



Genetic and functional insights into CDA-I prevalence and pathogenesis

Olijnik, Aude Anais; Roy, Noémi B.A.; Scott, Caroline; Marsh, Joseph A.; Brown, Jill; Lauschke, Karin; Ask, Katrine; Roberts, Nigel; Downes, Damien J.; Brolih, Sanja

Total number of authors:
29

Published in:
Journal of Medical Genetics

Link to article, DOI:
[10.1136/jmedgenet-2020-106880](https://doi.org/10.1136/jmedgenet-2020-106880)

Publication date:
2021

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

[Link back to DTU Orbit](#)

Citation (APA):
Olijnik, A. A., Roy, N. B. A., Scott, C., Marsh, J. A., Brown, J., Lauschke, K., Ask, K., Roberts, N., Downes, D. J., Brolih, S., Johnson, E., Xella, B., Proven, M., Hipkiss, R., Ryan, K., Frisk, P., Mäkk, J., Stattin, E. L. M., Sadasivam, N., ... Babbs, C. (2021). Genetic and functional insights into CDA-I prevalence and pathogenesis. *Journal of Medical Genetics*, 58, 185-195. <https://doi.org/10.1136/jmedgenet-2020-106880>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

ORIGINAL RESEARCH

Genetic and functional insights into CDA-I prevalence and pathogenesis

Aude-Anais Olijnik,¹ Noémi B A Roy,^{1,2,3} Caroline Scott,¹ Joseph A Marsh,⁴ Jill Brown,¹ Karin Lauschke ^{5,6} Katrine Ask,^{5,7} Nigel Roberts ⁸, Damien J Downes,¹ Sanja Brolih,⁸ Errin Johnson,⁹ Barbara Xella,¹ Melanie Proven,¹⁰ Ria Hipkiss,¹⁰ Kate Ryan,¹¹ Per Frisk,¹² Johan Mäkk,¹³ Eva-Lena Maria Stattin,¹⁴ Nandini Sadasivam,¹¹ Louisa McIlwaine,¹⁵ Quentin A Hill,¹⁶ Raffaele Renella,¹⁷ Jim R Hughes,¹ Richard J Gibbons,¹ Anja Groth,^{5,18} Peter J McHugh,⁸ Douglas R Higgs,¹ Veronica J Buckle,¹ Christian Babbs ¹

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2020-106880>).

For numbered affiliations see end of article.

Correspondence to

Dr Christian Babbs, MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, Oxford, Oxfordshire, UK; christian.babbs@imm.ox.ac.uk

A-AO and NBAR contributed equally.

Received 22 January 2020
Revised 5 March 2020
Accepted 2 April 2020



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Olijnik A-A, Roy NBA, Scott C, et al. *J Med Genet* Epub ahead of print: [please include Day Month Year]. doi:10.1136/jmedgenet-2020-106880

ABSTRACT

Background Congenital dyserythropoietic anaemia type I (CDA-I) is a hereditary anaemia caused by biallelic mutations in the widely expressed genes *CDAN1* and *C15orf41*. Little is understood about either protein and it is unclear in which cellular pathways they participate.

Methods Genetic analysis of a cohort of patients with CDA-I identifies novel pathogenic variants in both known causative genes. We analyse the mutation distribution and the predicted structural positioning of amino acids affected in Codanin-1, the protein encoded by *CDAN1*. Using western blotting, immunoprecipitation and immunofluorescence, we determine the effect of particular mutations on both proteins and interrogate protein interaction, stability and subcellular localisation.

Results We identify six novel *CDAN1* mutations and one novel mutation in *C15orf41* and uncover evidence of further genetic heterogeneity in CDA-I. Additionally, population genetics suggests that CDA-I is more common than currently predicted. Mutations are enriched in six clusters in Codanin-1 and tend to affect buried residues. Many missense and in-frame mutations do not destabilise the entire protein. Rather *C15orf41* relies on Codanin-1 for stability and both proteins, which are enriched in the nucleolus, interact to form an obligate complex in cells.

Conclusion Stability and interaction data suggest that *C15orf41* may be the key determinant of CDA-I and offer insight into the mechanism underlying this disease. Both proteins share a common pathway likely to be present in a wide variety of cell types; however, nucleolar enrichment may provide a clue as to the erythroid specific nature of CDA-I. The surprisingly high predicted incidence of CDA-I suggests that better ascertainment would lead to improved patient care.

INTRODUCTION

Congenital dyserythropoietic anaemia type I (CDA-I) (MIM 607465 and #224120) is an autosomal-recessive macrocytic anaemia, characterised by ineffective erythropoiesis and morphological abnormalities of erythroblasts.¹ The anaemia is usually mild to moderate; however, CDA-I is variable and some patients are dependent on blood

transfusions. There are also limb abnormalities in a number of cases.² Advances in genetic diagnosis over the last 20 years have led to the four major types of CDA (CDA-I to CDA-IV) being increasingly genetically defined and CDA-I has been shown to be caused by biallelic mutations in *CDAN1* and *C15orf41*.^{3,4} In the absence of proven pathogenic mutations in either of these genes, diagnosis of CDA-I relies on the presence of characteristic ultrastructural abnormalities of erythroblast nuclei detected by electron microscopy (EM). In affected cells, heterochromatin is abnormally electron dense with electron lucent patches, this has been termed spongy or ‘Swiss Cheese’ heterochromatin⁵ and is illustrated in figure 1.

The molecular basis of the heterochromatin defects remains unknown. However, the specificity of the abnormality strongly suggests that both proteins known to cause CDA-I function in a common pathway. Some progress has been made in understanding the function of both proteins, yet despite this, their specific functions and the pathway(s) in which they operate remain poorly understood.⁶ Codanin-1 is a relatively large protein (~132 kDa) and the absence of evolutionarily conserved domains, apart from a putative B-domain (see figure 3) or any structural information has been a hurdle to understanding its function. It is known that Codanin-1 plays a role in nucleosome assembly through regulation of the histone chaperone ASF1,^{7,8} is cell cycle regulated and may localise to heterochromatin.⁹ *C15orf41* is predicted to function as a nuclease⁴ but the specific activity remains to be shown. Given that both known causative proteins are widely expressed, the erythroid specific nature of CDA-I is surprising and it has been hypothesised that this may arise from the disruption of the connection between cell-cycle dynamics and terminal erythroid maturation.¹⁰

In this work, we report seven novel mutations causative of CDA-I, five of which are missense or in-frame changes in Codanin-1, present in patients with EM-positive CDA-I. By combining these novel mutations with previously reported missense mutations and in-frame deletions, we find that six regions of Codanin-1 are mutational hotspots for

Figure 1

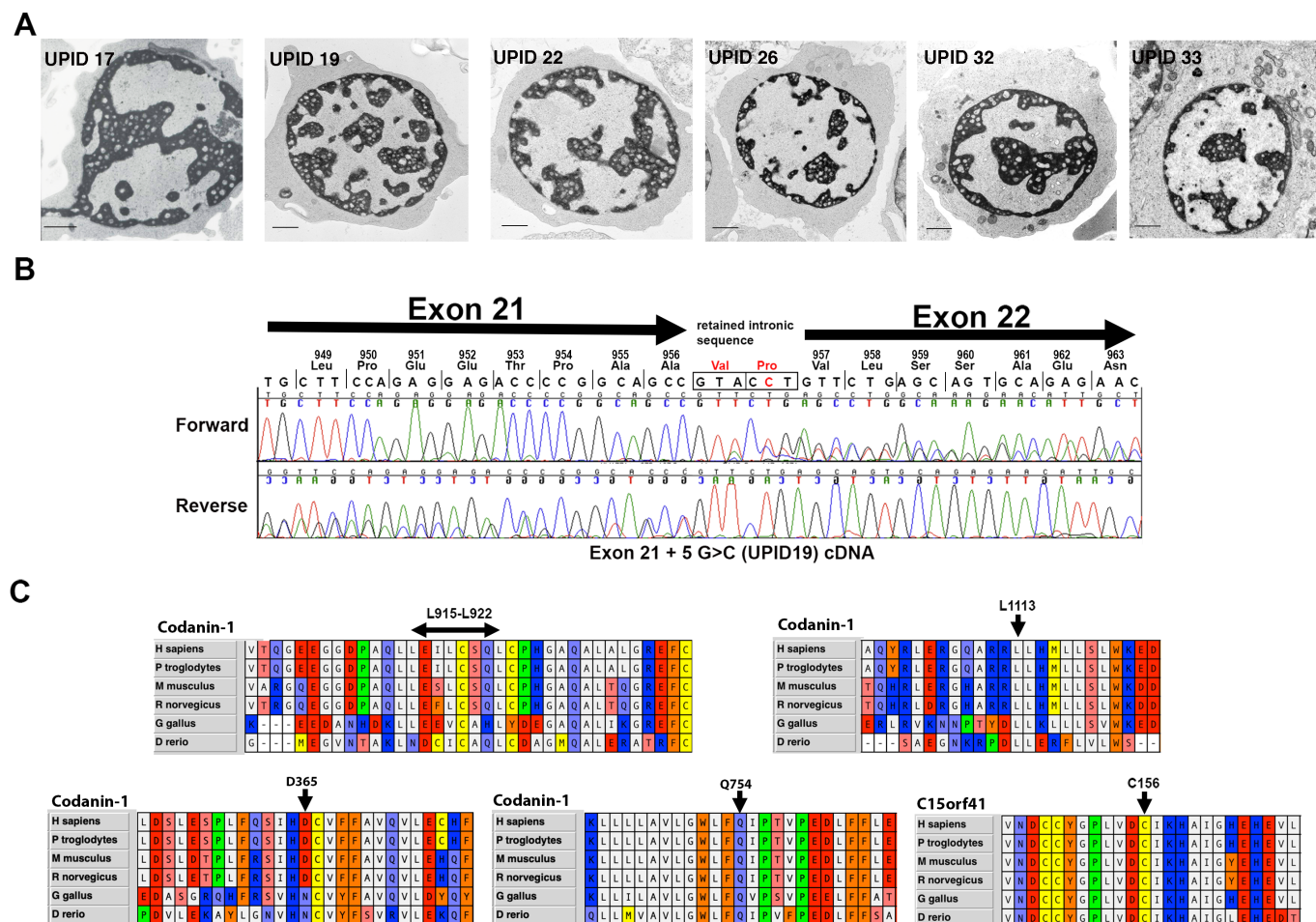


Figure 1 Novel pathogenic mutations. (A) Transmission electron microscopy showing the pathognomonic chromatin abnormalities in intermediate erythroblasts for patients as labelled. (B) Chromatograms showing aberrant splicing caused by the ivs21+5G>C variant in patient UPID19. (C) Alignments of Codanin-1 and C15orf41 protein sequences showing evolutionary conservation of amino acids in the regions flanking the missense changes as indicated.

CDA-I. Analysis of patients with CDA-I caused by *CDAN1* mutations suggests a number of the reported missense and in-frame mutations are unlikely to completely abrogate protein function. Supporting this, we find protein levels to be unaffected in primary patient erythroblasts. Finally, we report that both Codanin-1 and C15orf41 are enriched in the nucleolus and that they form a stable complex where C15orf41 appears to be the critical partner.

METHODS

Patient recruitment

Patients were recruited to the Unexplained Anaemia clinic, which is part of the Oxford Molecular Diagnostics Centre (<https://www.oxford-translational-molecular-diagnostics.org.uk/content/unexplained-anaemia>). Patients were assessed and analysed using the Oxford Red Cell Panel Targeted Resequencing strategy as previously described.¹¹

Differentiation of CD34⁺ cells

CD34⁺ cells were isolated (as described in online supplementary methods) and differentiated using a modified version of a published three-phase protocol¹² or using a two-phase liquid culture.^{13 14}

ChIP-seq

ChIP-seq was carried out using ChIP Kit (17–295, Upstate Cell Signaling, now Millipore) following the manufacturer's instructions. See online supplementary methods for full method.

Next generation Capture-C

3C Libraries were generated according to the published protocol.¹⁵ See online supplementary methods for a full description.

Library preparation and sequencing

Library preparation was performed with ILLUMINA'S TSCA v1.5 kit (FC-130–1001) using 250 ng genomic DNA, following manufacturer's instructions. Samples were pooled (average 26 samples) and loaded at 20 pM on a MiSeq using a v3 600-cycle reagent kit sequencing 2×301 paired-end reads (Illumina).

Sanger sequencing

Sequencing was performed on an ABI3730 DNA Analyser (Applied Biosystems, Foster City, California, USA) using 200–500 ng DNA template, BigDye (Applied Biosystems) reaction mix and 3.2 pmol of sequencing primer.

Optical mapping using Bionano Saphyr technology

High molecular weight DNA from patient derived lymphoblastoid cell lines was isolated using the Bionano plug lysis protocol, using the Bionano Prep Blood and Cell Culture DNA isolation kit. Genomic DNA was barcoded using the Bionano Prep DLS labelling kit, in accordance with the manufacturer's instructions. Labelled samples were analysed using the Saphyr device, and optical maps were assembled and structural variants analysis was conducted using Bionano Access software.

RNA isolation and RT-PCR

2×10^6 cells were pelleted and washed with PBS. The pellet was fixed in 500 μ l TRIzol (Invitrogen 15596018) and snap frozen on dry ice before storage at -80°C . RNA was extracted using Direct-zol RNA MiniPrep kit (Zymo Research R2050) following manufacturers' instructions (excepting for the DNase I treatment where the in-column DNA digestion was performed for 30 min instead of 15). Total RNA was quantified using Qubit RNA BR Assay kit (Invitrogen Q10211). For RT-qPCR 1 μ g of RNA was used to generate cDNA using High Capacity cDNA RT Kit (Applied Biosystems 4368814) according to manufacturers' instructions.

Western blots of patient cells and immunofluorescence

2.5×10^6 cells were pelleted, washed with ice cold PBS supplemented with 10% v:v protease inhibitor cocktail (Sigma, P8340) and subjected to subcellular fractionation using a modified published method.¹⁶ See online supplementary methods for full western blot protocol and immunofluorescence methods.

Mutation cluster analysis

The locations of the 41 known missense changes and single amino acid in-frame deletions were mapped onto the Codanin-1 protein. Over the 1227 amino length, each residue has an expected mutation probability of 0.0334. We compared this figure to the observed values using a χ^2 test over a 30 amino acid sliding window using a Bonferroni correction to account for the 1227 tests performed over the entirety of the protein.

Protein structure predictions

Relative solvent accessibility was predicted from the Codanin-1 amino acid sequence using four different methods: NetsurfP-2.0 (ref: <https://www.ncbi.nlm.nih.gov/pubmed/30785653>), I-TASSER (ref: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2849174/>), SPIDER3 (ref: <https://academic.oup.com/bioinformatics/article/33/18/2842/3738544>) and SPIDER3-Single (ref: <https://www.ncbi.nlm.nih.gov/pubmed/30368831>). We compared the relative solvent accessibility of all Codanin-1 residues associated with at least one pathogenic mutation to those other residues with at least one putatively benign mutation in gnomAD v2.1. For in-frame deletions bigger than one residue, the middle residue was used. Statistical significance was assessed using the Wilcoxon rank-sum test.

U-2-OS cells

Conditions for culture, transfection and collection of material for western blotting and immunoprecipitation from U-2-OS cells are given in online supplementary methods. The following primary antibodies were used: Codanin-1 (Abcam ab31236, ab28392), C15orf41 in figure 5A,B (Abcam, ab107352), ASF1, HA-tag (Roche 1867423).

Gel filtration

Size exclusion chromatography of cytosolic and nuclear extracts from U-2-OS cells was performed as previously described.^{7 17} See online supplementary methods for detailed description. The following primary antibodies were used: Codanin-1 (Abcam ab31236 and Abcam ab28392) and C15orf41 in figure 5C,D, online supplementary figure 2C, mouse monoclonal antibody raised in house using 1-159aa of C15orf41.

RESULTS

Novel pathogenic mutations

We screened a cohort of patients with unexplained anaemia, referred to the Oxford Molecular Diagnostics Centre and identified five novel in-frame and missense changes and a loss of function (LOF) frameshift mutation in *CDAN1* and one novel missense change in *C15orf41* in nine patients (figure 1 and table 1). All novel changes are extremely rare or absent from the gnomAD database (allele frequencies listed in table 1). Pathogenicity of novel missense and in-frame changes was confirmed by finding chromatin abnormalities by EM in at least one family member (figure 1A). Patient UPID23, who harbours an eight amino acid deletion, is the exception to this as no EM is available; however, UPID23 showed a clear response to treatment with interferon alpha, supporting the diagnosis of CDA-I (table 1). All missense and in-frame changes affect conserved residues, strongly suggesting that they affect protein function (figure 1C).

Patient UPID32 has EM positive CDA-I (figure 1A) with splenomegaly (~16 cm) but no distal limb abnormalities. There is no reported history of consanguinity and he was not responsive to two prolonged treatment periods with interferon alpha. No pathogenic mutations in the coding regions or intron/exon boundaries in *CDAN1* nor *C15orf41* could be identified using our targeted resequencing strategy¹¹ nor by whole exome sequencing (WES). Analysis of WES data revealed 705 regions of loss of heterozygosity (LOH) in this patient ranging from 1.7 kb to 10.7 Mb and encompassing a total of 740.8 Mb. One ~6 Mb LOH region (chr15: 39 876 412–45 779 810 bp) includes *CDAN1* (chr15: 43 015 760–43 029 417). To assay for a chromosomal rearrangement in this region, we analysed genomic DNA from patient UPID32 by optical mapping (figure 2). This revealed that there are no structural rearrangements within the ~6 Mb region of LOH, excluding this possibility as a cause of the patient's CDA-I. Interestingly, there is a homozygous inversion present between ~30.4 and ~32.7 Mb (figure 2A). This inversion has been reported as benign (ns3955659 in Ref. 18), suggesting that it is a population specific structural variation.

To determine whether the CDA-I in this patient resulted from a *cis*-acting mutation affecting the expression of *CDAN1* or *C15orf41*, we searched for any *cis*-acting regulatory regions using a combination of erythropoietic ATAC-seq,^{19 20} DNaseI-seq²¹ and chromosome conformation capture (3C). We determined chromatin interactions for the *CDAN1* and *C15orf41* promoters using NG Capture-C^{15 22} in cultured human erythroblasts. Neither promoter forms chromatin loops with distal open chromatin elements, suggesting that both genes are entirely promoter driven in intermediate erythroblasts (figure 2B,C), excluding the possibility of a long-range distal regulatory variant causing CDA-I in UPID32. We also captured from the alpha-globin gene promoters in the same experiment, as a positive control and identified the known enhancer interactions (online supplementary figure 3). Analysis of open chromatin showed that both gene promoters are accessible from early haematopoiesis to terminal

Table 1 Clinical information														
Patient ID	Sex	Ethnicity	Hb (g/L)	MCV (fL)	WBC (×10 ⁹ /L)	Platelets (×10 ⁹ /L)	Retics	Ferritin (μg/L)	EM	Gene	Alleles	Protein	Allele frequency	Reference
UPID17	F	Caucasian	105	118	9.3	596	4%	328	N	CDAN1	c.2015C>T c.2681_2682delAG	p.P672L p.E894VfsX108	8.5×10 ⁻⁵ 7.97×10 ⁻⁶	Dgany ³ Novel
UPID19	F	Caucasian	120	96	2.8	289	NA	588	Y	CDAN1	c.2015C>T c.2868+5G>C	p.P672L Ex21 +5G>C	8.5×10 ⁻⁵ 4.00×10 ⁻⁶	Dgany ³ Novel
UPID21	M	Caucasian	145	94	4.5	270	NA	NA	N	CDAN1	c.2015C>T c.3338T>C	p.P672L p.L1113P	8.5×10 ⁻⁵ 1.06×10 ⁻⁵	Dgany ³ Novel
UPID22	F	Caucasian	115	101.8	6.7	419	83×10 ⁹ /L	750	Y	CDAN1	c.2015C>T c.3338T>C	p.P672L p.L1113P	8.5×10 ⁻⁵ 1.06×10 ⁻⁵	Dgany ³ Novel
UPID23	F	Caucasian	138*	104*	NA	NA	NA	NA	N	CDAN1	c.2044C>T c.del2744_2767	p.R682X p.L915_L922del	NA 3.98×10 ⁻⁶	Dgany ³ Novel
UPID26	F	Caucasian	121	103	8.3	185	14%	645	Y	C15orf41	c.467G>A	p.C156Y	NA	Novel
UPID27	F	Caucasian	115	101	7.8	145	15%	369	Y	C15orf41	c.467G>A	p.C156Y	NA	Novel
UPID32†	M	Romany	90	87	4.3	244	64×10 ⁹ /L	NA	Y	NA	NA	NA	NA	NA
UPID33†	M	Caucasian/Indian	87	75.2	5.99	314	2.5%	2145	Y	CDAN1	c.1093G>A c.2261A>G	p.D365N p.Q754R	1.19×10 ⁻⁵ NA	Novel Novel

*These values are postinterferon treatment, preinterferon values were: Hb (g/L) 89 and MCV (fL) 103.

†Patient UPID 33 was 5 years of age at sampling and UPID 32 was 15 years of age at sampling.

EM, electron microscopy; Hb, Haemoglobin; MCV, Mean Corpuscular Volume
; NA, data not available ; WBC, White Blood Cells

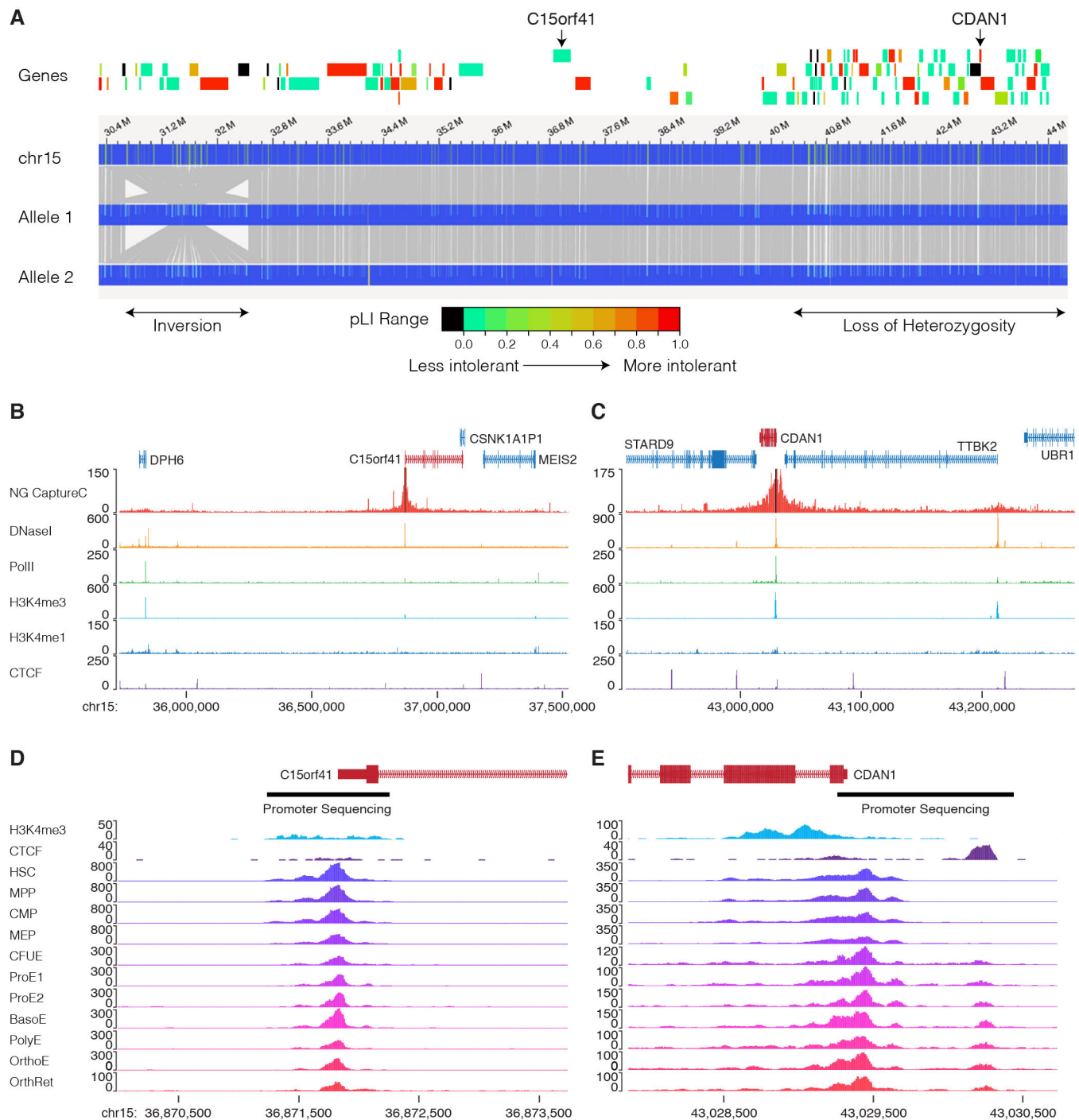


Figure 2 Evidence for a third CDA-I locus. (A) Optical mapping of ~35 Mb on chromosome 15 for patient UPID32. The locations of the inversion and the region of loss of heterozygosity are shown (lower). Protein coding genes are shown and shaded according to their constraint metric (gnomAD) and C15orf41 and CDAN1 are labelled. Optical mapping shows that there are no structural rearrangements over the CDAN1 or C15orf41 loci; however, there is a homozygous inversion between 30.4 Mb and 32.6 Mb. (B,C) Capture-C showing chromatin interactions of the CDAN1 and C15orf41 loci. In each case, the upper track shows Capture-C demonstrating chromatin interaction counts for each captured fragment (black bars). Lower tracks show normalised (RPKM) DNaseI-seq and ChIP seq for Polymerase II (PolII), H3K4me3, H3K4me1 and CTCF. These data show that the promoters of CDAN1 and C15orf41, despite being actively transcribed (as shown by PolII loading and DNaseI hypersensitivity), do not interact with any other DNaseI sensitive sites decorated with the H3K4me1 mark. This suggests that neither of these genes is controlled by distal cisacting regulatory elements in cultured erythroblasts. (D) Chromatin accessibility of the genomic landscape of CDAN1 and C15orf41 during erythroid differentiation (data from Refs. 18 and 19) shows the extent of the gene promoters. The regions of the promoters subjected to Sanger sequencing is shown by black bars. BasoE, basophilic erythroblast; CDA-I, congenital dyserythropoietic anaemia type I; CFUE, colony forming unit erythroid; CMP, common myeloid progenitor; HSC, haematopoietic stem cell; MEP, megakaryocyte erythroid precursor; MPP, multipotent progenitor; OrthoE, orthochromatic erythroblast; OrthRet, orthochromatic/reticulocyte; PolyE, polychromatic erythroblast; ProE1, proerythroblast stage 1; ProE2, proerythroblast stage 2.

erythroid maturation (figure 2). Sanger sequencing of both of these accessible regions found no likely pathogenic changes.

Given the clear phenotype of CDA-I in UPID32 and the absence of coding variants, *cis*-acting distal regulatory sites and structural variants disrupting *CDAN1* and *C15orf41*, it is likely that there is at least one further locus underlying CDA-I, identification of which would be likely to offer insight into the pathogenic mechanism.

CDA-I prevalence

We combined the allele frequencies of all known pathogenic missense and in-frame mutations in *CDAN1* and *C15orf41* together with all LOF mutations in ~140 000 individuals in the gnomAD database (<https://gnomad.broadinstitute.org/>). This shows the incidence of CDA-I caused by compound heterozygosity for two mutations in either gene is 4.83794×10^{-6} or ~4.84 cases per million live births, which is a fivefold increase on the currently accepted incidence. Mutations used are given in online supplementary tables 4 and 5.

This calculation is based on the likelihood of occurrence of compound heterozygosity for two missense mutations, or a missense mutation with a LOF mutation, and assumes that compound heterozygosity for two LOF mutations for either gene would be incompatible with life (see below). These calculations use a neutral model; however, the actual incidence of CDA-I is likely to be higher than this owing to population specific variation in allele frequencies.

Mutation clustering indicates functional domains

CDA-I mutations are present throughout *CDAN1* and *C15orf41*. However, the positioning of mutations likely to lead to LOF, such as truncations and frameshift mutations, is less important for determining protein function as all LOF changes result in a lack of protein resulting from nonsense mediated decay. CDA-I mutations are present throughout Codanin-1 and *C15orf41*, so by focusing on the location of missense and in-frame deletion mutations rather than truncation mutations, important subdomains within proteins can be identified. Even with the report of one novel *C15orf41* missense mutation here, there are a total of six *C15orf41* missense mutations, which is too few for a domain enrichment analysis. Codanin-1 however, has 37 missense and in-frame deletion mutations. This approach shows that six regions are significantly enriched for pathogenic mutations ($p < 0.001$) (figure 3A) and these are therefore more likely to play a role in affecting specific protein functions and hence in the pathogenesis of CDA-I. These six areas of Codanin-1 are 340-398aa, 594-626aa, 648-754aa, 852-910aa, 1014-1071aa and 1101-1142aa (regions A-F in figure 3A). The distribution of missense changes in the general population across the length of the Codanin-1 protein shows a uniform degree of variation (figure 3B), suggesting that there are no specific regions intolerant to missense variation. Therefore, we hypothesised that the six mutational clusters represent functional domains.

Analysis of Codanin-1 missense mutations

No three-dimensional structure is available for Codanin-1, nor are there related structures that can be used for homology modelling. However, we can still use predicted structural properties to investigate the possible molecular mechanism underlying the pathogenic mutations. In particular, sequence-based predictors of solvent accessibility, a measure of residue burial, have been used successfully to differentiate mutations by phenotype in the absence of a structure.²³ It is well known that, in general,

pathogenic mutations are enriched at buried positions within proteins, as mutations in these regions are much more likely to cause a LOF by disrupting protein folding and stability.²⁴ Therefore, mutations at residues with lower predicted relative solvent accessibility values should be more likely to destabilise the protein and cause a LOF.

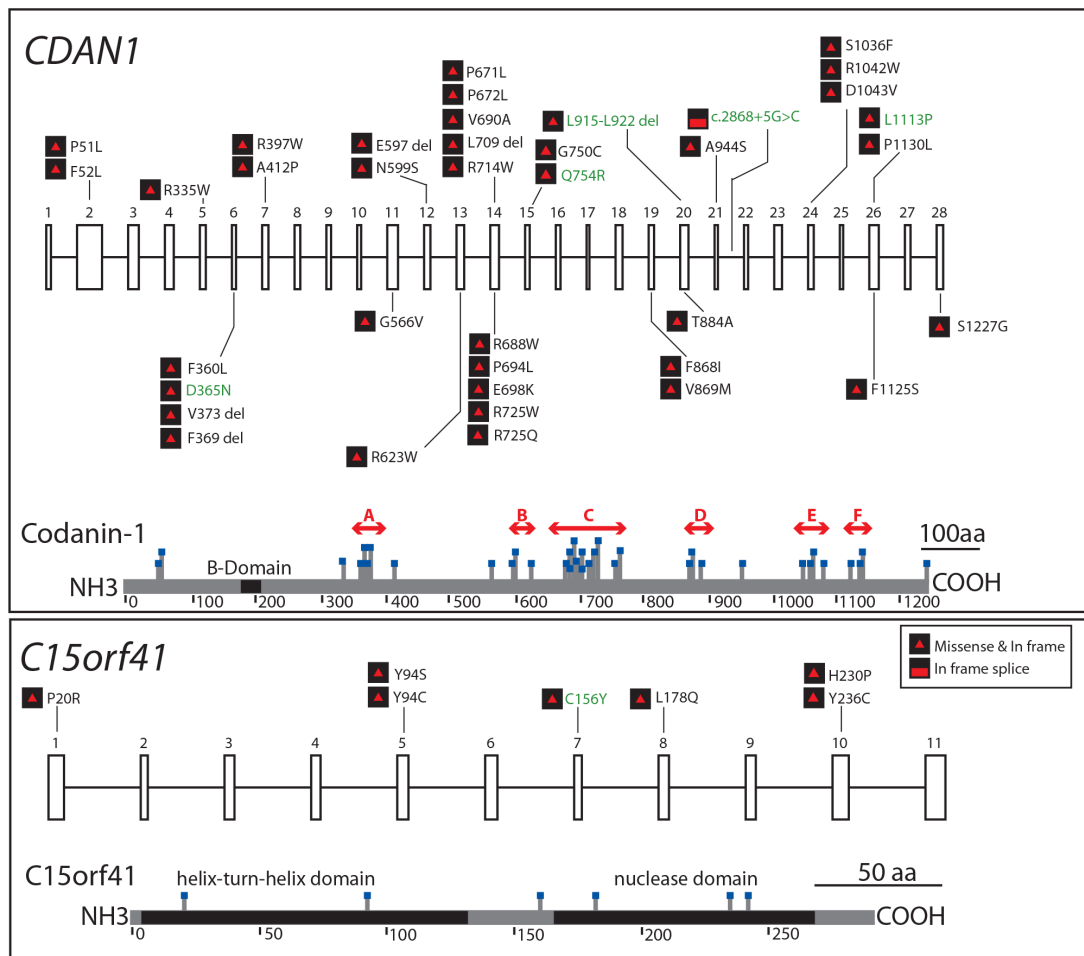
Interestingly, no homozygotes or compound heterozygotes for *CDAN1* LOF mutations have been identified,²⁵ suggesting that Codanin-1 may have a unique function and may be essential during development. This view is supported by evidence from *Cdan1* null mice, which show embryonic lethality prior to embryonic day 6.5.²⁶ Because of these observations, we can deduce that missense mutations found in the homozygous state or in compound heterozygosity with a LOF mutation are unlikely to cause a complete LOF. Following a search of the literature and review of our own data, we identified 13 missense and in-frame deletion mutations present in patients in the homozygous state or in combination with a LOF allele. We can assume these 13 changes are not causing complete LOF and so we termed them 'non-LOF or non-destabilising'. The remaining 24 in-frame and missense *CDAN1* mutations have an unknown effect on the protein and here are termed 'mutations of unknown effect'. The status of the latter group could be determined by functional testing or identification in homozygosity or in heterozygosity with a LOF allele. Mutations in each category are listed in online supplementary information. Interestingly, we find a highly significant tendency for the non-LOF mutations to be more solvent accessible (i.e. less buried) than the mutations of unknown effect, suggesting that at least some of the latter group are likely to cause complete destabilisation of the protein. Mutations in both groups tend to be more buried than the sites of putatively benign variants observed in gnomAD (figure 4A and online supplementary figure 1). While the non-LOF mutations tend to be slightly more buried than the gnomAD variants, there is no statistically significant difference between the groups. This suggests that the non-LOF mutations are either causing weaker destabilisation (i.e. they are hypomorphic) or their damaging effect is due to some other molecular mechanism.

Codanin-1 and C15orf41 expression and stability

We tested the effects of pathogenic mutations on Codanin-1 and *C15orf41* by two colour near-infrared quantitative western blot using protein extracted from erythroblasts cultured from the peripheral blood of healthy individuals and patients with *CDAN1* and *C15orf41* CDA-I using a well characterised *in vitro* culture system¹² (figure 4B). To detect Codanin-1, we validated a polyclonal antibody from Bethyl Laboratories (catalogue number A304-951A) by showing it cross-reacts with Codanin-1 conjugated to mCherry when overexpressed in HEK293T cells (online supplementary figure 2A). To detect *C15orf41*, we validated a polyclonal antibody supplied by Cusabio (catalogue number CSBPA897474LA01HU) by expressing a FLAG-tagged version of *C15orf41* in HEK293T cells and showing colocalisation of the *C15orf41* antibody with the FLAG antibody (Sigma F1804) by western blotting using the immunoprecipitated cell lysate (online supplementary figure 2B).

Quantification of Codanin-1 and *C15orf41* in three healthy control samples (NCO27, NCO28 and NCO29) showed protein levels decrease during terminal erythroid differentiation (figure 4C). Interestingly, the Codanin-1 and *C15orf41* protein levels in erythroblasts cultured from two patients with CDA-I with *CDAN1* mutations (UPID 15: p.F369del;D1043V and UPID 20: p.P672L; V993GfsTer13) and two with *C15orf41* mutations

A



B

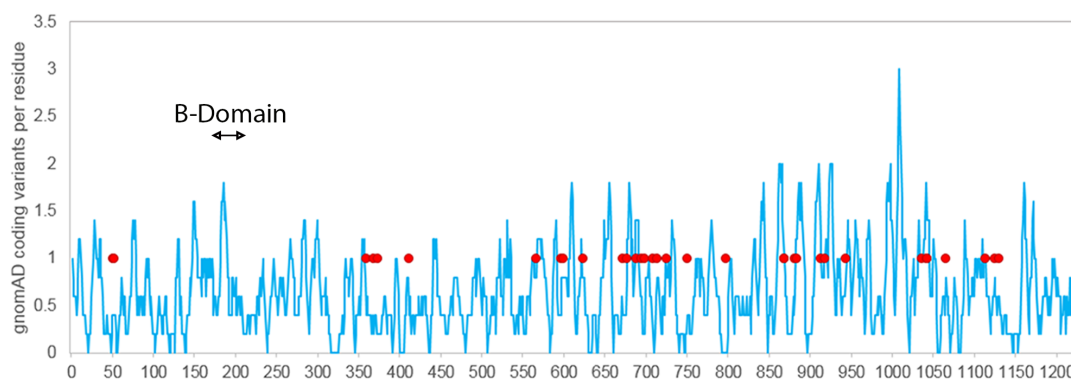


Figure 3 (A) Missense and in frame pathogenic variants in *CDAN1* and *C15orf41* causing CDA-I. Published variants are shown in black text and the five novel missense and in-frame changes reported here are in green. Exons in each gene are numbered. Missense and in-frame changes are denoted by protein position. The in-frame splicing change is shown according to the closest exon. Lower sections of each box show a representation of each protein with missense and in-frame deletions marked. For Codanin-1, we have identified mutation clusters A–F (red arrows), highlighting likely functional domains in the protein. (B) The distribution of missense variants in the general population in Codanin-1 showing missense and in-frame indels per residue, smoothed over a 10 residue window across the length of the protein (blue line). Pathogenic mutations are shown by closed red circles and the location of the putative B-domain is indicated. CDA-I, congenital dyserythropoietic anaemia type I.

(UPID25: p.Y94C;Y94C and UPID26: p.C156Y;C156Y) showed no reduction in either protein compared with the healthy control samples. The D1043V and P672L mutations are predicted to be non-LOF, and F369del is categorised as a mutation of unknown effect. We also measured *CDAN1* and *C15orf41* mRNA levels,

which were comparable to those of healthy controls (figure 4D), showing that increased transcription does not underlie the constant protein levels in these cases. To ensure patient and control samples were stage matched, erythroid cell counts at sequential stages of differentiation in culture showed similar

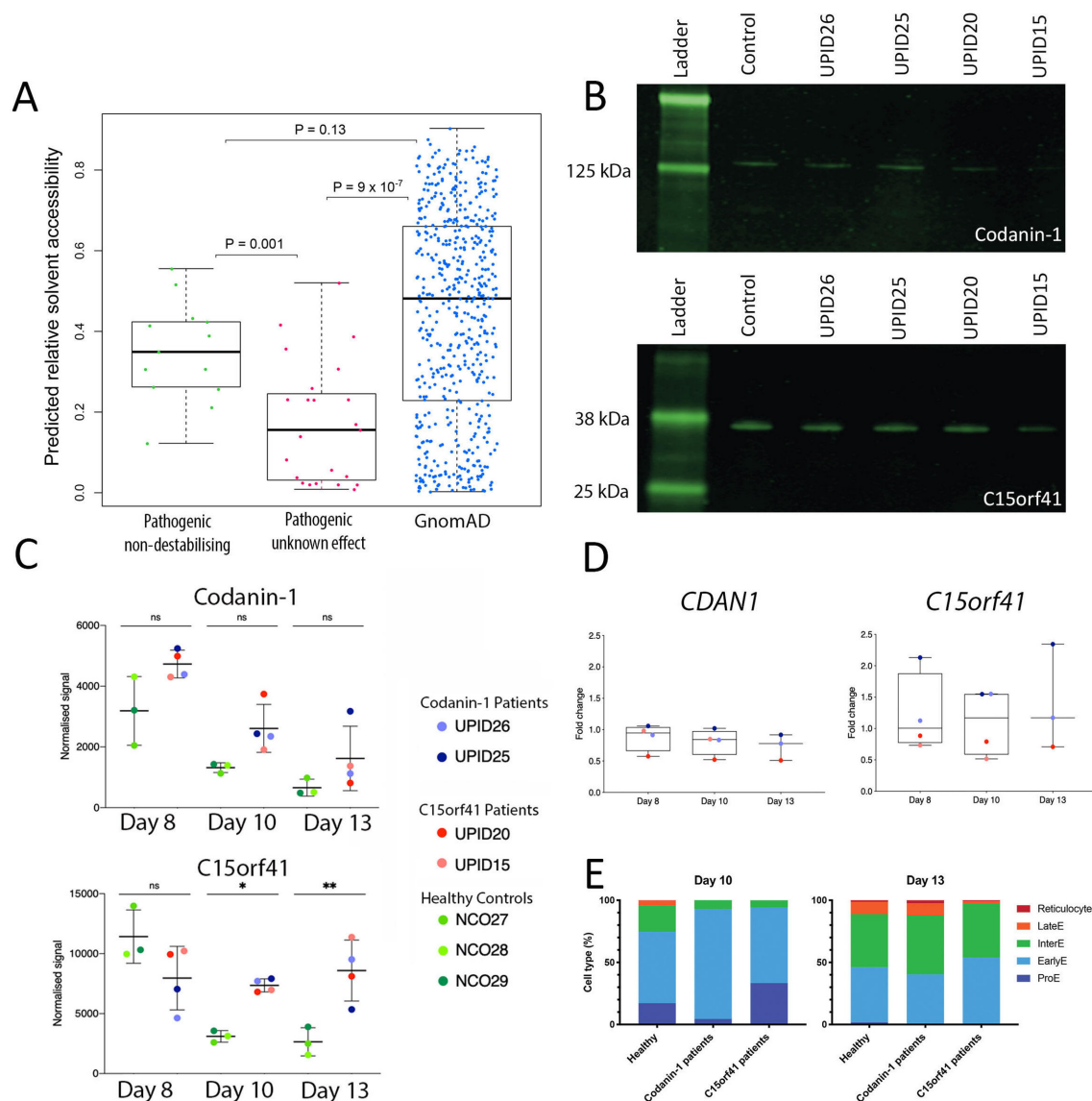


Figure 4 Effect of missense mutations on Codanin-1. (A) Pathogenic Codanin-1 mutations observed homozygously, or compound heterozygously with a loss of function mutation ('not LOF'), are predicted to be significantly more accessible to solvent (ie, less buried) by the programme NetSurfP-2.0 than other pathogenic mutations ('possibly LOF'), but are not significantly different from putatively benign variants observed in the human population (gnomAD). (B) Representative image of quantitative western blots using ex vivo cultured patient erythroblasts detecting Codanin-1 (upper) and C15orf41 (lower). (C) Quantification of western blots shown in C showing normalised values of Codanin-1 and C15orf41 proteins compared with erythroblasts cultured from three healthy individuals. Bars show the mean \pm SD. (D) Box and whiskers plot showing median, quartiles and outliers of mRNA levels of *CDAN1* and *C15orf41* in patients with CDA-I plotted relative to the levels seen in healthy controls (patient samples colour coded as in C). (E) Erythroid cell counts showing that altered maturation rates do not explain the lack of protein degradation seen in this experiment. * $P < 0.05$, ** $p < 0.005$, significance determined by one-way ANOVA with Sidak's multiple comparison test. ANOVA, analysis of variance; LOF, loss of function; ns, not significant.

numbers of morphologically identified erythroid cells were present at day 10 (intermediate erythroblasts) and day 13 (late stage erythroblasts) in cells from patients and healthy donors (figure 4E). Taken together, these data show that there is no overall destabilisation of Codanin-1 nor C15orf41 by the mutations we tested, confirming the prediction that D1043V and P672L changes do not destabilise the entire protein.

Protein interaction and nucleolar enrichment

Because mutations in *CDAN1* and *C15orf41* both lead to the specific chromatin abnormalities seen in CDA-I, we hypothesised both proteins function in the same pathway. To test whether the proteins interact, we performed coimmunoprecipitation experiments. Using cell lines expressing FLAG-HA

tagged Codanin-1, we found that the known interactor, ASF1 as well as C15orf41 copurified with FLAG-HA-Codanin-1 (figure 5A). Codanin-1 also copurified with FLAG-HA tagged C15orf41, confirming the interaction (figure 5B). Gel-filtration analysis revealed that the two proteins coelute in both nuclear and cytosolic extracts, implying that the majority of Codanin-1 and C15orf41 are found in a complex together. This suggested that the two proteins might form an obligate complex and we therefore addressed whether their stability might be interdependent. Whereas depletion of C15orf41 did not affect the level of Codanin-1, removal of Codanin-1 by several independent and validated siRNAs⁷ resulted in a concomitant reduction in C15orf41 levels. Immunofluorescence of the two proteins in

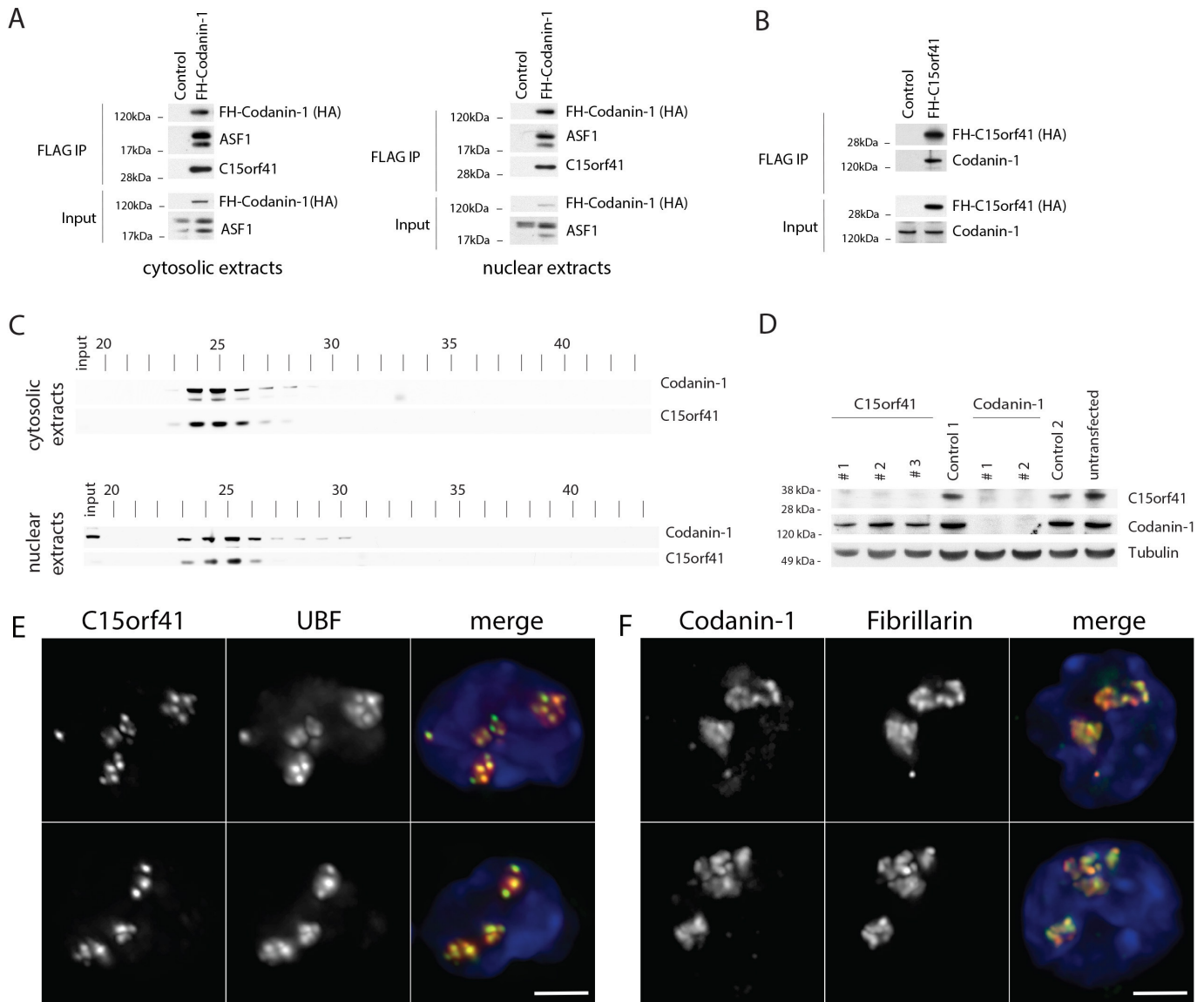


Figure 5 Codanin-1 and C15orf41 form a complex and are enriched in the nucleolus. (A) FLAG-HA Codanin-1 (FH-Codanin-1) was immunoprecipitated from cytosolic and nuclear extracts and analysed by western blot using HA, ASF1 and C15orf41 antibodies. (B) FLAG-HA-C15orf41 (FH-C15orf41) was immunoprecipitated from whole cell lysates and analysed by western blot using HA and Codanin-1 antibodies. (C) Cytosolic and nuclear extracts from U-2-OS cells were subjected to size-exclusion chromatography and analysed by western blot. (D) U-2-OS cells transfected with the indicated siRNAs were analysed by western blot (see also online supplementary figure 2C). (E,F) C15orf41 and Codanin-1 are enriched in the nucleolus in erythroblasts as shown by overlap with the nucleolar proteins UBF and fibrillarin (two representative examples of each are shown), nuclei are stained with DAPI (blue). Scale bars are 2 μ m.

erythroblasts shows they are pan-cellular but both are enriched in the nucleolus (figure 5E,F).

DISCUSSION

In this work, we report six mutations in *CDAN1* and one in *C15orf41* causative of CDA-I, thus increasing the diagnostic range of this disease. In cases with novel missense mutations, we show that patients have the pathognomonic chromatin abnormalities indicative of CDA-I, critical for demonstrating pathogenicity of novel missense and in-frame mutations. This increases the number of patients who can be diagnosed without recourse to EM examination of bone marrow biopsy, which is an invasive procedure. This is particularly helpful as EM is unavailable at the majority of referral centres.

Although causative mutations are spread throughout *CDAN1*, when only the locations of pathogenic missense and in-frame mutations in the Codanin-1 protein are considered, it becomes clear these mutations occur in six distinct domains. Because complete loss of Codanin-1 is incompatible with life,²⁶ any missense mutation that occurs homozygously, or in tandem with a LOF mutation, is unlikely to cause a complete LOF. When all reported patients are considered, 13 missense mutations are likely to impair protein function rather than obliterate it. Our analysis of solvent accessibility shows that these 13 mutations tend to be less buried than mutations of unknown effect. This suggests that the 13 non-LOF mutations are less likely to destabilise the entire protein and we show this is the case in erythroblasts from two patients with *CDAN1* CDA-I. It is interesting that

knockdown of Codanin-1 destabilises C15orf41, yet in patient cells, we find that protein levels remain unchanged for the mutations tested. This further suggests that a specific interaction is affected in patients with CDA-I. This hypothesis is supported by our previous data showing that the R714W missense mutation, present in the largest mutation cluster in Codanin-1 (cluster C), disrupts the interaction with ASF1.⁷ It is unclear which other protein interactions are affected by mutations in the six domains; however, identification of further interacting proteins would allow this to be tested. Given the enrichment of mutation clusters in the C-terminal region of the protein, it may be this region that interacts with C15orf41, although this will need to be directly tested.

We find that the two mutations in C15orf41 we tested do not destabilise this protein. Since both mutations we tested were present in the homozygous state, this suggests that C15orf41 is an essential protein.

Codanin-1 is cell cycle regulated⁹ and it is possible that some of the mutations affect post-translational modifications that regulate Codanin-1 levels. Analysis of synchronised cell populations may be able to address this question. Expression of *CDAN1* and *C15orf41* has been shown to be tightly correlated.²⁷ In our study, we observe a similar trend as both genes are downregulated in tandem during terminal erythroid differentiation. Russo and colleagues²⁷ also report no change in localisation of mutant proteins, further suggesting that mutations affect specific interactions.

We report enrichment of both proteins in the nucleolus. This observation has the potential to provide insight into the erythroid specific nature of CDA-I as mutations affecting ribosome biogenesis are known to give rise to tissue specific abnormalities, including severely impaired erythropoiesis in diseases such as Diamond Blackfan Anaemia and Shwachman-Diamond Syndrome.²⁸ Further studies will be required to establish whether ribosomes are affected in CDA-I. Insight into the biochemical activity of C15orf41 would greatly help with understanding the role of these proteins in the nucleolus and CDA-I disease pathology.

Identification of a third locus for CDA-I would also offer insight into disease pathology and broaden the diagnostic range. Here, we show that, unless there are distal cis-acting elements at earlier stages of erythroid maturation, it is extremely unlikely that mutations in distal regulatory elements or structural rearrangements cause CDA-I in UPID32. This observation argues for the presence of at least one further locus underlying this disease. We hope excluding the possibility of distal regulatory elements and defining the extent of promoter sequences for both causative genes is of use to other researchers.

As with many rare disorders, establishing an accurate estimate of incidence and prevalence of CDA-I is difficult. Over 300 cases have been reported.¹⁰ Most are sporadic cases from diverse regions such as Western Europe, North Africa and Asia,²⁹ while some series are accounted for by a founder effect, particularly in the Middle East.³⁰ Currently, the reported incidence of CDA-I suggests a frequency of ~1:1 000 000 live births⁶ although estimates as low as 0.24 cases per million live births have been made based on European epidemiological data.²⁵ However, it has been suggested that CDA-I may be more common than previously estimated.³¹ In support of this, analysis of allele frequencies in this work suggests that CDA-I may be present at a frequency of 1 per ~207 000 live births suggesting that the majority of CDA-I cases remain undiagnosed. Because treatment with interferon alpha is effective in many cases of CDA-I, improving ascertainment will be important in best treating this disease.

Here, we have expanded the diagnostic range for CDA-I and reassessed its prevalence. We prove the hypothesis that the two CDA-I causative proteins, Codanin-1 and C15orf41, are involved in a common pathway by showing that they interact and that Codanin-1 stabilises C15orf41. We postulate that mutation hotspots we have identified across Codanin-1 represent functional domains, some of which mediate Codanin-1's interaction with C15orf41. Further, our finding of enrichment of these proteins in nucleoli is an important step towards elucidating the erythroid-specific nature of the disease. Taken together, these results will improve CDA-I patient ascertainment and allow future focus to further understand the mechanisms behind this disease as well as the normal process of erythropoiesis.

Data accessibility: data accessibility

Erythroid sequencing data generated for this work are deposited with the GEO archive (GSE125753). Previously published open-chromatin data sets (GSE86393, GSE75384, GSE115684) were used.^{19 20 31}

Author affiliations

¹MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

²Department of Haematology, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

³NIHR Oxford Biomedical Research Centre and BRC/NHS Translational Molecular Diagnostics Centre, John Radcliffe Hospital, Oxford, UK

⁴MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

⁵Biotech Research and Innovation Centre (BRIC), Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

⁶National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark

⁷Eli Lilly Denmark, Herlev, Denmark

⁸Department of Oncology, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

⁹Sir William Dunn School of Pathology, University of Oxford, Oxford, UK

¹⁰Molecular Haematology Laboratory, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

¹¹Haematology Department, Manchester University NHS Foundation Trust, Manchester, UK

¹²Department of Women's and Children's Health, Uppsala University and Uppsala University Children's Hospital, Uppsala, Sweden

¹³Centre for Health Development, Västmanland Region, Uppsala University, Uppsala, Sweden

¹⁴Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden

¹⁵Department of Haematology, NHS Trust Greater Glasgow and Clyde, Glasgow, UK

¹⁶Department of Haematology, St James's University Hospital, Leeds, UK

¹⁷Pediatric Hematology-Oncology Laboratory, Lausanne University Hospital and University of Lausanne, Lausanne, VD, Switzerland

¹⁸The Novo Nordisk Center for Protein Research (CPR), Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

Twitter Nandini Sadasivam @nandinikori16

Contributors CB, VJB, DRH and NBR conceived the study and wrote the manuscript. A-AO cultured erythroblasts and performed quantitative western blots in figure 4 and antibody validation. CS cultured erythroblasts and analysed patient mutation data. JAM performed analysis of missense and in frame mutations. KL and KA performed immunoprecipitation and western blots in figure 5. NR performed capture-C from the gene promoters. DJD analysed capture-C, ChIP and ATAC data and generated figure 2. JB cultured patient cells, performed immunofluorescence in figure 5 and commented on the manuscript. SB performed C15orf41 and Codanin-1 antibody validation. EJ performed EM studies in figure 1. BX performed optical mapping studies to identify structural rearrangements in figure 2. MP and RH analysed patient DNA and identified pathogenic mutations shown in figure 1. KR, PF, JM, ES, NS, LMI, QAH, NBR and RR ascertained the CDA-I patients analysed in this paper, provided samples and analysed clinical data. JRH, RG, AG and PJMcH provided supervision, resources and analysed data and commented on the manuscript.

Funding This work was supported by MRC (MC_uu_12009), the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre Haematology Theme at Oxford University Hospitals NHS Trust and University of Oxford and the charities Blood Buddies, and the Congenital Anaemia Network. D.J.D. and J.H.

were funded by a Wellcome Trust Strategic Award (106130/Z/14/Z). EM work was undertaken at the DunnSchool EM Facility. J.M. is supported by an MRC Career Development Award(MR/M02122X/1) and is a Lister Institute Research Prize Fellow.

Competing interests JRH is a founder and shareholder of Nucleome Therapeutics.

Patient consent for publication Not required.

Ethics approval This study was approved by the Wales Research Ethics Committee (REC5) (13/WA/0371) with written consent from patients and/or parents.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Erythroid sequencing data generated for this work are deposited with the GEO archive (GSE125753). Previously published open-chromatin data sets (GSE86393, GSE75384, GSE115684) were also used.

ORCID iDs

Karin Lauschke <http://orcid.org/0000-0001-8016-7190>

Nigel Roberts <http://orcid.org/0000-0002-5091-3604>

Christian Babbs <http://orcid.org/0000-0002-1898-5878>

REFERENCES

- Heimpel H, Wendt F. Congenital dyserythropoietic anemia with karyorrhexis and multinuclearity of erythroblasts. *Helv Med Acta* 1968;34:103–15.
- Wickramasinghe SN, Wood WG. Advances in the understanding of the congenital dyserythropoietic anaemias. *Br J Haematol* 2005;131:431–46.
- Dgany O, Avidan N, Delaunay J, Krasnov T, Shalmon L, Shalev H, Eidelitz-Markus T, Kapelushnik J, Cattani D, Pariente A, Tulliez M, Crétien A, Schischmanoff P-O, Iolascon A, Fibach E, Koren A, Rössler J, Le Merrer M, Yaniv I, Zaizov R, Ben-Asher E, Olender T, Lancet D, Beckmann JS, Tamary H. Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1. *Am J Hum Genet* 2002;71:1467–74.
- Babbs C, Roberts NA, Sanchez-Pulido L, McGowan SJ, Ahmed MR, Brown JM, Sabry MA, Bentley DR, McVean GA, Donnelly P, Gileadi O, Ponting CP, Higgs DR, Buckle VJ, WGS500 Consortium. Homozygous mutations in a predicted endonuclease are a novel cause of congenital dyserythropoietic anemia type I. *Haematologica* 2013;98:1383–7.
- Heimpel H, Forteza-Vila J, Queisser W, Spiertz E. Electron and light microscopic study of the erythroblasts of patients with congenital dyserythropoietic anemia. *Blood* 1971;37:299–310.
- Roy NBA, Babbs C. The pathogenesis, diagnosis and management of congenital dyserythropoietic anaemia type I. *Br J Haematol* 2019;185:436–49.
- Ask K, Jasencakova Z, Menard P, Feng Y, Almouzni G, Groth A. Codanin-1, mutated in the anaemic disease CDAl, regulates ASF1 function in S-phase histone supply. *Embo J* 2012;31:2013–23.
- Tamary H, Marcoux N, Noy-Lotan S, Yaniv I, Dgany O. Codanin-1, the product of the gene mutated in congenital dyserythropoietic anemia type I (Cda I), binds to histone chaperone Asf1a and inhibits its nucleosome assembly activity. *Blood* 2010;116:1004–2.
- Noy-Lotan S, Dgany O, Lahmi R, Marcoux N, Krasnov T, Yissachar N, Ginsberg D, Motro B, Resnitzky P, Yaniv I, Kupfer GM, Tamary H. Codanin-1, the protein encoded by the gene mutated in congenital dyserythropoietic anemia type I (CDAN1), is cell cycle-regulated. *Haematologica* 2009;94:629–37.
- Iolascon A, Heimpel H, Wahlin A, Tamary H. Congenital dyserythropoietic anaemias: molecular insights and diagnostic approach. *Blood* 2013;122:2162–6.
- Roy NBA, Wilson EA, Henderson S, Wray K, Babbs C, Okoli S, Atoyebi W, Mixon A, Cahill MR, Carey P, Cullis J, Curtin J, Dreau H, Ferguson DJP, Gibson B, Hall G, Mason J, Morgan M, Proven M, Qureshi A, Sanchez Garcia J, Sirachainan N, Teo J, Tedgård U, Higgs D, Roberts D, Roberts I, Schuh A, Garcia JS, Roberts I. A novel 33-Gene targeted resequencing panel provides accurate, clinical-grade diagnosis and improves patient management for rare inherited anaemias. *Br J Haematol* 2016;175:318–30.
- Scott C, Downes DJ, Brown JM, Babbs C, Olijnik AA, Gosden M, Beagrie R, Schwessinger R, Fisher CA, Rose A, Ferguson DJP, Johnson E, Hill QA, Okoli S, Renella R, Ryan K, Brand M, Hughes JR, Roy NR DRH, Buckle VJ. Modelling erythropoiesis in congenital dyserythropoietic anaemia type I (CDA-I). *bioRxiv* 2019.
- Pope SH, Fibach E, Sun J, Chin K, Rodgers GP. Two-Phase liquid culture system models normal human adult erythropoiesis at the molecular level. *Eur J Haematol* 2000;64:292–303.
- Fibach E, Manor D, Oppenheim A, Rachmilewitz EA. Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood* 1989;73:100–3.
- Davies JOJ, Telenius JM, McGowan SJ, Roberts NA, Taylor S, Higgs DR, Hughes JR. Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Methods* 2016;13:74–80.
- Baghirova S, Hughes BG, Hendzel MJ, Schulz R. Sequential fractionation and isolation of subcellular proteins from tissue or cultured cells. *MethodsX* 2015;2:440–5.
- Groth A, Ray-Gallet D, Quivy J-P, Lukas J, Bartek J, Almouzni G. Human ASF1 regulates the flow of S phase histones during replicational stress. *Mol Cell* 2005;17:301–11.
- Levy-Sakin M, Pastor S, Mostovoy Y, Li L, Leung AKY, McCaffrey J, Young E, Lam ET, Hastie AR, Wong KHY, Chung CYL, Ma W, Sibert J, Rajagopalan R, Jin N, Chow EYC, Chu C, Poon A, Lin C, Naguib A, Wang W-P, Cao H, Chan T-F, Yip KY, Xiao M, Kwok P-Y. Genome maps across 26 human populations reveal population-specific patterns of structural variation. *Nat Commun* 2019;10:1025.
- Ludwig LS, Lareau CA, Bao EL, Nandakumar SK, Muus C, Ulirsch JC, Chowdhary K, Buenrostro JD, Mohandas N, An X, Aryee MJ, Regev A, Sankaran VG. Transcriptional states and chromatin accessibility underlying human erythropoiesis. *Cell Rep* 2019;27:3228–40.
- Corces MR, Buenrostro JD, Wu B, Greenside PG, Chan SM, Koenig JL, Snyder MP, Pritchard JK, Kundaje A, Greenleaf WJ, Majeti R, Chang HY. Lineage-Specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet* 2016;48:1193–203.
- Schwessinger R, Suci MC, McGowan SJ, Telenius J, Taylor S, Higgs DR, Hughes JR. Sasquatch: predicting the impact of regulatory SNPs on transcription factor binding from cell- and tissue-specific DNase footprints. *Genome Res* 2017;27:1730–42.
- Telenius JM, Downes DJ, Sergeantm, Oudelaar AM, McGowan S, Kerry J, Hanssen LLP, Schwessinger R, Eijsbouts CQ, Davies JOJ, Taylor S, Hughes JR. CaptureCompendium: a comprehensive toolkit for 3C analysis. *bioRxiv* 2020.
- Shaw ND, Brand H, Kupchinsky ZA, Bengani H, Plummer L, Jones TI, Erdin S, Williamson KA, Rainger J, Stortchevoi A, Samocha K, Curral BB, Dunican DS, Collins RL, Willer JR, Lek A, Lek M, Nassan M, Pereira S, Kammin T, Lucente D, Silva A, Seabra CM, Chiang C, An Y, Ansari M, Rainger JK, Joss S, Smith JC, Lippincott MF, Singh SS, Patel N, Jing JW, Law JR, Ferraro N, Verloes A, Rauch A, Steindl K, Zweier M, Scheer I, Sato D, Okamoto N, Jacobsen C, Tryggvstad J, Chernauek S, Schimmenti LA, Brasseur B, Cesaretti C, Garcia-Ortiz JE, Buitrago TP, Silva OP, Hoffman JD, Mühlbauer W, Ruprecht KW, Loeys BL, Shino M, Kaindl AM, Cho C-H, Morton CC, Meehan RR, van Heyningen V, Liao EC, Balasubramanian R, Hall JE, Seminara SB, Macarthur D, Moore SA, Yoshiura K-I, Gusella JF, Marsh JA, Graham JM, Lin AE, Katsanis N, Jones PL, Crowley WF, Davis EE, FitzPatrick DR, Talkowski ME. Smchd1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet* 2017;49:238–48.
- Yue P, Li Z, Mout J. Loss of protein structure stability as a major causative factor in monogenic disease. *J Mol Biol* 2005;353:459–73.
- Iolascon A, Esposito MR, Russo R. Clinical aspects and pathogenesis of congenital dyserythropoietic anaemias: from morphology to molecular approach. *Haematologica* 2012;97:1786–94.
- Renella R, Roberts NA, Brown JM, De Gobbi M, Bird LE, Hassanali T, Sharpe JA, Sloane-Stanley J, Ferguson DJP, Cordell J, Buckle VJ, Higgs DR, Wood WG. Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1[alpha] localization in erythroblasts. *Blood* 2011;117:6928–38.
- Russo R, Marra R, Andolfo I, De Rosa G, Rosato BE, Manna F, Gambale A, Raia M, Unal S, Barella S, Iolascon A. Characterization of two cases of congenital dyserythropoietic anemia type I shed light on the uncharacterized C15orf41 protein. *Front Physiol* 2019;10:621.
- Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* 2010;115:3196–205.
- Iolascon A, Delaunay J. Close to unraveling the secrets of congenital dyserythropoietic anemia types I and II. *Haematologica* 2009;94:599–602.
- Tamary H, Dgany O, Proust A, Krasnov T, Avidan N, Eidelitz-Markus T, Tchernia G, Geneviève D, Cormier-Daire V, Bader-Meunier B, Ferrero-Vacher C, Munzer M, Gruppo R, Fibach E, Konen O, Yaniv I, Delaunay J. Clinical and molecular variability in congenital dyserythropoietic anaemia type I. *Br J Haematol* 2005;130:628–34.
- Renella R, Wood WG. The congenital dyserythropoietic anaemias. *Hematol Oncol Clin North Am* 2009;23:283–306.