DNAJC21 mutations link a cancer prone bone marrow failure syndrome to corruption in 60S ribosome subunit maturation

Hemanth Tummala[1], Amanda J Walne[1], Mike Williams[2], Nicholas Bockett[1], Laura Collopy[1], Shirleny Cardoso[1], Alicia Ellison[1], Rob Wynn[3], Thierry Leblanc[4], Jude Fitzgibbon[5], David P Kelsell[6], David A van Heel[1], Elspeth Payne[7], Vincent Plagnol[8], Inderjeet Dokal[1*] and Tom Vulliamy[1*†]

Running Title: DNAJC21 mutations cause bone marrow failure

[1]Centre for Genomics and Child Health, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, E1 2AT, UK; [2]Clinical and Laboratory Haematology, Birmingham Children's Hospital, Birmingham, B4 6NH, UK; [3]Blood and Marrow Transplant Unit, Royal Manchester Children's Hospital, Manchester, M13 9WL, UK; [4]Pediatric Hematology, Hôpital Robert-Debré, APHP, Paris, 75019, France; [5]Barts Cancer Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK; [6]Centre for Cell Biology and Cutaneous Research, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, E1 2AT, UK; [7]UCL Cancer Institute, 72 Huntley Street, London, WC1E 6DD, UK; [8]UCL Genetics Institute, Gower Place, London, WC1E 6DD, UK.

*Joint senior authors

†Author of Correspondence:
Professor Tom Vulliamy
Blizard Institute, Centre for Genomics and Child Health
Barts and The London School of Medicine and Dentistry
Queen Mary University of London
4 Newark Street, London, E1 2AT, UK
Tel: +44 (0) 207 882 2623; Fax: +44 (0) 207 882 2195; email: t.vulliamy@qmul.ac.uk

A substantial number of cases with bone marrow failure (BMF) present with one or more extra-hematopoietic abnormality. This suggests a constitutional or inherited basis, and yet many of them do not fit the diagnostic criteria of the known BMF syndromes. Through exome sequencing, we have now identified a sub group of these cases, defined by germline biallelic mutations in *DNAJC21* (DNAJ homolog subfamily C member 21). They present with global bone marrow failure and one individual developed a hematological cancer (acute myeloid leukemia) in childhood. We show that the encoded protein associates with ribosomal RNA (rRNA) and plays a highly conserved role in the maturation of the 60S ribosomal subunit. Lymphoblastoid cells obtained from an affected individual exhibit increased sensitivity to the transcriptional inhibitor actinomycin D and reduced levels of rRNA. Characterisation of mutations revealed impairment in interactions with cofactors (PA2G4, HSPA8 and ZNF622) involved in 60S maturation. DNAJC21 deficiency resulted in cytoplasmic accumulation of the 60S nuclear export factor PA2G4, aberrant ribosome profiles and increased cell death. Collectively these findings demonstrate that mutations in *DNAJC21* cause a cancer prone BMF syndrome due to corruption of early nuclear rRNA biogenesis and late cytoplasmic maturation of the 60S subunit.

Introduction

Inherited bone marrow failure (BMF) syndromes are a heterogeneous group of life threatening disorders characterized by a hematopoietic defect in association with a range of variable extra hematopoietic features. Recognized syndromes include Fanconi anemia [MIM 227650],[1] dyskeratosis congenita [MIM305000],[2] Shwachman Diamond syndrome [MIM260400] [3] and Diamond Blackfan anemia [MIM105650].[4] Cells from Fanconi anemia cases exhibit increased chromosomal breakage due to specific genetic and/or functional defects causing genome instability.[1] Dyskeratosis congenita cases are characterized by diagnostic mucocutaneous features such as abnormal skin pigmentation, nail dystrophy and leukoplakia together with the presence of very short telomeres.[2] Shwachman Diamond syndrome is characterized by exocrine pancreatic insufficiency and Diamond Blackfan anemia cases are typically reported to have craniofacial abnormalities, including clefting of the lip or palate, thumb abnormalities, cardiac malformations, and short stature.[5] Importantly, these syndromes are also associated with an increased cancer risk.

In addition to the well-defined BMF syndromes, we have accrued a number of cases to our BMF registry who do not fit any definite diagnostic classification. These uncharacterized cases have BMF together with one or more somatic abnormality suggesting a constitutional or inherited basis, but do not fulfil the clinical criteria of the known BMF syndromes. With the advent of whole exome sequencing it is now possible to unify subsets of these cases through the identification of their underlying genetic defect. This is the case here, as we identify a cohort of simplex BMF cases harbouring biallelic mutations in *DNAJC21*. Through characterisation of the encoded protein, we define a cancer prone BMF syndrome caused by defective DNAJC21.

Materials and Methods

Study approval. All experiments were conducted with the approval of Barts and The London Hospital. Peripheral blood samples were obtained with written consent under the approval of our local research ethics committee (London – City and East).

DNA sequencing and plasmids. All samples were obtained with informed, written consent and the approval of our local ethics committee. Genomic DNA was extracted from peripheral blood using Gentra reagents (Qiagen). For exome sequencing, 50ng of genomic DNA underwent library preparation and exome capture using the Illumina Nextera Rapid capture exome kit. Sequencing was performed using the Illumina HiSeq 2000 system and 100bp paired end reads were generated and the data processed through the Illumina pipeline. Variants were called as described previously.[6] Resequencing of *DNAJC21* was performed as a part of a 32 gene panel, amplified using a TruSeq Custom Amplicon reagent which was then prepared for MiSeq sequencing according to the manufacturer's instructions (Illumina). All relevant calls were confirmed by Sanger sequencing. *DNAJC21* wildtype and mutant expression constructs were obtained by standard cDNA cloning and mutagenesis procedures into peGFP-C1 (Clontech) plasmid.

Nuclear rRNA coimmunoprecipitation assay. Dialyzed nuclear extracts from ~ 10^6 cells expressing eGFP alone or eGFP tagged wild type and mutant forms of DNAJC21 were subjected to immunoprecipitation with GFP TRAP agarose beads (ChromTek). After incubation and stringent washes, immune complexes were treated with DNase I-RNase free (Fermentas) for 30 min at 37°C in the presence of RNasein (Invitrogen). The reactions were stopped by addition of 2.5 mM EDTA and heating for 10 min at 65°C. Coimmunoprecipitated RNA was extracted by Trizol (Invitrogen), purified using RNeasy kit

columns (Qiagen) and analysed by reverse-transcription quantitative PCR for precursor *45S* rRNA using primers indicated in Supplemental Table 3.

Quantitative RTPCR

RNA from nuclear extracts was purified with TRIzol® reagent (Invitrogen) according to manufacturer's instructions. Reverse transcription was performed with SuperScript™ III Reverse Transcriptase (Invitrogen). Quantitative PCR was carried out on a real-time PCR detection system with SYBR® Green (Thermo Fischer) for *45S* precursor rRNA, mature *28S*, *18S* rRNA in case 4, both parents and 3 unrelated samples as controls. Data presented is normalized to *GAPDH* mRNA levels and expression levels of *SNORA63* and *SNORA68* were determined as internal standard controls.

RNAi studies. All siRNA and shRNA experiments were performed on HeLa cells cultured at 37°C in 5% CO_2 and DMEM (Cambrex), supplemented with L-glutamine, 10% FCS, and penicillin/streptomycin antibiotics. Transfections of *DNAJC21* siRNA (Sigma Cat No: EHU103331) was carried out using 50nM siRNA and Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol; a predesigned non target siRNA (Ambion) was used as negative control. Cells were harvested at 72 hours for subsequent analysis. Tet-pLKO-puro shRNA-expressing plasmids were obtained from Addgene and prepared using a protocol adopted from Addgene[7] to generate Tet-pLKO-puro -*DNAJC21* 3'UTRshRNA specific clones. The specific shRNA sequences of *DNAJC21* 3'UTR used in this study were obtained from the RNAi Consortium (TRC MISSION® shRNA library,) and the RNAi designer tool from Thermo Fischer. Target sequences are given in Supplemental Figure 4A. *DNAJC21* 3'UTR shRNA expression was induced by culturing cells in DMEM supplemented

with TET free serum (Clontech) for 72 hrs in the presence of the indicated concentrations of doxycycline (dox, Sigma, Supplemental Figure 4B) dissolved in deionized water.

Immuno-studies. For co-immunoprecipitation studies, approximately 250 µg cell lysate was used per immunoprecipitation in cells transfected with *eGFP* encoding wild type and mutant *DNAJC21* plasmids. Briefly, 72 hours post transfection lysates were precleared with rabbit control IgG (monoclonal SP137, Abcam) and further incubated with 10 µL anti-GFP antibody (Abcam, Cat. No. EPR14104) overnight at 4°C. For determination of nucleic acid mediated interaction, cell lysates were initially treated with 40 U DNase I-RNase free or 25 µg/mL RNase A for 30 min at 37 °C. The next day, 20 µL Protein A Dynabeads (Thermo Fischer) pre-equilibrated with IP buffer (25 mM Tris-HCl pH 7.9, 5 mM MgCl2, 10 % glycerol, 0.1 % NP-40, 1 mM DTT, 1X Complete EDTA-free protease inhibitor cocktail) containing 100 mM KCl was added and gently mixed for 2 h at 4 °C. Immune complexes were washed three times with IP buffer containing 150 mM KCl and two times with IP buffer containing 100 mM KCl, each for 5 min at 4 °C. Immunoprecipitates were boiled in SDS-sample buffer (25 mM TrisHCl pH 6.8, 2 % SDS, 10 % glycerol, 0.05 % bromophenol blue, 5 % 2-mercaptoethanol) for 5 min, separated by SDS-polyacrylamide gel electrophoresis (PAGE), transferred to polyvinylidene membranes and subject to western blot analysis using Western Breeze Chromogenic Kit (Thermo Fischer Cat. No. WB7106, WB7104). Antibodies against DNAJC21 (Abcam, Cat. No. ab86434, C-terminal epitope corresponding to amino acids 400-450; Protein tech Cat. No. 23411-1-AP, N terminal epitope corresponding to amino acids 43-288), PA2G4 (Abcam, Cat. No. ab119037), ZNF622 (Abcam, Cat. No. Ab57859), and HSPA8 (Abcam, Cat. No. EP1513Y) were used in the study. For immunocytochemical staining, HeLa cells and IMR-90 cells (primary lung fibroblasts obtained from ATCC® CCL-186™) were grown on coverslips, fixed in 4% paraformaldehyde, permeabilized using 0.1%

Triton-X-100 (TX100) in PBS, quenched in 50mM NH_4Cl and blocked in 10% goat serum, 1% BSA in PBS containing 0.05% TX100 for 1 hour. Cells were incubated with corresponding primary and secondary antibodies and mounted using vectashield containing DAPI (Vector Labs Cat. No. H-1200). Images were collected using an LSM710 laser scanning confocal microscope (Olympus) under relevant excitation and the emitted signals were visualized using 'ZEN' software (Zeiss). Nucleolar co-staining was determined by mouse monoclonal antibody against nucleophosmin (Abcam Cat. No. ab40696), and DNAJC21.

Sub-cellular fractionation. DNAJC21 knockdown and control cells were lysed in ice-cold HEPES containing 0.1% NP40 and centrifuged at 3,000 rpm and the supernatant was collected for extraction of cytoplasmic protein. The pellet containing nuclei were briefly washed three times in ice-cold HEPES containing 0.1% NP40 and to the pellet ice-cold radioimmunoprecipitation assay buffer (RIPA) containing a cocktail of protease and phosphatase inhibitors (Roche Cat. No. 04693116001) was added and sonicated twice for 10 s at 50% pulse to release nuclear proteins. The final mixture was shaken gently on ice for 15 min, and the nuclear fraction protein supernatant was obtained by centrifugation at 14,000 g for 15 min. Fractionated lysates were verified with antibodies against cytoplasmic GAPDH (Abcam Cat. No. EPR16891) and nuclear TATA binding protein (Abcam Cat. No. EPR3826) by immunoblotting as described above.

T cell isolation and culture. Blood samples were obtained from case 4 and her unaffected heterozygous parents. Peripheral blood mononuclear cells were separated using lymphocyte separation medium (Lonza) according to the manufacturer's instructions, washed twice with complete RPMI containing 15% FCS and resuspended at $1{\cdot}0 \times 10^6$ cells/ml in RPMI-1640

medium + 20% fetal calf serum (FCS; Invitrogen). T-cell mitosis was stimulated by the addition of 10 μg/ml of phytohaemagglutinin (PHA; Roche Applied Science). After 48 hrs the medium was replaced with fresh RPMI containing 10 units/ml of recombinant human interleukin 2 (Invitrogen), and harvested after 14 days. Cells were counted on a NucleoCounter YC-100 automated cell counter (ChaemoMetec). EBV-infected LCLs were established and grown in RPMI Media 1640 supplemented with penicillin and streptomycin, 2 mM L-glutamine, and 20% (vol/vol) FBS (Invitrogen).

Polysome analysis. For analysis of monosomes and polysomes, HeLa cells and LCLs were grown to confluence, harvested 10 minutes after treatment with 100 mg/mL cycloheximide (Sigma), and resuspended in PBS containing 100 μg/mL cycloheximide. Equal numbers of cells were lysed in 425 μl of hypotonic buffer [(5 mM Tris-HCl (pH 7.5), 2.5 mM $MgCl_2$, 1.5 mM KCl and 1x protease inhibitor cocktail (EDTA-free)], with 5 μl of 100 μg/ml CHX, 1 μl of 1 M DTT, 100 units of RNase inhibitor. Lysates were vortexed for 5 sec followed by addition of 25 μl of 10% Triton X-100 (final concentration 0.5%) and 25 μl of 10% sodium deoxycholate (final concentration 0.5%) and vortexed again for 5 sec. Lysates were then centrifuged at 16,000g for 2 minutes and supernatants were layered on 15% to 45% (wt/vol) sucrose gradients. The gradients were centrifuged at 36,000 rpm for 4 hours at 4°C in a Beckman SW41Ti rotor. The fractionated lysates were serially collected and polysome profile peaks were obtained by measuring absorbance at 254 nm. Proteins from these fractions were precipitated using 20% (vol/vol) trichloroacetic acid, separated on SDS–PAGE gels and transferred to polyvinylidene membranes for immunoblotting.

Cell death and rescue by wild type DNAJC21. After 72 hours of RNAi treatment, cells were washed in PBS; propidium iodide (PI) was added to the cell suspension to a final

concentration of 5μg/ml and analyzed on an LSRII Flow Cytometer (BD Biosciences). Plasmids encoding WT-eGFPDNAJC21 and peGFP-C1 alone were transfected into DNAJC21 3'UTR shRNA expressing stable cell lines. 48hrs post transfection cells were FAC sorted by gating for eGFP, collected and re-plated in 96 wells in equal number. It is important to note that the shRNA system targets the 3' UTR of the endogenous DNAJC21 mRNA. This UTR is absent from the plasmid-derived transcript, resulting in shRNA depletion of the endogenous DNAJC21 only upon dox induction. DNAJC21 3'UTR shRNA 1 (Supplemental Figure 3) was preferred as it showed DNAJC21 knockdown upon dox induction at a lower concentration 100ng/ml.

Structural analysis of the DNAJC21 p.Pro32Ala variant. The J domain crystal structure (PBD id: 1FAF) was obtained from the Protein Data Bank (PBD) and visualized using Swiss-PDB viewer.[8] The predicted effects of DNAJC21 p.Pro32Ala substitution were generated using this program.

Statistics. Analyses were performed with GraphPad Prism 5.0 software (GraphPad Software). Mann-Whitney U test, and 1-way ANOVA with Tukey's post hoc tests were used when appropriate as indicated within the text or Figure legends. Data are presented as the means ± SEM unless otherwise indicated. A 2-tailed P value of <0.05 was considered significant.

Results

Biallelic mutations in *DNAJC21* cause a bone marrow failure syndrome

In a cohort of 28 unrelated cases with BMF and non-specific somatic features from our BMF registry (Table S1), we sought an underlying genetic basis through exome sequencing. Three

of these cases, none of whom had any family history of disease, were found to have homozygous likely pathogenic variants in *DNAJC21* (RefSeq: NM_001012339.2, Table 1, Figures 1A and 1B). Specifically, these were a nonsense variant c.517C>T, p.Arg173* in case 1, a splice variant c.983+1G>T in case 2 and a missense variant c.94C>G, p.Pro32Ala in case 3. Targeted resequencing of *DNAJC21* in a second cohort of 23 similar cases (Table S2) identified one further individual (case 4) with a homozygous nonsense variant c.793G>T, p.Glu265* (Table 1). Sanger sequencing of parental DNA revealed that this variant segregated as an autosomal recessive trait. The four cases with biallelic *DNAJC21* mutations are similar in clinical presentation (Table 1): they have global BMF associated with peripheral pancytopenia and are characterised by intrauterine growth restriction and/or short stature. One of them (case 3) developed acute myeloid leukemia, sub-type megakaryocytic (AML-M7), aged 12 years (Table 1).

Three of the variants are predicted to cause loss of function (LOF, defined here as nonsense or splice variants) and the transcripts containing these nonsense and predicted frameshift mutations are likely to undergo nonsense mediated mRNA decay. In the fourth, a missense variant (p.Pro32Ala) disrupts the highly conserved histidine-proline-aspartic acid (HPD) motif that lies at the heart of a J domain, which defines a family of proteins (Figures 1B and 1C).[9] In silico analysis of the p.Pro32Ala substitution on the tertiary structure of the *E coli* J domain (PDB id: 1FAF) reveals the loss of the distinctive cyclic structure of proline in a loop between two α-helices (Figure 1D). This structural alteration to the HPD motif is likely to disrupt the interaction of the J domain with its cognate heat shock protein 70 and the subsequent stimulation of ATPase.[10] To our knowledge, none of the variants have been reported previously apart from the splice variant (c.983+1 G>A), which is present only in the

heterozygous state in the Exome Aggregation Consortium database at an allele frequency of 5/120,170.

The Exome Aggregation Consortium database estimates the frequency of LOF alleles in *DNAJC21* (combining frameshift, nonsense and splice donor/acceptor site variants) at about 0.1% (123 LOF alleles in ~ 120,000 chromosomes). Hence, a frequency of LOF homozygotes and compound heterozygotes is expected to be in the range of 1 per million individuals. Not including the likely damaging homozygous missense variant, the observed frequency of LOF homozygous cases in our combined cohort (3 out of 51 individuals) significantly exceeds this expectation (binomial test $p < 10^{-10}$). We conclude that the allelic series of mutations that we have identified here defines a subgroup of cases with constitutional BMF caused by defective *DNAJC21*.

DNAJC21 is implicated in rRNA biogenesis

DNAJC21 belongs to the family of DnaJ/heat shock protein 40 chaperone proteins [10] and is ubiquitously expressed in human tissues (Figure S1). We found that DNAJC21 was present in both cytoplasm and nucleus of HeLa and 293T cells. (Figure 2A) Within the nucleus, it localised primarily to the nucleolus in three different cell lines tested (Figure 2B). T lymphocytes obtained from one affected individual (case 4) revealed a lack of DNAJC21 immunoreactivity to an antibody recognising an N-terminal epitope, while the parental control T cells stained positive (Figure 2B).

To date, little is known about human DNAJC21 apart from a report that its depletion causes a 'miscellaneous' effect on ribosomal RNA (rRNA) processing in HeLa cells.[11] To investigate the possible relationship between DNAJC21 and rRNA, HeLa cells were treated with low

levels (2nM) of actinomycin D (ActD) that are sufficient to exert its inhibitory effect on rRNA synthesis.[12] This treatment resulted in a translocation of DNAJC21 from the cytoplasm to the nucleus (Figure 2C and D). We also observed that lymphoblastoid cells (LCLs) derived from case 4 showed an increased sensitivity to ActD treatment when compared to parent controls (Figure 2E). Quantitative RTPCR analysis of nuclear RNA extracts obtained from these cells showed a reduction in rRNA levels in the affected individual, when compared to both parents and unrelated controls (Figure 2F). These results indicate that DNAJC21 is involved in rRNA biogenesis.

Given its nucleolar localisation and response to ActD, we sought to investigate a role for DNAJC21 in *45S* precursor rRNA binding. We therefore performed native RNA co-immunoprecipitation experiments on nuclear extracts from HeLa cells expressing wild type and mutant forms of green fluorescent protein (GFP) tagged DNAJC21. Levels of ectopically expressed GFP tagged proteins were established by immunoblotting using an anti-GFP antibody (Figure 3A). Quantitative RTPCR analysis on GFP-TRAP immunoprecipitates showed that DNAJC21 associates with the precursor *45S* rRNA. The truncated DNAJC21 mutant (p.Arg173*) failed to bind to precursor *45S* rRNA (Figure 3B). Taken together these results support the notion that DNAJC21 is indeed a pre-rRNA processing factor,[11] as it associates with precursor *45S* rRNA and is involved in rRNA biogenesis.

DNAJC21 interacts with 60S ribosome maturation factors

The yeast orthologue of DNAJC21, Jjj1 (Figure S2A) has been studied in some detail. In addition to its role in rRNA processing in the nucleus,[13] Jjj1 has been shown to act in the final stages of 60S ribosome subunit maturation in the cytoplasm .[13-18] Specifically, Jjj1 is responsible for the eviction of the 60S nuclear export receptor, Arx1 (human orthologue -

proliferation associated protein 2G4, PA2G4/ Erbb3 binding protein, EBP1/ IRES specific cellular transacting factor 45, $ITAF_{45}$ (Figure S2B).[19-21] In order to execute this function, Jjj1 acts in concert with Rei1 (human orthologue - zinc finger protein 622, ZNF622, Figure S2C), binding to the 60S subunit. Together, Jjj1 and Rei1 recruit Ssa1p (human orthologue - heat shock 70kDa protein 8, HSPA8, Figure S2D) to stimulate ATPase activity. This takes place in the cytoplasm, allowing Arx1 to recycle back into the nucleus to aid in further rounds of 60S nuclear export (Figure S3).

To determine whether these interactions also occurred in humans, we expressed eGFP-tagged wild type and mutant forms of DNAJC21 in HeLa cells. Co-immunopreciptation experiments revealed the predicted interactions of DNAJC21 with PA2G4, ZNF622 and HSPA8 (Figure 3C). These interactions were not mediated by nucleic acids since they were detected in both DNaseI and RNaseA treated cell lysates (Figure 3C). In contrast, expression of the truncation mutant (p.Arg173*) revealed no detectable interaction with any of the aforementioned proteins (Figure 3D). We note that it is likely that *in vivo*, the transcript encoding p.Arg173* undergoes nonsense mediated decay and the variant is included in these studies as a negative control. The missense variant (p.Pro32Ala) did interact with PA2G4 and ZNF622, but failed to interact with HSPA8 (Figure 3D). This is consistent with studies in yeast, which have shown that mutation of the HPD motif in Jjj1 results in loss of interaction with its HSP70 binding partner, Ssa1p. [13, 15, 17]

As these mutants failed to interact with 60S maturation factors, we further investigated the role of DNAJC21 in ribosome biogenesis. Immunostudies revealed a dramatic accumulation of PA2G4 in the cytoplasm of *DNAJC21* knockdown cells as well as T lymphocytes from case 4, compared to the predominantly nuclear localisation observed in the relevant controls

(Figures 4A-C). This data suggests that loss of DNAJC21 perturbs PA2G4 traffic and is consistent with previous studies in yeast, which showed that Arx1 fails to traffic back to the nucleus in strains that lack Jjj1.[14, 16-18] However, there was no change in the relative abundance of PA2G4, ZNF622 and HSPA8 in case 4 LCLs compared to parent controls (Figure 4D). Sucrose density centrifugation analysis revealed abnormal polysome profiles in case 4 LCLs as well as in *DNAJC21* knockdown cells when compared to controls (Figure 4E). Furthermore, a stark increase in the amount of PA2G4 bound to 60S and 80S peak fractions is observed in case 4 LCLs (Figure 4E). Collectively, these observations indicate that human DNAJC21 plays a prominent role in the maturation of the 60S ribosome subunit.

Loss of DNAJC21 disrupts cell morphology and inhibits cell growth

Defective ribosome biogenesis could disrupt protein translation, impairing cellular growth. In line with this, knockdown of *DNAJC21* induced cell death in HeLa cells (Figure 5A and5 B). DNAJC21 knockdown cells also exhibited an unusual morphology with markedly elongated cellular morphology (Figure 5C). T cells obtained from case 4 also showed a significant impairment in growth rate following mitogenic stimulation with phytohemagglutinin and interleukin-2, relative to parental controls (Figure 5D, $p = <0.001$, Student's t-Test).

To confirm that the cell death, cytoplasmic accumulation of PA2G4 and altered morphology were a consequence of DNAJC21 deficit, we performed experiments using a doxycycline (dox) inducible shRNA system that targeted the 3'-untranslated region of *DNAJC21*. Cell lines expressing *DNAJC21* 3'UTR shRNA were transfected with either eGFP-DNAJC21 or the eGFP vector alone. Knockdown of the endogenous protein and the ectopic expression of eGFP tagged wild type DNAJC21 (which escapes the effect of the shRNA) were verified by immunoblotting (Figure 5E). As expected, re-introduction of wild type DNAJC21 rescued

cell viability, restored normal PA2G4 traffic and morphology (Figure 5F and 5G), demonstrating that the observed effects were a specific consequence of DNAJC21 depletion.

Discussion

Although the inherited BMF syndromes have several distinguishing hallmarks, their heterogeneous presentation and overlapping features have often confounded their clinical diagnosis. However, as the underlying disease associated genes have been discovered, our understanding of BMF pathogenesis has improved considerably. Now, in the era of next generation sequencing it is possible to identify subgroups of cases, unified by their underlying genetic defects. In this study we have used whole exome sequencing to identify a distinct subgroup of cases presenting with constitutional BMF unified by defective DNAJC21.

We have shown that DNAJC21 associates with precursor *45S* rRNA. Cells from affected case harbouring biallelic *DNAJC21* mutations had reduced levels of precursor *45S* and mature *28S* rRNA and exhibited increased sensitivity to actinomycin D, implicating its role in rRNA biogenesis. In addition, DNAJC21 participates in coordinating nucleo-cytoplasmic shuttling of PA2G4, to aid in nuclear export of 60S subunit. Our findings from *DNAJC21* RNAi studies in HeLa cells and cells from an affected case provide compelling evidence in support of the hypothesis that loss of DNAJC21 perturbs an essential late cytoplasmic step of the 60S subunit maturation, thereby corrupting ribosome biogenesis. Previous studies reported that PA2G4 binds to FG repeat nucleoporins through its methionine amino peptidase activity and aid in 60S nuclear export.[22] In eukaryotic cells, PA2G4 is shown to regulate diverse functions such as cell growth and differentiation, interaction with transcription factors,[21] ribosome biogenesis and rRNA processing [19, 23] as well as IRES-mediated translation.[20] Based on these

reports it is conceivable that cytoplasmic accumulation of PA2G4 in absence of DNAJC21 would affect these wide range of diverse functions in human cells. Together, our data clearly establishes that the function of DNAJC21 in ribosome biogenesis is highly conserved from yeast to humans.[13-18]

Altered or reduced ribosome biogenesis has long been implicated in bone marrow failure syndromes.[5] We now add DNAJC21 to the growing list of ribosomopathies, which are a class of disease caused by mutations that affect the biosynthesis and or function of the ribosome.[24] Ribosomopathies are intriguingly enigmatic as they initially present with pathognomonic features of too few cells in the marrow but may (in ~ 10-20% of cases) later give rise to hyper-proliferative disorders such as acute myeloid leukaemia.[5,25] It is notable that this holds true for one of the four cases presented here, although with such a small number of affected individuals it is difficult to extrapolate this proportion to a larger cohort. While it is likely that these cancers arise from a selection pressure resulting in the acquisition of mutations, we also believe that a dearth of the 60S ribosome subunit could exert pressure on cells to use defective ribosomes. Over time, these cells may become error prone leading eventually to transition from a hypo- to a hyper-proliferative phenotype.[25, 26]

We have identified two distinct areas of operation of DNAJC21, one in nucleolar rRNA biogenesis and a second in cytoplasmic recycling of nuclear export factor PA2G4 for 60S subunit maturation. We note that the yeast orthologue of human DNAJC21 (Jjj1) acts just upstream of the release of Tif6 by Sdo1 in 60S maturation,[27] and that the human orthologue of Sdo1 is SBDS the protein defective in Shwachman Diamond syndrome.[28-31] By analogy with SBDS, it is therefore more likely that the disease phenotype in our cases is caused by the late 60S maturation deficit. This is also consistent with the suggestion that mutations that affect the overall structural integrity of the ribosome are less likely to be tolerated. [26]

It is striking that two out of the four cases presented here have homozygous nonsense variants and we were unable to detect DNAJC21 in the lymphocytes of an affected individual. It therefore appears that individuals who lack DNAJC21 remain viable. This is in contrast to Shwachman Diamond syndrome cases, where no homozygotes for a recurrent nonsense variant have been reported, suggesting that null mutations in SBDS are embryonic lethal.[32] We do note, however, that the small numbers of DNAJC21 cases presented here have global BMF by the age of 12 years, while in Shwachman Diamond syndrome the BMF is more variable and often only involves an isolated neutropenia.[33]

In conclusion, our study defines a cancer prone bone marrow failure syndrome caused by mutations in *DNAJC21*. Characterisation of the encoded protein shows that it plays a highly conserved role in ribosome biogenesis. We have therefore identified a distinct ribosomopathy, highlighting the significance of corrupted ribosomes in the etiology of both hypo-proliferative and hyper-proliferative disorders.

Accession codes

Variants identified in *DNAJC21* were deposited in ClinVar database [NCBI/NIH/NLH] under accession numbers SCV000257530 SCV000257531 SCV000257532 SCV000257533.

Web Resources

Variants were compared against the following databases: dbSNP (http://www.ncbi.nlm.nih.gov/SNP/), the 1000 Genomes (http://www.1000genomes.org/data) and the ExaC database (http://www.exac.broadinstitute.org, accessed January 2016). Conserved protein domains were identified using

http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. mRNA expression array data was extracted from BioGPS (http://www.biogps.org) and GTex (http://www.gtexportal.org). The DNAJC213'UTR shRNA targets are obtained from TRC shRNA mission library (https://www.broadinstitute.org/rnai/trc/lib) and RNAi designer tool (https://rnaidesigner.thermofisher.com/rnaiexpress/setOption.do?designOption=sirna). The protocol for Addgene shRNA production was provided at http://www.addgene.org/static/data/41/67/165920fc-af64-11e0-90fe-003048dd6500.pdf.Crystal structure of J domain (PBD id: 1FAF) was obtained from protein data bank (http://www.rcsb.org) and visualized using Swiss-PDB viewer (http://spdbv.vital-it.ch/).

Online Mendelian Inheritance in Man (http://www.omim.org).

Acknowledgements

We would like to thank all the families and clinicians who contributed to this research. This work was funded by The Medical Research Council UK (MR/K000292/1), Children with Cancer UK (2013/144), Barts and The London Charity (845/1796) and Leukaemia Lymphoma Research/Bloodwise (14032). Authors would like to thank Frederic Baleydier for providing case samples, Gary Warnes for support in FACS analysis and staff at the Barts and The London Genome Centre for Sanger sequencing analysis. The authors declare no conflict of interest.

References

1. Kottemann, M.C., & Smogorzewska, A. (2013).Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature, 493, 356-363.
2. Dokal, I. (2011). Dyskeratosis congenita. Hematology Am. Soc. Hematol. Educ. Program

480-486.

3. Shimamura, A. (2006). Shwachman-Diamond syndrome. Semin. Hematol, 43, 178-188.
4. Vlachos, A., Ball, S., Dahl, N., Alter, B.P., Sheth, S., Ramenghi, U., Meerpohl, J., Karlsson, S., Liu, J.M., Leblanc T., et al. (2008). International Consensus Conference. Diagnosing and treating Diamond Blackfan anaemia: results of an international clinical consensus conference. Br. J. Haematol, 142, 859-876.
5. Ruggero, D. & Shimamura, A. (2014). Marrow failure: a window into ribosome biology. Blood, 124, 2784-2792.
6. Tummala, H., Walne, A., Collopy, L., Cardoso, S., de la Fuente, J., Lawson, S., Powell, J., Cooper, N., Foster, A., Mohammed, S., et al. (2015). Poly(A)-specific ribonuclease deficiency impacts telomere biology and causes dyskeratosis congenita. J. Clin. Invest, 125, 2151-2160.
7. Wiederschain, D., Wee, S., Chen, L., Loo, A., Yang, G., Huang, A., Chen, Y., Caponigro, G., Yao, Y.M., Lengauer, C., et al. (2009). Single-vector inducible lentiviral RNAi system for oncology target validation. Cell Cycle, 8, 498-504.
8. Guex, N. & Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis, 18, 2714–2723.
9. Tsai, J. & Douglas, M.G. (1996). A conserved HPD sequence of the J-domain is necessary for YDJ1 stimulation of Hsp70 ATPase activity at a site distinct from substrate binding. J. Biol Chem, 271, 9347-9354.
10. Qiu, X.B., Shao, Y.M., Miao, S., & Wang, L. (2006). The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. Cell Mol. Life Sci, 63, 2560-2570.
11. Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou,V., Mullier, R., Wacheul, L., Lafontaine, D.L. (2013).The complexity of human ribosome

biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Mol. Cell, 51, 539-551.

12. Iapalucci-Espinoza, S. & Franze-Fernández, M.T. (1979). Effect of protein synthesis inhibitors and low concentrations of actinomycin D on ribosomal RNA synthesis. FEBS Lett 107, 281-284.
13. Albanèse, V., Reissmann, S. & Frydman, J. (2010). A ribosome-anchored chaperone network that facilitates eukaryotic ribosome biogenesis. J. Cell Biol, 189: 69-81.
14. Greber, B.J., Boehringer, D., Montellese, C. & Ban, N. (2012). Cryo-EM structures of Arx1 and maturation factors Rei1 and Jjj1 bound to the 60S ribosomal subunit. Nat. Struct. Mol. Biol, 19, 1228-1233.
15. Meyer, A.E., Hung, N.J., Yang, P., Johnson, A.W. & Craig, E.A. (2007).The specialized cytosolic J-protein, Jjj1, functions in 60S ribosomal subunit biogenesis. Proc. Natl. Acad. Sci. USA, 104, 1558-1563.
16. Demoinet, E., Jacquier, A., Lutfalla, G. & Fromont-Racine, M. (2007).The Hsp40 chaperone Jjj1 is required for the nucleo-cytoplasmic recycling of preribosomal factors in Saccharomyces cerevisiae. RNA, 13, 1570-1581.
17. Sahi, C. & Craig, E.A. (2007). Network of general and specialty J protein chaperones of the yeast cytosol. Proc. Natl. Acad. Sci. U S A, 104, 7163-7168.
18. Meyer, A.E., Hoover, L.A. & Craig, E.A. (2010). The cytosolic J-protein, Jjj1, and Rei1 function in the removal of the pre-60 S subunit factor Arx1. J. Biol. Chem, 285, 961-968.
19. Squatrito, M., Mancino, M., Donzelli, M., Areces, L.B. & Draetta, G.F. (2004). EBP1 is a nucleolar growth-regulating protein that is part of pre-ribosomal ribonucleoprotein complexes. Oncogene, 23, 4454-4465.

20. Pilipenko, E.V., Pestova, T.V., Kolupaeva, V.G., Khitrina, E.V., Poperechnaya ,A.N., Agol, V.I., Hellen, C.U. (2000). A cell cycle-dependent protein serves as a template-specific translation initiation factor. Genes Dev, 14, 2028-2045.
21. Xia, X., Cheng, A., Lessor, T., Zhang, Y., Hamburger, A.W. (2001). Ebp1, an ErbB-3 binding protein, interacts with Rb and affects Rb transcriptional regulation. J Cell Physiol, 187, 209-217.
22. Bradatsch, B., Katahira, J., Kowalinski, E., Bange, G., Yao, W., Sekimoto, T., Baumgärtel, V., Boese, G., Bassler, J., Wild, K., et al. (2007). Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. Mol Cell, 27, 767-779.
23. Squatrito, M., Mancino, M., Sala, L., Draetta, G.F. (2006). Ebp1 is a dsRNA-binding protein associated with ribosomes that modulates eIF2alpha phosphorylation. Biochem Biophys Res Commun 344, 859-868.
24. Narla, A. & Ebert, B.L. (2010). Ribosomopathies: Human disorders of ribosome dysfunction. Blood, 115, 3196-3205.
25. Sulima, S.O., Patchett, S., Advani, V.M., De Keersmaecker, K., Johnson, A.W., Dinman, J.D. (2014).Bypass of the pre-60S ribosomal quality control as a pathway to oncogenesis. Proc. Natl. Acad. Sci. USA, 111, 5640-5645.
26. De Keersmaecker, K., Sulima, S.O., Dinman, J.D. (2015). Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. Blood, 125, 1377-1382.
27. Thomson, E., Ferreira-Cerca, S., Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. J Cell Sci, 126, 4815-4821.
28. Menne, T.F., Goyenechea, B., Sánchez-Puig, N., Wong, C.C., Tonkin, L.M., Ancliff, P.J., Brost, R.L., Costanzo, M., Boone, C., Warren, A.J. (2007). The Shwachman-Bodian-

Diamond syndrome protein mediates translational activation of ribosomes in yeast. Nat. Genet, 39, 486-495.

29. Weis, F., Giudice, E., Churcher, M., Jin, L., Hilcenko, C., Wong, C.C., Traynor, D., Kay, R.R., Warren, A.J. (2015). Mechanism of eIF6 release from the nascent 60S ribosomal subunit. Nat. Struct. Mol. Biol 22, 914-919.
30. Wong, C.C., Traynor, D., Basse, N., Kay, R.R., Warren, A.J. (2011). Defective ribosome assembly in Shwachman-Diamond syndrome. Blood, 118, 4305–4312.
31. Lo, K.Y., Li, Z., Bussiere, C., Bresson, S., Marcotte, E.M., Johnson, A.W. (2010). Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. Mol. Cell, 39, 196-208.
32. Shammas, C., Menne, T.F., Hilcenko, C., Michell, S.R., Goyenechea, B., Boocock, G.R., Durie, P.R., Rommens, J.M., Warren, A.J. (2005). Structural and mutational analysis of the SBDS protein family. Insight into the leukemia-associated Shwachman-Diamond Syndrome. J. Biol. Chem, 280, 19221-19229.
33. Dror, Y. & Freedman, M.H. (2002). Shwachman-diamond syndrome. Br. J. Haematol 118, 701-713.
34. Cawthon, R. M. (2009). Telomere length measurement by a novel monochrome multiplex quantitation PCR method. Nucleic Acids Res 37, e21.

Figures

Figure 1.

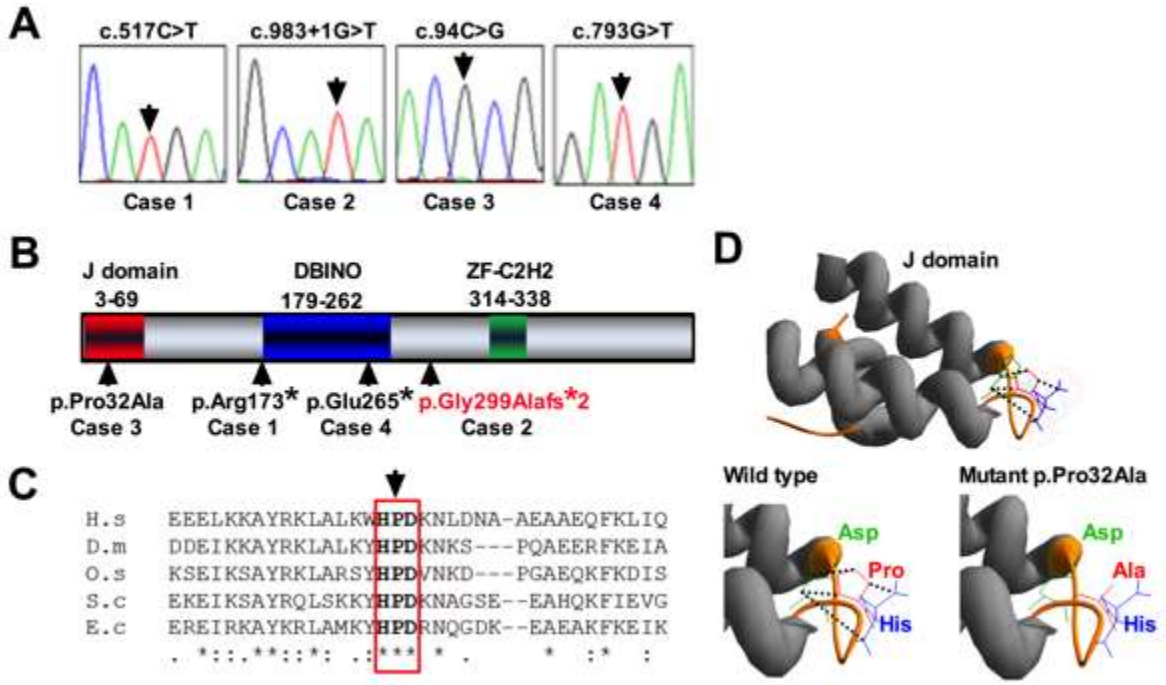


Figure 2.

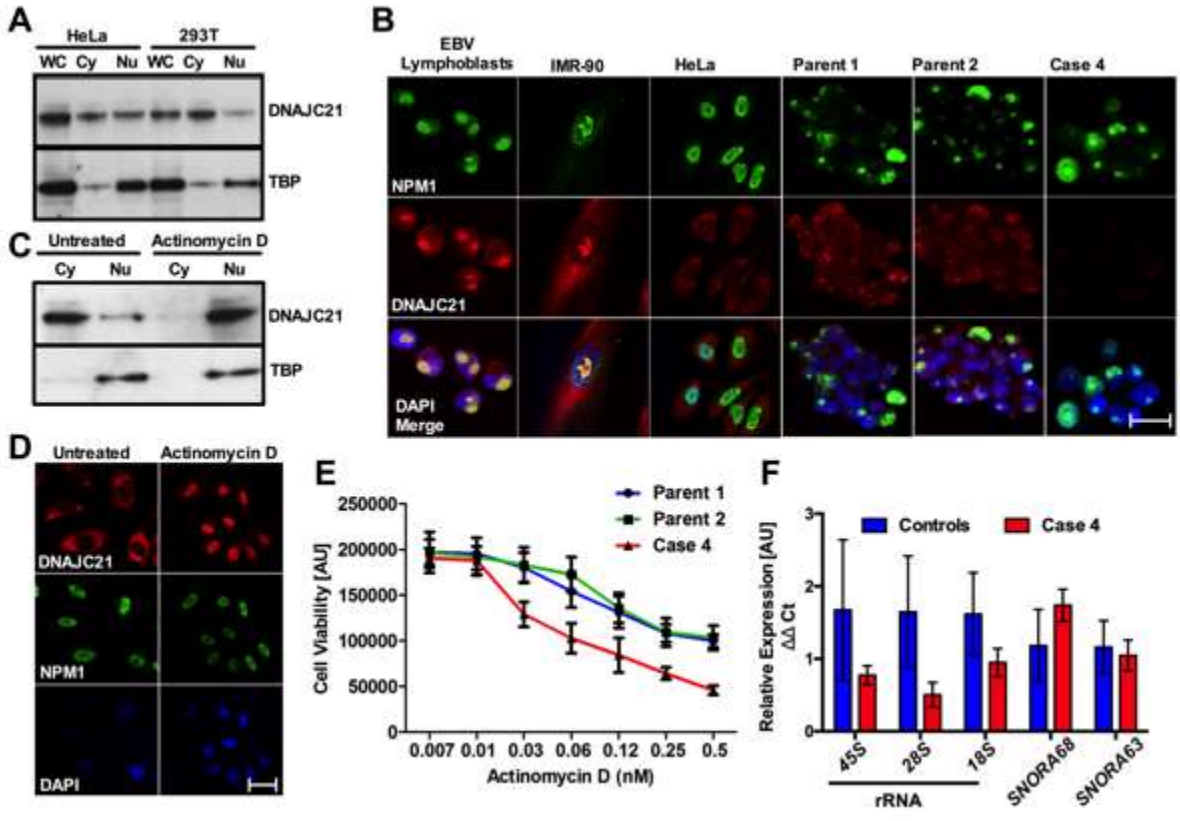


Figure 3

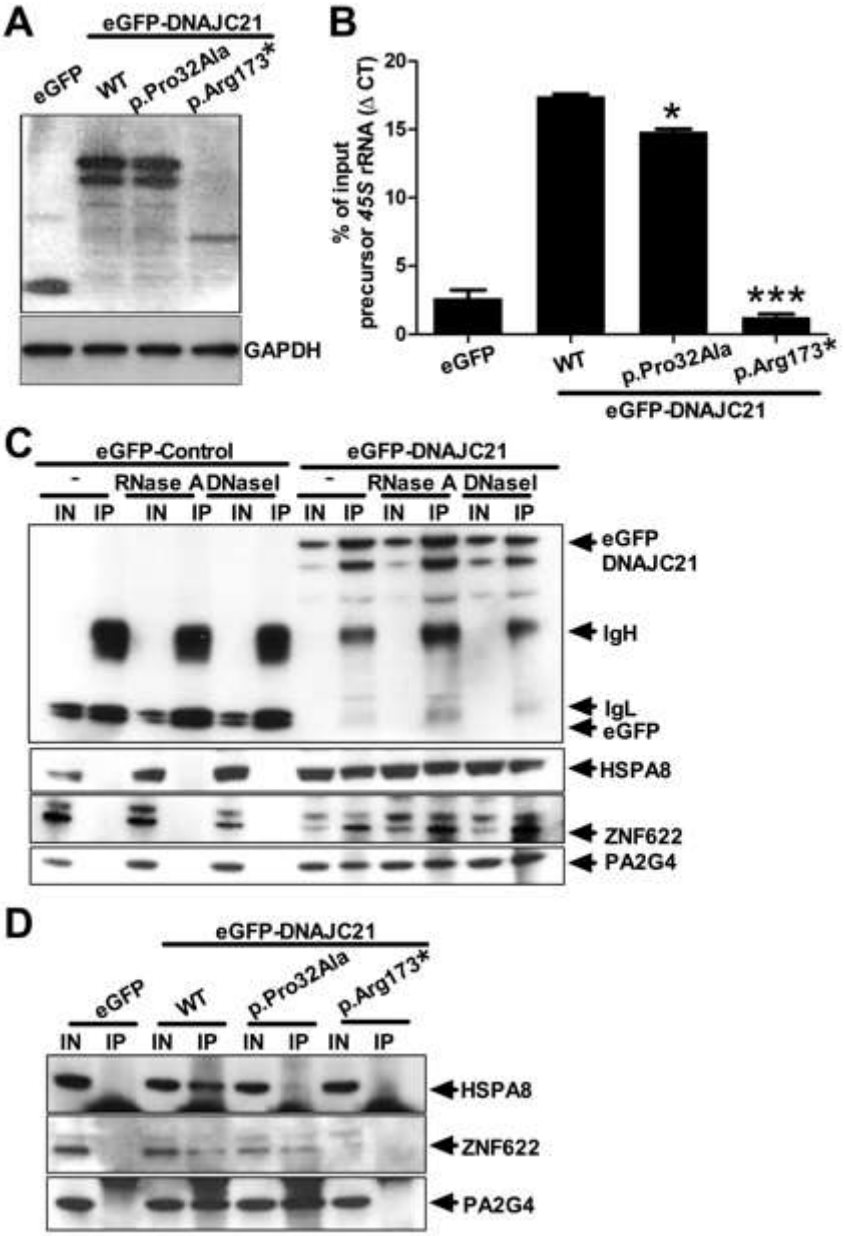

A
eGFP-DNAJC21
eGFP
WT
p.Pro32Ala
p.Arg173*
GAPDH
B
% of input precursor 45S rRNA (Δ CT)
20
15
10
5
0
*
***
eGFP
WT
p.Pro32Ala
p.Arg173*
eGFP-DNAJC21
C
eGFP-Control
eGFP-DNAJC21
-
RNase A
DNaseI
IN
IP
eGFP
DNAJC21
IgH
IgL
eGFP
HSPA8
ZNF622
PA2G4
D
eGFP-DNAJC21
eGFP
WT
p.Pro32Ala
p.Arg173*
IN
IP
HSPA8
ZNF622
PA2G4

Figure 4.

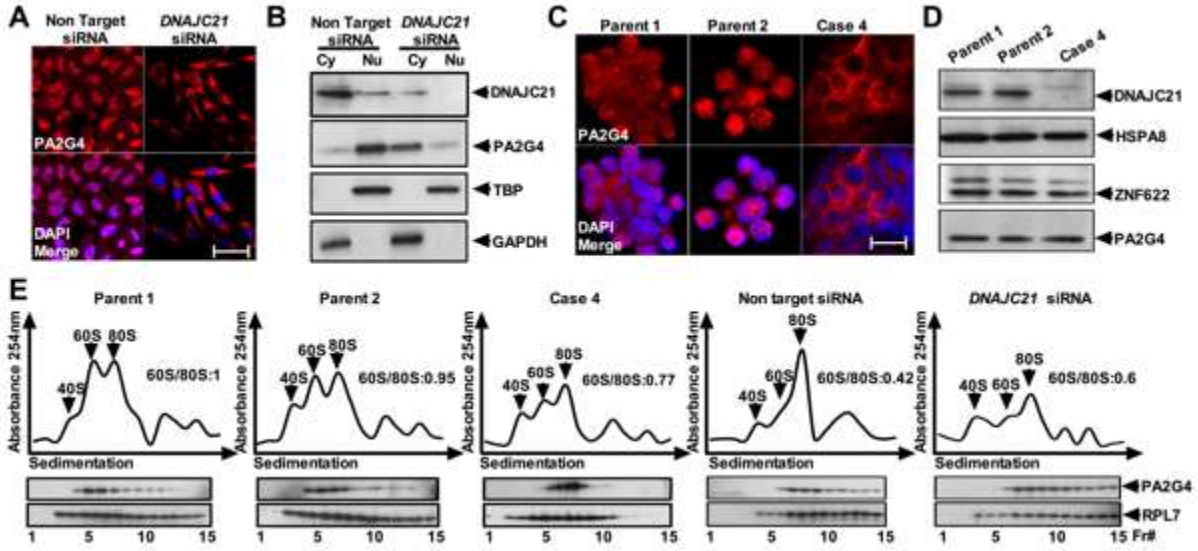

A
Non Target siRNA
DNAJC21 siRNA
PA2G4
DAPI
Merge
B
Non Target siRNA
DNAJC21 siRNA
Cy
Nu
Cy
Nu
DNAJC21
PA2G4
TBP
GAPDH
C
Parent 1
Parent 2
Case 4
PA2G4
DAPI
Merge
D
Parent 1
Parent 2
Case 4
DNAJC21
HSPA8
ZNF622
PA2G4
E
Parent 1
Parent 2
Case 4
Non target siRNA
DNAJC21 siRNA
Absorbance 254nm
40S
60S
80S
60S/80S:1
60S/80S:0.95
60S/80S:0.77
60S/80S:0.42
60S/80S:0.6
Sedimentation
PA2G4
RPL7
1
5
10
15
Fr#

Figure 5.

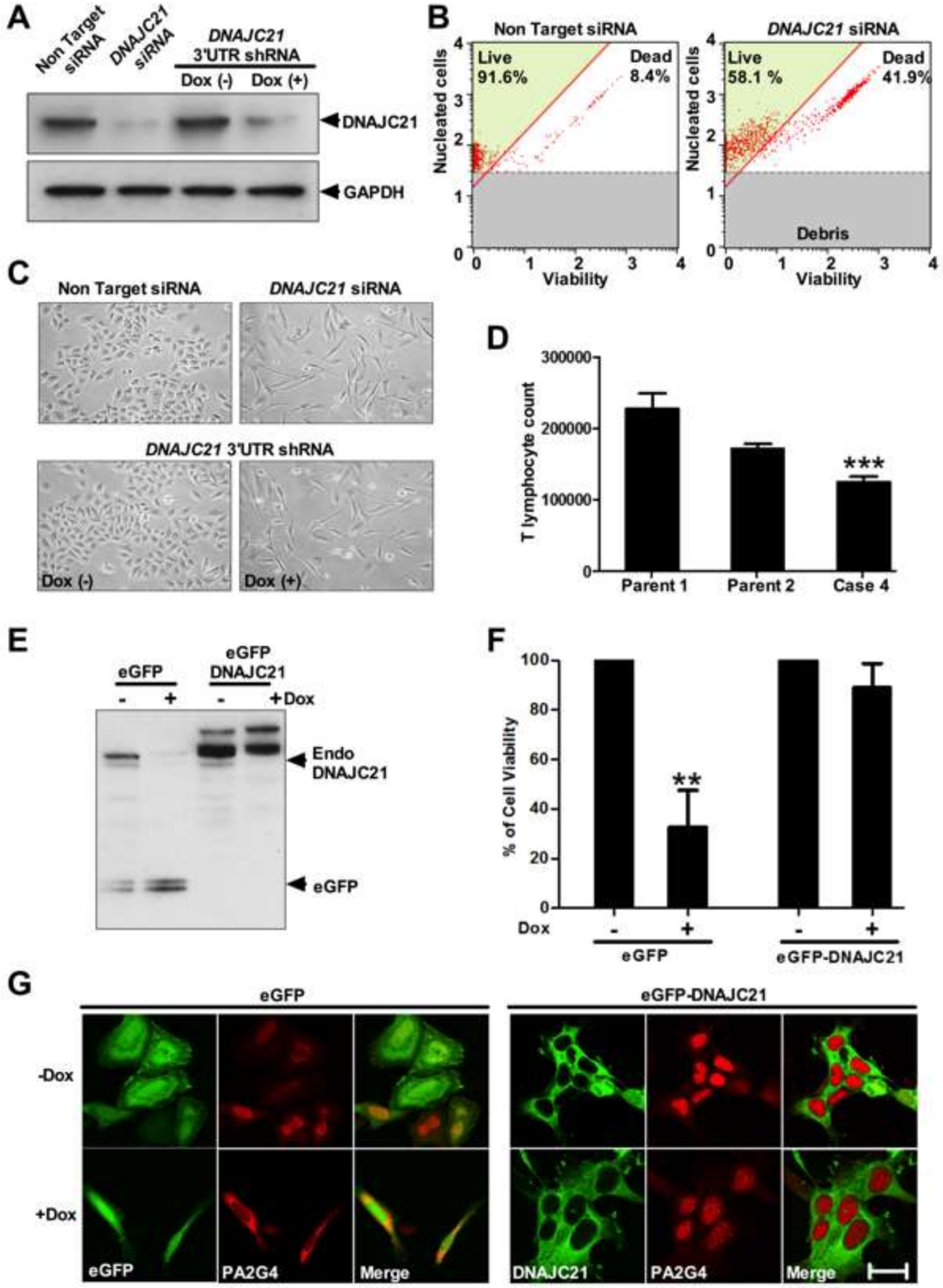

A
Non Target siRNA
DNAJC21 siRNA
DNAJC21 3'UTR shRNA
Dox (-)
Dox (+)
DNAJC21
GAPDH
B
Non Target siRNA
DNAJC21 siRNA
Live 91.6%
Dead 8.4%
Live 58.1 %
Dead 41.9%
Debris
Nucleated cells
Viability
C
Non Target siRNA
DNAJC21 siRNA
DNAJC21 3'UTR shRNA
Dox (-)
Dox (+)
D
T lymphocyte count
300000
200000
100000
0
***
Parent 1
Parent 2
Case 4
E
eGFP
eGFP DNAJC21
- + - +Dox
Endo DNAJC21
eGFP
F
% of Cell Viability
100
80
60
40
20
0
**
Dox
- + - +
eGFP
eGFP-DNAJC21
G
eGFP
eGFP-DNAJC21
-Dox
+Dox
eGFP
PA2G4
Merge
DNAJC21
PA2G4
Merge

Figure Legends

Figure 1. Biallelic mutations in *DNAJC21*

(**A**) Sanger sequencing traces and the genotype of each case are given. Arrows indicate the mutated nucleotides. (**B**) The positions of the alterations caused by the mutations observed in our cases are shown in the context of the conserved functional domains predicted in the DNAJC21 amino acid sequence. Residue numbers are given for the different domains. J domain refers to DNAJ domain that defines this group of proteins; DBINO refers to the DNA-binding domain found on global transcription activator SNF2L1 proteins and chromatin re-modelling proteins. The frameshift p.Gly299Alafs*2 is shown in red as it is predicted to arise from deletion of exon 7 due to the splice site mutation c.983+1G>T in case 2. (**C**) Conservation of the HPD motif in the J domain of DNAJC21. Alignment of residues 16-50 of the human DNAJC21 J domain with four other DnaJ proteins was generated with ClustalW. The arrow indicates the proline residue that was mutated in case 3. The aligned sequences are human (H.s) DNAJC21 (Q5F1R6), Drosophila melanogaster (D.m) DNAJ-1 (Q53ZT0), rice (O.s) Oryza sativa DNAJ homologue (Q948S9), Saccharomyces cerevisiae (S.c) DnaJ-related protein SCJ1 (P25303) and Escherichia coli (E.c) chaperone protein DnaJ (P08622). Asterisks indicate positions that have a single fully conserved residue, colons indicate conservation between groups of strongly similar properties, and periods indicate conservation between groups of weakly similar properties. (**D**) E coli DnaJ domain topology (PDB id:1FAF) visualised using swiss PDB viewer. Ribbon diagram depicts J domain (grey), position of HPD motif (orange loop). The hydrogen bonds between residues in the HDP motif are indicated by dotted line (black). Mimicking the case 3 mutant by in silico analysis revealed loss of the proline's cyclic ring disrupting hydrogen bonds (dotted line,black) when substituted with alanine in the J domain.

Figure 2. Nucleolar localisation of DNAJC21 and its role in rRNA biogenesis

(**A**) DNAJC21 is present in both cytoplasmic (Cy) and nuclear fractions (Nu); WC, whole cell. Antibody against TBP was used as control for nuclear fractions. (**B**) Immunocytochemistry shows the subcellular localisation of DNAJC21 in different cell types. A lack of DNAJC21 immunostaining is observed in T cells from case 4, while the parental T cells stain positive. Nucleophosmin (NPM1) is used as a control. Images display NPM1 (green), DNAJC21 (red) and DAPI (blue). Scale bar 20 µm. (**C and D**) DNAJC21 translocates to the nucleus following actinomycin D treatment. NPM1 is used as positive control. (**E**) LCLs from Case 4 and asymptomatic heterozygous parent controls (1 and 2) were plated in the presence of increasing concentrations of actinomycin D for 48hrs and assayed for cell viability by staining with Neutral Red. Assays were performed in octuplets per experiment and repeated for a minimum of two independent experiments. (**F**) Case 4 LCLs show reduced levels of expression for rRNA in nuclear extracts analysed when compared to both parents and 3 unrelated samples as controls. Expression levels of *SNORA63* and *SNORA68* are determined as internal controls. All genes are normalized relative to expression levels of *GAPDH* mRNA. Data represent mean ± SD, $n = 2$, performed in triplicates.

Figure 3. DNAJC21 associates with rRNA and interacts with 60S ribosome maturation factors

(**A**) Expression levels of eGFP tagged DNAJC21 wild type and mutants were detected by immunoblotting using anti-GFP. GAPDH is used as loading control. (**B**) Immunoprecipitation was performed with GFPTRAP agarose beads and nuclear extracts of HeLa cells, expressing wildtype and mutant forms of DNAJC21. EGFP alone expressing cells were used as control for the co immunoprecipitation. RNA was extracted from the

immunoprecipitates and analysed by RTPCR. Data represent mean ± SD, $n = 2$ independent experiments performed in duplicate. (*$P < 0.05$, ***$P < 0.0001$ Mann-Whitney U test). (**C**) Co-immunoprecipitations were performed in HeLa cells expressing eGFP or eGFP-DNAJC21. Lysates were treated with DNase I or RNase A prior to immunoprecipitation where indicated. Co-immunoprecipitated proteins were detected by western blot analysis with eGFP, HSPA8, ZNF622 and PA2G4 specific antibodies. (**D**) Co-immunopreciptation from cells expressing eGFP tagged wild type and mutant forms of DNAJC21 reveal the variable interaction with co factors involved in 60S ribosome maturation. IN, 10% input; IP, immunoprecipitate; WT, wild type.

Figure 4. Loss of DNAJC21 impairs traffic of PA2G4 corrupting ribosome biogenesis

(**A, B, C**) Cytoplasmic accumulation of PA2G4 was observed in DNAJC21 knockdown cells and T lymphocytes from case 4 when compared to relevant controls. Images (**A** and **C**) show DAPI (blue), PA2G4 (red). Scale bars 35 µm (panel **A**) and 20 µm (panel **C**). Panels indicate representative of images taken from different fields of view in three separate experiments. Panel **B** shows immunoblotting of cytoplasmic (C) and nuclear (N) fractions. (**D**) Immunoblotting of LCL lysates from case 4 and parental controls for DNAJC21 and 60S maturation factors HSPA8, PA2G4 and ZNF622. (**E**) Cycloheximide treated lysates from case 4 and parent control LCLs as well as *DNAJC21* and non-target siRNA treated HeLa cells were fractionated on 15% to 45% sucrose gradients by ultracentrifugation. Absorbance at 254 nM across the gradient fraction is shown (top panels). Ribosomal subunit ratios (60S:80S) are indicated. TCA precipitated proteins from equal aliquots of each fraction were immunoblotted for PA2G4 and RPL7 as a marker for 60S and 80S subunits and polysomes. Data presented is representative of two independent experiments (n=2)

Figure 5. Loss of DNAJC21 inhibits cell growth

(A) After 72 hours of *DNAJC21* siRNA by doxycycline treatment to induce shRNA targeting the 3'UTR of *DNAJC21* (100ng/ml), whole-cell extracts were subjected to western blot analysis using DNAJC21 and GAPDH specific antibodies. **(B)** FACS plots show increased cell death in cells treated with *DNAJC21* siRNA. **(C)** Phase contrast microscopy images of *DNAJC21* RNAi treated cells exhibiting abnormal morphology when compared to control. All panels are representative of images taken from different fields of view in two separate experiments. (**D**) T lymphocytes from case 4 show impaired growth when compared to parental controls. ***, p = <0.001, Student's t-Test. (**E**) Knockdown of endogenous DNAJC21 by dox administration is detected by western blotting using a mix of both DNAJC21 and GFP antibodies in stably transduced *DNAJC21* 3'UTR shRNA HeLa cells. Arrows indicate positions of endogenous DNAJC21 and eGFP control. (**F**) Cell viability relative to an untreated dox (-) culture, was measured by GFP fluorescence. This revealed the rescue of cell viability in the presence of eGFP-DNAJC21, upon dox induction of the shRNA targeting the endogenous transcript. Data represent mean ± SEM; $^{**}P < 0.01$ 1-way ANOVA with Tukey's post hoc test. (**G**) In cells expressing eGFP alone, confocal imaging revealed cytoplasmic accumulation of PA2G4 and abnormal morphology upon dox administration. However, in cells expressing eGFP-DNAJC21 no defect in PA2G4 shuttling to nucleus or cell morphology was observed upon dox addition (+) when compared to an untreated dox (-) culture. Images display eGFP (green) and PA2G4 (red). Panels are representative of images taken from different fields of view in three separate experiments. Scale bar 20 μm.

Table 1. Features of cases with biallelic mutations in *DNAJC21*

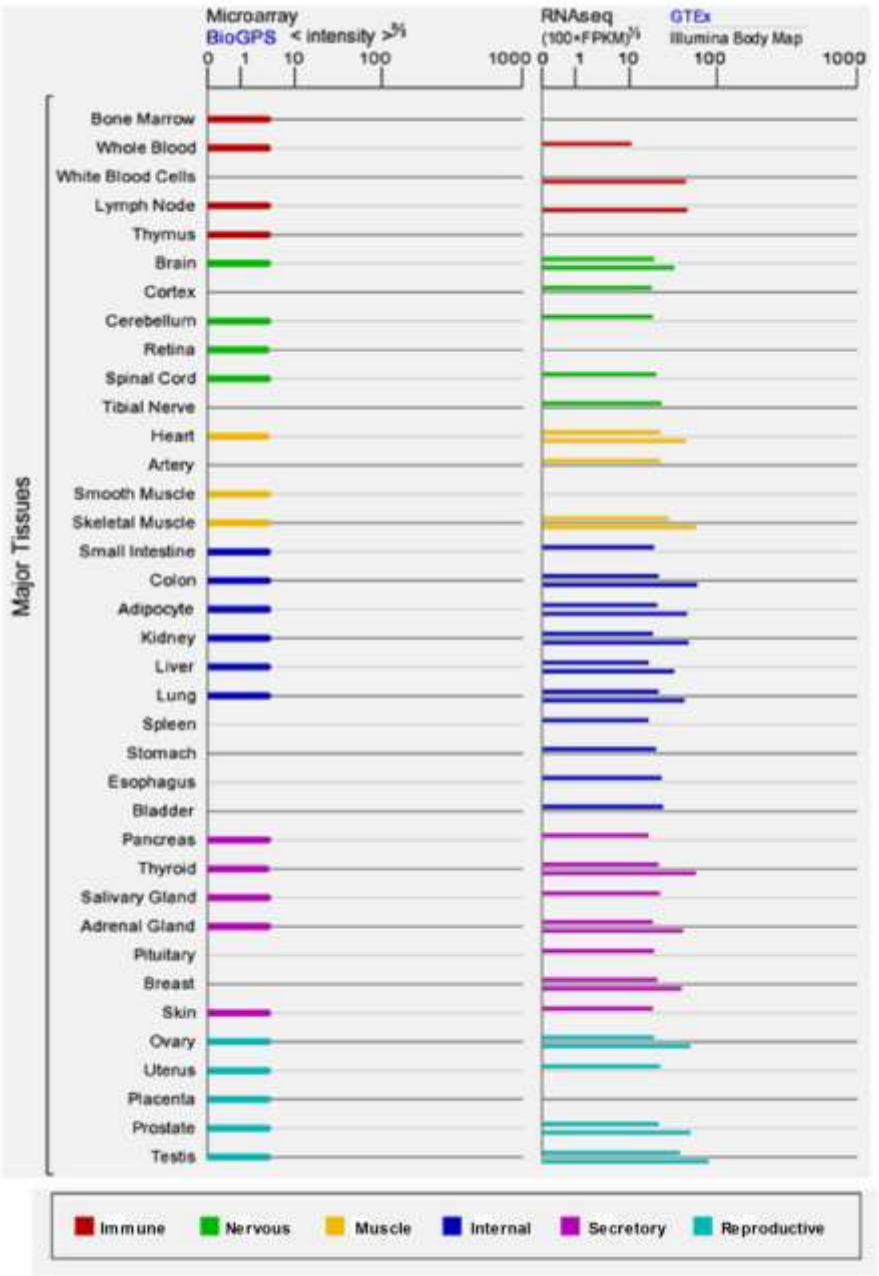
Cases	Case 1	Case 2	Case 3	Case 4
Genotype (c.)	517C>T	983+1G>T	94C>G	793G>T
Protein (p.)	Arg173*	Gly299Alafs*2[a]	Pro32Ala	Glu265*
Feature				
Gender	F	F	M	F
Age when sample received (yrs)	3	6	12	6
Country of origin	France	Algeria	Pakistan	Pakistan
Parents first cousins	Y	Y	Y	N
Bone marrow failure[b]	Y	Y	Y	Y
IUGR and/or short stature	Y	Y	Y[c]	Y
Microcephaly	Y	Y	N	N
Other features	N	Y[d]	Y[e]	Y[f]
Chromosomal breakage[g]	N	N	N	N
Telomere length [h]	Normal	Normal	Normal	Normal

F = female; M = male; Y = yes; N = no; IUGR = intrauterine growth restriction; a = inferred from splice site variant. b = the bone marrow failure was global and was associated with a reduction in all cell lineages in the peripheral blood (pancytopenia). Case 3 developed acute myeloid leukaemia sub-type megakaryocytic (AML-M7); c = growth hormone defect; d = dental abnormalities (microdontia), hyperkeratosis, retinal dystrophy with poor vision; e = skin pigmentation abnormalities on feet and dental abnormalities; f = dysphagia/oral ulceration; g = following treatment with diepoxybutane or mitomycin C; h = measured by monochrome multiplex quantitative PCR.[34]

Supplemental Information

Supplemental Figures

Supplemental Figure 1. Expression profile of *DNAJC21*



DNAJC21 is shown to be ubiquitously expressed, as observed by tissue specific mRNA expression profiles. Data obtained from two independent sources: BioGPS (http://www.biogps.org), left hand panel and GTex (http://www.gtexportal.org), right hand panel.

Supplemental Figure 2a. DnaJ homolog subfamily C member 21 (DNAJC21, UniProt reference Q5F1R6) is the human orthologue of yeast J protein Jjj1 (P53863)

Supplemental Figure 2b. Proliferation-associated protein 2G4 (PA2G4, UniProt reference Q9UQ80) is the human orthologue of yeast probable metalloprotease Arx1 (Q03862)

```
PA2G4           ----MSGEDEQ---QEQTIAEDLVVTKYKMGGDIANRVLRSLVEAS---------SSGVS 44
Arx1            MALAISHEDTQILLKDKNILQESVLNKYRTAGQIAQTALKYVTSLINDSYHSKTTQRQLT 60
                    :* ** *   :::.* :: *:.**: .*:**: .*: :..           .  ::

PA2G4           VLSLCEKGDAMIMEETGKIFKKEKEMKKGIAFPTSISVNNCVCHFSPLKSDQD------- 97
Arx1            VPELCLLTDSFILTRLEQYYKN-KVNERGIAIPTTIDIDQISGGWCPEIDDTQNLLNWNK 119
                * .**   *::*: .  : :*: *  ::***:**:*.:::    :.*  .* :

PA2G4           -----------YILKEGDLVKIDLGVHVDGFIANVAHTFVVDVAQGTQ--------VTGR 138
Arx1            GKDSTFASSVTGTLRPGDLVKITLGVHIDGYTSEVSHTMVIYPVDETKPILQPTGPLLGG 179
                             *: ****** ****:**: ::*:**:*:  .: *:        : *

PA2G4           KADVIKAAHLCAEAALRLVKPGNQNTQVTEAWN----------------KVAHSFNCTPI 182
Arx1            KADAVAAAHIAMETVVALLACALTPEKLPASLGGTSSGITGQLIRTIVDTIARSYNCGVV 239
                ***.: ***:. *:.: *:  .    ::. : .                .:*:*:**  :

PA2G4           EGMLSHQLKQHVIDGEKTIIQN-------------------------------------PT 206
Arx1            PGSRVRRIRRFLAGQNEGIVAEREYKGVVWTESHQEADLLSNTDAKDLTVVDRGQSTPFT 299
                 *   :::::.: . :: *: :                                      *

PA2G4           DQQKKDHEKAEFEVHEVYAVDVLVSSGE-----------------GKAKDAG---QRTTI 246
Arx1            NVSAIPSDDFVVQSGEVYLIDLKMASLEHCTKKGLVTLETVDSYTGKSHKAGELIARPGA 359
                : .    :.  .:  *** :*: ::* *                **::.**    *.

PA2G4           YKRDPSKQYGLKMKTSRAFFSEVERR----------FDAMPFTLRAFE-------DEKKA 289
Arx1            YVRDFAQTHILKLKTSRQLLTKIDKQGVYPFKLSHLSSNFPFVHENEEELQSLKKDLKSF 419
                * ** :: : **:**** ::::::::           . :**. .  *       * *.

PA2G4           RMGVVECAK-------------------------------------------HELLQP--- 304
Arx1            RLGMSEISNNYLCVESPIQIARWVPWDHILKATNPNGNLSYDATSTLTLPGHELPLPKLG 479
                *:*: * ::                                             ***  *

PA2G4           ---------FNVLYEKEGEFVAQFKFTVLLMPNGP--------MRITSG--PFEPDLYKS 345
Arx1            VSAIKLKSLMNSTKESISLPVARECNTIVLCDSSVSTTDRPELLRLTGGSKTCQPSWIHS 539
                         :*   *. .  **:   *::*  ..         :*:*.*  . :*.  :*

PA2G4           --EMEVQDAELKALLQSSASRKTQKK----KKKKASKTAENATSGETLEENEAGD 394
Arx1            QHELNPQDSIVQGIFQLATLAKDKRFGLLLKETQPMKQKSVETSNGGVEETMKM- 593
                  *:: **: ::.::* ::  * ::     *:.:. *  .  **.  :**.
```

Supplemental Figure 2c. Zinc finger protein 622 (ZNF622, UniProt reference Q969S3) is the human orthologue of yeast cytoplasmic 60S subunit biogenesis factor Rei1 (P38344)

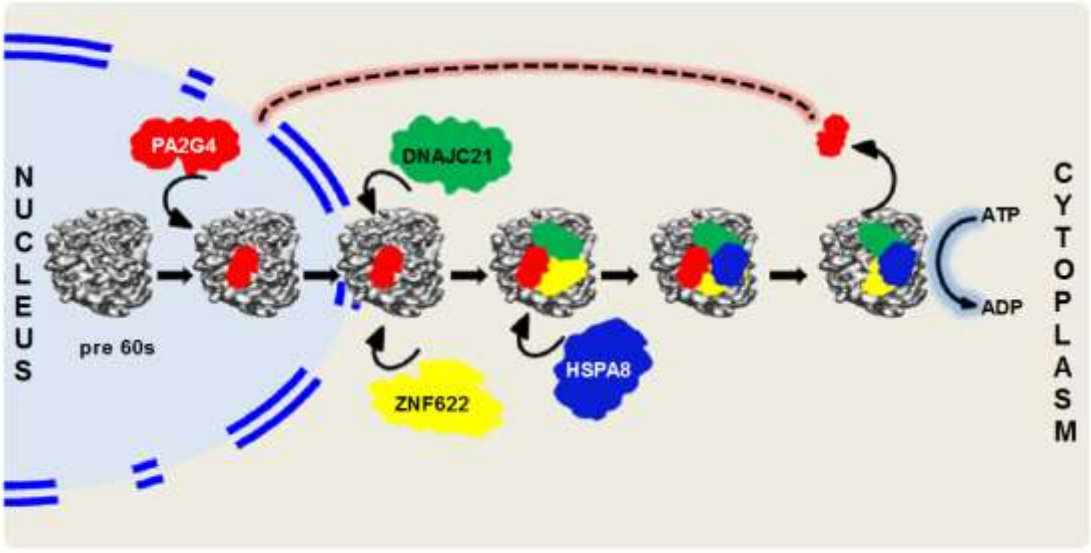
```
ZNF622          ---MATYTCITCRVAFRDADMQRAHYKTDWHRYNLRRKVASMAPVTAEGFQERVRAQRAV 57
Rei1            MSSSGVYTCNSCVLTFDSSDEQRAHMKSDWHRYNLKRRVAQLPPISFETFDSKVSAAAAS 60
                    ..*** :* ::* .:* **** *:*******:*:**.:.*:: * *:.:* *  *

ZNF622          AEEESKGSATYCTVCSKKFASFNAYENHLKSRRHVELEKKAVQAVNRKVEMMNEKNLEKG 117
Rei1            TSKSAEK------------------EKPVTKKELKRREKQALLEKKKKLLEIARAN---- 98
                :.:.::                   *: :..:.  . **:*:   ::*:  : . *

ZNF622          LGVDSVDKDAMNAAIQQAIKAQPSMSPKKAPPAPAKEARNVVAVGTGGRGTHDRDPSEKP 177
Rei1            -------------MLENMQKSQEGNTP--------------------------------- 112
                             :::  *:* . :*

ZNF622          PRLQWFEQQAKKLAKQQEEDSEEEEEDLDGDDWEDIDSDEELECEDTEAMDDVVEQDAEE 237
Rei1            --------DLSKLSLQENEENKEKEE----------PKKEEPEQLTEEEMAERVMQEKLR 154
                        : .**: *::*:.:*:**           ..** *    * * : * *:  .

ZNF622          EEAEEGPPLGAIPITDCLFCSH--HSSSLMKNVAHMTKDHSFFIPDIEYLSDIKGLIKYL 295
Rei1            NRVD-------IPLEQCLFCEHNKHFKDVEENLEHMFRTHGFYIPEQKYLVDKIGLVKYM 207
                :..:       **: :****.*  * ..: :*: ** : *.*:**: :** *  **:**:

ZNF622          GEKVGVGKICLWCNEKGKSFYSTEAVQAHMNDKSHCKLFTDG-DAALEFADFYDFRSSYP 354
Rei1            SEKIGLGNICIVCNYQGR---TLTAVRQHMLAKRHCKIPYESEDERLEISEFYDFTSSYA 264
                .**:*:*:**: ** :*:   :  **: **  * ***:  :. *  **:::**** ***.

ZNF622          DHKEGEDPNKAEELPSEKNLEYDDET-----------MELILPSGARVGHRSLMRYYKQ 402
Rei1            NFNSNTTPDNEDDWEDVGSDEAGSDDEDLPQEYLYNDGIELHLPTGIKVGHRSLQRYYKQ 324
                :.:..  *:: ::  .  . * ..:             :** **:* :****** *****

ZNF622          RFGLSRAVAVAKNRKAVGRVLQQYRALGWTGSTGAALMRERDMQYVQRMKSKWMLKTGMK 462
Rei1            DLKPEVILTEGQGTLVAAETRSFLPAFDKKG-----VQTQQRVWQTERFDKKRLDKRSAK 379
                 :  .  :: .:.  ..... .   *:. .*     :  :: :  .:*:..* : * . *

ZNF622          NNATKQMHFRVQVRF 477
Rei1            -FVNNQPHYRDQLLQ 393
                  ..:* *:* *:
```

Supplemental Figure 2d. Heat shock cognate 71 kDa protein HSPA8 (UniProt reference P11142) is the human orthologue of yeast heat shock protein Ssa1 (P10591)

```
HSPA8           MSKGPAVGIDLGTTYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVA 60
Ssa1            MSK--AVGIDLGTTYSCVAHFANDRVDIIANDQGNRTTPSFVAFTDTERLIGDAAKNQAA 58
                ***  *************. * :.:*:*************:****************.*

HSPA8           MNPTNTVFDAKRLIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVS 120
Ssa1            MNPSNTVFDAKRLIGRNFNDPEVQADMKHFPFKLIDVDGKPQIQVEFKGETKNFTPEQIS 118
                ***:************.*:*. **:****:** :::  *:*::***:*****.* **::*

HSPA8           SMVLTKMKEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAA 180
Ssa1            SMVLGKMKETAESYLGAKVNDAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAA 178
                **** **** **:*** .*.:****************************************

HSPA8           IAYGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRMVNH 240
Ssa1            IAYGLDKK-GKEEHVLIFDLGGGTFDVSLLSIEDGIFEVKATAGDTHLGGEDFDNRLVNH 237
                ******** * *.:***************:*:*********:***************:***

HSPA8           FIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGIDFYTSITRA 300
Ssa1            FIQEFKRKNKKDLSTNQRALRRLRTACERAKRTLSSSAQTSVEIDSLFEGIDFYTSITRA 297
                ** *****:***:* *:**:******************:*:*:*****:************

HSPA8           RFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQKLLQDFFNGKELN 360
Ssa1            RFEELCADLFRSTLDPVEKVLRDAKLDKSQVDEIVLVGGSTRIPKVQKLVTDYFNGKEPN 357
                ***** *****.*******.**********:.:************:***: *:***** *

HSPA8           KSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLSLGIETAGGVMTVLIKRNTTI 420
Ssa1            RSINPDEAVAYGAAVQAAILTGDESSKTQDLLLLDVAPLSLGIETAGGVMTKLIPRNSTI 417
                :*******************:**:*.:.********:************** ** **:**

HSPA8           PTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFELTGIPPAPRGVPQIEVTFDI 480
Ssa1            PTKKSEIFSTYADNQPGVLIQVFEGERAKTKDNNLLGKFELSGIPPAPRGVPQIEVTFDV 477
                ***::: *:**:**********:***** ************:*****************:

HSPA8           DANGILNVSAVDKSTGKENKITITNDKGRLSKEDIERMVQEAEKYKAEDEKQRDKVSSKN 540
Ssa1            DSNGILNVSAVEKGTGKSNKITITNDKGRLSKEDIEKMVAEAEKFKEEDEKESQRIASKN 537
                *:*********:*.***.******************:** ****:* ****: ::::***

HSPA8           SLESYAFNMKATVEDEKLQGKINDEDKQKILDKCNEIINWLDKNQTAEKEEFEHQQKELE 600
Ssa1            QLESIAYSLKNTIS--EAGDKLEQADKDTVTKKAEETISWLDSNTTASKEEFDDKLKELQ 595
                .*** *:.:* *:.  :  .*::: **:.: .*.:* *.***.* **.****:.: ***:

HSPA8           KVCNPIITKLYQS---AGGMPGGMPGGFPGGGAPPSGGASSGPTIEEVD 646
Ssa1            DIANPIMSKLYQAGGAPGGAAGGAPGGFPGG-APPAP-EAEGPTVEEVD 642
                .:.***::****:   .** .** ******* ***:   :.***:****
```

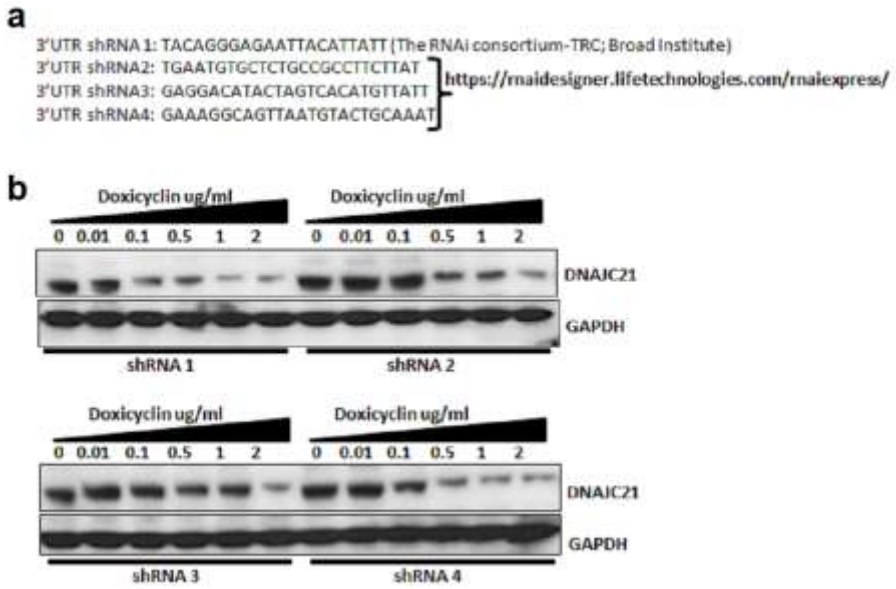
ClustalW alignment shows the homology of human (a) DNAJC21, (b) PA2G4, (c) HSPA8 and (d) ZNF622 to yeast counterparts, Jjj1, Arx1, SSa1 and Rei1. Asterisks indicate positions that have a single fully conserved residue, colons indicate conservation between groups of strongly similar properties, and periods indicate conservation between groups of weakly similar properties.

Supplemental Figure 3. Cytoplasmic eviction of PA2G4 from 60S ribosome subunit upon stimulation of ATPase activity



A schematic model of the PA2G4 release process depicted on the basis of studies reported in yeast by Greber et al 2012. In the nucleus, PA2G4 binds and exports the pre 60S ribosome particle. In the cytoplasm, DNAJC21 and ZNF622 bind the pre-60S particle, establishing a network of interactions that enable the recruitment of HSPA8. The stimulation of the HSPA8 ATPase by DNAJC21 leads to dissociation of PA2G4, enabling it to recycle back into nucleus (red dotted arrow).

Supplemental Figure 4. Inducible knockdown of *DNAJC21*



(a) Sequences of shRNA's targeting the 3'UTR of *DNAJC21*. (**b**) HeLa cell lines transduced with four independent inducible shRNA's targeting *DNAJC21 3'UTR* were cultured by the addition of doxycycline to the growth media. 72 hours post-induction cells were analysed by Western blot using a DNAJC21 antibody. GAPDH served as a loading control.

References

- Greber BJ, Boehringer D, Montellese C & Ban N. Cryo-EM structures of Arx1 and maturation factors Rei1 and Jjj1 bound to the 60S ribosomal subunit. 2012. *Nat. Struct. Mol. Biol* **19**, 1228-1233.

Supplemental Table 1. Somatic features of patients with BMF (exome sequencing)

Age	Sex	Somatic features
2	F	Developmental delay, failure to thrive
14	M	Alopecia, nail pitting
37	M	Hypopigmented skin, nail ridging
11	F	Immunodeficiency, enteropathy
4	M	Developmental delay, thin hair, café au lait spot, recurrent infections
25	F	Leukoplakia, asceptic necrosis
1	M	Developmental delay
7	M	Developmental delay, hypopigmented skin patches
6	F	Short stature, microcephaly, dental abnormalities, hyperkeratosis, retinal dystrophy (case 2)
11	F	Short stature, IUGR, hair loss, abnormal faces, extensive caries
66	F	Short stature, hair loss, dental loss, pulmonary disease
4	F	Dysmorphic features
41	F	Nail dystrophy
1	M	Dysmorphic features, skin lesions, hepatomegaly, splenomegaly
30	F	Short stature, psychiatric problems, nail dystrophy
22	M	Abnormal skin pigmentation, hair loss, dental caries
12	M	Short stature, abnormal skin pigmentation, dental abnormalities (case 3)
28	M	Short stature, cardiac hypertrophy, skeletal abnormality, cirrhosis
32	M	Marfanoid, syndactyly, learning difficulties, epilepsy, immunodeficiency
15	F	Dysmorphic face, learning difficulties, epilepsy, misaligned teeth
16	M	Nail dystrophy
11	F	Low birth weight, telangiectasia, white hair
3	F	Short stature, microcephaly (case 1)
57	M	Pulmonary fibrosis
2	M	Global developmental delay
22	F	Developmental delay, SS. Small eyes
10	M	Mental retardation, discoloured teeth, hyperkeratotic nails, ichthyotic legs
43	F	Grey hair

IUGR = intrauterine growth restriction

Supplemental Table 2. Somatic features of patients with BMF (targeted resequencing)

Age	Sex	Somatic features
39	M	Nail dystrophy
61	M	Idiopathic pulmonary fibrosis, abnormal hands
6	F	Short stature, oral ulceration/dysphagia (case 4)
42	F	Short stature, abnormal skin pigmentation
30	M	Abnormal skin pigmentation, short thumbs, splenomegaly
36	F	Short stature
7	F	Developmental delay, autistic features, alopecia, crumbling nails,
58	F	Bilateral deafness, infertility
60	M	Early grey
12	M	Short stature, skeletal abnormality
50	M	Nail dystrophy
2	M	Abnormal liver function, very short telomeres
19	F	Short stature, learning difficulties, moles, nail dystrophy, cerebellar hypoplasia
24	M	Avascular necrosis, eczema, squint
58	F	Hair loss, no eyebrows, nail dystrophy
33	F	Short stature, abnormal faces
32	M	Learning difficulties, inflammatory polyps
45	F	Nail dystrophy
14	M	Poor teeth
7	M	Enteropathy, hepatosplenomegaly, café au lait spot, pancreatic atrophy
42	M	Short stature, dysmorphic faces, stubby hands, nail dystrophy
63	F	Skin pigmentation, nail ridging, extensive caries
13	F	Colectomy, recurrent infections, liver dysfunction, poor dentition

Supplementary Table 3. rRNA primers

45S_fwd	CGCCGCGCTCTACCTTACCTA
45S_rev	TAGGAGAGGAGCGAGCGACCA
18S_fwd	CTGCCCTATCAACTTTCGATGGTAG
18S_rev	CCGTTTCTCAGGCTCCCTCTC
28S_fwd	TGTCGGCTCTTCCTATCATTGT
28S_rev	ACCCAGCTCACGTTCCCTATTA