doi:10.1126/science.1219179 (2012).
164 Hashimoto, D. *et al.* Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* **38**, 792-804, doi:10.1016/j.immuni.2013.04.004 (2013).
165 Aziz, A., Soucie, E., Sarrazin, S. & Sieweke, M. H. MafB/c-Maf deficiency enables self-renewal of differentiated functional macrophages. *Science* **326**, 867-871, doi:10.1126/science.1176056 (2009).
166 Yona, S. *et al.* Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* **38**, 79-91, doi:10.1016/j.immuni.2012.12.001 (2013).
167 Mossadegh-Keller, N. *et al.* M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* **497**, 239-243, doi:10.1038/nature12026 (2013).
168 Davies, L. C. *et al.* A quantifiable proliferative burst of tissue macrophages restores homeostatic macrophage populations after acute inflammation. *Eur J Immunol* **41**, 2155-2164, doi:10.1002/eji.201141817 (2011).
169 Jenkins, S. J. *et al.* Local macrophage proliferation, rather than recruitment from the blood, is a signature of TH2 inflammation. *Science* **332**, 1284-1288, doi:10.1126/science.1204351 (2011).
170 Leuschner, F. *et al.* Rapid monocyte kinetics in acute myocardial infarction are sustained by extramedullary monocytopoiesis. *J Exp Med* **209**, 123-137, doi:10.1084/jem.20111009 (2012).
171 Jenkins, S. J. *et al.* IL-4 directly signals tissue-resident macrophages to proliferate beyond homeostatic levels controlled by CSF-1. *J Exp Med* **210**, 2477-2491,

doi:10.1084/jem.20121999 (2013).
172 Franklin, R. A. *et al.* The cellular and molecular origin of tumor-associated macrophages. *Science* **344**, 921-925, doi:10.1126/science.1252510 (2014).
173 Murray, P. J. & Wynn, T. A. Protective and pathogenic functions of macrophage subsets. *Nature reviews. Immunology* **11**, 723-737, doi:10.1038/nri3073 (2011).
174 Tatano, Y., Shimizu, T. & Tomioka, H. Unique macrophages different from M1/M2 macrophages inhibit T cell mitogenesis while upregulating Th17 polarization. *Scientific reports* **4**, 4146, doi:10.1038/srep04146 (2014).
175 Wynn, T. A., Chawla, A. & Pollard, J. W. Macrophage biology in development, homeostasis and disease. *Nature* **496**, 445-455, doi:10.1038/nature12034 (2013).
176 Sasmono, R. T. *et al.* A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse. *Blood* **101**, 1155-1163, doi:10.1182/blood-2002-02-0569 (2003).
177 Savill, J. Phagocyte clearance of cells dying by apoptosis and the regulation of glomerular inflammation. *Adv Nephrol Necker Hosp* **31**, 21-28 (2001).
178 De Domenico, I., McVey Ward, D. & Kaplan, J. Regulation of iron acquisition and storage: consequences for iron-linked disorders. *Nature reviews. Molecular cell biology* **9**, 72-81, doi:10.1038/nrm2295 (2008).
179 Corna, G. *et al.* Polarization dictates iron handling by inflammatory and alternatively activated macrophages. *Haematologica* **95**, 1814-1822, doi:10.3324/haematol.2010.023879 (2010).
180 Hotamisligil, G. S. Inflammation and metabolic disorders. *Nature* **444**, 860-867, doi:10.1038/nature05485 (2006).
181 Mackaness, G. B. The influence of immunologically committed lymphoid cells on macrophage activity in vivo. *J Exp Med* **129**, 973-992 (1969).
182 Pace, J. L., Russell, S. W., Schreiber, R. D., Altman, A. & Katz, D. H. Macrophage activation: priming activity from a T-cell hybridoma is attributable to interferon-gamma. *Proc Natl Acad Sci U S A* **80**, 3782-3786 (1983).
183 Celada, A., Gray, P. W., Rinderknecht, E. & Schreiber, R. D. Evidence for a gamma-interferon receptor that regulates macrophage tumoricidal activity. *J Exp Med* **160**, 55-74 (1984).
184 Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *Journal of immunology* **136**, 2348-2357 (1986).
185 Doyle, A. G. *et al.* Interleukin-13 alters the activation state of murine macrophages in vitro: comparison with interleukin-4 and interferon-gamma. *Eur J Immunol* **24**, 1441-1445, doi:10.1002/eji.1830240630 (1994).
186 Stein, M., Keshav, S., Harris, N. & Gordon, S. Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. *J Exp Med* **176**, 287-292 (1992).
187 Mantovani, A. *et al.* The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol* **25**, 677-686, doi:S1471-4906(04)00295-9 [pii] 10.1016/j.it.2004.09.015 (2004).
188 Hsieh, C. S., Macatonia, S. E., O'Garra, A. & Murphy, K. M. T cell genetic background determines default T helper phenotype development in vitro. *J Exp Med* **181**, 713-721 (1995).
189 Mills, C. D., Kincaid, K., Alt, J. M., Heilman, M. J. & Hill, A. M. M-1/M-2 macrophages and the Th1/Th2 paradigm. *Journal of immunology* **164**, 6166-6173 (2000).
190 Gordon, S. & Taylor, P. R. Monocyte and macrophage heterogeneity. *Nature reviews. Immunology* **5**, 953-964, doi:10.1038/nri1733 (2005).
191 Murdoch, C., Giannoudis, A. & Lewis, C. E. Mechanisms regulating the recruitment of macrophages into hypoxic areas of tumors and other ischemic tissues. *Blood* **104**, 2224-2234,

doi:10.1182/blood-2004-03-1109
2004-03-1109 [pii] (2004).
192 Bystrom, J. *et al.* Resolution-phase macrophages possess a unique inflammatory phenotype that is controlled by cAMP. *Blood* **112**, 4117-4127, doi:10.1182/blood-2007-12-129767 (2008).
193 Stout, R. D. *et al.* Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences. *Journal of immunology* **175**, 342-349 (2005).
194 Martinez, F. O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *Journal of immunology* **177**, 7303-7311, doi:177/10/7303 [pii] (2006).
195 Martinez, F. O. *et al.* Genetic programs expressed in resting and IL-4 alternatively activated mouse and human macrophages: similarities and differences. *Blood* **121**, e57-69, doi:10.1182/blood-2012-06-436212 (2013).
196 Pello, O. M. *et al.* Role of c-MYC in alternative activation of human macrophages and tumor-associated macrophage biology. *Blood* **119**, 411-421, doi:10.1182/blood-2011-02-339911 (2012).
197 Pello, O. M. *et al.* In vivo inhibition of c-MYC in myeloid cells impairs tumor-associated macrophage maturation and pro-tumoral activities. *PloS one* **7**, e45399, doi:10.1371/journal.pone.0045399 (2012).
198 Xue, J. *et al.* Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity* **40**, 274-288, doi:10.1016/j.immuni.2014.01.006 (2014).
199 Gautier, E. L. *et al.* Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat Immunol* **13**, 1118-1128, doi:10.1038/ni.2419 (2012).
200 Medzhitov, R. Toll-like receptors and innate immunity. *Nature reviews. Immunology* **1**, 135-145, doi:10.1038/35100529 (2001).
201 Hagemann, T. *et al.* "Re-educating" tumor-associated macrophages by targeting NF-kappaB. *J Exp Med* **205**, 1261-1268, doi:jem.20080108 [pii]
10.1084/jem.20080108 (2008).
202 Gordon, S. & Martinez, F. O. Alternative activation of macrophages: mechanism and functions. *Immunity* **32**, 593-604, doi:10.1016/j.immuni.2010.05.007 (2010).
203 Bouhlel, M. A. *et al.* PPARgamma activation primes human monocytes into alternative M2 macrophages with anti-inflammatory properties. *Cell Metab* **6**, 137-143, doi:S1550-4131(07)00166-0 [pii]
10.1016/j.cmet.2007.06.010 (2007).
204 Ruffell, D. *et al.* A CREB-C/EBPbeta cascade induces M2 macrophage-specific gene expression and promotes muscle injury repair. *Proc Natl Acad Sci U S A* **106**, 17475-17480, doi:10.1073/pnas.0908641106 (2009).
205 Ivashkiv, L. B. Epigenetic regulation of macrophage polarization and function. *Trends Immunol* **34**, 216-223, doi:10.1016/j.it.2012.11.001 (2013).
206 De Santa, F. *et al.* The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* **130**, 1083-1094, doi:10.1016/j.cell.2007.08.019 (2007).
207 Ghisletti, S. *et al.* Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317-328, doi:10.1016/j.immuni.2010.02.008 (2010).
208 Hartmann, S. *et al.* Macrophages in T cell/histiocyte rich large B cell lymphoma strongly express metal-binding proteins and show a bi-activated phenotype. *International*

journal of cancer. Journal international du cancer **133**, 2609-2618, doi:10.1002/ijc.28273 (2013).
209 Murray, P. J. & Wynn, T. A. Obstacles and opportunities for understanding macrophage polarization. *Journal of leukocyte biology* **89**, 557-563, doi:10.1189/jlb.0710409 (2011).
210 Schneemann, M. & Schoedon, G. Species differences in macrophage NO production are important. *Nat Immunol* **3**, 102, doi:10.1038/ni0202-102a (2002).
211 Choi, H. S., Rai, P. R., Chu, H. W., Cool, C. & Chan, E. D. Analysis of nitric oxide synthase and nitrotyrosine expression in human pulmonary tuberculosis. *American journal of respiratory and critical care medicine* **166**, 178-186 (2002).
212 Schroder, K. *et al.* Conservation and divergence in Toll-like receptor 4-regulated gene expression in primary human versus mouse macrophages. *Proc Natl Acad Sci U S A* **109**, E944-953, doi:10.1073/pnas.1110156109 (2012).
213 Murray, P. J. *et al.* Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. *Immunity* **41**, 14-20, doi:10.1016/j.immuni.2014.06.008 (2014).
214 Zhu, Z., Wang, D. C., Popescu, L. I. & Wang, X. Single-cell transcriptome in the identification of disease biomarkers: opportunities and challenges. *Journal of translational medicine* **12**, 212, doi:10.1186/s12967-014-0212-3 (2014).
215 Stougaard, M., Lohmann, J. S., Zajac, M., Hamilton-Dutoit, S. & Koch, J. In situ detection of non-polyadenylated RNA molecules using Turtle Probes and target primed rolling circle PRINS. *BMC biotechnology* **7**, 69, doi:10.1186/1472-6750-7-69 (2007).
216 Zhang, Q. W. *et al.* Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. *PloS one* **7**, e50946, doi:10.1371/journal.pone.0050946 (2012).
217 Forssell, J. *et al.* High macrophage infiltration along the tumor front correlates with improved survival in colon cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **13**, 1472-1479, doi:10.1158/1078-0432.CCR-06-2073 (2007).
218 Piras, F. *et al.* The predictive value of CD8, CD4, CD68, and human leukocyte antigen-D-related cells in the prognosis of cutaneous malignant melanoma with vertical growth phase. *Cancer* **104**, 1246-1254, doi:10.1002/cncr.21283 (2005).
219 Greaves, P. *et al.* Expression of FOXP3, CD68, and CD20 at diagnosis in the microenvironment of classical Hodgkin lymphoma is predictive of outcome. *J Clin Oncol* **31**, 256-262, doi:10.1200/JCO.2011.39.9881 (2013).
220 Qian, B. Z. *et al.* CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature*, doi:10.1038/nature10138 (2011).
221 Guilloton, F. *et al.* Mesenchymal stromal cells orchestrate follicular lymphoma cell niche through the CCL2-dependent recruitment and polarization of monocytes. *Blood* **119**, 2556-2567, doi:10.1182/blood-2011-08-370908 (2012).
222 Lin, E. Y., Nguyen, A. V., Russell, R. G. & Pollard, J. W. Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. *J Exp Med* **193**, 727-740 (2001).
223 Sarkar, S. *et al.* Therapeutic activation of macrophages and microglia to suppress brain tumor-initiating cells. *Nat Neurosci* **17**, 46-55, doi:10.1038/nn.3597 (2014).
224 Wu, Y. & Zheng, L. Dynamic education of macrophages in different areas of human tumors. *Cancer microenvironment : official journal of the International Cancer Microenvironment Society* **5**, 195-201, doi:10.1007/s12307-012-0113-z (2012).
225 Deng, L. *et al.* A novel mouse model of inflammatory bowel disease links mammalian target of rapamycin-dependent hyperproliferation of colonic epithelium to inflammation-associated tumorigenesis. *The American journal of pathology* **176**, 952-967, doi:10.2353/ajpath.2010.090622 (2010).
226 Munari, F. *et al.* Tumor-associated macrophages as major source of APRIL in gastric MALT lymphoma. *Blood*, doi:blood-2010-06-293266 [pii]

10.1182/blood-2010-06-293266 (2011).
227 Langowski, J. L. *et al.* IL-23 promotes tumour incidence and growth. *Nature* **442**, 461-465, doi:10.1038/nature04808 (2006).
228 Kryczek, I. *et al.* B7-H4 expression identifies a novel suppressive macrophage population in human ovarian carcinoma. *J Exp Med* **203**, 871-881, doi:10.1084/jem.20050930 (2006).
229 Ojalvo, L. S., Whittaker, C. A., Condeelis, J. S. & Pollard, J. W. Gene expression analysis of macrophages that facilitate tumor invasion supports a role for Wnt-signaling in mediating their activity in primary mammary tumors. *Journal of immunology* **184**, 702-712, doi:jimmunol.0902360 [pii]
10.4049/jimmunol.0902360 (2010).
230 Coussens, L. M., Tinkle, C. L., Hanahan, D. & Werb, Z. MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. *Cell* **103**, 481-490 (2000).
231 Gocheva, V. *et al.* IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion. *Genes Dev* **24**, 241-255, doi:gad.1874010 [pii]
10.1101/gad.1874010 (2010).
232 Vasiljeva, O. *et al.* Tumor cell-derived and macrophage-derived cathepsin B promotes progression and lung metastasis of mammary cancer. *Cancer research* **66**, 5242-5250, doi:10.1158/0008-5472.CAN-05-4463 (2006).
233 Wyckoff, J. *et al.* A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. *Cancer research* **64**, 7022-7029, doi:10.1158/0008-5472.CAN-04-1449 (2004).
234 DeNardo, D. G. *et al.* CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell* **16**, 91-102, doi:S1535-6108(09)00216-5 [pii]
10.1016/j.ccr.2009.06.018 (2009).
235 Solinas, G. *et al.* Tumor-conditioned macrophages secrete migration-stimulating factor: a new marker for M2-polarization, influencing tumor cell motility. *Journal of immunology* **185**, 642-652, doi:10.4049/jimmunol.1000413 (2010).
236 Torroella-Kouri, M. *et al.* Identification of a subpopulation of macrophages in mammary tumor-bearing mice that are neither M1 nor M2 and are less differentiated. *Cancer research* **69**, 4800-4809, doi:10.1158/0008-5472.CAN-08-3427 (2009).
237 Wu, K., Kryczek, I., Chen, L., Zou, W. & Welling, T. H. Kupffer cell suppression of CD8+ T cells in human hepatocellular carcinoma is mediated by B7-H1/programmed death-1 interactions. *Cancer research* **69**, 8067-8075, doi:10.1158/0008-5472.CAN-09-0901 (2009).
238 Kuang, D. M. *et al.* Activated monocytes in peritumoral stroma of hepatocellular carcinoma foster immune privilege and disease progression through PD-L1. *J Exp Med* **206**, 1327-1337, doi:10.1084/jem.20082173 (2009).
239 Curiel, T. J. *et al.* Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nature medicine* **10**, 942-949, doi:10.1038/nm1093 (2004).
240 Du, R. *et al.* HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. *Cancer Cell* **13**, 206-220, doi:10.1016/j.ccr.2008.01.034 (2008).
241 Stockmann, C. *et al.* Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis. *Nature* **456**, 814-818, doi:10.1038/nature07445 (2008).
242 Lin, E. Y. *et al.* Vascular endothelial growth factor restores delayed tumor progression in tumors depleted of macrophages. *Mol Oncol* **1**, 288-302, doi:10.1016/j.molonc.2007.10.003 (2007).
243 Lewis, C. & Murdoch, C. Macrophage responses to hypoxia: implications for tumor

progression and anti-cancer therapies. *The American journal of pathology* **167**, 627-635, doi:10.1016/S0002-9440(10)62038-X (2005).
244 Doedens, A. L. *et al.* Macrophage expression of hypoxia-inducible factor-1 alpha suppresses T-cell function and promotes tumor progression. *Cancer research* **70**, 7465-7475, doi:10.1158/0008-5472.CAN-10-1439 (2010).
245 Pucci, F. *et al.* A distinguishing gene signature shared by tumor-infiltrating Tie2-expressing monocytes, blood "resident" monocytes, and embryonic macrophages suggests common functions and developmental relationships. *Blood* **114**, 901-914, doi:blood-2009-01-200931 [pii]
10.1182/blood-2009-01-200931 (2009).
246 Mazzieri, R. *et al.* Targeting the ANG2/TIE2 axis inhibits tumor growth and metastasis by impairing angiogenesis and disabling rebounds of proangiogenic myeloid cells. *Cancer Cell* **19**, 512-526, doi:10.1016/j.ccr.2011.02.005 (2011).
247 Casazza, A. *et al.* Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor immunity. *Cancer Cell* **24**, 695-709, doi:10.1016/j.ccr.2013.11.007 (2013).
248 Huang, Y. *et al.* Vascular normalizing doses of antiangiogenic treatment reprogram the immunosuppressive tumor microenvironment and enhance immunotherapy. *Proc Natl Acad Sci U S A* **109**, 17561-17566, doi:10.1073/pnas.1215397109 (2012).
249 Lewis, C. E. & Pollard, J. W. Distinct role of macrophages in different tumor microenvironments. *Cancer research* **66**, 605-612, doi:66/2/605 [pii]
10.1158/0008-5472.CAN-05-4005 (2006).
250 Biswas, S. K. *et al.* A distinct and unique transcriptional program expressed by tumor-associated macrophages (defective NF-kappaB and enhanced IRF-3/STAT1 activation). *Blood* **107**, 2112-2122, doi:2005-01-0428 [pii]
10.1182/blood-2005-01-0428 (2006).
251 Laoui, D. *et al.* Tumor-associated macrophages in breast cancer: distinct subsets, distinct functions. *The International journal of developmental biology* **55**, 861-867, doi:10.1387/ijdb.113371dl (2011).
252 Movahedi, K. *et al.* Different tumor microenvironments contain functionally distinct subsets of macrophages derived from Ly6C(high) monocytes. *Cancer research* **70**, 5728-5739, doi:10.1158/0008-5472.CAN-09-4672 (2010).
253 Grugan, K. D. *et al.* Tumor-associated macrophages promote invasion while retaining Fc-dependent anti-tumor function. *Journal of immunology* **189**, 5457-5466, doi:10.4049/jimmunol.1201889 (2012).
254 Savage, N. D. *et al.* Human anti-inflammatory macrophages induce Foxp3+ GITR+ CD25+ regulatory T cells, which suppress via membrane-bound TGFbeta-1. *Journal of immunology* **181**, 2220-2226 (2008).
255 Porta, C. *et al.* Tolerance and M2 (alternative) macrophage polarization are related processes orchestrated by p50 nuclear factor kappaB. *Proc Natl Acad Sci U S A* **106**, 14978-14983, doi:0809784106 [pii]
10.1073/pnas.0809784106 (2009).
256 Stout, R. D., Watkins, S. K. & Suttles, J. Functional plasticity of macrophages: in situ reprogramming of tumor-associated macrophages. *Journal of leukocyte biology* **86**, 1105-1109, doi:10.1189/jlb.0209073 (2009).
257 Sutterwala, F. S., Noel, G. J., Salgame, P. & Mosser, D. M. Reversal of proinflammatory responses by ligating the macrophage Fcgamma receptor type I. *J Exp Med* **188**, 217-222 (1998).
258 Zhang, Y. *et al.* Immune complex/Ig negatively regulate TLR4-triggered inflammatory response in macrophages through Fc gamma RIIb-dependent PGE2 production. *Journal of immunology* **182**, 554-562 (2009).

259 Griffin, D. O., Holodick, N. E. & Rothstein, T. L. Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+ CD27+ CD43+ CD70. *J Exp Med* **208**, 67-80, doi:10.1084/jem.20101499 (2011).
260 Wong, S. C. *et al.* Macrophage polarization to a unique phenotype driven by B cells. *Eur J Immunol* **40**, 2296-2307, doi:10.1002/eji.200940288 (2010).
261 Andreu, P. *et al.* FcRgamma activation regulates inflammation-associated squamous carcinogenesis. *Cancer Cell* **17**, 121-134, doi:10.1016/j.ccr.2009.12.019 (2010).
262 Levesque, M. C. Translational Mini-Review Series on B Cell-Directed Therapies: Recent advances in B cell-directed biological therapies for autoimmune disorders. *Clinical and experimental immunology* **157**, 198-208, doi:10.1111/j.1365-2249.2009.03979.x (2009).
263 Lund, F. E. & Randall, T. D. Effector and regulatory B cells: modulators of CD4+ T cell immunity. *Nature reviews. Immunology* **10**, 236-247, doi:10.1038/nri2729 (2010).
264 Bouaziz, J. D. *et al.* Therapeutic B cell depletion impairs adaptive and autoreactive CD4+ T cell activation in mice. *Proc Natl Acad Sci U S A* **104**, 20878-20883, doi:10.1073/pnas.0709205105 (2007).
265 O'Neill, S. K. *et al.* Expression of CD80/86 on B cells is essential for autoreactive T cell activation and the development of arthritis. *Journal of immunology* **179**, 5109-5116 (2007).
266 Shah, S. & Qiao, L. Resting B cells expand a CD4+CD25+Foxp3+ Treg population via TGF-beta3. *Eur J Immunol* **38**, 2488-2498, doi:10.1002/eji.200838201 (2008).
267 Harris, D. P. *et al.* Reciprocal regulation of polarized cytokine production by effector B and T cells. *Nat Immunol* **1**, 475-482, doi:10.1038/82717 (2000).
268 Glennie, M. J., French, R. R., Cragg, M. S. & Taylor, R. P. Mechanisms of killing by anti-CD20 monoclonal antibodies. *Mol Immunol* **44**, 3823-3837, doi:10.1016/j.molimm.2007.06.151 (2007).
269 Clynes, R., Takechi, Y., Moroi, Y., Houghton, A. & Ravetch, J. V. Fc receptors are required in passive and active immunity to melanoma. *Proc Natl Acad Sci U S A* **95**, 652-656 (1998).
270 Minard-Colin, V. *et al.* Lymphoma depletion during CD20 immunotherapy in mice is mediated by macrophage FcgammaRI, FcgammaRIII, and FcgammaRIV. *Blood* **112**, 1205-1213, doi:10.1182/blood-2008-01-135160 (2008).
271 Cittera, E. *et al.* The CCL3 family of chemokines and innate immunity cooperate in vivo in the eradication of an established lymphoma xenograft by rituximab. *Journal of immunology* **178**, 6616-6623, doi:178/10/6616 [pii] (2007).
272 Leidi, M. *et al.* M2 macrophages phagocytose rituximab-opsonized leukemic targets more efficiently than m1 cells in vitro. *Journal of immunology* **182**, 4415-4422, doi:182/7/4415 [pii]
10.4049/jimmunol.0713732 (2009).
273 Chao, M. P. *et al.* Anti-CD47 antibody synergizes with rituximab to promote phagocytosis and eradicate non-Hodgkin lymphoma. *Cell* **142**, 699-713, doi:10.1016/j.cell.2010.07.044 (2010).
274 Steidl, C., Farinha, P. & Gascoyne, R. D. Macrophages predict treatment outcome in Hodgkin's lymphoma. *Haematologica* **96**, 186-189, doi:96/2/186 [pii]
10.3324/haematol.2010.033316 (2011).
275 Pulford, K. A., Sipos, A., Cordell, J. L., Stross, W. P. & Mason, D. Y. Distribution of the CD68 macrophage/myeloid associated antigen. *Int Immunol* **2**, 973-980 (1990).
276 Gordon, S. & Pluddemann, A. Tissue macrophage heterogeneity: issues and prospects. *Seminars in immunopathology* **35**, 533-540, doi:10.1007/s00281-013-0386-4 (2013).
277 Vacca, A. *et al.* Angiogenesis extent and macrophage density increase simultaneously with pathological progression in B-cell non-Hodgkin's lymphomas. *Br J Cancer* **79**, 965-970, doi:10.1038/sj.bjc.6690154 (1999).

278 Hasselblom, S. *et al.* Expression of CD68+ tumor-associated macrophages in patients with diffuse large B-cell lymphoma and its relation to prognosis. *Pathol Int* **58**, 529-532, doi:PIN2268 [pii]
10.1111/j.1440-1827.2008.02268.x (2008).
279 Cai, Q. C. *et al.* High expression of tumor-infiltrating macrophages correlates with poor prognosis in patients with diffuse large B-cell lymphoma. *Med Oncol*, doi:10.1007/s12032-011-0123-6 (2011).
280 Brandt, S. *et al.* The combined expression of the stromal markers fibronectin and SPARC improves the prediction of survival in diffuse large B-cell lymphoma. *Experimental hematology & oncology* **2**, 27, doi:10.1186/2162-3619-2-27 (2013).
281 Bradshaw, A. D. Diverse biological functions of the SPARC family of proteins. *Int J Biochem Cell Biol* **44**, 480-488, doi:10.1016/j.biocel.2011.12.021 (2012).
282 Meyer, P. N. *et al.* The stromal cell marker SPARC predicts for survival in patients with diffuse large B-cell lymphoma treated with rituximab. *Am J Clin Pathol* **135**, 54-61, doi:10.1309/AJCPJX4BJV9NLQHY (2011).
283 Wada, N. *et al.* Tumour-associated macrophages in diffuse large B-cell lymphoma: a study of the Osaka Lymphoma Study Group. *Histopathology* **60**, 313-319, doi:10.1111/j.1365-2559.2011.04096.x (2012).
284 Epron, G. *et al.* Monocytes and T cells cooperate to favor normal and follicular lymphoma B-cell growth: role of IL-15 and CD40L signaling. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, doi:10.1038/leu.2011.179 (2011).
285 Evans, R. Phagocytosis of murine lymphoma cells by macrophages. I. Factors affecting in vitro phagocytosis. *Immunology* **20**, 67-74 (1971).
286 Evans, R. Phagocytosis of murine lymphoma cells by macrophages. II. Differences between opsonic and cytotoxic activity of mice immunized with lymphoma cells. *Immunology* **20**, 75-83 (1971).
287 Gemsa, D., Leser, H. G., Deimann, W. & Resch, K. Suppression of T lymphocyte proliferation during lymphoma growth in mice: role of PGE2-producing suppressor macrophages. *Immunobiology* **161**, 385-391 (1982).
288 Savill, J., Dransfield, I., Gregory, C. & Haslett, C. A blast from the past: clearance of apoptotic cells regulates immune responses. *Nature reviews. Immunology* **2**, 965-975, doi:10.1038/nri957 (2002).
289 Ronchetti, A. *et al.* Immunogenicity of apoptotic cells in vivo: role of antigen load, antigen-presenting cells, and cytokines. *Journal of immunology* **163**, 130-136 (1999).
290 Truman, L. A. *et al.* CX3CL1/fractalkine is released from apoptotic lymphocytes to stimulate macrophage chemotaxis. *Blood* **112**, 5026-5036, doi:blood-2008-06-162404 [pii]
10.1182/blood-2008-06-162404 (2008).
291 Ogden, C. A. *et al.* Enhanced apoptotic cell clearance capacity and B cell survival factor production by IL-10-activated macrophages: implications for Burkitt's lymphoma. *Journal of immunology* **174**, 3015-3023 (2005).
292 Reimann, M. *et al.* Tumor stroma-derived TGF-beta limits myc-driven lymphomagenesis via Suv39h1-dependent senescence. *Cancer Cell* **17**, 262-272, doi:10.1016/j.ccr.2009.12.043 (2010).
293 Haabeth, O. A. *et al.* Inflammation driven by tumour-specific Th1 cells protects against B-cell cancer. *Nat Commun* **2**, 240, doi:ncomms1239 [pii]
10.1038/ncomms1239 (2011).
294 Kononen, J. *et al.* Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* **4**, 844-847 (1998).
295 Zettl, A. *et al.* Immunohistochemical analysis of B-cell lymphoma using tissue microarrays identifies particular phenotypic profiles of B-cell lymphomas. *Histopathology* **43**, 209-219, doi:1702 [pii] (2003).

296 Hedvat, C. V. *et al.* Application of tissue microarray technology to the study of non-Hodgkin's and Hodgkin's lymphoma. *Hum Pathol* **33**, 968-974, doi:S0046817702000977 [pii] (2002).
297 Nakane, P. K. & Pierce, G. B., Jr. Enzyme-labeled antibodies: preparation and application for the localization of antigens. *J Histochem Cytochem* **14**, 929-931 (1966).
298 Wilfinger, W. W., Mackey, K. & Chomczynski, P. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques* **22**, 474-476, 478-481 (1997).
299 Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* **7**, 3, doi:1471-2199-7-3 [pii]
10.1186/1471-2199-7-3 (2006).
300 Dunleavy, K. *et al.* Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. *Blood* **113**, 6069-6076, doi:blood-2009-01-199679 [pii]
10.1182/blood-2009-01-199679 (2009).
301 Hernandez-Ilizaliturri, F. J. *et al.* Higher response to lenalidomide in relapsed/refractory diffuse large b-cell lymphoma in nongerminal center b-cell-like than in germinal center b-cell-like phenotype. *Cancer*, doi:10.1002/cncr.26135 (2011).
302 Hans, C. P. *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* **103**, 275-282, doi:10.1182/blood-2003-05-1545
2003-05-1545 [pii] (2004).
303 Muris, J. J. *et al.* Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *J Pathol* **208**, 714-723, doi:10.1002/path.1924 (2006).
304 Natkunam, Y. *et al.* LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. *J Clin Oncol* **26**, 447-454, doi:JCO.2007.13.0690 [pii]
10.1200/JCO.2007.13.0690 (2008).
305 Nyman, H., Jerkeman, M., Karjalainen-Lindsberg, M. L., Banham, A. H. & Leppa, S. Prognostic impact of activated B-cell focused classification in diffuse large B-cell lymphoma patients treated with R-CHOP. *Mod Pathol* **22**, 1094-1101, doi:10.1038/modpathol.2009.73 (2009).
306 Choi, W. W. *et al.* A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **15**, 5494-5502, doi:1078-0432.CCR-09-0113 [pii]
10.1158/1078-0432.CCR-09-0113 (2009).
307 Meyer, P. N. *et al.* Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. *J Clin Oncol* **29**, 200-207, doi:JCO.2010.30.0368 [pii]
10.1200/JCO.2010.30.0368 (2011).
308 Visco, C. *et al.* Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K* **26**, 2103-2113, doi:10.1038/leu.2012.83 (2012).
309 Moskowitz, C. H. *et al.* Cell of origin, germinal center versus nongerminal center, determined by immunohistochemistry on tissue microarray, does not correlate with outcome in patients with relapsed and refractory DLBCL. *Blood* **106**, 3383-3385, doi:2005-04-1603 [pii]

10.1182/blood-2005-04-1603 (2005).
310 Ott, M. M., Horn, H., Kaufmann, M. & Ott, G. The Hans classificator does not predict outcome in diffuse large B cell lymphoma in a large multicenter retrospective analysis of R-CHOP treated patients. *Leuk Res* **36**, 544-545, doi:10.1016/j.leukres.2012.01.022 (2012).
311 Gutierrez-Garcia, G. *et al.* Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Blood* **117**, 4836-4843, doi:10.1182/blood-2010-12-322362 (2011).
312 Fu, K. *et al.* Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. *J Clin Oncol* **26**, 4587-4594, doi:JCO.2007.15.9277 [pii]
10.1200/JCO.2007.15.9277 (2008).
313 Thieblemont, C. *et al.* The germinal center/activated B-cell subclassification has a prognostic impact for response to salvage therapy in relapsed/refractory diffuse large B-cell lymphoma: a bio-CORAL study. *J Clin Oncol* **29**, 4079-4087, doi:10.1200/JCO.2011.35.4423 (2011).
314 de Jong, D. *et al.* Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications--a study from the Lunenburg Lymphoma Biomarker Consortium. *J Clin Oncol* **25**, 805-812, doi:10.1200/JCO.2006.09.4490 (2007).
315 Seki, R. *et al.* Prognostic impact of immunohistochemical biomarkers in diffuse large B-cell lymphoma in the rituximab era. *Cancer Sci* **100**, 1842-1847, doi:CAS1268 [pii]
10.1111/j.1349-7006.2009.01268.x (2009).
316 Horn, H. *et al.* MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. *Blood* **121**, 2253-2263, doi:10.1182/blood-2012-06-435842 (2013).
317 Montes-Moreno, S. *et al.* Gcet1 (centerin), a highly restricted marker for a subset of germinal center-derived lymphomas. *Blood* **111**, 351-358, doi:blood-2007-06-094151 [pii]
10.1182/blood-2007-06-094151 (2008).
318 Barrans, S. L., Fenton, J. A., Banham, A., Owen, R. G. & Jack, A. S. Strong expression of FOXP1 identifies a distinct subset of diffuse large B-cell lymphoma (DLBCL) patients with poor outcome. *Blood* **104**, 2933-2935, doi:10.1182/blood-2004-03-1209 (2004).
319 Natkunam, Y. *et al.* The oncoprotein LMO2 is expressed in normal germinal-center B cells and in human B-cell lymphomas. *Blood* **109**, 1636-1642, doi:blood-2006-08-039024 [pii]
10.1182/blood-2006-08-039024 (2007).
320 van Imhoff, G. W. *et al.* Prognostic impact of germinal center-associated proteins and chromosomal breakpoints in poor-risk diffuse large B-cell lymphoma. *J Clin Oncol* **24**, 4135-4142, doi:10.1200/JCO.2006.05.5897 (2006).
321 Maeshima, A. M. *et al.* Bcl-2, Bcl-6, and the International Prognostic Index are prognostic indicators in patients with diffuse large B-cell lymphoma treated with rituximab-containing chemotherapy. *Cancer science*, doi:10.1111/j.1349-7006.2012.02382.x (2012).
322 Mounier, N. *et al.* Rituximab plus CHOP (R-CHOP) overcomes bcl-2--associated resistance to chemotherapy in elderly patients with diffuse large B-cell lymphoma (DLBCL). *Blood* **101**, 4279-4284, doi:10.1182/blood-2002-11-3442 (2003).
323 Wilson, K. S. *et al.* CHOP-R therapy overcomes the adverse prognostic influence of BCL-2 expression in diffuse large B-cell lymphoma. *Leuk Lymphoma* **48**, 1102-1109, doi:10.1080/10428190701344881 (2007).
324 Salles, G. *et al.* Prognostic significance of immunohistochemical biomarkers in diffuse large B-cell lymphoma: a study from the Lunenburg Lymphoma Biomarker Consortium. *Blood*, doi:blood-2011-04-345256 [pii]
10.1182/blood-2011-04-345256 (2011).
325 Challa-Malladi, M. *et al.* Combined genetic inactivation of beta2-Microglobulin and

CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell* **20**, 728-740, doi:10.1016/j.ccr.2011.11.006 (2011).
326 Hasselblom, S. *et al.* The number of tumour-infiltrating TIA-1+ cytotoxic T cells but not FOXP3+ regulatory T cells predicts outcome in diffuse large B-cell lymphoma. *Br J Haematol* **137**, 364-373, doi:BJH6593 [pii]
10.1111/j.1365-2141.2007.06593.x (2007).
327 Tzankov, A. *et al.* Correlation of high numbers of intratumoral FOXP3+ regulatory T cells with improved survival in germinal center-like diffuse large B-cell lymphoma, follicular lymphoma and classical Hodgkin's lymphoma. *Haematologica* **93**, 193-200, doi:haematol.11702 [pii]
10.3324/haematol.11702 (2008).
328 Yhaka, R. a. G. R. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299-314 (1996).
329 Sander, B. *et al.* The reliability of immunohistochemical analysis of the tumor microenvironment in follicular lymphoma: a validation study from the Lunenburg Lymphoma Biomarker Consortium. *Haematologica* **99**, 715-725, doi:10.3324/haematol.2013.095257 (2014).
330 Wahlin, B. E. *et al.* T cells in tumors and blood predict outcome in follicular lymphoma treated with rituximab. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17**, 4136-4144, doi:10.1158/1078-0432.CCR-11-0264 (2011).
331 Lee, A. M. *et al.* Number of CD4+ cells and location of forkhead box protein P3-positive cells in diagnostic follicular lymphoma tissue microarrays correlates with outcome. *J Clin Oncol* **24**, 5052-5059, doi:10.1200/JCO.2006.06.4642 (2006).
332 Yang, Z. Z., Novak, A. J., Stenson, M. J., Witzig, T. E. & Ansell, S. M. Intratumoral CD4+CD25+ regulatory T-cell-mediated suppression of infiltrating CD4+ T cells in B-cell non-Hodgkin lymphoma. *Blood* **107**, 3639-3646, doi:2005-08-3376 [pii]
10.1182/blood-2005-08-3376 (2006).
333 Yang, Z. Z., Novak, A. J., Ziesmer, S. C., Witzig, T. E. & Ansell, S. M. Attenuation of CD8(+) T-cell function by CD4(+)CD25(+) regulatory T cells in B-cell non-Hodgkin's lymphoma. *Cancer research* **66**, 10145-10152, doi:66/20/10145 [pii]
10.1158/0008-5472.CAN-06-1822 (2006).
334 Galani, I. E. *et al.* Regulatory T cells control macrophage accumulation and activation in lymphoma. *International journal of cancer. Journal international du cancer* **127**, 1131-1140, doi:10.1002/ijc.25132 (2010).
335 Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264, doi:10.1093/biostatistics/4.2.249 (2003).
336 Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193 (2003).
337 Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**, 14863-14868 (1998).
338 Abbas, A. R. *et al.* Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* **6**, 319-331, doi:10.1038/sj.gene.6364173 (2005).
339 Basso, K. Toward a systems biology approach to investigate cellular networks in normal and malignant B cells. *Leukemia* **23**, 1219-1225, doi:10.1038/leu.2009.4 (2009).
340 Van Loo, P. *et al.* T-cell/histiocyte-rich large B-cell lymphoma shows transcriptional features suggestive of a tolerogenic host immune response. *Haematologica* **95**, 440-448, doi:10.3324/haematol.2009.009647 (2010).
341 Tome, M. E. *et al.* A redox signature score identifies diffuse large B-cell lymphoma

patients with a poor prognosis. *Blood* **106**, 3594-3601, doi:10.1182/blood-2005-02-0487 (2005).
342 Segal, E., Friedman, N., Kaminski, N., Regev, A. & Koller, D. From signatures to models: understanding cancer using microarrays. *Nat Genet* **37 Suppl**, S38-45, doi:10.1038/ng1561 (2005).
343 Poulsen, C. B. *et al.* Prognostic significance of metallothionein in B-cell lymphomas. *Blood* **108**, 3514-3519, doi:10.1182/blood-2006-04-015305 (2006).
344 Greenberg, M. E. *et al.* Oxidized phosphatidylserine-CD36 interactions play an essential role in macrophage-dependent phagocytosis of apoptotic cells. *J Exp Med* **203**, 2613-2625, doi:10.1084/jem.20060370 (2006).
345 Feng, J. *et al.* Induction of CD36 expression by oxidized LDL and IL-4 by a common signaling pathway dependent on protein kinase C and PPAR-gamma. *Journal of lipid research* **41**, 688-696 (2000).
346 Huang, S. C. *et al.* Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages. *Nat Immunol* **15**, 846-855, doi:10.1038/ni.2956 (2014).
347 Clement-Ziza, M. *et al.* Evaluation of methods for amplification of picogram amounts of total RNA for whole genome expression profiling. *BMC genomics* **10**, 246, doi:10.1186/1471-2164-10-246 (2009).
348 Svaren, J. *et al.* EGR1 target genes in prostate carcinoma cells identified by microarray analysis. *The Journal of biological chemistry* **275**, 38524-38531, doi:10.1074/jbc.M005220200 (2000).
349 Doig, T. N. *et al.* Coexpression analysis of large cancer datasets provides insight into the cellular phenotypes of the tumour microenvironment. *BMC genomics* **14**, 469, doi:10.1186/1471-2164-14-469 (2013).
350 Mueller, C. G. *et al.* Critical role of monocytes to support normal B cell and diffuse large B cell lymphoma survival and proliferation. *Journal of leukocyte biology* **82**, 567-575, doi:10.1189/jlb.0706481 (2007).
351 Lin, Y. *et al.* Immunosuppressive CD14+HLA-DR(low)/- monocytes in B-cell non-Hodgkin lymphoma. *Blood* **117**, 872-881, doi:10.1182/blood-2010-05-283820 (2011).
352 Agre, P. & Kozono, D. Aquaporin water channels: molecular mechanisms for human diseases. *FEBS letters* **555**, 72-78 (2003).
353 Ishibashi, K. *et al.* Cloning and functional expression of a new aquaporin (AQP9) abundantly expressed in the peripheral leukocytes permeable to water and urea, but not to glycerol. *Biochemical and biophysical research communications* **244**, 268-274, doi:10.1006/bbrc.1998.8252 (1998).
354 Van den Bossche, J., Malissen, B., Mantovani, A., De Baetselier, P. & Van Ginderachter, J. A. Regulation and function of the E-cadherin/catenin complex in cells of the monocyte-macrophage lineage and DCs. *Blood* **119**, 1623-1633, doi:10.1182/blood-2011-10-384289 (2012).
355 Godin-Ethier, J., Hanafi, L. A., Duvignaud, J. B., Leclerc, D. & Lapointe, R. IDO expression by human B lymphocytes in response to T lymphocyte stimuli and TLR engagement is biologically inactive. *Mol Immunol* **49**, 253-259, doi:10.1016/j.molimm.2011.08.017 (2011).
356 Greaves, P. & Gribben, J. G. The role of B7 family molecules in hematologic malignancy. *Blood* **121**, 734-744, doi:10.1182/blood-2012-10-385591 (2013).
357 Loke, P. & Allison, J. P. PD-L1 and PD-L2 are differentially regulated by Th1 and Th2 cells. *Proc Natl Acad Sci U S A* **100**, 5336-5341, doi:10.1073/pnas.0931259100 (2003).
358 Semnani, R. T., Mahapatra, L., Moore, V., Sanprasert, V. & Nutman, T. B. Functional and phenotypic characteristics of alternative activation induced in human monocytes by interleukin-4 or the parasitic nematode Brugia malayi. *Infection and immunity* **79**, 3957-3965, doi:10.1128/IAI.05191-11 (2011).

359 Caro, P. *et al.* Metabolic signatures uncover distinct targets in molecular subsets of diffuse large B cell lymphoma. *Cancer Cell* **22**, 547-560, doi:10.1016/j.ccr.2012.08.014 (2012).
360 Hathcock, K. S., Laszlo, G., Pucillo, C., Linsley, P. & Hodes, R. J. Comparative analysis of B7-1 and B7-2 costimulatory ligands: expression and function. *J Exp Med* **180**, 631-640 (1994).
361 Orlikowsky, T. W. *et al.* Expression and regulation of B7 family molecules on macrophages (MPhi) in preterm and term neonatal cord blood and peripheral blood of adults. *Cytometry. Part B, Clinical cytometry* **53**, 40-47, doi:10.1002/cyto.b.10033 (2003).
362 Alvaro, T. *et al.* The presence of STAT1-positive tumor-associated macrophages and their relation to outcome in patients with follicular lymphoma. *Haematologica* **91**, 1605-1612 (2006).
363 Seymour, J. F., Talpaz, M., Cabanillas, F., Wetzler, M. & Kurzrock, R. Serum interleukin-6 levels correlate with prognosis in diffuse large-cell lymphoma. *J Clin Oncol* **13**, 575-582 (1995).
364 Lech-Maranda, E. *et al.* Plasma TNF-alpha and IL-10 level-based prognostic model predicts outcome of patients with diffuse large B-Cell lymphoma in different risk groups defined by the International Prognostic Index. *Archivum immunologiae et therapiae experimentalis* **58**, 131-141, doi:10.1007/s00005-010-0066-1 (2010).
365 Lech-Maranda, E. *et al.* Elevated IL-10 plasma levels correlate with poor prognosis in diffuse large B-cell lymphoma. *European cytokine network* **17**, 60-66 (2006).

Appendix

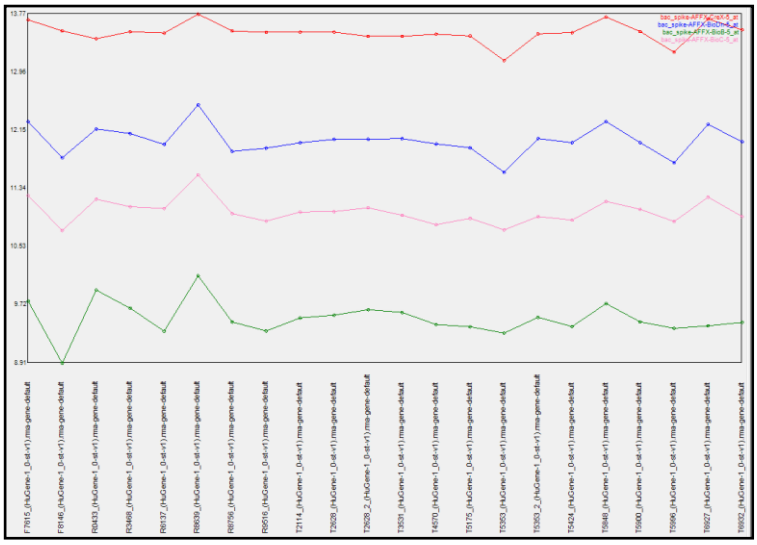

Figure 1: Hibridisation intensity of control probesets
The hibridisation intensity (varying from 8.91 to 13.77) for the spiked-in control probes are represented for each sample. bioB, bioC and bioD are genes in the biotin synthesis pathway of E. coli, and cre is the recombinase gene from P1 bacteriophage. The intensity pattern for these 4 controls should show the increase in target concentration (from bioB to CreX). Other patterns would be a sign of bad hybridization.

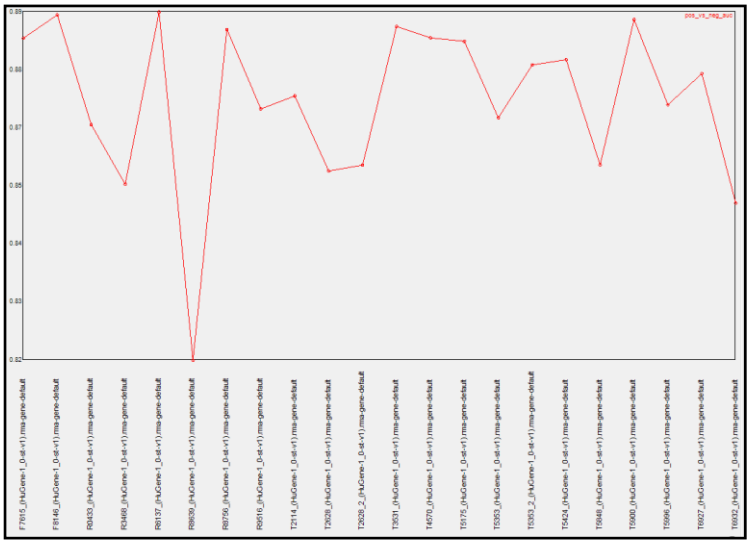

Figure 2: Ratio of hibridisation intensity between positive and negative probes
This metrics helps to assess the ability to distinguish true signal from noise. All arrays analysed had positive versus negative AUC (area under the curve) value of ≥ 0.8 and hence passed the quality criteria.

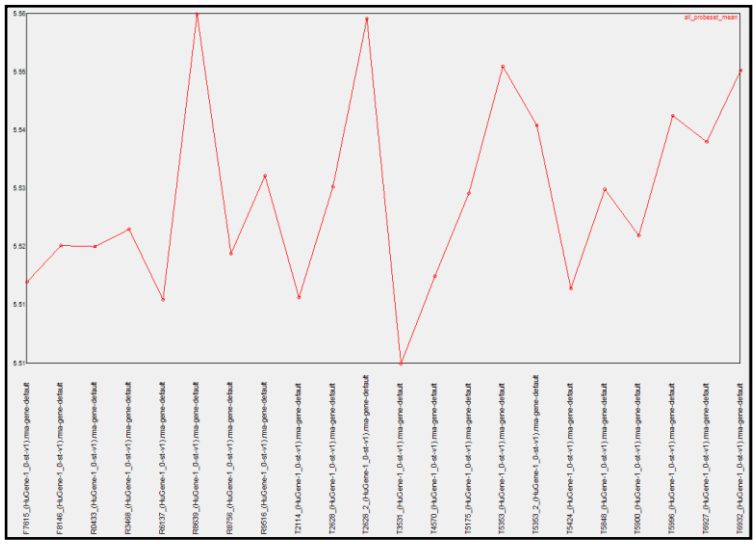

Figure 3: Mean hybridisation intensities including all probesets
The mean value of the hibridisation intensity of all probes for each sample is graphically represented to demonstrate the minimal variability (< 0.05) across samples.

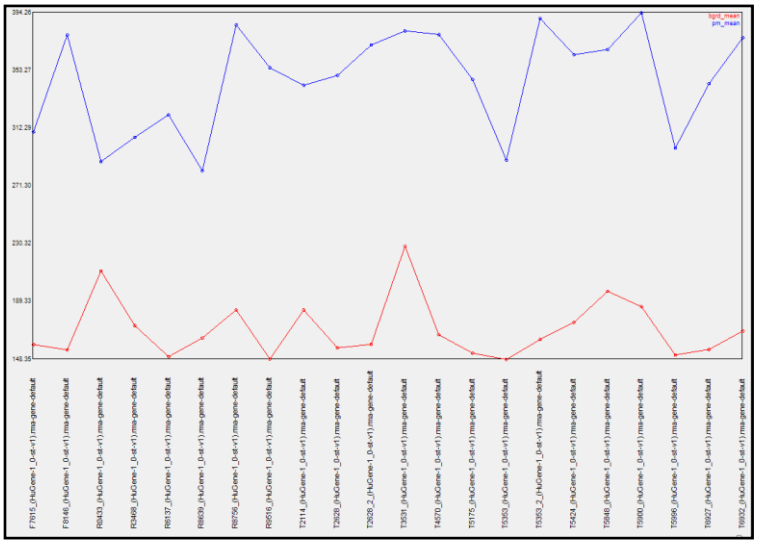

Figure 4: Mean hybridisation intensities for perfect-match and background probes
The number of expressed genes is estimated using the differences in intensity between the PM (perfect match) and the MM (mismatch) probes. The hibridisation intensity of the last is a surrogate of the background intensity of the array.

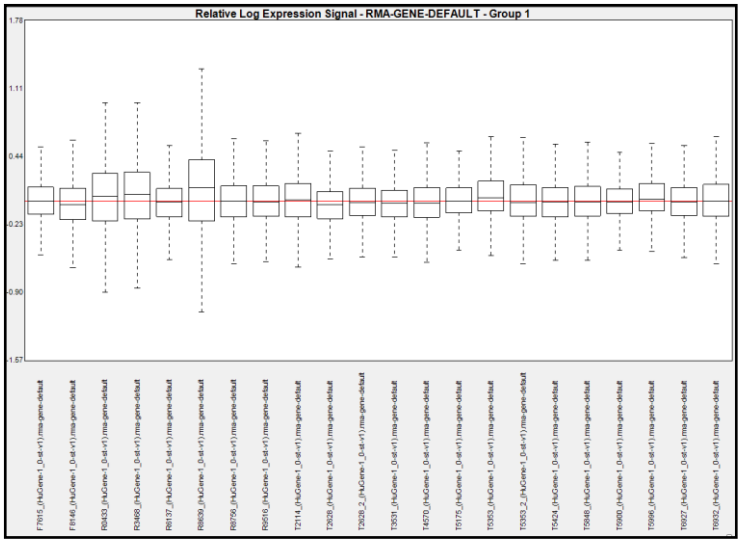

Figure 5: Relative Log Expression signal; mean and standard deviation represented
The Relative Log Expression (RLE) values for each probe-set correspond to the ratio between the expression of a probe-set and the median expression of this probe-set across all arrays of the experiment. It is assumed that most probe-sets are not changed across the arrays, so it is expected that these ratios are around 0 on a log scale. The boxplots presenting the distribution of these log-ratios should then be centered near 0 and have similar spread.

Table 1: Differentially expressed probesets between DLBCL TAM and reactive LN control macrophages

HGNC symbol	Probeset	Gb accession	Description	Log^2 Fold Change	Average Expression	Adjusted p-value
ACP2	7947815	NM_001610	acid phosphatase 2, lysosomal	2,334	6,881	0,014
ACSL1	8103951	NM_001995	acyl-CoA synthetase long-chain family member 1	1,956	7,704	0,027
ADAMDEC 1	8145317	NM_014479	ADAM-like, decysin 1	2,495	7,126	0,048
AFF3	8054254	NM_002285	AF4/FMR2 family, member 3	-1,215	7,170	0,023
AGTRAP	7897745	NM_020350	angiotensin II receptor-associated protein	1,195	7,762	0,004
ALOX15	8011680	NM_001140	arachidonate 15-lipoxygenase	-2,513	5,078	0,003
ANKRD22	7934898	NM_144590	ankyrin repeat domain 22	3,882	7,733	0,002
APCDD1	8020141	NM_153000	adenomatosis polyposis coli down-regulated 1	-1,046	6,500	0,008
APOBEC3A	8073056	NM_145699	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A	2,035	7,630	0,012
APOC1	8029536	NM_001645	apolipoprotein C-I	2,863	8,138	0,007
APP	8069644	NM_000484	amyloid beta (A4) precursor protein	-1,024	10,501	0,048
AQP9	7983910	NM_020980	aquaporin 9	3,351	6,814	0,003
ASAH2	7927599	NM_019893	N-acylsphingosine amidohydrolase (non-lysosomal ceramidase) 2	-1,075	4,266	0,027
BATF	7975793	NM_006399	basic leucine zipper transcription factor, ATF-like	1,178	7,054	0,041
BCL2A1	7990818	NM_001114735	BCL2-related protein A1	1,679	8,598	0,012
C14orf145	7980496	NM_152446	chromosome 14 open reading frame 145	-1,077	6,799	0,039
C15orf48	7983478	NM_032413	chromosome 15 open reading frame 48	2,552	6,356	0,033
C1QA	7898793	NM_015991	complement component 1, q subcomponent, A chain	2,492	6,529	0,003
C1QB	7898805	NM_000491	complement component 1, q subcomponent, B chain	2,203	8,667	0,009

C1QC	7898799	NM_001114101	complement component 1, q subcomponent, C chain	2,616	8,946	0,011
C3AR1	7960874	NM_004054	complement component 3a receptor 1	3,925	8,092	0,000
C5AR1	8029907	NM_001736	complement component 5a receptor 1	1,853	7,883	0,009
C6orf192	8129649	NM_052831	chromosome 6 open reading frame 192	-1,323	5,350	0,045
CA8	8150978	NM_004056	carbonic anhydrase VIII	-1,784	5,272	0,025
CASP5	7951385	NM_004347	caspase 5, apoptosis-related cysteine peptidase	2,071	5,486	0,039
CCL2	8006433	NM_002982	chemokine (C-C motif) ligand 2	2,478	7,539	0,027
CCL3	8014369	NM_002983	chemokine (C-C motif) ligand 3	1,353	11,035	0,050
CCL4	8006602	NM_002984	chemokine (C-C motif) ligand 4	1,875	9,678	0,024
CCRL2	8079407	NM_003965	chemokine (C-C motif) receptor-like 2	1,829	5,826	0,022
CCRL2	8093304	NM_003965	chemokine (C-C motif) receptor-like 2	1,645	5,699	0,039
CD14	8114612	NM_000591	CD14 molecule	2,781	10,282	0,000
CD151	7937508	NM_004357	CD151 molecule (Raph blood group)	1,160	5,542	0,045
CD163	7960794	NM_004244	CD163 molecule	2,843	7,676	0,011
CD1A	7906339	NM_001763	CD1a molecule	-1,345	4,813	0,049
CD1C	7906348	NM_001765	CD1c molecule	-1,918	8,858	0,024
CD1E	7906355	NM_030893	CD1e molecule	-2,314	7,937	0,039
CD207	8052916	NM_015717	CD207 molecule, langerin	-1,464	5,357	0,008
CD274	8154233	NM_014143	CD274 molecule	2,344	7,215	0,038
CD58	7918902	NM_001779	CD58 molecule	1,159	7,583	0,041
CD81	7937802	NM_004356	CD81 molecule	1,087	9,165	0,049
CDCA7L	8138489	NM_018719	cell division cycle associated 7-like	-1,344	7,529	0,049

CDH1	7996837	NM_004360	cadherin 1, type 1, E-cadherin (epithelial)	-2,589	7,756	0,034
CEBPB	8063386	NM_005194	CCAAT/enhancer binding protein (C/EBP), beta	1,004	6,393	0,012
CHML	7925500	NM_001821	choroideremia-like (Rab escort protein 2)	-1,088	6,799	0,049
CLEC4D	7953749	NM_080387	C-type lectin domain family 4, member D	1,886	5,733	0,025
CLEC4E	7960900	NM_014358	C-type lectin domain family 4, member E	2,809	6,340	0,008
CLEC6A	7953737	NM_001007033	C-type lectin domain family 6, member A	2,587	6,965	0,012
CLEC7A	7961120	NM_197947	C-type lectin domain family 7, member A	1,310	7,239	0,049
CREG1	7922051	NM_003851	cellular repressor of E1A-stimulated genes 1	1,103	8,021	0,017
CRYM	8000117	NM_001888	crystallin, mu	-1,047	6,327	0,012
CSF1R	8115076	NM_005211	colony stimulating factor 1 receptor	1,309	9,169	0,044
CSF3R	7914950	NM_156039	colony stimulating factor 3 receptor (granulocyte)	1,080	8,104	0,024
CTSD	7945666	NM_001909	cathepsin D	1,027	9,086	0,018
CTSL1	8156228	NM_001912	cathepsin L1	3,099	7,806	0,012
CXCL10	8101126	NM_001565	chemokine (C-X-C motif) ligand 10	3,277	10,407	0,002
CXCL9	8101118	NM_002416	chemokine (C-X-C motif) ligand 9	3,191	9,471	0,008
CXCR2	8048227	NM_001557	chemokine (C-X-C motif) receptor 2	-1,180	4,605	0,032
DAB2	8111772	NM_001343	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	-1,079	8,450	0,028
DENND5B	7962151	NM_144973	DENN/MADD domain containing 5B	-1,048	5,736	0,049
DMXL2	7988789	NM_001174116	Dmx-like 2	1,394	8,152	0,048
DOCK4	8142345	NM_014705	dedicator of cytokinesis 4	1,694	5,670	0,027
DRAM1	7958019	NM_018370	DNA-damage regulated autophagy modulator 1	2,036	6,641	0,011
EMR1	8025103	NM_001974	egf-like module containing, mucin-like, hormone receptor-like 1	1,956	5,578	0,024

ENG	8164269	NM_000118	endoglin	1,218	7,740	0,037
EPB41L3	8022118	NM_012307	erythrocyte membrane protein band 4.1-like 3	1,740	7,318	0,029
FAM160A1	8097801	NM_001109977	family with sequence similarity 160, member A1	-1,823	6,435	0,014
FBXO16	8150002	NM_172366	F-box protein 16	-1,064	4,163	0,012
FCAR	8031374	NM_002000	Fc fragment of IgA, receptor for	1,004	5,965	0,045
FCER1A	7906443	NM_002001	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide	-1,671	6,511	0,028
FCER2	8033420	NM_002002	Fc fragment of IgE, low affinity II, receptor for (CD23)	-2,420	5,829	0,002
FCGR1A	7905047	NM_000566	Fc fragment of IgG, high affinity Ia, receptor (CD64)	3,810	8,473	0,001
FCGR1A	7905060	NM_000566	Fc fragment of IgG, high affinity Ia, receptor (CD64)	3,794	8,801	0,001
FCGR1B	7919133	NM_001017986	Fc fragment of IgG, high affinity Ib, receptor (CD64)	3,780	8,830	0,000
FCGR2A	7906757	NM_001136219	Fc fragment of IgG, low affinity IIa, receptor (CD32)	2,313	9,679	0,008
FCGR3A	7921868	NM_000569	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	4,184	8,330	0,000
FCGR3A	7921873	NM_000569	Fc fragment of IgG, low affinity IIIa, receptor (CD16a)	3,672	8,706	0,001
FCRL1	7921319	NM_052938	Fc receptor-like 1	-1,571	6,400	0,024
FLT3	7970737	NM_004119	fms-related tyrosine kinase 3	-1,317	7,854	0,024
FNBP1L	7903092	NM_001024948	formin binding protein 1-like	-1,247	5,677	0,025
FPR1	8038899	NM_002029	formyl peptide receptor 1	2,412	9,463	0,006
FPR2	8030860	NM_001462	formyl peptide receptor 2	2,402	6,295	0,036
GAS7	8012605	NM_201433	growth arrest-specific 7	1,007	6,058	0,027
GBP1	7917516	NM_002053	guanylate binding protein 1, interferon-inducible, 67kDa	2,295	9,863	0,022
GBP2	7917532	NM_004120	guanylate binding protein 2, interferon-inducible	2,679	7,752	0,003
GBP4	7917561	NM_052941	guanylate binding protein 4	2,285	8,113	0,018

GBP5	7917576	NM_052942	guanylate binding protein 5	2,902	8,467	0,017
GCH1	7979269	NM_000161	GTP cyclohydrolase 1	1,445	8,073	0,019
GIMAP5	8137257	NM_018384	GTPase, IMAP family member 5	2,486	8,087	0,011
GNG10	7946559	NM_001017998	guanine nucleotide binding protein (G protein), gamma 10	1,014	10,540	0,039
GOLPH3L	7919780	NM_018178	golgi phosphoprotein 3-like	-1,122	6,923	0,039
GPNMB	8131844	NM_001005340	glycoprotein (transmembrane) nmb	2,532	6,225	0,024
GPR84	7963770	NM_020370	G protein-coupled receptor 84	2,550	7,451	0,003
HAPLN3	7991224	NM_178232	hyaluronan and proteoglycan link protein 3	1,425	6,794	0,017
HEATR5B	8051464	NM_019024	HEAT repeat containing 5B	-1,065	7,857	0,041
HHAT	7909510	NM_001170580	hedgehog acyltransferase	-1,245	5,940	0,024
HK3	8115957	NM_002115	hexokinase 3 (white cell)	1,247	5,866	0,031
HLA-DPB2	8118607	NR_001435	major histocompatibility complex, class II, DP beta 2 (pseudogene)	-1,005	6,332	0,013
HLA-DRB5	8125436	NM_002125	major histocompatibility complex, class II, DR beta 5	-2,316	7,834	0,048
HMOX1	8072678	NM_002133	heme oxygenase (decycling) 1	1,686	9,062	0,014
HNMT	8045499	NM_006895	histamine N-methyltransferase	1,516	6,333	0,023
HOMER2	7991034	NM_199330	homer homolog 2 (Drosophila)	-1,886	5,973	0,000
HSPA6	7906764	NM_002155	heat shock 70kDa protein 6 (HSP70B')	1,490	8,163	0,048
IDO1	8146092	NM_002164	indoleamine 2,3-dioxygenase 1	3,337	8,055	0,018
IER3	8124848	NM_003897	immediate early response 3	1,345	9,484	0,029
IER3	8179704	NM_003897	immediate early response 3	1,345	9,484	0,029
IER3	8178435	NM_003897	immediate early response 3	1,363	9,777	0,030
IFITM3	7945371	NM_021034	interferon induced transmembrane protein 3 (1-8U)	1,109	10,160	0,017

IFNGR2	8068280	NM_005534	interferon gamma receptor 2 (interferon gamma transducer 1)	1,168	9,628	0,012
IL15RA	7931899	NM_002189	interleukin 15 receptor, alpha	1,150	7,958	0,049
IL1R1	8043995	NM_000877	interleukin 1 receptor, type I	-1,469	5,217	0,036
IL4I1	8038487	NM_172374	interleukin 4 induced 1	1,253	7,750	0,010
IRF1	8114010	NM_002198	interferon regulatory factor 1	1,233	9,154	0,024
ITGAM	7995096	NM_001145808	integrin, alpha M (complement component 3 receptor 3 subunit)	1,328	7,632	0,049
KCNA5	7953278	NM_002234	potassium voltage-gated channel, shaker-related subfamily, member 5	-1,252	6,083	0,018
KCNJ10	7921533	NM_002241	potassium inwardly-rectifying channel, subfamily J, member 10	1,665	5,874	0,044
KCNK10	7980605	NM_021161	potassium channel, subfamily K, member 10	-1,009	4,494	0,029
KCTD5	7992685	NM_018992	potassium channel tetramerisation domain containing 5	-1,047	7,454	0,012
LDOC1	8175539	NM_012317	leucine zipper, down-regulated in cancer 1	-1,018	5,605	0,009
LEPREL1	8092707	NM_018192	leprecan-like 1	-1,389	6,795	0,040
LGALS3	7974461	NR_003225	lectin, galactoside-binding, soluble, 3	1,185	7,348	0,008
LILRA3	8039226	NM_006865	leukocyte immunoglobulin-like receptor, subfamily A (without TM domain), member 3	2,490	8,123	0,008
LILRA5	8039236	NM_021250	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 5	1,580	7,254	0,027
LILRB2	8039212	NM_005874	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2	1,543	8,983	0,008
LPAR1	8163257	NM_057159	lysophosphatidic acid receptor 1	1,030	5,414	0,039
LRP1	7956301	NM_002332	low density lipoprotein receptor-related protein 1	1,327	6,491	0,027
MAFB	8066266	NM_005461	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	2,056	8,373	0,008
MT1DP	7995813	NR_027781	metallothionein 1D (pseudogene)	1,076	7,464	0,033

MT1E	7995797	NM_175617	metallothionein 1E	2,336	7,332	0,011
MT1G	8001531	NM_005950	metallothionein 1G	2,595	7,651	0,034
MT1X	7995838	NM_005952	metallothionein 1X	1,480	8,731	0,028
MT2A	7995783	NM_005953	metallothionein 2A	2,092	11,737	0,001
MT2A	8095376	NM_005953	metallothionein 2A	2,058	10,375	0,003
MT2A	8095362	NM_005953	metallothionein 2A	2,170	11,164	0,004
NET1	7925954	NM_001047160	neuroepithelial cell transforming 1	-1,100	6,981	0,012
NKG7	8038809	NM_005601	natural killer cell group 7 sequence	1,569	9,396	0,029
NR1H3	7939751	NM_005693	nuclear receptor subfamily 1, group H, member 3	1,767	6,327	0,011
NRP1	7932985	NM_003873	neuropilin 1	-1,393	7,770	0,046
ODF3B	8077116	NM_001014440	outer dense fiber of sperm tails 3B	1,140	7,987	0,019
P2RX4	7959267	NM_002560	purinergic receptor P2X, ligand-gated ion channel, 4	1,132	7,298	0,048
P2RX7	7959251	NM_002562	purinergic receptor P2X, ligand-gated ion channel, 7	1,567	6,087	0,011
PAIP2B	8052940	NM_020459	poly(A) binding protein interacting protein 2B	-1,410	5,023	0,033
PDCD1LG2	8154245	NM_025239	programmed cell death 1 ligand 2	2,371	6,513	0,040
PHLPP2	8002571	NM_015020	PH domain and leucine rich repeat protein phosphatase 2	-1,061	5,303	0,024
PILRA	8134814	NM_013439	paired immunoglobin-like type 2 receptor alpha	1,525	9,426	0,012
PIM1	8119161	NM_002648	pim-1 oncogene	1,126	8,017	0,046
PLA2G7	8126784	NM_001168357	phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma)	2,161	8,046	0,005
PLD4	7977319	NM_138790	phospholipase D family, member 4	-1,650	8,756	0,050
PLXNA4	8142997	NM_020911	plexin A4	-1,691	6,803	0,002
PLXND1	8090591	NM_015103	plexin D1	1,215	7,168	0,025

PON2	8141076	NM_000305	paraoxonase 2	-1,174	5,805	0,010
PPA1	7934133	NM_021129	pyrophosphatase (inorganic) 1	1,435	9,731	0,006
PSTPIP2	8023043	NM_024430	proline-serine-threonine phosphatase interacting protein 2	2,138	8,337	0,011
PTGER2	7974366	NM_000956	prostaglandin E receptor 2 (subtype EP2), 53kDa	1,717	5,792	0,027
RAB20	7972805	NM_017817	RAB20, member RAS oncogene family	1,204	6,712	0,029
RALGPS2	7907657	NM_152663	Ral GEF with PH domain and SH3 binding motif 2	-1,382	6,865	0,008
RARRES3	7940775	NM_004585	retinoic acid receptor responder (tazarotene induced) 3	1,892	8,028	0,039
RASSF4	7927186	NM_032023	Ras association (RalGDS/AF-6) domain family member 4	1,325	7,808	0,017
RBMS3	8078330	NM_001003793	RNA binding motif, single stranded interacting protein 3	-1,503	5,134	0,008
RGL1	7908125	NM_015149	ral guanine nucleotide dissociation stimulator-like 1	1,635	8,193	0,013
RGS7	7925457	NM_002924	regulator of G-protein signaling 7	-1,226	6,730	0,014
RNF125	8020806	NM_017831	ring finger protein 125	-1,218	6,090	0,050
S100A12	7920238	NM_005621	S100 calcium binding protein A12	2,292	8,013	0,039
S100A8	7920244	NM_002964	S100 calcium binding protein A8	1,921	11,019	0,008
S100A9	7905571	NM_002965	S100 calcium binding protein A9	2,113	10,269	0,004
SCARB1	7967544	NM_005505	scavenger receptor class B, member 1	-1,064	7,289	0,029
SCO2	8077099	NM_005138	SCO cytochrome oxidase deficient homolog 2 (yeast)	1,026	8,108	0,041
SCRN1	8138824	NM_014766	secernin 1	-1,076	8,928	0,049
SECTM1	8019486	NM_003004	secreted and transmembrane 1	1,375	8,783	0,049
SERPINA1	7981068	NM_001002236	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	1,678	10,490	0,008
SERPINB9	8123609	NM_004155	serpin peptidase inhibitor, clade B (ovalbumin), member 9	1,016	9,627	0,049
SIDT1	8081710	NM_017699	SID1 transmembrane family, member 1	-1,337	6,135	0,040

SIGLEC14	8038885	NM_001098612	sialic acid binding Ig-like lectin 14	2,056	7,581	0,017
SIRPB2	8064451	NM_001122962	signal-regulatory protein beta 2	1,214	5,938	0,037
SLC11A1	8048283	NM_000578	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	2,145	6,406	0,006
SLC1A3	8104930	NM_004172	solute carrier family 1 (glial high affinity glutamate transporter), member 3	2,417	6,048	0,028
SLC25A37	8145281	NM_016612	solute carrier family 25, member 37	1,245	7,226	0,025
SLC37A2	7944931	NM_198277	solute carrier family 37 (glycerol-3-phosphate transporter), member 2	1,226	6,455	0,030
SLC38A1	7962516	NM_030674	solute carrier family 38, member 1	-1,160	9,610	0,022
SLC39A10	8047174	NM_001127257	solute carrier family 39 (zinc transporter), member 10	-1,011	6,640	0,046
SLC41A2	7965964	NM_032148	solute carrier family 41, member 2	-1,575	6,389	0,044
SLC47A1	8005603	NM_018242	solute carrier family 47, member 1	-1,206	4,810	0,011
SLC7A7	7977786	NM_003982	solute carrier family 7 (cationic amino acid transporter, y+ system), member 7	1,744	8,381	0,049
SOD2	8130556	NM_001024465	superoxide dismutase 2, mitochondrial	1,972	10,526	0,001
SPTBN1	8041995	NM_003128	spectrin, beta, non-erythrocytic 1	-1,152	6,773	0,040
STAB1	8080344	NM_015136	stabilin 1	1,209	6,413	0,011
STEAP4	8140840	NM_024636	STEAP family member 4	2,319	6,190	0,039
SYCP2L	8116874	NM_001040274	synaptonemal complex protein 2-like	-1,626	4,402	0,008
TASP1	8065018	NM_017714	taspase, threonine aspartase, 1	-1,055	6,334	0,009
TBC1D2B	7985224	NM_144572	TBC1 domain family, member 2B	1,195	7,121	0,048
TBC1D4	7972021	NM_014832	TBC1 domain family, member 4	-1,211	7,739	0,044
TCF7L2	7930537	NM_001146274	transcription factor 7-like 2 (T-cell specific, HMG-box)	1,638	7,738	0,044
TCN2	8072360	NM_000355	transcobalamin II	1,373	6,211	0,044

TIMP1	8167185	NM_003254	TIMP metallopeptidase inhibitor 1	1,730	10,423	0,033
TLR10	8099826	NM_030956	toll-like receptor 10	-1,026	7,802	0,044
TLR2	8097903	NM_003264	toll-like receptor 2	1,619	7,192	0,024
TLR4	8157524	NR_024168	toll-like receptor 4	2,322	7,912	0,002
TMEM176A	8137264	NM_018487	transmembrane protein 176A	2,514	8,524	0,008
TNFAIP2	7977046	NM_006291	tumor necrosis factor, alpha-induced protein 2	1,284	7,874	0,029
TNFSF10	8092169	NM_003810	tumor necrosis factor (ligand) superfamily, member 10	1,662	8,924	0,039
TREM1	8126303	NM_018643	triggering receptor expressed on myeloid cells 1	1,597	7,886	0,031
TYMP	8077103	NM_001113756	thymidine phosphorylase	1,395	8,677	0,011
UBD	8124650	NM_006398	ubiquitin D	1,694	7,030	0,032
UBD	8178295	NM_006398	ubiquitin D	1,570	6,943	0,039
UPP1	8132725	NM_003364	uridine phosphorylase 1	1,429	6,242	0,018
VCAN	8106743	NM_004385	versican	2,091	8,315	0,021
VSIG6	7981732	ENST00000338567	V-set and immunoglobulin domain containing 6	-2,017	7,491	0,032
ZNF585B	8036389	NM_152279	zinc finger protein 585B	-1,049	5,589	0,036
ZNF660	8079198	NM_173658	zinc finger protein 660	-1,024	3,907	0,011

Table 2. Toppfun Analysis

GENE ONTOLOGY – CELLULAR FUNCTIONS				
#	**Name**	**FDR**	**# genes included/ annotated**	**Genes**
1	immune response	$5.3X10^{-32}$	77/1416	CEBPB, IDO1, TNFSF10, CXCL10, GIMAP5, LILRB2, HLADRB5, CLEC7A, GBP1, GBP2, IRF1, GCH1, BCL2A1, HMOX1, ITGAM, APOBEC3A, S100A8, S100A9, S100A12, CLEC4E, SERPINB9, FCAR, FCER1A, RNF125, FCER2, GBP4, FCGR1A, FCGR1B, ADAMDEC1, GBP5, FCGR2A, FCGR3A, IFITM3, CXCL9, C1QA, C1QB, C1QC, CCL2, CCL3, C3AR1, CCL4, C5AR1, LILRA5, ALOX15, SECTM1, CLEC6A, UBD, BATF, SIGLEC14, NR1H3, APP, AQP9, LGALS3, TLR10, IFNGR2, CD1A, CD1C, CD1E, MT2A, CSF1R, SLC11A1, CD14, P2RX7, CLEC4D, SCARB1, TLR2, TLR4, TREM1, CD58, CD274, CD81, IL1R1, CDH1, CTSL, CXCR2, PHLPP2, PDCD1LG2
2	defence response	$1.1\ X10^{-27}$	74/1515	CEBPB, IDO1, CXCL10, GIMAP5, LILRB2, HLADRB5, CLEC7A, GBP1, GBP2, IRF1, GCH1, HMOX1, ITGAM, CD163, APOBEC3A, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, FCER1A, RNF125, GBP4, FCGR1A, FCGR1B, GBP5, FCGR2A, FCGR3A, IER3, IFITM3, KCNJ10, CXCL9, C1QA, C1QB, C1QC, CCL2, CCL3, C3AR1, CCL4, C5AR1, LILRA5, ALOX15, PLA2G7, LILRA3, CLEC6A, UBD, BATF, FPR2, CD207, SIGLEC14, CASP5, CCRL2, NR1H3, APP, LGALS3, TLR10, IFNGR2, MT2A, CSF1R, SLC11A1, CD14, CSF3R, P2RX7, CLEC4D, SCARB1, TLR2, TLR4, TREM1, CD58, IL1R1, CTSL, CXCR2, PHLPP2
3	innate immune response	$2.5\ X10^{-18}$	48/883	GIMAP5, HLA-DRB5, CLEC7A, GBP1, GBP2, IRF1, GCH1, ITGAM, APOBEC3A, S100A8, S100A9, S100A12, FCER1A, RNF125, GBP4, FCGR1A, FCGR1B, GBP5, FCGR2A, FCGR3A, IFITM3, C1QA, C1QB, C1QC, CCL2, CCL3, LILRA5, CLEC6A, UBD, SIGLEC14, NR1H3, APP, LGALS3, TLR10, IFNGR2, MT2A, CSF1R, SLC11A1, CD14, P2RX7, CLEC4D, SCARB1, TLR2, TLR4, TREM1, CD58, CTSL, PHLPP2
4	inflammatory response	$4.9\ X10^{-18}$	40/599	CEBPB, IDO1, CXCL10, HLA-DRB5, CLEC7A, HMOX1, ITGAM, CD163, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, FCER1A, FCGR1A, FCGR2A, IER3, KCNJ10, CXCL9, CCL2, CCL3, C3AR1, CCL4, C5AR1, ALOX15, PLA2G7, FPR2, CASP5, CCRL2, NR1H3, TLR10, CSF1R, SLC11A1, CD14, P2RX7, TLR2, TLR4, IL1R1, CXCR2
5	response to biotic stimulus	$3.8\ X10^{-16}$	42/760	CEBPB, IDO1, CXCL10, LILRB2, HLA-DRB5, CLEC7A, GBP1, GBP2,IRF1, GCH1, PTGER2, HNMT, APOBEC3A, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, GBP4, FCGR1A, FCGR3A, IER3, IFITM3, CXCL9, CCL2, CCL3, CCL4, C5AR1, CLEC6A, BATF, CD207, NR1H3, IFNGR2, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, TREM1, SOD2
6	response to other organism	$3.8\ X10^{-18}$	41/726	CEBPB, IDO1, CXCL10, LILRB2, HLA-DRB5, CLEC7A, GBP1, GBP2, IRF1, GCH1, PTGER2, APOBEC3A, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, GBP4, FCGR1A, FCGR3A, IER3, IFITM3, CXCL9, CCL2, CCL3, CCL4, C5AR1, CLEC6A, BATF, CD207, NR1H3, IFNGR2, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, TREM1, SOD2
7	response to external biotic stimulus	$3.8\ X10^{-16}$	41/726	CEBPB, IDO1, CXCL10, LILRB2, HLA-DRB5, CLEC7A, GBP1, GBP2, IRF1, GCH1, PTGER2, APOBEC3A, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, GBP4, FCGR1A, FCGR3A, IER3, IFITM3, CXCL9, CCL2, CCL3, CCL4, C5AR1, CLEC6A, BATF, CD207, NR1H3, IFNGR2, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, TREM1, SOD2

8	response to wounding	1.2 X10^{-14}	51/1255	CEBPB, IDO1, CXCL10, HLA-DRB5, CLEC7A, IRF1, HMOX1, ITGAM, CD163, NRP1, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, FCER1A, FCGR1A, FCGR2A, IER3, KCNJ10, CXCL9, CCL2, CCL3, C3AR1, CCL4, C5AR1, ALOX15, PLA2G7, FPR2, CASP5, CCRL2, NR1H3, APP, SLC7A7, SLC1A3, TLR10, CSF1R, SLC11A1, CD14, P2RX7, TIMP1, SCARB1, TLR2, TLR4, TREM1, CD58, IL1R1, ENG, SOD2, CXCR2
9	positive regulation of immune system process	1.7 X10^{-14}	39/732	IL15RA, IDO1, CXCL10, GIMAP5, LILRB2, HLA-DRB5, CLEC7A, IRF1, BCL2A1, HMOX1, ITGAM, FCER1A, FCER2, FCGR1A, FCGR2A, FCGR3A, C1QA, C1QB, C1QC, CCL2, CCL3, C3AR1, CCL4, C5AR1, PLA2G7, NR1H3, LGALS3, TLR10, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, CD274, CD81, CTSL, CXCR2, PDCD1LG2
10	regulation of immune system process	2.7 X10^{-13}	48/1212	IL15RA, IDO1, CXCL10, GIMAP5, LILRB2, HLA-DRB5, CLEC7A, TMEM176A, IRF1, BCL2A1, HMOX1, ITGAM, FCER1A, FCER2, FCGR1A, FCGR1B, FCGR2A, FCGR3A, C1QA, C1QB, C1QC, CCL2, CCL3, C3AR1, CCL4, MAFB, C5AR1, ALOX15, PLA2G7, FLT3, NR1H3, LGALS3, TLR10, IFNGR2, CSF1R, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, CD274, CD81, CDH1, CTSL, CXCR2, PHLPP2, PDCD1LG2
11	response to IFN-γ	2.7 X10^{-13}	18/123	HLA-DRB5, GBP1, GBP2, IRF1, GCH1, GBP4, FCGR1A, FCGR1B, GBP5, IFITM3, CCL2, CCL3, UBD, NR1H3, IFNGR2, MT2A, SLC11A1, CD58
12	response to cytokine	1.1 X10^{-12}	34/629	IL15RA, CXCL10, HLA-DRB5, GBP1, GBP2, IRF1, GCH1, HNMT, ACSL1, SERPINA1, GBP4, FCGR1A, FCGR1B, GBP5, IFITM3, CCL2, CCL3, ALOX15, FLT3, UBD, AFF3, CCRL2, NR1H3, IFNGR2, MT1X, MT2A, CSF1R, SLC11A1, CD14, CSF3R, TIMP1, CD58, IL1R1, CXCR2
13	response to bacterium	1.7 X10^{-12}	29/451	CEBPB, IDO1, CXCL10, LILRB2, HLA-DRB5, GBP2, GCH1, PTGER2, STAB1, S100A8, S100A9, S100A12, SERPINA1, SERPINB9, GBP4, FCGR1A, FCGR3A, CCL2, CCL3, C5AR1, NR1H3, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, TREM1, SOD2
14	cell chemotaxis	2.0 X10^{-12}	21/210	DOCK4, CXCL10, ITGAM, NRP1, S100A8, S100A9, S100A12, FCGR2A, CXCL9, CCL2, CCL3, C3AR1, CCL4, C5AR1, PLA2G7, LPAR1, LGALS3, CSF3R, TREM1, ENG, CXCR2
15	positive regulation of response to stimulus	4.6 X10^{-12}	53/1584	IDO1, TNFSF10, CXCL10, GIMAP5, HLA-DRB5, DAB2, CLEC7A, IRF1, BCL2A1, HMOX1, ITGAM, NRP1, S100A8, S100A9, S100A12, FCER1A, FCER2, FCGR1A, FCGR2A, FCGR3A, CXCL9, C1QA, C1QB, C1QC, CCL2, CCL3, C3AR1, CCL4, C5AR1, ALOX15, SECTM1, PLA2G7, FLT3, TCF7L2, CLEC6A, UBD, FPR1, NR1H3, LPAR1, LGALS3, TLR10, CSF1R, SLC11A1, CD14, P2RX4, P2RX7, SCARB1, TLR2, TLR4, CD81, ENG, CTSL, CXCR2
16	response to lipopolysaccharide	5.8 X10^{-12}	22/250	CEBPB, IDO1, CXCL10, LILRB2, GBP2, GCH1, PTGER2, S100A8, S100A9, SERPINA1, FCGR3A, CCL2, CCL3, C5AR1, NR1H3, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, SOD2
17	response to molecule of bacterial origin	2.4 X10^{-11}	22/269	CEBPB, IDO1, CXCL10, LILRB2, GBP2, GCH1, PTGER2, S100A8, S100A9, SERPINA1, FCGR3A, CCL2, CCL3, C5AR1, NR1H3, SLC11A1, CD14, P2RX7, SCARB1, TLR2, TLR4, SOD2
18	cytokine production	3.1 X10^{-11}	29/451	CEBPB, IDO1, GIMAP5, LILRB2, HLA-DRB5, CLEC7A, IRF1, HMOX1, S100A8, S100A9, S100A12, CLEC4E, FCER1A, RNF125, FCGR2A, CCL2, CCL3, C3AR1, CCL4, CLEC6A, BATF, TLR10, CSF1R, SLC11A1, CD14, P2RX7, TLR2, TLR4, TREM1, CD58, CD274
21	leukocyte activation	6.8 X10^{-11}	33/695	IL15RA, IDO1, GIMAP5, LILRB2, HLA-DRB5, CLEC7A, IRF1, BCL2A1, HMOX1, ITGAM, S100A12, FCER1A, FCGR2A, FCGR3A, CCL2, CCL3, MAFB, FLT3, UBD, BATF, NR1H3, LGALS3, CD1C, MT1G, SLC11A1, P2RX7, TLR2, TLR4, CD274, CD81, CD151, CXCR2, PDCD1LG2

27	immune effector process	6.0×10^{-10}	30/628	CXCL10, GIMAP5, HLA-DRB5, CLEC7A, GBP1, IRF1, HMOX1, APOBEC3A, SERPINB9, FCER1A, FCER2, FCGR1A, FCGR2A, FCGR3A, IFITM3, CXCL9, C1QA, C1QB, C1QC, CCL2, CCL3, BATF, CD207, LGALS3, CD1C, SLC11A1, P2RX7, TLR2, TLR4, TREM1
GENE ONTOLOGY – MOLECULAR FUNCTIONS				
1	immunoglobulin binding	3.9×10^{-8}	8/23	FCAR, FCER1A, FCER2, FCGR1A, FCGR1B, FCGR2A, FCGR3A, LGALS3
2	receptor activity	1.3×10^{-5}	42/1617	IL15RA, LILRB2, DAB2, CLEC7A, GPR84, PLXND1, PTGER2, CD163, NRP1, STAB1, CLEC4E, FCER1A, FCGR1A, FCGR1B, FCGR2A, C3AR1, C5AR1, AGTRAP, FLT3, LILRA3, FPR1, FPR2, CCRL2, NR1H3, LPAR1, TLR10, IFNGR2, CSF1R, CD14, CSF3R, P2RX4, P2RX7, SCARB1, TLR2, TLR4, PLXNA4, TREM1, LRP1, EMR1, IL1R1, ENG, CXCR2
3	RAGE receptor binding	2.0×10^{-5}	5/12	S100A8, S100A9, S100A12, FPR1, FPR2
4	signaling pattern recognition receptor activity	6.7×10^{-5}	5/16	CLEC7A, CD14, SCARB1, TLR2, TLR4
5	pattern recognition receptor activity	6.7×10^{-5}	5/16	CLEC7A, CD14, SCARB1, TLR2, TLR4
6	receptor binding	7.2×10^{-5}	36/1405	DOCK4, CEBPB, TNFSF10, CXCL10, LILRB2, DAB2, CLEC7A, S100A8, S100A9, S100A12, KCNA5, FCER2, KCNJ10, CXCL9, C1QC, CCL2, CCL3, CCL4, GPNMB, HOMER2, SECTM1, FLT3, TCF7L2, FPR1, FPR2, CCRL2, NR1H3, APP, TYMP, P2RX4, P2RX7, TIMP1, LRP1, CD58, IL1R1, ENG
7	immunoglobulin receptor activity	7.8×10^{-5}	4/8	FCER1A, FCGR1A, FCGR1B, FCGR2A
8	signalling receptor activity	1.1×10^{-4}	35/1390	IL15RA, LILRB2, CLEC7A, GPR84, PLXND1, PTGER2, NRP1, FCER1A, FCGR1A, FCGR1B, FCGR2A, C3AR1, C5AR1, AGTRAP, FLT3, FPR1, FPR2, CCRL2, NR1H3, LPAR1, TLR10, IFNGR2, CSF1R, CD14, CSF3R, P2RX4, P2RX7, SCARB1, TLR2, TLR4, PLXNA4, EMR1, IL1R1, ENG, CXCR2
9	signal transducer activity	1.2×10^{-4}	40/1730	IL15RA, LILRB2, CLEC7A, GPR84, PLXND1, HMOX1, PTGER2, NRP1, S100A9, FCER1A, FCGR1A, FCGR1B, FCGR2A, C3AR1, C5AR1, GNG10, AGTRAP, SECTM1, FLT3, FPR1, FPR2, CCRL2, NR1H3, LPAR1, RGS7, TLR10, IFNGR2, CSF1R, CD14, CSF3R, P2RX4, P2RX7, SCARB1, TLR2, TLR4, PLXNA4, EMR1, IL1R1, ENG, CXCR2
10	molecular transducer activity	1.2×10^{-4}	40/1730	IL15RA, LILRB2, CLEC7A, GPR84, PLXND1, HMOX1, PTGER2, NRP1, S100A9, FCER1A, FCGR1A, FCGR1B, FCGR2A, C3AR1, C5AR1, GNG10, AGTRAP, SECTM1, FLT3, FPR1, FPR2, CCRL2, NR1H3, LPAR1, RGS7, TLR10, IFNGR2, CSF1R, CD14, CSF3R, P2RX4, P2RX7, SCARB1, TLR2, TLR4, PLXNA4, EMR1, IL1R1, ENG, CXCR2
11	IgG binding	3.1×10^{-4}	4/12	FCGR1A, FCGR1B, FCGR2A, FCGR3A

12	transmembrane signaling receptor activity	$3.1 X10^{-4}$	32/1228	IL15RA, LILRB2, GPR84, PLXND1, PTGER2, NRP1, FCER1A, FCGR1A, FCGR1B, FCGR2A, C3AR1, C5AR1, AGTRAP, FLT3, FPR1, FPR2, CCRL2, LPAR1, TLR10, IFNGR2, CSF1R, CD14, CSF3R, P2RX4, P2RX7, TLR2, TLR4, PLXNA4, EMR1, IL1R1, ENG, CXCR2
13	glycosaminoglycan binding	$5.0 X10^{-4}$	11/202	CXCL10, ITGAM, NRP1, STAB1, CCL2, GPNMB, APP, VCAN, TLR2, HAPLN3, ENG
14	IgE binding	$5.0 X10^{-4}$	3/5	FCER1A, FCER2, LGALS3
15	lipopolysaccharide receptor activity	$5.0 X10^{-4}$	3/5	SCARB1, TLR2, TLR4
PATHWAYS				
1	Staphylococcus aureus infection	$5.5 X10^{-10}$	13/59	HLA-DRB5, ITGAM, FCAR, FCGR1A, FCGR2A, FCGR3A, C1QA, C1QB, C1QC, C3AR1, C5AR1, FPR1, FPR2
2	Hematopoietic cell lineage	$7.9 X10^{-7}$	12/88	HLA-DRB5, ITGAM, FCER2, FCGR1A, FLT3, CD1A, CD1C, CD1E, CSF1R, CD14, CSF3R, IL1R1
3	Interferon gamma signaling	$8.9 X10^{-4}$	10/71	HLA-DRB5, GBP1, GBP2, IRF1, GBP4, FCGR1A, FCGR1B, GBP5, IFNGR2, MT2A
4	Tuberculosis	$2.0 X10^{-4}$	13/182	CEBPB, HLA-DRB5, CLEC7A, ITGAM, CLEC4E, FCGR1A, FCGR2A, FCGR3A, IFNGR2, CD14, TLR2, TLR4, CTSD
5	Phagosome	$2.3 X10^{-4}$	12/158	HLA-DRB5, CLEC7A, ITGAM, FCAR, FCGR1A, FCGR2A, FCGR3A, CD14, SCARB1, TLR2, TLR4, CTSL
6	Leishmaniasis	$9.3 X10^{-4}$	8/76	HLA-DRB5, ITGAM, FCGR1A, FCGR2A, FCGR3A, IFNGR2, TLR2, TLR4
7	Peptide ligand-binding receptors	$9.3 X10^{-4}$	12/189	CXCL10, CXCL9, CCL2, CCL3, C3AR1, CCL4, C5AR1, FPR1, FPR2, CCRL2, APP, CXCR2
8	Chemokine receptors bind chemokines	$9.3 X10^{-4}$	7/56	CXCL10, CXCL9, CCL2, CCL3, CCL4, CCRL2, CXCR2
9	Interferon Signaling	$2.0 X10^{-3}$	11/174	HLA-DRB5, GBP1, GBP2, IRF1, GBP4, FCGR1A, FCGR1B, GBP5, IFITM3, IFNGR2, MT2A
10	pyrimidine deoxyribonucleosides degradation	$3.6 X10^{-3}$	3/6	UPP1, SCO2, TYMP

CO-EXPRESSION				
1	THRLBCL, Van Loo et al.	4.5×10^{-60}	63/373	PDCD1LG2, CREG1, DMXL2, ACSL1, ANKRD22, DOCK4, TLR2, TLR4, ACP2, FCGR1A, CD274, KCNJ10, TNFSF10, CEBPB, SOD2, TNFAIP2, CD163, PILRA, AGTRAP, MT1X, APOBEC3A, MT1G, HNMT, MT1E, P2RX4, HK3, GPR84, SLC1A3, CXCL10, IDO1, C1QB, LILRB2, C1QA, GCH1, C1QC, C3AR1, IL15RA, GBP1, RASSF4, GBP5, SLC7A7, LRP1, PSTPIP2, PTGER2, S100A9, CCRL2, S100A8, SECTM1, IFNGR2, CASP5, RAB20, NR1H3, EPB41L3, RGL1, CD14, LGALS3, CTSD, FPR1, FPR2, CTSL, CSF1R, SERPINA1, CSF3R
2	cHL EBV +/-, Chetaille et al.	4.9×10^{-38}	46/348	CREG1, GPNMB, DMXL2, DOCK4, TLR2, ACP2, FCGR1A, TNFSF10, RNF125, CD163, HMOX1, TCF7L2, PILRA, TCN2, MT1G, HK3, PLA2G7, SLC1A3, CXCL10, IDO1, C1QB, LILRB2, C1QA, GCH1, C3AR1, MAFB, IL15RA, GBP1, RASSF4, SLC7A7, SLC47A1, GAS7, PSTPIP2, PTGER2, SCO2, CCRL2, SECTM1, APOC1, CASP5, RAB20, CXCL9, RGL1, CD14, CTSD, CTSL, SERPINA1
3	NPM mutated AML, Verhaak et al.	1.6×10^{-36}	48/427	FCER1A, DMXL2, ACSL1, TLR4, LILRA5, TNFSF10, SOD2, TNFAIP2, SCRN1, BCL2A1, CD163, TMEM176A, HMOX1, PILRA, APOBEC3A, HNMT, HK3, IER3, CXCL10, C1QB, LILRB2, C1QA, GCH1, C3AR1, MAFB, C5AR1, SIDT1, SLC38A1, TYMP, SCO2, SECTM1, APP, PLXND1, TREM1, CCL4, EPB41L3, AQP9, CD14, SPTBN1, EMR1, CTSD, CD1C, FPR1, CTSL, VCAN, SERPINA1, CSF3R, LILRA3
5	Lymph-node signature, Rosenwald et al.	3.4×10^{-22}	25/155	IL15RA, CEBPB, CXCL10, DAB2, GBP1, IRF1, CREG1, FCGR1A, IFITM3, CXCL9, CCL2, C3AR1, NKG7, FPR1, FPR2, MT2A, CSF1R, CD14, CSF3R, TIMP1, CD81, IL1R1, ENG, CTSL, SOD2
6	Low-dose radiotherapy in follicular lymphoma, Knoops et.al.	4.9×10^{-22}	23/121	TNFSF10, CXCL10, LILRB2, GBP1, HMOX1, EPB41L3, CREG1, ANKRD22, FCGR1A, ADAMDEC1, GBP5, CXCL9, C1QA, C1QB, GPNMB, PLA2G7, TCN2, UBD, NR1H3, APOC1, RAB20, TLR4, CTSD
7	CD molecules, Zola et al.	3.1×10^{-21}	32/345	TNFSF10, LILRB2, SIRPB2, ITGAM, CD163, NRP1, FCAR, FCER2, FCGR1A, FCGR2A, FCGR3A, FLT3, LILRA3, CD207, TLR10, CD1A, CD1C, CD1E, CSF1R, CD14, CSF3R, TLR2, TLR4, LRP1, CD58, CD274, CD81, CD151, IL1R1, ENG, CDH1, PDCD1LG2
CANCER MODULES				
1	Genes in module 84	1.7×10^{-27}	53/531	CREG1, GPNMB, ACSL1, DAB2, TLR2, HSPA6, FCGR3A, FCGR2A, TNFSF10, TNFAIP2, KCNA5, BCL2A1, CD163, RARRES3, ITGAM, IRF1, HK3, PLA2G7, IER3, C1QB, LILRB2, GCH1, C1QC, C3AR1, MAFB, HLA-DRB5, GBP1, GBP2, SLC7A7, IL1R1, TBC1D2B , S100A9, S100A8, SECTM1, APOC1, IGLJ3, CCL3, IGHM, CCL4, EPB41L3, CCL2, IGKC, AQP9, CXCL9, RGL1, ENG, CD14, LGALS3, FPR1, CSF1R, NKG7, CDH1, VCAN, UBD, SERPINA1, CSF3R
2	Genes in module 5	3.2×10^{-18}	39/423	GPNMB, ACSL1, DAB2, TLR2, ACP2, PON2, FCGR3A, FCGR2A, TNFSF10, TNFAIP2, CD163, RARRES3, HMOX1, ITGAM, IRF1, IER3, C1QB, GBP1, SLC7A7, UPP1, IL1R1, TBC1D2B, S100A9, S100A8, SECTM1, APOC1, PLXND1, STAB1, CCL3, IGHM, EPB41L3, CCL2, ENG, CD14, PIM1, NKG7, VCAN, CD151, SERPINA1, CSF3R
3	Genes in module 46	6.6×10^{-18}	37/386	CREG1, FCER1A, NRP1, TIMP1, DAB2, TLR2, FCGR3A, FCGR2A, TNFSF10, CEBPB, IFITM3, IRF1, CXCL10, LILRB2, C1QA, C3AR1, IL15RA, HLA-DRB5, C5AR1, GBP1, GBP2, LRP1, IL1R1, IFNGR2, CCL3, CCL4, CCL2, AQP9, CXCL9, FLT3, CD14, CD1C, BATF, UBD, CSF3R, CD81, LILRA3

4	Genes in module 75	4.4 $X10^{-15}$	34/391	CREG1, FCER1A, NRP1, TIMP1, DAB2, TLR2, FCGR3A, FCGR2A, TNFSF10, IFITM3, IRF1, CXCL10, LILRB2, C1QA, C3AR1, IL15RA, HLA-DRB5, C5AR1, GBP1, GBP2, LRP1, IL1R1, CCL3, CCL4, CCL2, AQP9, CXCL9, FLT3, CD14, CD1C, BATF, UBD, CSF3R, CD81
5	Genes in module 170	4.5 $X10^{-13}$	18/100	TIMP1, DAB2, FCGR1A, FCGR2A, TNFSF10, IFITM3, IER3, C3AR1, GBP1, GBP2, IL1R1, CCL2, CXCL9, CD14, FPR1, CTSL, CSF1R, CD151
6	Genes in module 64	4.3 $X10^{-12}$	34/506	FCER1A, NRP1, TLR2, FCGR1A, FCER2, FCGR3A, FCGR2A, TNFSF10, CD163, ITGAM, P2RX7, HK3, CXCL10, ADAMDEC1, LILRB2, C3AR1, IL15RA, HLA-DRB5, C5AR1, IL1R1, S100A9, CCL3, CCL4, CCL2, AQP9, CXCL9, FLT3, CD14, LGALS3, EMR1, CD1C, FPR1, NKG7, CSF3R
7	Genes in module 79	4.3 $X10^{-12}$	17/100	TIMP1, DAB2, FCGR1A, FCGR2A, IFITM3, IER3, C3AR1, GBP1, GBP2, IL1R1, CCL2, CXCL9, CD14, FPR1, CTSL, CSF1R, CD151
8	Genes in module 45	3.8 $X10^{-12}$	36/563	ACSL1, TLR2, SLC11A1, HSPA6, FCGR3A, FCGR2A, TNFSF10, SOD2, TNFAIP2, KCNA5, BCL2A1, RARRES3, ITGAM, IRF1, HK3, IER3, LILRB2, GCH1, GBP1, GBP2, SLC7A7, TBC1D2B, S100A9, SCO2, S100A8, SECTM1, STAB1, IGHM, AQP9, FLT3, CD14, PIM1, NKG7, VCAN, SERPINA1, CSF3R, CD58
9	Genes in module 128	3.8 $X10^{-12}$	17/98	TIMP1, DAB2, FCGR1A, FCGR2A, IFITM3, IER3, C3AR1, GBP1, GBP2, IL1R1, CCL2, CXCL9, CD14, FPR1, CTSL, CSF1R, CD151
10	Genes in module 76	2.2 $X10^{-11}$	15/79	TLR2, CEBPB, PLA2G7, CXCL10, C3AR1, ALOX15, IL1R1, S100A9, S100A12, CCL3, CCL4, CCL2, CXCL9, CD14, FPR1
GENE FAMILY				
1	CD molecules	4.5 $X10^{-22}$	28/276	ITGAM, CD163, NRP1, FCER2, FCGR1A, FCGR2A, FCGR3A, C5AR1, LILRA5, FLT3, CD207, TLR10, CD1A, CD1C, CD1E, CSF1R, CD14, CSF3R, TLR2, TLR4, LRP1, CD58, CD274, CD81, CD151, CDH1, CXCR, PDCD1LG2
2	Metallothioneins	8.1 $X10^{-6}$	5/20	MT1DP, MT1E, MT1G, MT1X, MT2A

The cutpoint FDR p-value considered was 1.0 $X10^{-3}$. From the top 20 hits for each category, the most relevant data is provided

Thermal Cycler Programming

- First strand cDNA synthesis

Program 1: 65°C for 2 min, hold at 4°C

Program 2: 4°C for 2 min, 25°C for 30 min, 42°C for 15 min, 70°C for 15 min, hold at 4°C

- Second strand cDNA synthesis

Program 3: 4°C for 1 min, 25°C for 10 min, 50°C for 30 min, 80°C for 20 min, hold at 4°C

- SPIA amplification

Program 4: 4°C for 1 min, 47°C for 75 min, 95°C for 5 min, hold at 4°C

- cDNA fragmentation

Program 5: 37°C for 30min; 95°C for 2 min, hold at 4°C

- Biotin labelling

Program 6: 37°C for 60min, 70°C for 10 min, hold at 4°C