Classification: BIOLOGICAL SCIENCES / Medical Sciences

Pseudouridylation defect due to *DKC1* and *NOP10* mutations cause nephrotic syndrome with cataracts, hearing impairment and enterocolitis

Eszter Balogh[a,b,1], Jennifer C. Chandler[c,1], Máté Varga[a,d,1], Mona Tahoun[e,1], Dóra K. Menyhárd[f], Gusztáv Schay[g], Tomas Goncalves[h], Renáta Hamar[d], Regina Légrádi[a], Ákos Szekeres[b], Olivier Gribouval[i], Robert Kleta[j], Horia Stanescu[j], Detlef Bockenhauer[j], Andrea Kerti[a,b], Hywel Williams[k], Veronica Kinsler[l], Wei-Li Di[m], David Curtis[n], Maria Kolatsi-Joannou[c], Hafsa Hammid[c], Anna Szőcs[o], Kristóf Perczel[a], Erika Maka[p], Gergely Toldi[b], Florentina Sava[a], Christelle Arrondel[i], Magdolna Kardos[q], Attila Fintha[q], Ahmed Hossain[r], Felipe D'Arco[s], Mario Kaliakatsos[t], Jutta Koeglmeier[u], William Mifsud[v], Mariya Moosajee[w], Ana Faro[x], Eszter Jávorszky[a,b], Gábor Rudas[o], Marwa H. Saied[e], Salah Marzouk[e], Kata Kelen[b], Judit Götze[b], George Reusz[b], Tivadar Tulassay[b], François Dragon[r], Géraldine Mollet[i], Susanne Motameny[y], Holger Thiele[y], Guillaume Dorval[i], Peter Nürnberg[y], András Perczel[f], Attila J Szabó[b,z], David A. Long[c], Kazunori Tomita[h,aa], Corinne Antignac[i,bb], Aoife M. Waters[c,2], Kálmán Tory[a,b,2]

[a] MTA-SE Lendület Nephrogenetic Laboratory, Budapest, Hungary, HU 1083

[b] I[st] Department of Pediatrics, Semmelweis University, Budapest, Hungary, HU 1083

[c] Developmental Biology and Cancer Programme, UCL Great Ormond Street Institute of Child Health, London, UK WC1N 1EH

[d] Department of Genetics, ELTE Eötvös Loránd University, Budapest, Hungary, HU 1117

[e] Clinical and Chemical Pathology Department, Faculty of Medicine Alexandria University, Egypt, EG 21500

[f] MTA-ELTE Protein Modeling Research Group and Laboratory of Structural Chemistry and Biology, Eötvös Loránd University, Budapest, Hungary, HU 1117

[g] Department of Biophysics and Radiation Biology, Semmelweis University, Budapest, Hungary, HU 1094

[h] Chromosome Maintenance Research Group, UCL Cancer Institute, London, UK WC1E 6DD

[i] Laboratory of Hereditary Kidney Diseases, Imagine Institute, INSERM, UMR 1163, Université de Paris, Paris, France, FR 75015

[j] Division of Medicine, Royal Free Hospital, Rowland Hill Street, London, UK NW3 2QG

[k] GOSgene, Experimental and Personalised Medicine, UCL Great Ormond Street Institute of Child Health, London, UK WC1N 1EH

[l] Genetics & Genomic Medicine, UCL Great Ormond Street Institute of Child Health, London, UK WC1N 1EH

[m] Infection, Immunity, Inflammatory & Physiological Medicine, UCL Great Ormond Street Institute of Child Health, London, UK WC1N 1EH

[n] UCL Genetics Institute, Darwin Building, Gower Street, London, UK WC1E 6AD

[o] Magnetic Resonance Imaging Research Center, Semmelweis University, Budapest, Hungary, HU 1083

[p] Department of Ophthalmology, Semmelweis University, Budapest, Hungary, HU 1085

[q] 2nd Department of Pathology, Semmelweis University, Budapest, Hungary, HU 1091

[r] Département des sciences biologiques and Centre d'excellence en recherche sur les maladies orphelines, Université du Québec à Montréal, Montréal, Québec, Canada, CA QC, H2X 3Y7

[s] Department of Neuro-radiology, Great Ormond Street Hospital, London, UK WC1N 3JH

[t] Department of Neurology, Great Ormond Street Hospital, London, UK WC1N 3JH

[u] Department of Gastroenterology, Great Ormond Street Hospital, London, UK WC1N 3JH

[v] Department of Histopathology, Great Ormond Street Hospital, London, UK WC1N 3JH

[w] Institute of Ophthalmology, UCL, London, UK EC1V 9EL

[x] Department of Cell and Developmental Biology, Division of Biosciences, University College London, London, UK WC1E 6BT

[y] Cologne Center for Genomics and Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany, DE 50931

[z] MTA-SE Pediatrics and Nephrology Research Group, Budapest, Hungary, HU 1083

[aa] Department of Life Sciences, Brunel University London, UK UB8 3PH

[bb] Assistance Publique – Hôpitaux de Paris, Hôpital Necker-Enfants Malades, Service de Génétique moléculaire, Paris, France, FR 75015

[1] These authors contributed equally to this work

[2] Co-last authors

Correspondence:

Máté Varga, mvarga@ttk.elte.hu, +36 30 5420541

Department of Genetics, Eötvös Loránd University, Pázmány Péter sétány 1/C, Budapest, 1117, Hungary

Aoife Waters, aoife.waters@ucl.ac.uk; +44 20 7905 2206

Developmental Biology and Cancer Programme, UCL Great Ormond Street Institute of Child Health, London, WC1N 1EH, UK

Kálmán Tory, tory.kalman@med.semmelweis-univ.hu, +36 20 8258166

I[st] Department of Pediatrics, Semmelweis University, Bókay J. u. 53., Budapest, 1083, Hungary

Abstract

RNA modifications play a fundamental role in cellular function. Pseudouridylation, the most abundant RNA modification, is catalysed by the H/ACA small ribonucleoprotein (snoRNP) complex that shares four core proteins, dyskerin (DKC1), NOP10, NHP2 and GAR1. Mutations in *DKC1*, *NOP10* or *NHP2* cause Dyskeratosis Congenita (DC), a disorder characterized by telomere attrition. Here we report a novel phenotype comprising nephrotic syndrome, cataracts, sensorineural deafness, enterocolitis and early lethality in two pedigrees; males with *DKC1* p.Glu206Lys and in two children with homozygous *NOP10* p.Thr16Met. Females with heterozygous *DKC1* p.Glu206Lys developed cataracts and sensorineural deafness, but nephrotic syndrome in only one case of skewed X-inactivation. We found telomere attrition in both pedigrees but no mucocutaneous abnormalities suggestive of DC. Both mutations fall at the dyskerin-NOP10 binding interface in a region distinct from those implicated in DC, impair the dyskerin-NOP10 interaction and disrupt the catalytic pseudouridylation site. Accordingly, we found reduced pseudouridine levels in the rRNA of the patients. Zebrafish *dkc1* mutants recapitulate the human phenotype and show reduced 18S pseudouridylation, ribosomal dysregulation and a cell-cycle defect in the absence of telomere attrition. We therefore propose that this novel human disorder is the consequence of defective snoRNP pseudouridylation and ribosomal dysfunction.

Significance Statement

Isomerization of uridine to pseudouridine is the most abundant RNA modification in eukaryotes. In ribosomal (r)RNA, this process of pseudouridylation is catalyzed by a ribonucleoprotein complex. Mutations of this complex were formerly identified in mucocutaneous and developmental abnormalities, resulting from telomere attrition. Here we identified complementary mutations in two proteins of the complex, affecting the highly conserved pseudouridylation catalytic site, associated with a novel phenotype characterized by renal, ocular, intestinal and auditory features, alongside reduced pseudouridine in rRNA and telomere attrition. Using a zebrafish model, we provide supporting evidence that this phenotype results from ribosomal dysfunction arising from a pseudouridylation defect of rRNAs. Together this describes a novel phenotype associated with the disruption of the most abundant RNA modification.

Main text

Pseudouridylation, the isomerization of uridine (U) to pseudouridine (Ψ), is the most common modification of RNA. It can be catalyzed by single protein pseudouridine synthases (PUSs) that act independently to recognize the substrate uridine or by the H/ACA small nucleolar ribonucleoprotein (snoRNP) complex (1, 2). Each complex is composed of a unique guide RNA and four core proteins, NOP10, NHP2, GAR1 and the catalytically active dyskerin (DKC1) (3). Knockdown of dyskerin results in a >50% reduction in rRNA pseudouridylation, indicating the primary role of the H/ACA snoRNP complex in the pseudouridylation of rRNAs (4).

Through its association with the telomerase RNA, *TERC*, the H/ACA snoRNP complex also plays a critical role in telomere synthesis (5). It is this function which links the snoRNPs (DKC1 (6, 7), NHP2 (8, 9) and NOP10 (10)) to the disease Dyskeratosis Congenita (DC, OMIM: 305000, 224230, 613987), characterised by mucocutaneous abnormalities and bone marrow failure and its more severe form, Hoyeraal-Hreidarsson syndrome (HH) with intrauterine growth retardation, microcephaly, cerebellar hypoplasia and in rare cases enteropathy. The role of defective telomere biogenesis in DC has been further corroborated by the identification of mutations in other genes *TERT*, *TERC*, *TINF2* and *RTEL1* (7, 11-14), which are implicated in telomere maintenance but not in pseudouridylation. Indeed, progressive telomere attrition in the autosomal dominant *TERT* or *TERC*-related DC results in disease anticipation (15, 16) and wild-type (WT) offspring of telomerase deficient mice with shortened telomeres develop an occult DC phenotype (17). The exhaustion of cellular renewal in *DKC1*-linked DC is rescued by the overexpression of *TERC*, reflecting the driving role of *TERC* level in the pathogenesis (18).

While human disorders associated with PUS-mediated pseudouridylation defects have been described (19-21), a human phenotype related to defective H/ACA snoRNP complex-mediated pseudouridylation has been lacking. Herein, we describe a novel human phenotype and early lethality in two unrelated pedigrees. Using linkage analysis and whole exome sequencing, novel mutations within *DKC1* and *NOP10* were identified in two families. A combination of structural and *in vivo* analysis demonstrates that a pseudouridylation defect of rRNA is what drives the distinction of this novel phenotype from classic DC.

The two unrelated pedigrees presented with an infantile-onset disorder characterized by steroid-resistant nephrotic syndrome, cataracts (prior to steroid treatment), sensorineural deafness and enterocolitis (Table 1, Figure 1). In the first pedigree, the disorder segregated in an X-linked pattern (FamA, Figure 1v) and in an autosomal recessive mode in the second (FamB, Figure 1x). All six affected males in FamA and the two affected females in FamB died in early childhood (Table 1). To identify the causative genetic loci, we performed linkage analysis in both families (Figure 1w, y). A single locus of 5.1 Mb at the telomeric end of the X chromosome long arm segregated with the disease in FamA, assuming that the causative mutation led to germline mosaicism in I:2. By sequencing the locus-specific coding regions of the affected IV:14 and the haploidentical but unaffected II:9 males, we found a single difference in the sequences: a point mutation in *DKC1* (c.616 G>A, p.Glu206Lys; Figure S1a). We considered this variant to be causative as it appeared *de novo*, segregated with disease in generation II corresponding to the assumed maternal germline mosaicism (Figure 1v), affected a universally conserved glutamic acid (replaced by aspartic acid in some nonvertebrates, Figure S1b) and the resulting change to lysine was predicted to be pathogenic by Mutation Taster, Polyphen-2 and SIFT. This variant was absent in gnomAD and in 555 alleles of 368 ethnically matched controls. The mutation is thus classified as pathogenic based on the ACMG/AMP criteria with evidence levels PS2, PM1-2, PP1-3 (22).

In FamB, whole exome sequencing and linkage analysis identified a mutation in *NOP10* (c.47C>T, p.Thr16Met; Figure S1a) within a homozygous region of 2.6 Mb in 15q14 (Figure 1y). The affected threonine is conserved down to yeast (*Saccharomyces cerevisiae*, Figure S1b), the change to methionine has not been found in the general population and was predicted as pathogenic by MutationTaster and SIFT. This variant is therefore also classified as likely pathogenic based on the ACMG/AMP criteria with evidence levels PM1-2, PP1 and PP3 (22).

There was minimal overlap between the novel DKC1 p.Glu206Lys- and NOP10 p.Thr16Met-related phenotype (Figures 1l and 2h) and the developmental anomalies of HH or the diagnostic mucocutaneous triad of DC (dyskeratosis, leukoplakia and nail dystrophy) (7); though the latter typically appear between the ages of five and ten years (23), beyond the lifetime of the children presented here (Table 1). Furthermore, with the exception of enterocolitis, cerebellar hypoplasia and progressive bone marrow failure found in a subset of the affected children (Table 1), the prominent features of this novel phenotype (Figure 1d-u) have not been reported in DC or HH. In contrast to the X-linked recessive transmission of *DKC1*-linked DC (24), females heterozygous for dyskerin p.Glu206Lys also developed cataracts and hearing impairment, necessitating artificial lens implantation and hearing aid in the second decade of life (Figure 1, Table 1), as well as maxillary and mandibular hypoplasia, pigmentary retinopathy, microphthalmia, pineal hypoplasia, mild cerebellar vermis atrophy and failure to thrive (Figures 1 and S2, Table 1). Pigmentary retinopathy was diagnosed in three heterozygous females (III:3, IV:4 and IV:12) resulting in no loss of visual acuity, but a flat electroretinogram. The index female in FamA (IV:4) developed a phenotype comparable to that of the males, requiring a hearing device at the age of one year, cataract operation at five and renal replacement therapy by six years. However, no enterocolitis or bone marrow failure presented (up to her present age of 15 years). We found by allele-specific qPCR a highly skewed X-

inactivation in skin cells and fibroblasts explaining her severe phenotype (Figure 2g). In accordance with the survival advantage of the cells expressing the WT allele in *DKC1*-linked DC (25, 26), her X-inactivation in leukocytes tended to be skewed towards the mutant allele by the second decade of life (Figure 2g).

We measured telomere length by Southern blot, monochrome multiplex quantitative PCR (MM-qPCR) and flow- fluorescent *in-situ* hybridization (FISH) and found the telomeres to be shortened in both families (Figure 2a-f), as with DC (27). This indicates that the identified mutations reduce the telomerase activity of the H/ACA snoRNP complex similarly to other DC-related *DKC1* and *NOP10* mutations. However, as the novel phenotypic features could not be attributed to telomere shortening, we aimed to unravel what distinguishes the effect of these novel mutations from those associated with DC.

We first conducted structural analyses of the snoRNP complex. Dyskerin, NOP10 and NHP2 directly associate with the guide H/ACA small nucleolar RNA (snoRNA) (Figure 3a), which binds and orientates the substrate RNA, specifically its target uridine, within the active pocket of dyskerin. Previously described cases of DC are associated with mutations that mostly concern amino acids implicated in the binding of the guide snoRNA (1, 28). The majority of these mutations are thus located at sites that interact with the guide RNA, such as the N- or C-termini of dyskerin (or Ser121 of the TruB domain), the Arg34 residue of NOP10 – reported in the single family with a NOP10 mutation (10) – or Arg61, Pro87 and Val126 of NHP2 (Figure 3a; spheres); with the only known exceptions of R158 and S280 of dyskerin. In contrast, dyskerin Glu206 and NOP10 Thr16 are located at the dyskerin-NOP10 interface, remote from both RNA components (Figure 3a). We found the subcellular localization of the p.Glu206Lys dyskerin to be preserved, similar to the most common DC-related mutation, p.Ala353Val (Figure S3). Reciprocal co-immunoprecipitation analysis demonstrated that the complex formation capacity of p.Glu206Lys dyskerin-NOP10 and

p.Thr16Met NOP10-dyskerin (Figures 3b) was maintained. In native patient protein, where we observed a reduced level of p.Thr16Met NOP10, suggesting a possible effect on protein stability, the binding interaction with dyskerin was still maintained (Figure S4). However, both mutations do alter the hydrogen-bonding between dyskerin and NOP10, reflected in their dissociation constant, as reported by pressure tuning fluorescence spectroscopy (Figure 3c-f) and the conformational changes (Figure 3g) seen in molecular dynamics (MD) simulations (Figure S5). MD simulations of the WT and two mutant states of the human snoRNP complex (containing all four protein components, the guide- and substrate-RNA chains) were carried out using homology modeling. The complex was modeled based on crystal-defined structures of full and partial H/ACA snoRNPs from *Pyrococcus furiosus* and *Saccharomyces cerevisiae,* as the structure of the human enzyme complex has not yet been determined. Significantly, comparison of the MD derived structures of the WT and mutant complexes revealed that structural changes at the dyskerin/NOP10 interaction surface result in a rearrangement of the pseudouridylation pocket (which sits over 20 Å away), detaching the catalytic Asp125 of dyskerin from the uridine of the bound substrate RNA (Figure 3h). Two different interaction paths can be found between the mutation sites and the catalytic core: one connecting Tyr15 of NOP10 to Arg248 of dyskerin, a residue that plays a role in fixing the phosphate group of the uridine substrate (via the Tyr15(NOP10)↔Arg247(dyskerin) Hbond) (29) and another intricate network of H-bonds and hydrophobic interactions linking the 16-18 segment of NOP10 and the loop 122-131 carrying the catalytic Asp125 of dyskerin: Thr16 and Lys18 of NOP10 form H-bonds with Glu206 and Glu208 of dyskerin, the latter of which is also H-bound to Arg211 (dyskerin). This residue is stacked against Phe2 of NOP10 that reaches into a pocket also comprising dyskerin Arg158, Leu213, Gln244 and Glu245. Glu245 is H-bond to His31 of NOP10, a residue also coordinated by Thr129 of dyskerin, which is located in the loop carrying the active Asp125. This gear-wheel like contact surface of the two proteins allows for communication between far lying sites; the subtle changes elicited by the mutations at the

dyskerin/NOP10 interface therefore have the capacity to influence both substrate binding and catalysis in the substrate binding pocket of dyskerin. Together, these findings indicate that both dyskerin p.Glu206Lys and NOP10 p.Thr16Met alter the pseudouridylation capacity of the snoRNP complex, suggesting the pathogenic commonality underlying this phenotype.

As both mutations act by disrupting the catalytic center of dyskerin, we aimed to gain greater insight into the resulting pathogenesis and generated mutations in zebrafish *dkc1* targeting different regions of the gene (*elu1* and *elu8*, Figure S6). *In situ* hybridisation characterized the expression of *dkc1* during zebrafish development, confirming its ubiquitous expression up to 12 hours post fertilization (dpf), with strong expression at sphere stage, suggesting the presence of a maternal component (Figure S7). Both *elu1* and *elu8* homozygous null *dkc1* mutants die at five dpf with a phenotype equivalent to the human phenotype. We confirmed the specificity of the null phenotype with a translation-blocking anti-sense morpholino (Figure S6).

Ocular sections of *dkc1*$^{elu1/elu1}$ larvae showed opaque lenses due to the persistence of nucleated fiber cells akin to zebrafish cataract models (30) and a high abundance of cells with a neuroepithelial morphology, characteristic of progenitor cells (Figure 4a). We observed increased staining for proliferation markers in the retina and optic tecta (Figure 4a), areas with high *dkc1* expression (Figure S7), indicative of a cell-cycle defect. Inner ear development was impaired, with the cylindrical projections from the otocyst walls remaining unfused and the intestinal compartments of the gut undifferentiated (Fig 4b). Indeed, several tissues showed reduced expression of differentiation markers in *dkc1*$^{elu1/elu1}$ larvae (Figure 4d). Development of the pronephros was severely hypoplastic, with reduced Wt1-positive podocyte number (Figure 4b), though no increased filtration of 500kDa FITC-dextran was observed at this stage (Figure S8a).

A hematopoietic defect was also observed (Figures 4d and S8c), as described in previous *dkc1* and *nop10* null zebrafish models (31, 32). Lack of *dkc1* also caused defective jaw-cartilage development (Figure S8b) and a disorganized pineal gland (Figure 4c), features of the female *DKC1* p.Glu206Lys patients (Figure S2).

Null mutants showed rescue with zygotic injection of WT human *DKC1* mRNA (Figures 4e and S6d). A far weaker rescue was achieved by *DKC1* p.Glu206Lys mRNA, indicating its pathogenicity with some limited function (Figures 4e and S6d). A hypomorphic (*elu2*) allele was also generated (Figure S6a): these *dkc1*$^{elu2/elu2}$ fish were viable, albeit with significant growth retardation (Figure 5j). The combined analysis of two null mutants replicating the human phenotype, a morphant and the limited phenotypic rescue with *DKC1* p.Glu206Lys mRNA provides strong support for the phenotypic relevance of this model.

Telomere biogenesis in the zebrafish is similar to that of humans (33). We observed no telomere shortening in the *dkc1*$^{elu1/elu1}$ animals (Figure 5a), similar to *nop10* or *nola1* (*GAR1*) deficient larvae (31, 32). Our structural analysis, demonstrating the effect of p.Glu206Lys and p.Thr16Met on the pseudouridylation pocket, suggested a potential role for the defective pseudouridylation of the rRNA and consequential ribosomal dysfunction. The abundance of processed 18S rRNA was low in *dkc1*$^{elu1/elu1}$ larvae (Figure 5b), as with previous findings in *nop10* and *nola1* (*GAR1*) mutants and *dkc1* morphants (31, 32). This is in accordance with the reduced pseudouridylation of 18S rRNA in *dkc1*$^{elu1/elu1}$ and *dkc1*$^{elu8/elu8}$ larvae (Figure 5c), as well as in the PBMCs of patient IV:3, FamB (Figure 5d), as detected by Immuno-Northern blot. A pseudouridylation defect was also apparent in the decreased Ψ/U ratio (detected by HPLC-MS) in the PBMCs of the index female with skewed X-inactivation (FamA, IV:4) (Figure 5e). In contrast, pseudouridylation of 18S rRNA in the fibroblasts and the Ψ/U ratio in the skin of these patients was not different from that of controls (Figure S9), suggestive of the tissue-specific effect of *DKC1*-mutations on

pseudouridylation (34). The difference in the pseudouridylation defect may reflect contribute to the differential survival advantage of the mutant allele in fibroblasts and PBMCs in the index girl (FamA, IV:4) (Figure 2g).

Transcriptomic analysis of the *dkc1*$^{elu1/elu1}$ larvae also highlighted defects in ribosome biogenesis (Figure 5f) and the mutant phenotype was recapitulated in WT embryos treated with the translational inhibitor cycloheximide (35) (Figure S6c). This observed phenotype is also highly reminiscent to that described in homozygous mutants of genes encoding ribosomal proteins (31, 36-38). Together these results suggest ribosomal dysfunction to be a main driver of the phenotype.

In accordance with previous zebrafish models of ribosomopathies (31, 32, 38), we found dysregulated Tp53 expression in mutants. Western blot analysis suggested the stabilization and accumulation of Tp53 protein in mutants (Figure 5g), in line with previous results from disrupted ribosomal biogenesis models (39). Further analysis however, showed the transcription of the full-length isoform to be down-regulated and the truncated, anti-apoptotic *Δ113p53* isoform up-regulated (Figure 5h). The latter isoform inhibits the classic Tp53-response in zebrafish (40), fitting with the sustained proliferative state we describe (Figure 4a). However, the *dkc1*$^{elu1/elu1}$ phenotype was not rescued on a *tp53*-null background (Figure 5i), with the exception of a partial rescue of hematopoiesis (Figure S8c), as has been previously reported in *dkc1* (31) and *nop10* (32) deficient larvae. This indicates that the major phenotypic features observed upon loss of Dkc1 function are not mediated by Tp53.

Previous studies suggest a more fundamental role for the snoRNP complex than that of hematopoiesis or telomere maintenance alone: *Dkc1*-deficient mice die *in utero* (26) and *nop10* and *nola1*/*GAR1* mutant

and *dkc1* morphant zebrafishes die at 5-10 dpf, all with normal telomere length (32, 33). Though reduced dyskerin expression in mice (41) and some DC-related *DKC1* mutations (34) were shown to affect rRNA pseudouridylation, other DC-related *DKC1* mutations did not (18, 42), or exerted only a subtle effect (4). Here we show that mutations affecting the dyskerin-NOP10 interaction and the pseudouridylation pocket of the H/ACA snoRNP complex cause a novel phenotype with early lethality. Several non-classical phenotypic features have been described in telomere biology disorders (43). This phenotype presents the novel features of nephrotic syndrome and cataracts which have not yet reported to be associated with mutations in the H/ACA snoRNP complex.

Given the large number of the patients with H/ACA snoRNP complex-related disorders and the >70 identified *DKC1* mutations, the finding of a novel phenotype related to the H/ACA snoRNP complex might seem unexpected. However, none of the previous *DKC1* mutations affect the highly conserved pseudouridylation catalytic site and instead affect the N- and C-terminal dyskerin residues which show low conservation or are even absent in the *Pyrococcus furiosus* ortholog, Cbf5 (29). Given the fundamental role of the H/ACA snoRNP complex in development, reflected by the early lethality of the knockout animal models (26, 32, 33), alongside the lack of missense variants affecting the dyskerin catalytic site, or of *DKC1* and *NOP10* loss-of-function mutations in patients with telomeropathies, a severe phenotype related to disrupted pseudouridylation seems likely. This is corroborated by the pseudouridylation defect in *DKC1*-deficient cells (4), but no significant defect in DC (4, 18, 42).

We believe that the novel phenotype we describe is the first recorded example of that more severe case. Although at present we only describe two families, the causality of these mutations has strong support in: 1) the significant LOD scores; 2) the size of the *DKC1* family allowing for the genetic tracing of the germline

mosaic, *de novo* p.Glu206Lys mutation for three generations prior to the index case; 3) the well characterised role of both mutated amino acids in the interaction of DKC1 and NOP10 (29); and 4) the shared functional pathophysiology of telomere shortening. Similar examples of pleiotropy have been described in the *CEP290* (44), *PMM2* (45, 46), *REN* (47, 48) or *LMNA* (49, 50) genes. Indeed, it is unlikely that this novel phenotype represents the most severe of the H/ACA snoRNP complex syndromes. Loss-of-function mutations are expected to result in an even more severe, potentially in-utero lethal disorder, making the pleiotropy of the H/ACA snoRNPs even more pervasive.

Although telomere shortening was apparent, the disorder was lethal before the potential appearance of classic DC symptoms. Since the *dkc1* mutant zebrafish recapitulated the human phenotype with normal telomere length but ribosomal dysfunction, consequence of impaired 18S rRNA pseudouridylation, a feature also observed in the patients, we conclude that a pseudouridylation defect is the principal driver of this novel phenotype. Given the fundamental role of H/ACA snoRNP complex in targeting uridines not just in rRNAs but also in snRNAs and mRNAs (1, 4, 51, 52), the varied site-specific impact felt on each uridine residue (34) and the additive consequences of the altered dyskerin-NOP10 interaction we describe, this phenotype may culminate from a pseudouridylation defect implicating multiple RNAs. There is increasing evidence for the broader pathogenicity of defective RNA processing in human pathologies (53). Mutations in the KEOPS-complex, involved in the modification of tRNAs, have recently been identified in Galloway-Mowat syndrome (54) and a role for defective rRNA methylation has now been described in DC (55). The findings presented here propose that defective pseudouridylation is a new mechanism for injury in the human kidney, eye and cochlea, adding to our understanding of RNA processing in human disease.

Materials and methods

Identification of the causal mutations

All patients or their legal representatives gave written informed consent and the study was approved by the Local Research Ethic Committees (TUKEB 1154/0-2010-1018EKU and 05/Q0508/6). For Family A (FamA), linkage analysis was performed using the Human Mapping 250k NspI array (Affymetrix) and parametric LOD scores were calculated with Multipoint Engine for Rapid Likelihood Interference software (56), assuming X-linked inheritance. Females with an uncertain phenotype in generation IV were not included in the linkage analysis. Extended haplotype analysis using five microsatellite markers (DXS8011, DXS8103, DXS8061, DXS8087, DXS1073) flanking the *DKC1* locus was performed in all the available members of FamA. PCR products with fluorescent primers were separated by capillary electrophoresis (3130 Genetic Analyzer) and analyzed using GeneMapper analysis software (Life Technologies).

In the affected individual IV:14 and the haploidentical but unaffected individual II:9 of Family A, enrichment of the linkage interval (Xq28–Xqter:150,474,451–155,270,560 bp; GRCh37/hg19) and subsequent sequencing were performed as described elsewhere (57). In brief, we used the Roche NimbleGen 385K custom sequence capture array with a capacity of up to 5 Mb of target sequence. The exon-based design included all exons of the protein-coding genes, including 100bp of flanking intron sequence and 1 kb of the promoter regions and known miRNAs within the critical interval. Final coverage of the design was 99.2%. Enrichment factors of >700-fold were achieved for the two samples. From the enriched DNA samples, we generated paired-end libraries and sequenced them on an Illumina GA IIx device with a read length of 2 x 36 bp. The average coverage of the 447,699 bp of target sequences was >360x in both samples. For data handling and analysis, we used our in-house database and analysis tool kit VARBANK 2.0 (https://varbank.ccg.uni-koeln.de/varbank2/) with default settings for variant filtering.

These parameters guarentee to focus on rare high-quality variants altering the protein sequence or showing impairments of splice and translation initiation sites. Only one such variant was found in the affected individual that was absent from the unaffected one, namely *DKC1:*c.616G>A, p.Glu206Lys. The median of 31 pathogenicity rank scores calculated for this variant was 0.83, thus, strongly supporting causality. The segregation of the *DKC1* variant was validated by Sanger sequencing.

For Family B (FamB), whole exome sequencing of the first index patient (V:2) was conducted by GOSgene (BGI), from genomic DNA (gDNA) extracted from peripheral blood; exome capture was performed using the Agilent SureSelect v4 (51 Mb). Homozygosity mapping of candidate variants was conducted on index patients (V:2 and IV:3), parents and siblings via Sanger sequencing. Segregation analysis confirmed an autosomal recessive mode of inheritance for a mutation in *NOP10* (c.47C>T, p.Thr16Met). This was confirmed through linkage analysis using the Cyto SNP Microarray technique and Infinium assay HD Ultra kit (Illumina, #WG-901-4007). Sequencing files were assembled in Genome Studio and regions of shared homozygosity analysed in Homozygosity Mapper.

Telomere length analysis

Three independent methods were used to assay the length of the telomeres: Southern blot, MM-qPCR and flow-fluorescence *in situ* hybridization (FISH) (SI Appendix).

X-inactivation measurements

RNA was isolated from peripherial blood (NucleoSpin RNA Blood, Macherey-Nagel), skin and fibroblasts (Total RNA Mini Kit Tissue, Geneaid) and was reverse transcribed (BIO-65042, Bioline; K1671, Thermo Scientific) following the manufacturer's instructions. The X-inactivation ratio was assessed by allele-

specific qPCR using a Taqman probe (LightCycler TaqMan Master, Roche, Table S1). A dilution series of Flag-tagged WT and E206K *DKC1* plasmids were used for standardization; all measurements were performed three times.

Homology modeling and Molecular Dynamics simulations

Homology models for the WT and mutant H/ACA snoRNP complexes were constructed based on full and partial H/ACA sRNPs from *Pyrococcus furiosus* and *Saccharomyces cerevisiae* (PDB id: 3hay (58); 3hjw (59); 3lwq (60); 2lbw (61); 3u28 (62)) using Schrödinger Modeling Suite (63) and then subjected to 600 ns molecular dynamics simulations using GROMACS (64), with the AMBER-ff99SBildnp* forcefield (65) and the parametrization of Steinbrecher et al (66) for the phosphate moieties. The simulations were solvated by approximately 35880 OPC water molecules (67), the total charge of the system was neutralized and physiological salt concentration was set using Na^+ and Cl^- ions. Energy minimization of starting structures was followed by relaxation of constraints on protein atoms in three steps, with an additional NVT step (all of 200 ps) to stabilize pressure. Trajectories of 600 ns NPT simulations at 325K (to enhance sampling) and 1 bar were recorded for further analysis (collecting snapshots at every 4 ps). Clustering of the equilibrium trajectory was carried out based on the backbone structure of the interaction surface of dyskerin and NOP10 (containing the entire NOP10 model and residues of dyskerin that have an atom reaching within 6 Å of NOP10) using a 1 Å cutoff.

Co-immunoprecipitation studies

Total protein was harvested using Pierce™ IP Lysis Buffer (Thermo Fisher, #87787) and protein concentration was determined using the BCA Protein Assay Kit (Thermo Fisher, #23225). One mg of protein lysate was incubated with 4 ug of target antibody for 8 hours at 4°C, followed by an overnight incubation

with Dynabeads™ Protein G (Thermo Fisher, #10003D). Immunoprecipitated samples were subject to immunoblot, signal was detected using Pierce™ ECL Western Blotting Substrate (Thermo Fisher, #32106). In addition, primary keratinocytes were isolated from skin biopsy and cultured as described in Nowak et al (68). Co-immunoprecipitation studies were conducted as described above with the amendment of 500 ug of protein lysate, 2ug of NOP10 antibody (Abcam, #ab134902) and host-specific Sheep Anti-Rabbit IgG Dynabeads™ (Invitrogen, #11203D).

Protein extraction for pressure tuning fluorescence spectroscopy

Total protein was harvested by 150 mM NaCl, 20 mM Tris, 1% Triton-X supplemented with 0.1% protease inhibitor (Protease Inhibitor Cocktail, Sigma-Aldrich). Lysates were incubated with monoclonal anti-flag (F3165, Sigma Aldrich) or anti-V5/NOP10 (ab182008/ab134902, Abcam) antibodies) and subsequently with Protein G beads (Dynabeads Protein G for Immunoprecipitation, Thermo Fisher Scientific). Dyskerin and NOP10 variants were eluted by competition with Flag and V5 peptides (B23111, B23511, Biotool; F3290, Sigma Aldrich) for 30 minutes, repeated 5 times. Eluate concentration was measured by spectrophotometry (DC Protein Assay, Bio-Rad) and verified by SDS-PAGE, using anti-V5/anti-NOP10 (F3165, Sigma Aldrich, ab182008/ab134902, Abcam) as primary and anti-mouse/anti-goat IgG-HRP as secondary (sc-2005, sc-2357, SCBT) antibodies.

Pressure tuning fluorescence spectroscopy

Pressure tuning fluorescence spectroscopy was used to determine the dissociation constant of the NOP10-dyskerin complex (K_d) and the pKd (pKd = -log(Kd)) (69, 70) as well as the volume change accompanying dissociation (ΔV, ml/mol). Briefly, the method involves the use of a fluorescent dye which binds to the solvent-accessible, hydrophobic regions of proteins, hydrostatic pressure is then varied to induce

dissociation. By fitting a dissociation model to the fluorescence data, the K_d and ΔV values are determined as follows:

$$I(p) = I_0 + \Delta I \cdot \frac{\sqrt[2]{K_d^2 e^{-2\frac{p\Delta V}{RT}} + 16CK_d e^{-\frac{p\Delta V}{RT}}} - K_d e^{-\frac{p\Delta V}{RT}}}{8C}$$

Where I(p) and I_0 are the resulting and original fluorescence intensities, respectively, ΔI is the maximal intensity change and C is the total protein concentration. This method yields a specific volume and by determining the width of pressure range in the 5%-95% dissociation regime, structural heterogeneity was qualitatively assessed. A nonlinear least squares was fitted to the data and statistical difference between WT and mutant complexes was assessed via a Welch-test (Statistica 13.4). The interaction surface was directly related to the pKd through the general concept of the equilibrium constant:

$$K = A \cdot e^{-\frac{\Delta E}{kT}} = A \cdot e^{-\frac{a*S+b}{kT}}$$

Where ΔE is the complex formation energy, A is the pre-exponential factor, S is the interaction surface in Ångström (Å) and a, b are the parameters of the linear function connecting ΔE and S.

Generation of zebrafish mutants

CRISPR/Cas9 mediated mutagenesis of the *dkc1* gene was carried out in the WT *tuebingen* (*tue*) strain as previously described (71); targeted regions (gRNA sequences) are listed in Table S1. Genotyping was conducted using PCR on gDNA samples isolated from fin clips; the allele-specific forward primers for genotyping, sequencing of exon 7 and 11 and exon-specific reverse primers are listed in Table S1. An anti-sense, translation-blocking morpholino (MO) was ordered from GeneTools (Portland, OR) to confirm specificity of the null *elu1* and *elu8* phenotype. WT (*tue*) and mutant *dkc1* and *tp53* fish lines were maintained in the animal facility of ELTE Eötvös Loránd University according to standard protocols (72, 73). All zebrafish protocols were approved by the Hungarian National Food Chain Safety Office (Permit

Number: XIV-I-001/515-4/2012) and the Pest County Governmental Office (Permit Number: PE/EA/2023-7/2017).

Zebrafish rescue experiments

The CDS for human *DKC1* and *DKC1* E206K were cloned into a pCS2+ vector, linearized with *KpnI* and *in vitro* transcribed using the mMessage mMachine SP6 kit (Ambion). Zebrafish embryos from a *dkc1*$^{elu1/+}$ in-cross were injected with *DKC1* mRNA at 1-2 cell stage and phenotypically scored at 4 dpf.

Zebrafish histology

In situ hybridization experiments, Neutral Red and hematoxylin-eosin histological stainings, immunostaining, Acridin Orange labeling and filtration tests using fluorescent dextrans were performed according to standard protocols (SI Appendix).

Zebrafish growth measurements

Offspring of incrossed *dkc1*$^{elu2/+}$ heterozygous fish were housed and raised at standard stocking densities. Measurements were taken at four months of age. Body length was measured as standard length, from the tip of the head to the end of the trunk and wet body weight was determined using a standardised method (74), to avoid anesthesia.

Analysis of 28S/18S rRNA ratio and pseudouridylation

Total RNA was isolated from 4.5-5 dpf zebrafish by TRIzol Reagent (Invitrogen). The RNA quality was assessed on Agilent 2100 Bioanalyzer using the RNA 6000 Pico kit (Agilent Technologies) according to the manufacturers' instructions. For comparison of 18S and 28S rRNA abundance in mutant and WT fish,

densitometry of the 18S and 28S peaks was performed using ImageJ and their ratio was compared by Student's t-test (Statistica software version 13.2).

Pseudouridylation was assessed by immuno-northern blotting as previously described (75). In brief, 3 μg of total RNA was loaded per well on a 1.5% denaturing agarose gel and imaged to assess for degradation. Gel blotting was performed overnight onto a nylon membrane (Amersham HybondN -GE Healthcare Life Sciences), followed by UV cross-linking. Nylon membranes were blocked with 5% non-fat milk for 1 hour, then incubated with an anti-pseudouridine (D-347-3, MBL, dil. 1:500 in PBS) primary antibody for 1.5 hours, followed by an HRP-conjugated anti-mouse IgG (sc -2005, SCBT, dil. 1:1000 in PBS) secondary for 1 hour, all at room temperature. Bands were visualized by chemiluminescence (Western Blotting Luminol Reagent, sc-2048, SCBT) on Molecular Imager VersaDoc MP 5000 System (Bio-Rad). Densitometry was performed using ImageJ processing program. Experiments were repeated three times. To assess the 18S rRNA pseudouridylation, the ratio of the normalized 18S rRNA PU and the normalized 18S rRNA was compared between the mutant and the sibling zebrafish by Mann-Whitney U test.

Whole transcriptome analysis

Total RNA was extracted from whole zebrafish embryos at 36 hpf, using TRIzol reagent (Thermo Fisher, 15596018) and three biological replicate samples were prepared for mutants and siblings, respectively. The samples were sequenced on Illumina MiSeq platform by Microsynth (Microsynth AG, Switzerland) according to standard lllumina protocols. Raw sequence data was deposited to the Sequence Read Archive (SRA), accession number PRJNA548449. Raw reads were pre-processed with FASTQC (76), followed by analysis based on the 'new tuxedo pipeline' (HISAT, StringTie and Ballgown) using default settings (77). Gene ontology term analysis and visualization was performed using the emapplot function of the

clusterProfiler package (78).

Determination of pseudouridine and uridine content via HPLC-MS

RNA was isolated from peripherial blood (NucleoSpin RNA Blood, Macherey-Nagel) and skin (Total RNA Mini Kit Tissue, Geneaid) and their pseudouridine and uridine content were measured as described previously (79). Briefly, for the determination a HPLC-MS/MS system was used including an Agilent 1100 modular HPLC (Agilent Technologies, USA) and an MDS Sciex API 4000 Triple-Quad mass spectrometer (Applied Bioscience, USA) equipped with the TurboV-Spray source. For the gradient elution on the Phenomenex Luna C18, 5 µm 3.0 x 150 mm column (Gen-Lab, Hungary), water (A) and mixture of water and methanol (50 V/V%, B) were used, supplemented both with ammonium acetate (25 mM). At the initial stage of the separation the B was kept at 0% for 1 minute and increased up to 100% for 5 minutes, which was hold for 3 minutes before the column equilibration. The flow rate was 400 µl/min and the column was thermostated at 55°C. For the detection, the mass spectrometer operated in the positive mode, where the ESI ion source gas temperature was set at 525°C, the ionisation voltage at 5000 V. During the multiple reaction monitoring measurements, the precursor ion was set to m/z 245.2 while the monitored fragment ion was m/z 113.1 using 50 V and 20 eV for the declustering potential and collision energy, respectively.

Data availability statement

All relevant data are included in the main text and SI Appendix. Zebrafish transcriptomic datasets generated for this study can be found in the NCBI SRA, accession number PRJNA548449.

Acknowledgments

The authors would like to thank Rózsa Csohány, Mária Bernáth, Dániel Silhavy, Kata Gyulay, Noémi Borbély, Mirjam Pencz, Zoltán Varga and Julianna Víg for their help and Moin Saleem for providing the podocyte cell line. Financial support for this work was provided by: MTA-SE Lendulet Research Grant (LP2015-11/2015) of the Hungarian Academy of Sciences, Pfizer's Research Grant and the NRDI-K109718, KH125566 grants to KTor; NRDI-K116305 to DKM; the Ministry of Human Capacities in Hungary in the frame of Institutional Excellence Program for Higher Education, NRDI-FK124230 grant financed by the Hungarian National Research, Development and Innovation Office and the ÚNKP-17-4 New National Excellence Program and the ELTE Institutional Excellence Program (1783-3/2018/FEKUTSRAT) to MV; and a MedinProt Synergy IV grant to MV and GS; an MRC Clinician Scientist fellowship (MR/K010654/1) to AMW; Kidney Research UK Innovation and Project grant (Paed_RP_011_20170929) and the National Institute for Health Research (NIHR) Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London to AMW, JCC and DAL; an Egyptian Cultural Affairs and Mission Sector, Ministry of High Education grant to MT; Cancer Research UK (C36439/A12097) to KTom; a Chercheur-boursier award of the Fonds de recherche en santé du Québec and the Natural Sciences and Engineering Research Council of Canada (RGPIN 249792) to FD. Authors thank the NIIF institute for providing computational facilities.

Declaration of Interests

The authors declare no competing interests.

Web Resources

BIOPKU, http://www.biopku.org/home/home.asp

Ensemble, https://www.ensembl.org/index.html

gnomAD, https://gnomad.broadinstitute.org/

HGMD, http://www.hgmd.cf.ac.uk/ac/index.php

HGVS, https://varnomen.hgvs.org

LOVD, https://www.lovd.nl/3.0

Mutalyzer, https://mutalyzer.nl/

Mutation Taster, http://www.mutationtaster.org/

OMIM, http://www.omim.org/

PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/

SIFT, https://sift.bii.a-star.edu.sg/

References

1. A. M. Kiss, B. E. Jady, E. Bertrand, T. Kiss, Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol* **24**, 5797-5807 (2004).
2. M. Penzo, L. Montanaro, Turning Uridines around: Role of rRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function. *Biomolecules* **8** (2018).
3. M. Khanna, H. Wu, C. Johansson, M. Caizergues-Ferrer, J. Feigon, Structural study of the H/ACA snoRNP components Nop10p and the 3' hairpin of U65 snoRNA. *RNA* **12**, 40-52 (2006).
4. S. Schwartz *et al.*, Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148-162 (2014).
5. K. Collins, J. R. Mitchell, Telomerase in the human organism. *Oncogene* **21**, 564-579 (2002).
6. N. S. Heiss *et al.*, X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* **19**, 32-38 (1998).
7. I. Dokal, T. Vulliamy, P. Mason, M. Bessler, Clinical utility gene card for: Dyskeratosis congenita - update 2015. *Eur J Hum Genet* **23** (2015).
8. T. Vulliamy *et al.*, Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc Natl Acad Sci U S A* **105**, 8073-8078 (2008).
9. M. Benyelles *et al.*, NHP2 deficiency impairs rRNA biogenesis and causes pulmonary fibrosis and Hoyeraal-Hreidarsson syndrome. *Hum Mol Genet* 10.1093/hmg/ddaa011 (2020).
10. A. J. Walne *et al.*, Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. *Hum Mol Genet* **16**, 1619-1629 (2007).
11. T. Vulliamy *et al.*, The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**, 432-435 (2001).
12. A. Marrone *et al.*, Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal-Hreidarsson syndrome. *Blood* **110**, 4198-4205 (2007).
13. S. A. Savage *et al.*, TINF2, a component of the shelterin telomere protection complex, is mutated in dyskeratosis congenita. *Am J Hum Genet* **82**, 501-509 (2008).
14. T. Vulliamy *et al.*, Telomere length measurement can distinguish pathogenic from non-pathogenic variants in the shelterin component, TIN2. *Clin Genet* **81**, 76-81 (2012).
15. T. Vulliamy *et al.*, Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in TERC. *Nat Genet* **36**, 447-449 (2004).
16. M. Armanios *et al.*, Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc Natl Acad Sci U S A* **102**, 15960-15964 (2005).
17. L. Y. Hao *et al.*, Short telomeres, even in the presence of telomerase, limit tissue renewal capacity. *Cell* **123**, 1121-1131 (2005).
18. J. M. Wong, K. Collins, Telomerase RNA level limits telomere maintenance in X-linked dyskeratosis congenita. *Genes Dev* **20**, 2848-2858 (2006).
19. Y. Bykhovskaya, K. Casas, E. Mengesha, A. Inbal, N. Fischel-Ghodsian, Missense mutation in pseudouridine synthase 1 (PUS1) causes mitochondrial myopathy and sideroblastic anemia (MLASA). *Am J Hum Genet* **74**, 1303-1308 (2004).
20. R. Shaheen *et al.*, A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. *Hum Genet* **135**, 707-713 (2016).

21. A. P. M. de Brouwer *et al.*, Variants in PUS7 Cause Intellectual Disability with Speech Delay, Microcephaly, Short Stature, and Aggressive Behavior. *Am J Hum Genet* **103**, 1045-1052 (2018).
22. S. Richards *et al.*, Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405-424 (2015).
23. I. Dokal, Dyskeratosis congenita in all its forms. *Br J Haematol* **110**, 768-779 (2000).
24. C. Sirinavin, A. A. Trowbridge, Dyskeratosis congenita: clinical features and genetic aspects. Report of a family and review of the literature. *J Med Genet* **12**, 339-354 (1975).
25. T. J. Vulliamy, S. W. Knight, I. Dokal, P. J. Mason, Skewed X-inactivation in carriers of X-linked dyskeratosis congenita. *Blood* **90**, 2213-2216 (1997).
26. J. He *et al.*, Targeted disruption of Dkc1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice. *Oncogene* **21**, 7740-7744 (2002).
27. B. P. Alter *et al.*, Very short telomere length by flow fluorescence in situ hybridization identifies patients with dyskeratosis congenita. *Blood* **110**, 1439-1447 (2007).
28. C. Trahan, C. Martel, F. Dragon, Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* **19**, 825-836 (2010).
29. R. Rashid *et al.*, Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell* **21**, 249-260 (2006).
30. A. C. Morris, The genetics of ocular disorders: insights from the zebrafish. *Birth Defects Res C Embryo Today* **93**, 215-228 (2011).
31. Y. Zhang, K. Morimoto, N. Danilova, B. Zhang, S. Lin, Zebrafish models for dyskeratosis congenita reveal critical roles of p53 activation contributing to hematopoietic defects through RNA processing. *PLoS One* **7**, e30188 (2012).
32. T. C. Pereboom, L. J. van Weele, A. Bondt, A. W. MacInnes, A zebrafish model of dyskeratosis congenita reveals hematopoietic stem cell formation failure resulting from ribosomal protein-mediated p53 stabilization. *Blood* **118**, 5458-5465 (2011).
33. M. C. Carneiro *et al.*, Short Telomeres in Key Tissues Initiate Local and Systemic Aging in Zebrafish. *PLoS Genet* **12**, e1005798 (2016).
34. C. Bellodi *et al.*, H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation. *Cell Rep* **3**, 1493-1502 (2013).
35. M. Poulain, T. Lepage, Mezzo, a paired-like homeobox protein is an immediate target of Nodal signalling and regulates endoderm specification in zebrafish. *Development* **129**, 4901-4914 (2002).
36. G. Golling *et al.*, Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nat Genet* **31**, 135-140 (2002).
37. R. N. Kettleborough *et al.*, A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* **496**, 494-497 (2013).
38. N. Danilova, K. M. Sakamoto, S. Lin, Ribosomal protein L11 mutation in zebrafish leads to haematopoietic and metabolic defects. *Br J Haematol* **152**, 217-228 (2011).
39. Y. Zhang, H. Lu, Signaling to p53: ribosomal proteins find their way. *Cancer Cell* **16**, 369-377 (2009).
40. L. Guo *et al.*, Detection of the 113p53 protein isoform: a p53-induced protein that feeds back on the p53 pathway to modulate the p53 response in zebrafish. *Cell Cycle* **9**, 1998-2007 (2010).
41. D. Ruggero *et al.*, Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **299**, 259-262 (2003).

42. B. W. Gu *et al.*, Impaired Telomere Maintenance and Decreased Canonical WNT Signaling but Normal Ribosome Biogenesis in Induced Pluripotent Stem Cells from X-Linked Dyskeratosis Congenita Patients. *PLoS One* **10**, e0127414 (2015).
43. M. Kirwan, I. Dokal, Dyskeratosis congenita: a genetic disorder of many faces. *Clin Genet* **73**, 103-112 (2008).
44. L. Baala *et al.*, Pleiotropic effects of CEP290 (NPHP6) mutations extend to Meckel syndrome. *Am J Hum Genet* **81**, 170-179 (2007).
45. G. Matthijs *et al.*, Mutations in PMM2, a phosphomannomutase gene on chromosome 16p13, in carbohydrate-deficient glycoprotein type I syndrome (Jaeken syndrome). *Nat Genet* **16**, 88-92 (1997).
46. O. R. Cabezas *et al.*, Polycystic Kidney Disease with Hyperinsulinemic Hypoglycemia Caused by a Promoter Mutation in Phosphomannomutase 2. *J Am Soc Nephrol* **28**, 2529-2539 (2017).
47. O. Gribouval *et al.*, Mutations in genes in the renin-angiotensin system are associated with autosomal recessive renal tubular dysgenesis. *Nat Genet* **37**, 964-968 (2005).
48. M. Zivna *et al.*, Dominant renin gene mutations associated with early-onset hyperuricemia, anemia, and chronic kidney failure. *Am J Hum Genet* **85**, 204-213 (2009).
49. D. Fatkin *et al.*, Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med* **341**, 1715-1724 (1999).
50. M. Eriksson *et al.*, Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423**, 293-298 (2003).
51. T. M. Carlile *et al.*, Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143-146 (2014).
52. X. Li *et al.*, Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat Chem Biol* **11**, 592-597 (2015).
53. M. Penzo, A. N. Guerrieri, F. Zacchini, D. Trere, L. Montanaro, RNA Pseudouridylation in Physiology and Medicine: For Better and for Worse. *Genes (Basel)* **8** (2017).
54. D. A. Braun *et al.*, Mutations in KEOPS-complex genes cause nephrotic syndrome with primary microcephaly. *Nat Genet* **49**, 1529-1538 (2017).
55. D. Nachmani *et al.*, Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. *Nat Genet* 10.1038/s41588-019-0502-z (2019).
56. G. R. Abecasis, S. S. Cherny, W. O. Cookson, L. R. Cardon, Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* **30**, 97-101 (2002).
57. N. Jurkute, Leu, C., Pogoda, H.-M., Arno, G., Robson, A.G., Nürnberg, G., Altmüller, J., Thiele, H., Motameny, S., Toliat, M.R., Powell, K., Höhne, W., Michaelides, M., Webster, A.R., Moore, A.T., Hammerschmidt, M., Nürnberg, P., Yu-Wai-Man, P, Votruba, M. , SSBP1 mutations in dominant optic atrophy with variable retinal degeneration. *Ann Neurol*, in press (2019).
58. J. Duan, L. Li, J. Lu, W. Wang, K. Ye, Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell* **34**, 427-439 (2009).
59. B. Liang *et al.*, Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nat Struct Mol Biol* **16**, 740-746 (2009).
60. J. Zhou, B. Liang, H. Li, Functional and structural impact of target uridine substitutions on the H/ACA ribonucleoprotein particle pseudouridine synthase. *Biochemistry* **49**, 6276-6281 (2010).
61. B. K. Koo *et al.*, Structure of H/ACA RNP protein Nhp2p reveals cis/trans isomerization of a conserved proline at the RNA and Nop10 binding interface. *J Mol Biol* **411**, 927-942 (2011).

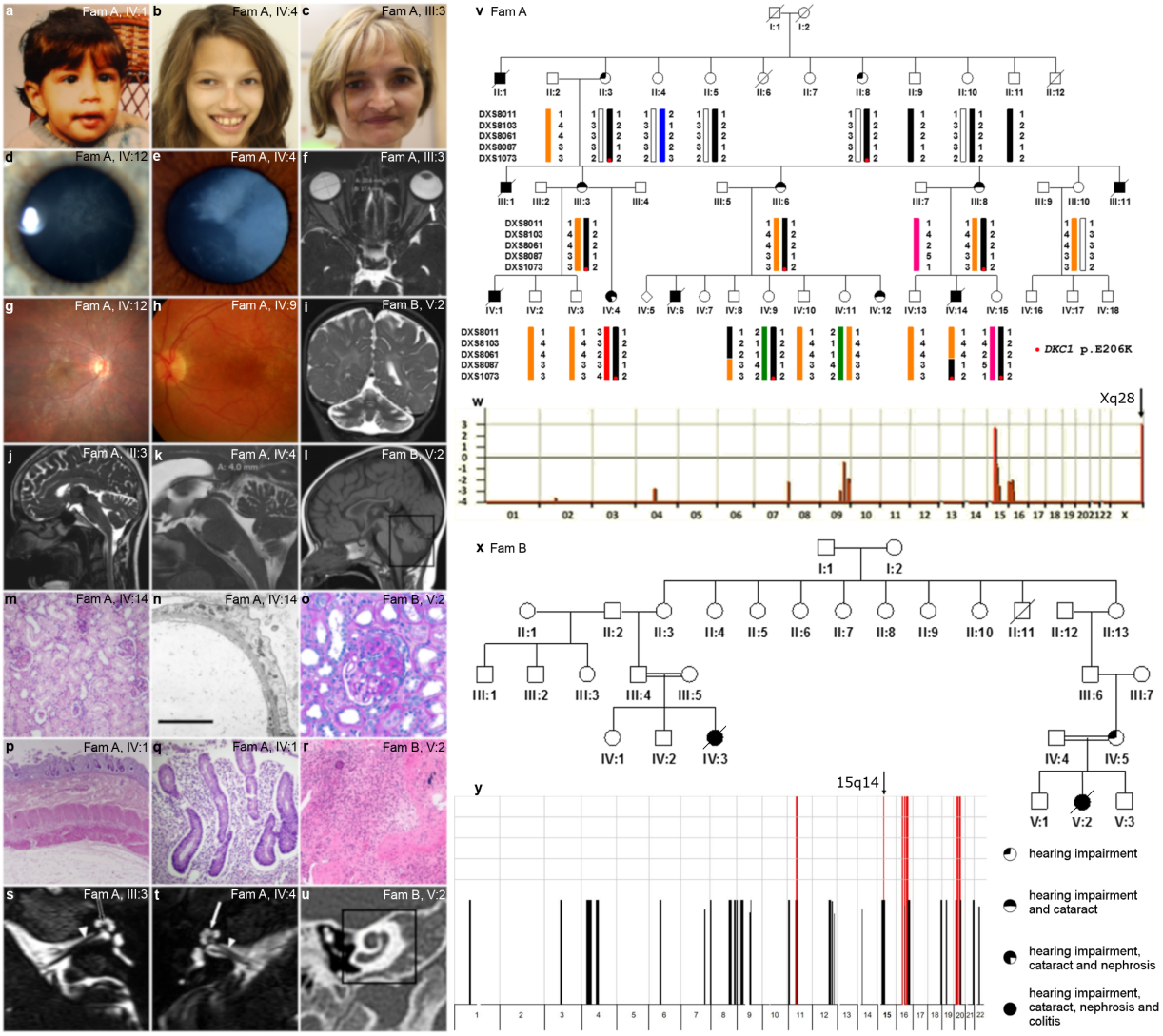
62. S. Li *et al.*, Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase. *Genes Dev* **25**, 2409-2421 (2011).
63. J. C. Shelley *et al.*, Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des* **21**, 681-691 (2007).
64. S. Pronk *et al.*, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845-854 (2013).
65. A. E. Aliev *et al.*, Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins* **82**, 195-215 (2014).
66. T. Steinbrecher, J. Latzer, D. A. Case, Revised AMBER parameters for bioorganic phosphates. *J Chem Theory Comput* **8**, 4405-4412 (2012).
67. S. Izadi, R. Anandakrishnan, A. V. Onufriev, Building Water Models: A Different Approach. *J Phys Chem Lett* **5**, 3863-3871 (2014).
68. K. Nowak *et al.*, Absence of gamma-Chain in Keratinocytes Alters Chemokine Secretion, Resulting in Reduced Immune Cell Recruitment. *J Invest Dermatol* **137**, 2120-2130 (2017).
69. G. Schay, L. Smeller, A. Tsuneshige, T. Yonetani, J. Fidy, Allosteric effectors influence the tetramer stability of both R- and T-states of hemoglobin A. *J Biol Chem* **281**, 25972-25983 (2006).
70. G. Schay *et al.*, Without Binding ATP, Human Rad51 Does Not Form Helical Filaments on ssDNA. *J Phys Chem B* **120**, 2165-2178 (2016).
71. L. E. Jao, S. R. Wente, W. Chen, Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A* **110**, 13904-13909 (2013).
72. P. Alestrom *et al.*, Zebrafish: Housing and husbandry recommendations. *Laboratory animals* 10.1177/0023677219869037, 23677219869037 (2019).
73. M. Westerfield, *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio).* (University of Oregon Press, Eugene, ed. 4th, 2000).
74. S. Krejszeff *et al.*, Dynamic of efficiency of MS-222 as an anaesthetic for tench Tinca tinca (L.) larvae. *Commun Agric Appl Biol Sci* **78**, 227-230 (2013).
75. E. Mishima *et al.*, Immuno-Northern Blotting: Detection of RNA Modifications by Using Antibodies against Modified Nucleosides. *PLoS One* **10**, e0143756 (2015).
76. P. J. Cock, C. J. Fields, N. Goto, M. L. Heuer, P. M. Rice, The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res* **38**, 1767-1771 (2010).
77. M. Pertea, D. Kim, G. M. Pertea, J. T. Leek, S. L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* **11**, 1650-1667 (2016).
78. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
79. G. la Marca *et al.*, Implementing tandem mass spectrometry as a routine tool for characterizing the complete purine and pyrimidine metabolic profile in urine samples. *J Mass Spectrom* **41**, 1442-1452 (2006).

Table 1. Clinical characteristics of the affected family members

Pt (sex)	visual inv (age at 1st cat surgery)	hearing impairment (age at HA)	nephrotic syndrome (age at pres)	enterocolitis (age at pres)	other	Age at last follow-up/ death (†)
Fam A						
II:1 (M)	cat	nd	+	+ (infancy)	FTT	† 3 yrs
II:3 (F)	-	+	-	-	-	60 yrs
II:8 (F)	glau	+	-	-	-	58 yrs
III:1 (M)	cat	nd	+	+ (6 mos)	BMF	† 2.5 yrs
III:3 (F)	cat (16 yrs), RP	+ (17 yrs)	-	-	HK, Mand, Max, MO, FTT	48 yrs
III:6 (F)	cat (14 yrs)	+ (8 yrs)	-	-	CA, Mand, Max, PH, FTT	40 yrs
III:8 (F)	cat (17 yrs)	+ (10 yrs)	-	-	FTT	45 yrs
III:11 (M)	nd	+	+	+ (6 mos)	FTT	† 4.5 yrs
IV:1 (M)	cat (2 yrs)	+ (2 yrs)	+ (>2 yrs)	+	-	† 2.5 yrs
IV:4 (F)	cat (4 yrs), RP	+ (1.5 yrs)	+ (2 yrs)	-	CA, MO, PH, FTT	15 yrs
IV:6 (M)	cat (6 mos)	+ (6 mos)	+ (2.5 yrs)	+ (6 mos)	BMF, FTT	† 2.5 yrs
IV:9 (F)	-	-	-	-	-	20 yrs
IV:12 (F)	cat (10 yrs), RP	+	PU (10 yrs)	-	-	11 yrs
IV:14 (M)	cat (3 yrs)	+ (3.5 yrs)	+ (2.5 yrs) FSGS	+ (5 yrs)	BMF, FTT	† 7.5 yrs
IV:15 (F)	-	-	-	-	-	15 yrs
Fam B						
IV:3 (F)	cat	+ (birth)	+ FSGS	+		† 3 yrs
V:2 (F)	cat	+ (birth)	+ MPGN	+	CH, HM	† 3 yrs

BMF: bone marrow failure, CA: mild cerebellar atrophy, cat: cataracts, CH: cerebellar hypoplasia, F: female, FSGS: focal segmental glomerulosclerosis, FTT: failure to thrive, glau: glaucoma, HA: indication of hearing aid, HK: hyperkeratosis, HM: hypomyelination, inv: involvement, M: male, mos: months, Mand: mandibular hypoplasia, Max: maxillary hypoplasia, MO: microphthalmia, MPGN: mesangial proliferative glomerulonephritis, nd: no data available, PH: pineal hypoplasia, pres: presentation, PU: non-nephrotic proteinuria, RP: retinitis pigmentosa, yrs: years

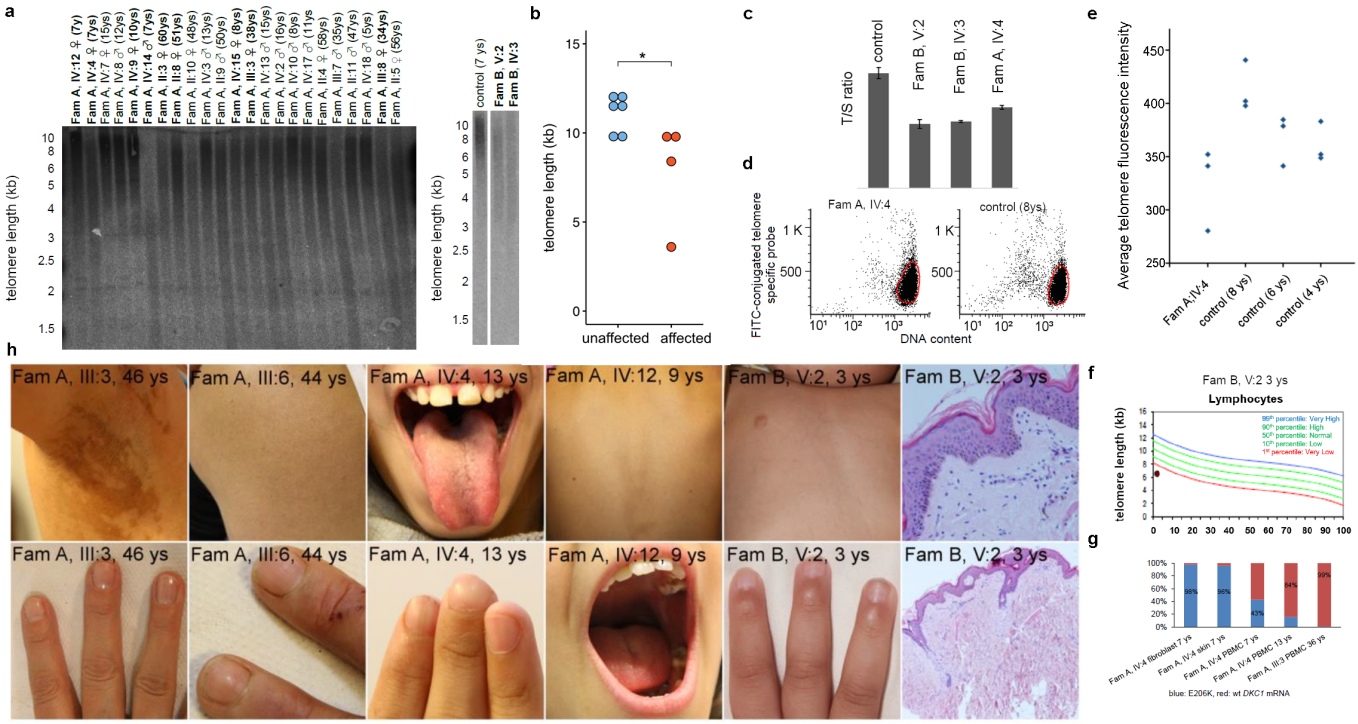
Figure 1. Phenotype and genetic identification of the two affected families



Affected (a) males (n=6) and (b-c) females (n=9) in FamA had no dysmorphic features apart from maxillary and mandibular hypoplasia in adult females. Affected males in FamA (a, v) and the two affected females (n=2) from the consanguineous FamB (x) developed nephrotic syndrome with focal segmental glomerulosclerosis (m, o), diffuse podocyte foot process effacement (n) and enterocolitis with extensive chronic nonspecific inflammation (p-r). FamB, Patient V:2, developed progressive hypomyelination (i) and

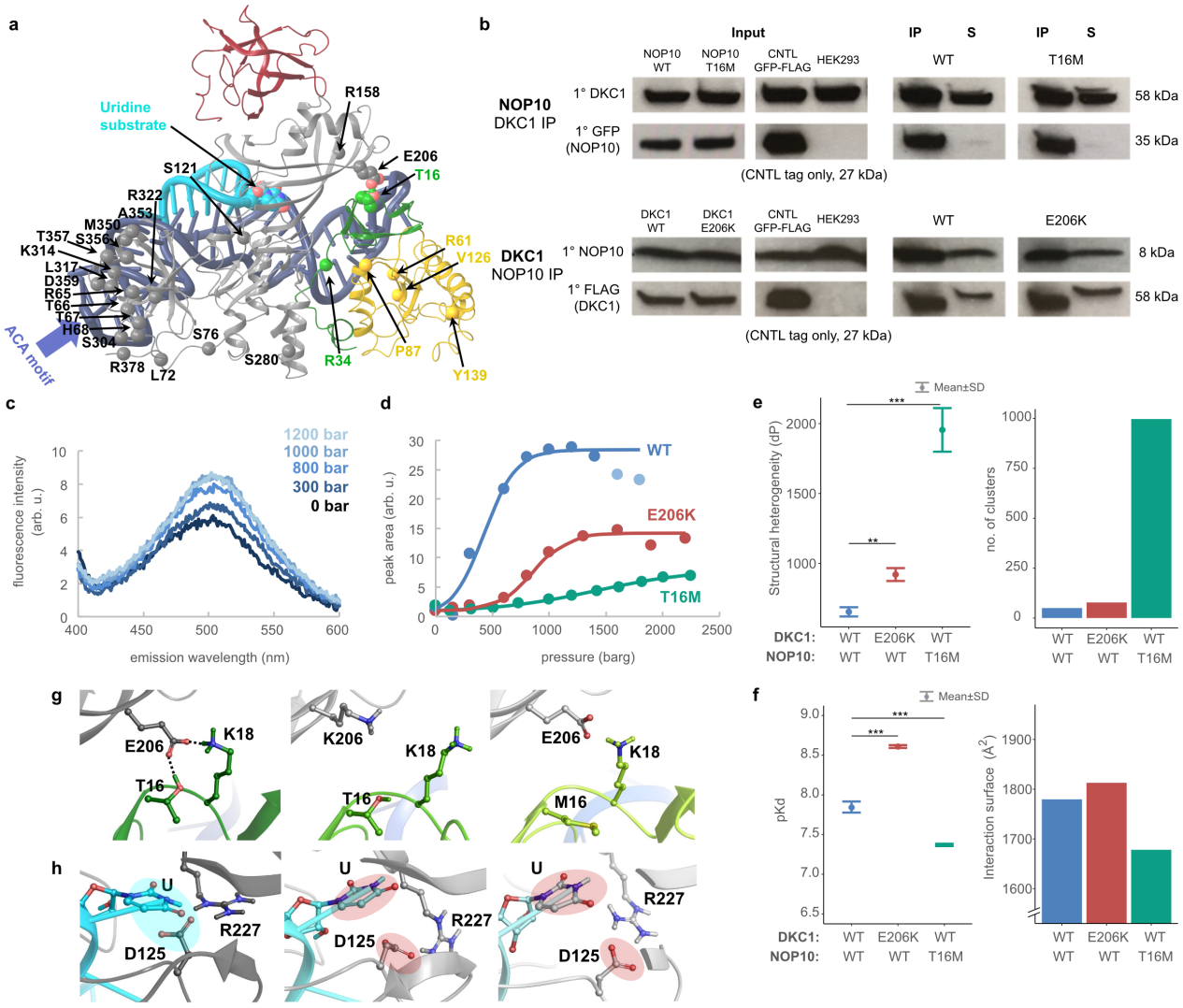
cerebellar hypoplasia (l), both IV:3 and V:2 passed away in infancy. Cataracts (d, e) and sensorineural hearing impairment with a modiolus (s: black arrow, u), cochlea (t: white arrow, u) and cochlear nerve (s, t: white arrow) of normal morphology were present, even in the affected females of FamA, who had a normal lifespan and developed pigmentary retinopathy (g-h), microphthalmia (f) and pineal hypoplasia (j-k, Figure S2). We found linkage to a 5.1 Mb region at Xq28 in the descendants of II:3 (w), assuming germline mosaicism in I:2 (v) with a LOD score of 3.01. Through targeted sequencing in an unaffected and an affected but haploidentical male (II:9 and IV:14), we found a single *de novo* mutation in *DKC1* (c.616 G>A, p.Glu206Lys), which segregated among the six haploidentical members in the generation II with disease occurrence (v). Homozygosity mapping in FamB (x) showed a haploidentical homozygous region of 2.6 Mb at 15q14 with a LOD score of 3.03 (y). Exome sequencing of patient V:2 and targeted segregation analysis of IV:3 and V:2, parents and siblings revealed a homozygous missense mutation in *NOP10* (c.47C>T, p.Thr16Met).

Figure 2. Absent or subtle symptoms of Dyskeratosis Congenita despite telomere shortening



We found telomere shortening in the affected individuals by (a) Southern blot (heterozygous individuals in FamA and affected individuals in FamB are in bold), (b) showing a significant difference between the severely affected children (FamA, IV:4, IV:14, FamB IV:3, V:2) and six age-matched controls ($p = 0.038$). Telomere attrition was also shown by (c) MM-qPCR and (d-f) Flow-FISH. (g) The severely affected index female (FamA, IV:4) showed a highly skewed X-inactivation in the fibroblasts and skin, with the ratio of her PBMCs expressing the mutant *DKC1* allele decreasing with age. FB: fibroblast; blue:E206K, red:wt *DKC1* mRNA. (h) Despite telomere shortening, no nail dystrophy or leukoplakia was observed, with one heterozygous female from FamA (III:3) and one female from FamB (V:2) being diagnosed with mild dyskeratosis after genetic diagnosis. ys: years

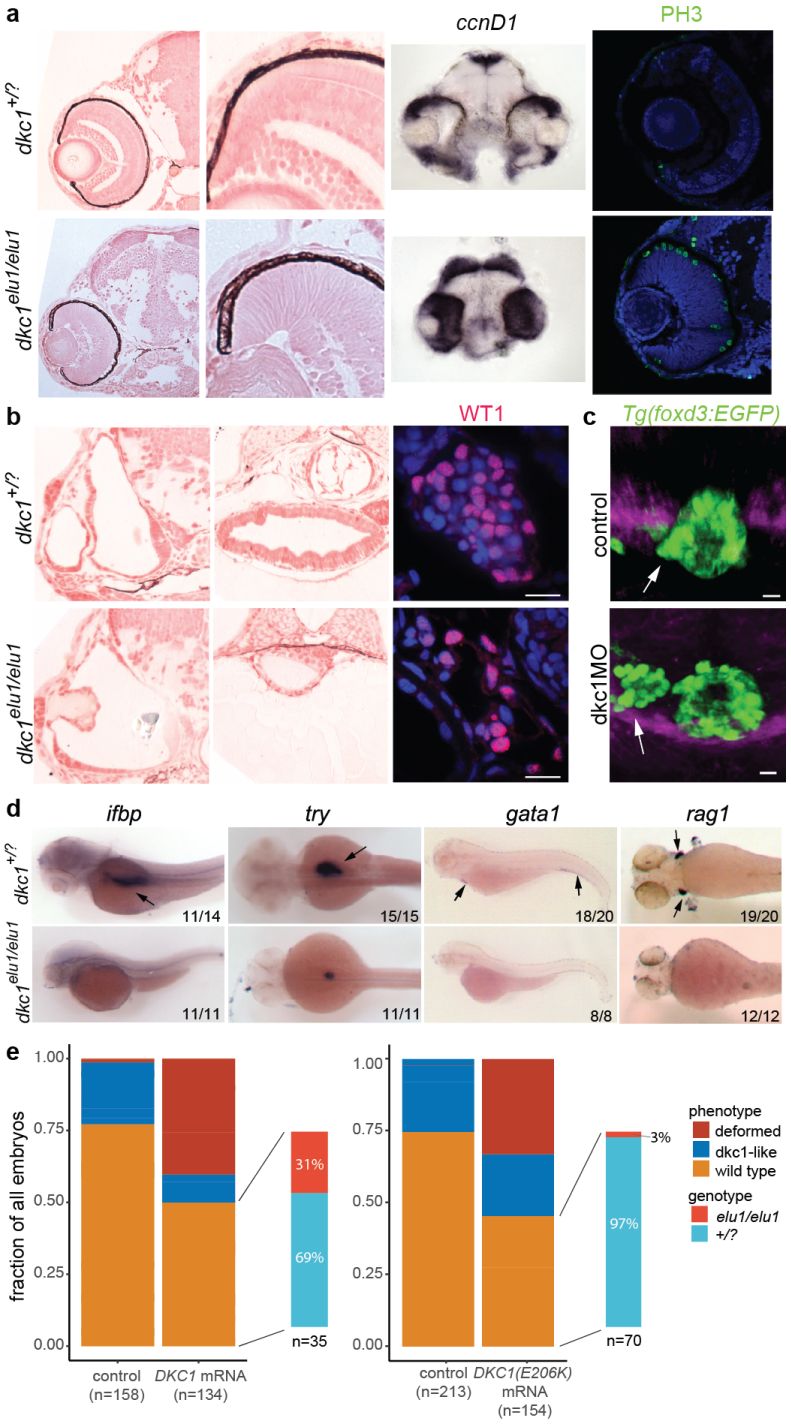
Figure 3. Dyskerin p.Glu206Lys and NOP10 p.Thr16Met mutations alter the pseudouridylation pocket of the H/ACA snoRNP complex



(a) Homology modelling of human H/ACA snoRNP dyskerin (amino acids 60-380), grey; NOP10, green; GAR1, red; NHP2, yellow; guide snoRNA, dark blue; substrate RNA, light blue. The Glu206 (dyskerin) and Thr16 (NOP10) are in CPK format (carbon atoms are shown in the colour of the backbone, oxygen in red and nitrogen in blue). The C-alpha atoms of residues associated with Dyskeratosis Congenita and Hoyeraal-Hreidarsson syndrome are shown in spheres. (b) Co-Immunoprecipitation of native dyskerin from HEK293

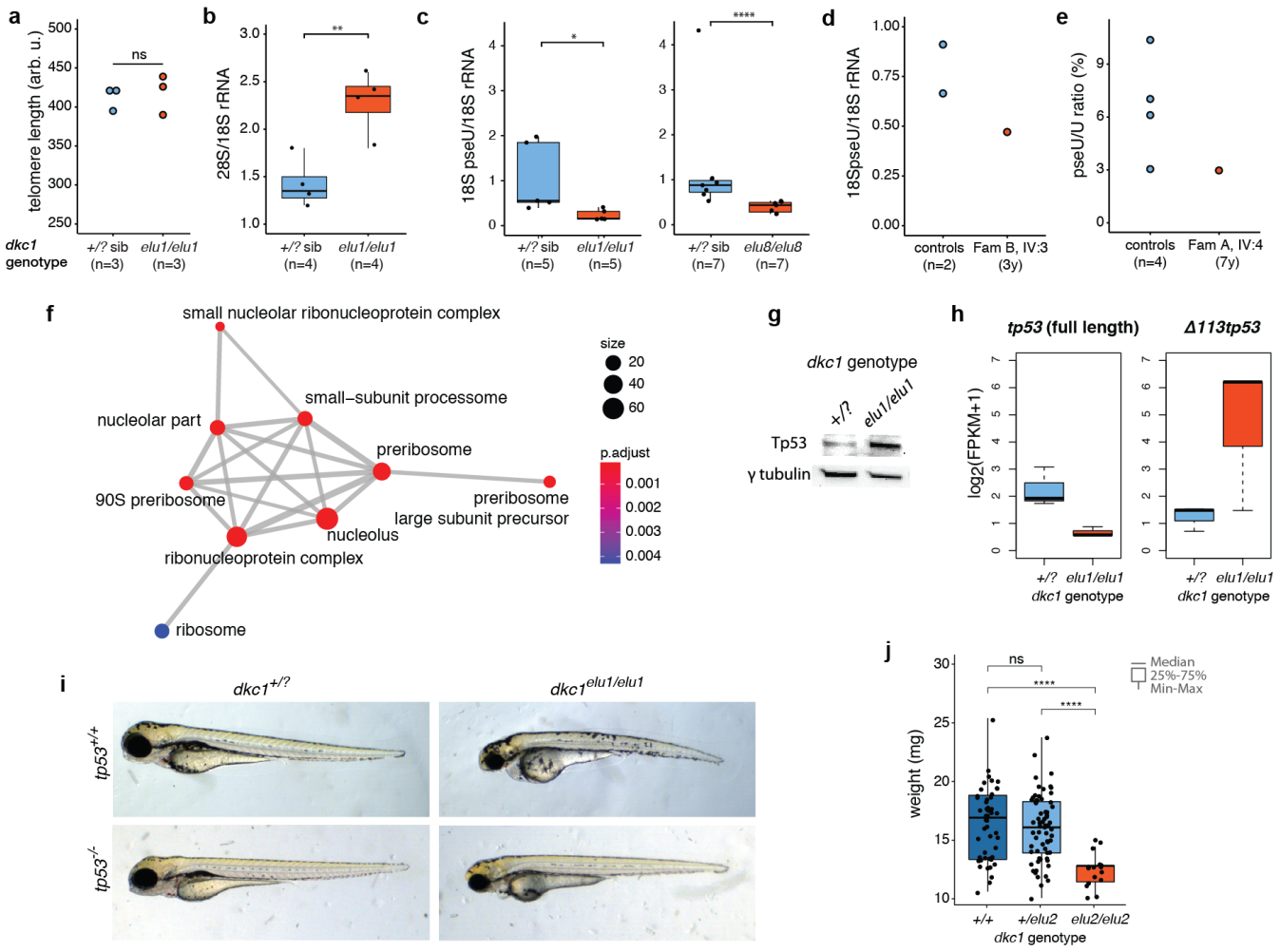
cells transfected with GFP-tagged WT and T16M NOP10 and reciprocal immunoprecipitation of native NOP10 from HEK293 cells transfected with Flag-tagged WT and E206K dyskerin. Immunoblots show that both mutant proteins immunoprecipitate with their native counterpart (see also Figure S4). (c-d) Pressure-tuning fluorescence spectra of WT and mutant dyskerin-NOP10 complexes, where mutant complexes show an altered stability. (e) Structural heterogeneity indicates significant structural difference between mutants and WT (Welch-test: DKC1, $p = 1.8x10^{-2}$; NOP10 $p = 4.04\times10^{-11}$). (f) Disassociation constants of WT and mutant complexes show both mutations cause a parallel change in pKd (left, Welch-test: DKC1, $p = 9.54\times10^{-3}$; NOP10, $p = 1.55 \times 10^{-7}$) and the interaction surface (right). (g-h) Conformational changes induced by the mutations (g) WT dyskerin Glu206 interacts with the NOP10 Thr16-Leu17-Lys18 segment forming hydrogen-bonds (left), these H-bonds are disrupted by both the dyskerin E206K (middle) and NOP10 T16M (right) mutations. (h) The WT interaction (left) between the substrate uridine and the catalytic D125 of dyskerin is uncoupled by both dyskerin E206K (middle) and NOP10 T16M (right). IP, immunoprecipitated protein; S, supernatant; CNTL, cells transfected with empty expression plasmid, expressing GFP-Flag, size of 28kDa; HEK293, non-transfected cells; and 1°, primary antibody. Statistical significance denoted by asterisks.

Figure 4. The phenotype of *dkc1*$^{elu1/elu1}$ larvae recapitulates the human phenotype



(a) Histological analysis of *dkc1*$^{elu1/elu1}$ mutant larvae shows microphthalmia and cataracts. Both the eyes and the optic tectum of the mutants are abnormal and contain a high prevalence of cells with neuroepithelial character. Expression of cell-cycle markers *ccnD1* and PH3 in the retinae and the tecta of 2 dpf and 3 dpf larvae, respectively, can be observed throughout these tissues instead of being restricted to the proliferative regions of the ciliary marginal zone and the mediolateral edges, suggesting defective cell cycle. (All pictures show coronal sections.) (b) Further histological analysis shows i) deformed semicircular canals, ii) undifferentiated gut iii) and hypoplastic pronephros with a reduced number of WT1-positive podocytes in the mutant animals (scale bar = 10 μm). (c) When Dkc1-function is abrogated in *Tg(foxd3:EGFP)* animals using a synthetic MO oligo, parapineal migration is impaired and the pineal-parapineal complex appears immature at 3 dpf. (White arrows denote the parapineal). (d) Markers of tissue differentiation demonstrate a lack of differentiation in the intestines (*ifbp*), pancreas (*try*) and the major blood lineages (*gata1* and *rag1*). (Black arrows denote area of expression.) (e) Injection of i) human WT *DKC1* mRNA resulted in phenotypic rescue of the mutant larvae, as shown by the genotyping of larvae showing a WT phenotype. In contrast, injection of ii) human Glu206Lys *DKC1* mRNA elicited a much milder rescue, demonstrating the hypomorphic nature of this allele.

Figure 5. Ribosomal dysfunction in *dkc1* zebrafish mutants due to defective pseudouridylation



(a) Telomere length is normal in *dkc1*$^{elu1/elu1}$ larvae at 4dpf as measured by flow-FISH (n = 3 pooled samples of 10 larvae each, p = 0.7). (b) The 28S/18S rRNA ratio is increased in 4 dpf *dkc1*$^{elu1/elu1}$ larvae, suggesting impaired 18S rRNA processing (p = 0.0033). (c-d) Immunonorthern-blot demonstrates a reduced pseudouridylation of 18S rRNA in *dkc1*$^{elu1/elu1}$, *dkc1*$^{elu8/elu8}$ 4 dpf larvae (*+/?* vs. *elu1/elu1*: p = 0.016, *+/?* vs. *elu8/elu8*: p = 0.00058) (c) and in the leukocytes of patient FamB IV:3 (d). (*+/?:* heterozygous or homozygous WT fish). (e) The female with skewed X-inactivation shows a decreased PU/U ratio in the leukocytes as determined by HPLC-MS. (f) Gene ontology analysis of differentially regulated genes from

36 hours post fertilization (hpf) *dkc1*$^{elu1/elu1}$ larvae demonstrates an upregulation of genes associated with ribosome assembly and function. Size of the circles indicate the number of genes associated with certain terms, color indicates the level of enrichment: red indicates high enrichment, blue indicates low. (g) Western blot suggests the stabilisation of Tp53 in the affected cells. (h) Transcriptomic analysis shows that the truncated, anti-apoptotic *tp53* isoform (*Ä113p53*) is upregulated in mutants, while the canonical, full-length, pro-apoptotic isoform shows decreased expression; measured as FPKM. (i) The phenotype of the *dkc1*$^{elu1/elu1}$ zebrafish mutants is Tp53 independent, as it is not rescued on a *tp53*$^{-}$ background. (j) Homozygous carriers of the missense (c.567_568insGTG) hypomorphic allele (*dkc1*$^{elu2/elu2}$) are viable, but show significant growth retardation compared with their siblings (n = 130) (*+/+* vs. *elu2/elu2*: $p = 1.9 \times 10^{-9}$, *+/elu2* vs. *elu2/elu2*: $p = 1.6 \times 10^{-9}$).

PNAS
www.pnas.org

Supplementary Information Appendix for

Pseudouridylation defect due to *DKC1* and *NOP10* mutations cause nephrotic syndrome with cataracts, hearing impairment and enterocolitis

Eszter Balogh, Jennifer C. Chandler, Máté Varga, Mona Tahoun, Dóra K. Menyhárd, Gusztáv Schay, Tomas Goncalves, Renáta Hamar, Regina Légrádi, Ákos Szekeres, Olivier Gribouval, Robert Kleta, Horia Stanescu, Detlef Bockenhauer, Andrea Kerti, Hywel Williams, Veronica Kinsler, Wei-Li Di, David Curtis, Maria Kolatsi-Joannou, Hafsa Hammid, Anna Szőcs, Kristóf Perczel, Erika Maka, Gergely Toldi, Florentina Sava, Christelle Arrondel, Magdolna Kardos, Attila Fintha, Ahmed Hossain, Felipe D'Arco, Mario Kaliakatsos, Jutta Koeglmeier, William Mifsud, Mariya Moosajee, Ana Faro, Eszter Jávorszky, Gábor Rudas, Marwa H. Saied, Salah Marzouk, Kata Kelen, Judit Götze, George Reusz, Tivadar Tulassay, François Dragon, Géraldine Mollet, Susanne Motameny, Holger Thiele, Guillaume Dorval, Peter Nürnberg, András Perczel, Attila J Szabó, David A. Long, Kazunori Tomita, Corinne Antignac, Aoife M. Waters, Kálmán Tory

Máté Varga, mvarga@ttk.elte.hu

Aoife Waters, aoife.waters@gosh.nhs.uk

Kálmán Tory, tory.kalman@med.semmelweis-univ.hu

This PDF file includes:

Figures S1 to S9

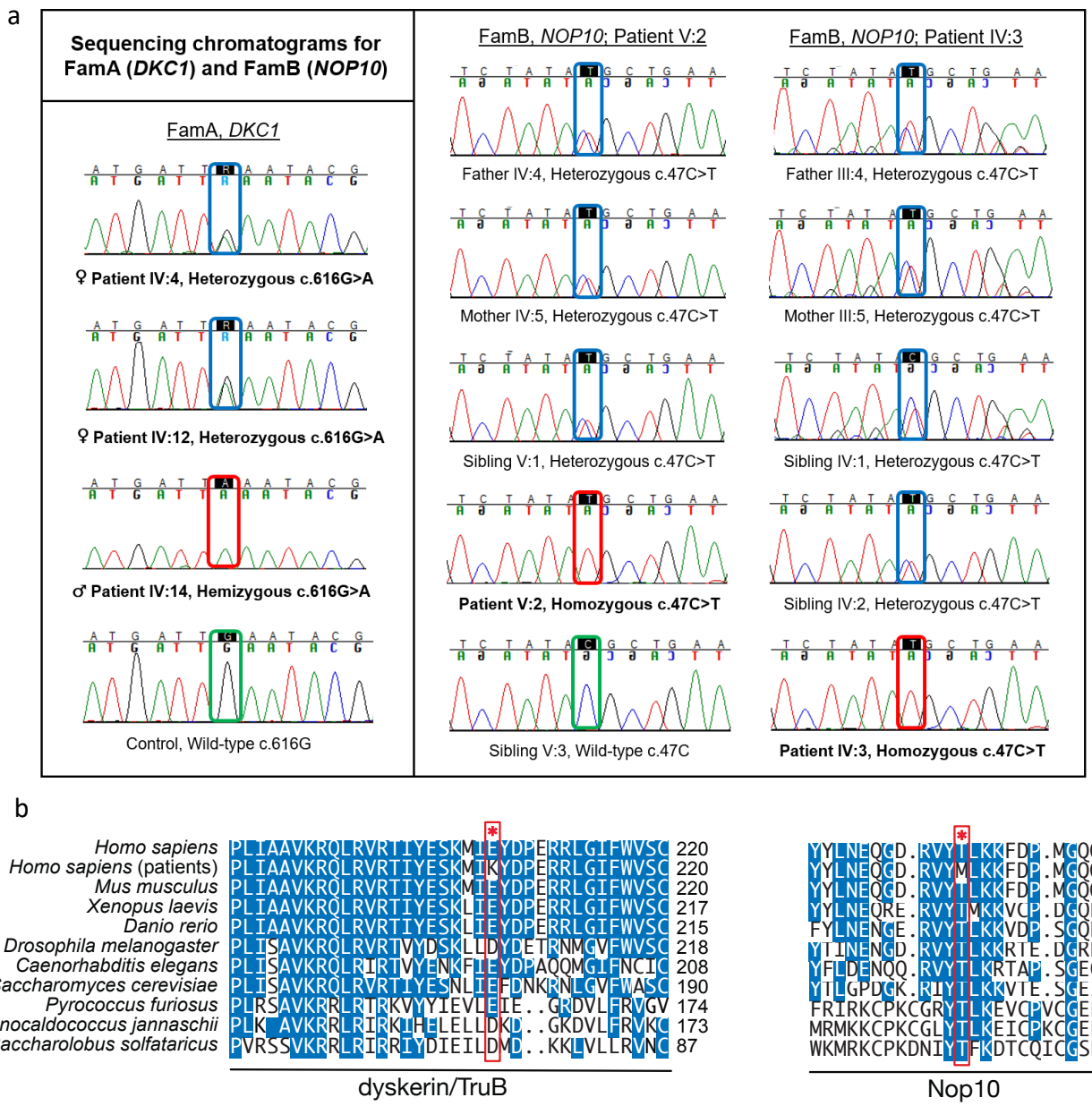
Supplementary methods

Table S1

SI Web Resource

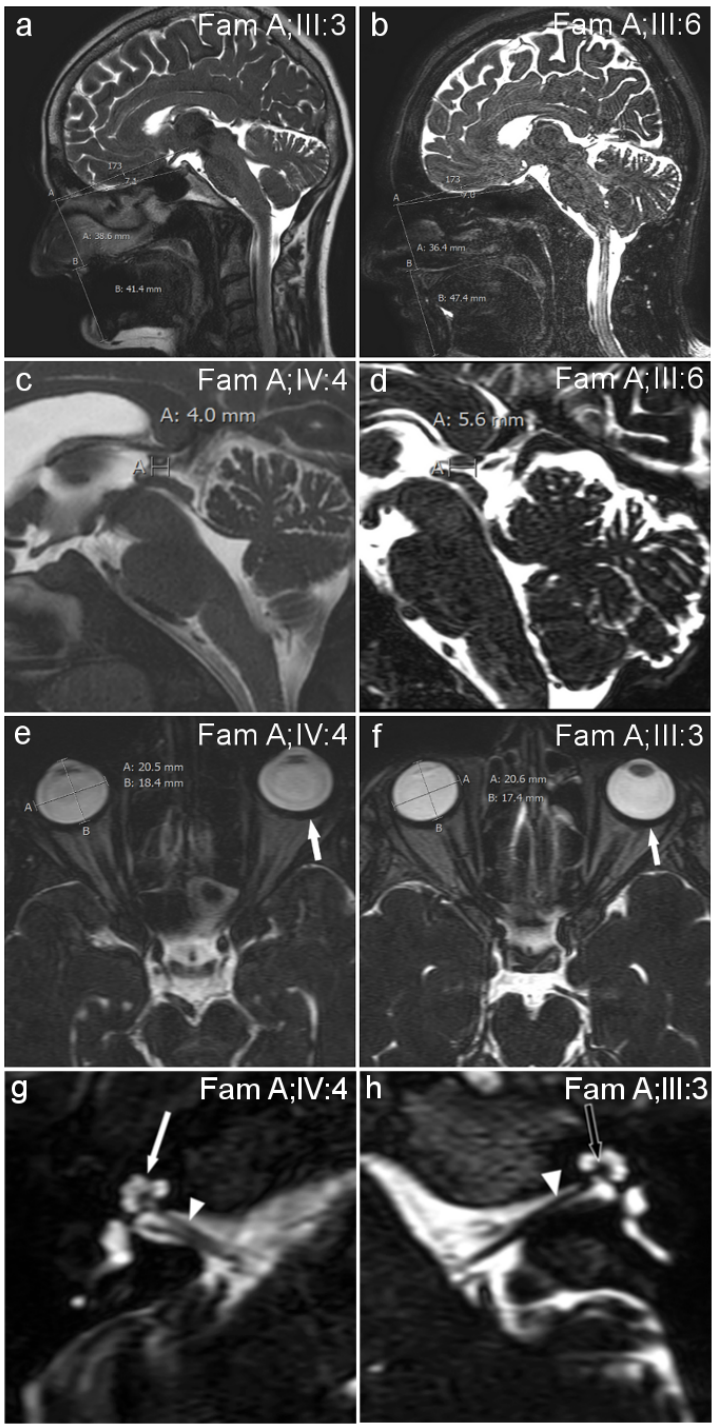
SI References

Figure S1. Sequencing chromatograms and evolutionary conservation of the identified mutations.



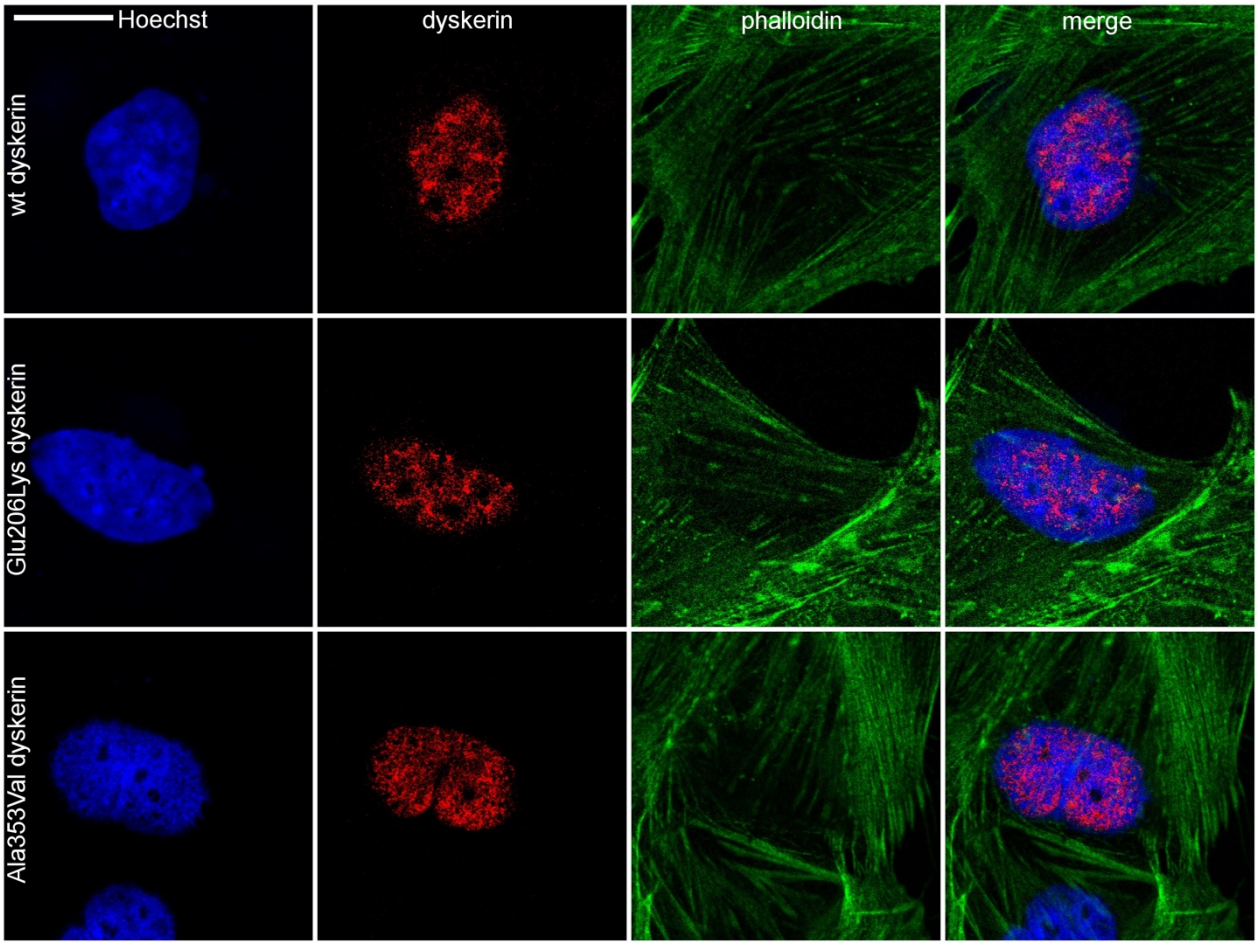
(a) Sequencing chromatograms of the *DKC1* c.616 G>A, p.Glu206Lys (FamA) and the *NOP10* c.47C>T, p.Thr16Met (FamB) mutations. Affected individuals are in bold. (b) Evolutionary conservation of DKC1 Glu206 and NOP10 Thr16.

Figure S2. Phenotypic features of females with heterozygous *DKC1* p.Glu206Lys mutation.



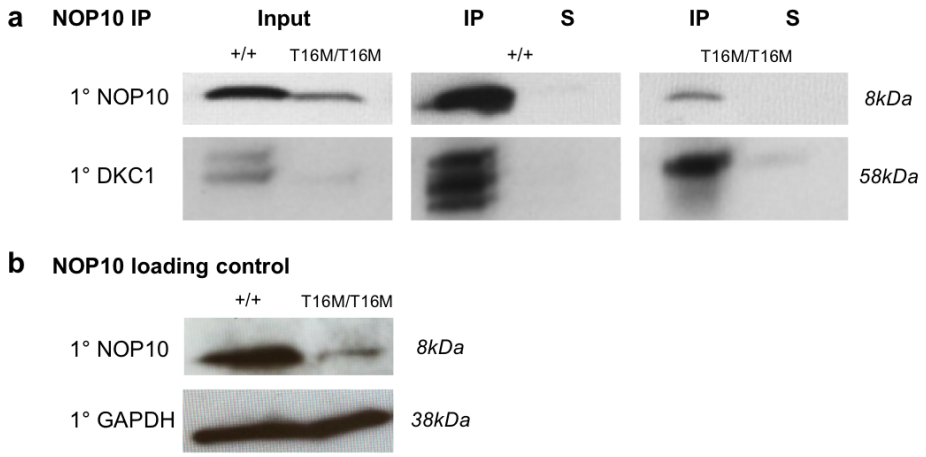
(a-b) Both the middle and lower facial heights are shortened on the sagittal T2 DRIVE MRI sequences in two adult females indicating maxillary and mandibular hypoplasia (reference values for the middle and lower third facial heights: 50 ± 2.4 mm and 61.3 ± 3.3 mm (mean ± SD)). (c-d) The sagittal T2 DRIVE images demonstrate subnormal dimensions of the pineal glands. Also note the cerebellar vermis, which – in contrast to HH syndrome – shows no signs of hypoplasia, but mild atrophy. (e-f) Microphthalmia was found on the axial T2 DRIVE images (reference values for the transverse (A) and anteroposterior (B) diameters: 24.156 ± 1.9 mm and 23.422 ± 1.9 mm (median ± SD), respectively) with mildly thickened sclera (arrows) (1). (g-h) The axial T2 DRIVE images show the internal acoustic canals of two females with impaired hearing. The cochlea (white arrow) shows normal dimensions and structure, the modiolus (black arrow) and cochlear nerve (arrowheads) are of normal appearance.

Figure S3. The p.Glu206Lys mutation does not alter the localization of dyskerin.



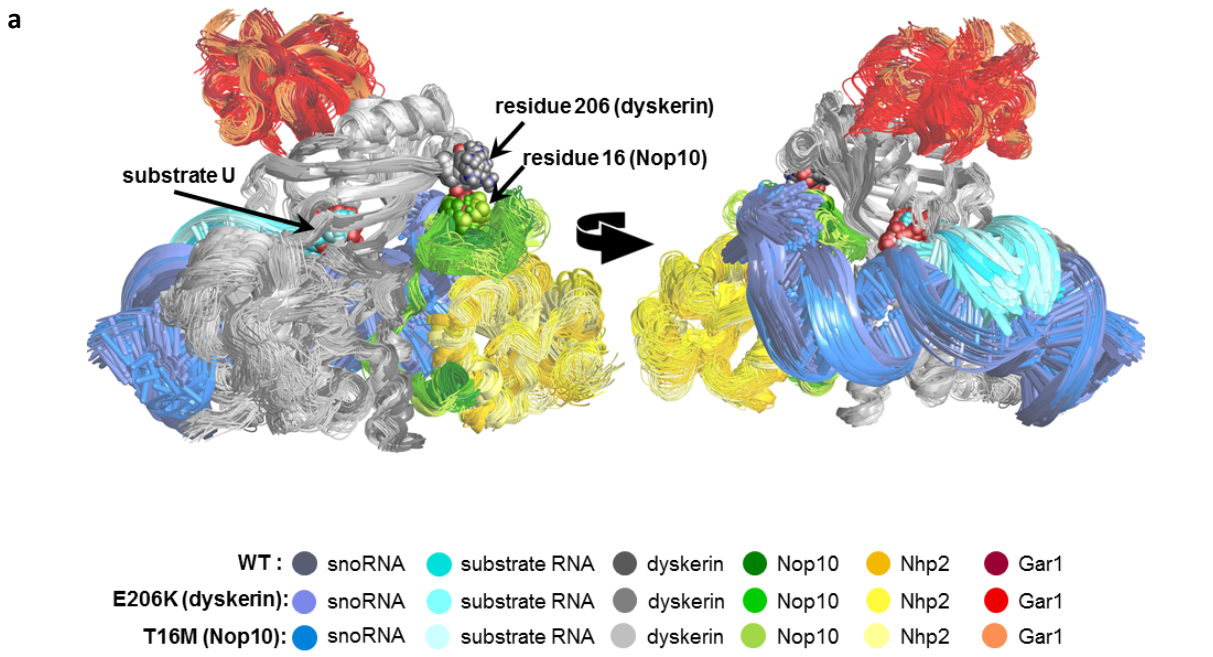
The nuclear localization of Flag-tagged p.Glu206Lys dyskerin is similar to that of the wild-type dyskerin and the dyskerin mutant most frequently described in dyskeratosis congenita (p.Ala353Val); scale bar = 5 μm

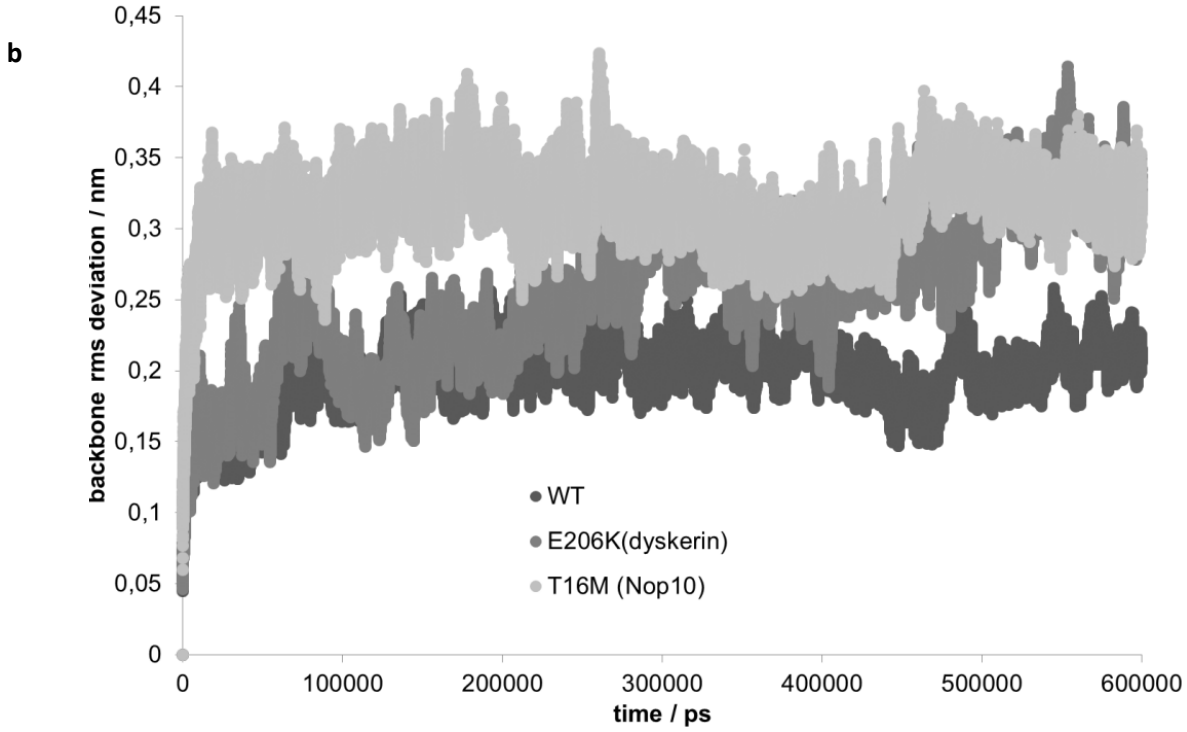
Figure S4. Co-Immunoprecipitation of NOP10 from endogenous keratinocyte protein extracts from FamB V:2 (T16M/T16M, ♀ at 3 years) and age-matched control (♀ at 5 years).



(a) Primary antibody probing for NOP10 (1° NOP10) and DKC1 (1° DKC1) shows that the p.Thr16Met NOP10 immunoprecipitates with DKC1, thus the NOP10 mutation does not result in a loss of DKC1 binding *in vivo*. (b) Loading control with Input keratinocyte protein, showing primary antibody probing for NOP10 (1° NOP10) and GAPDH (1° GAPDH). Labels as follows: Input, total extracted protein; IP, immunoprecipitated protein; S, supernatant; sizes indicated in kDa.

Figure S5. Details of the molecular dynamics simulations of wild-type (WT), E206K (dyskerin) and T16M (NOP10) mutants of pseudouridinase box H/ACA snoRNPs.





c

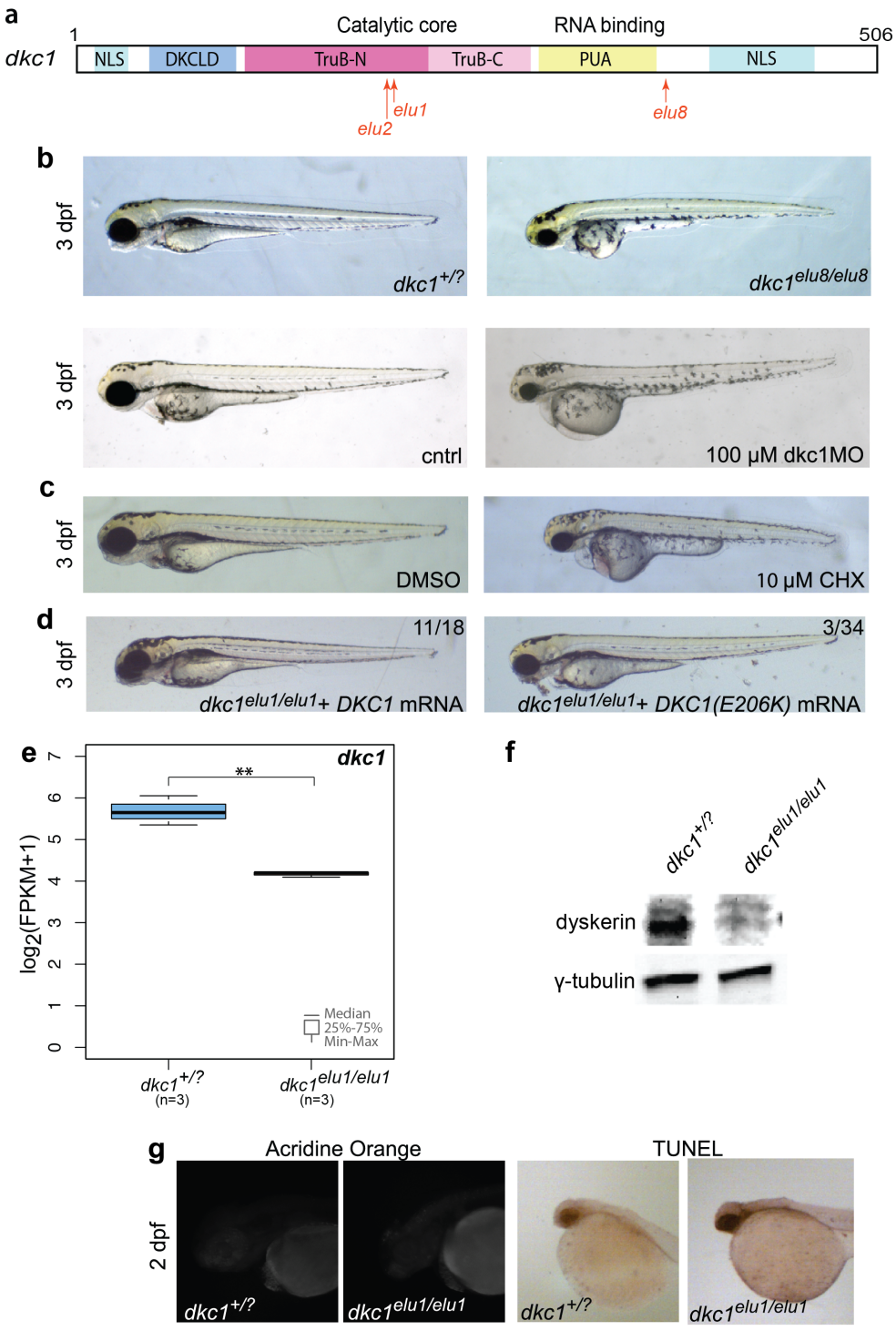
clustering trajectories based on:		WT	DKC1 (E206K)	Nop10 (T16M)
backbone of all 4 proteins of the snoRNP	total	68	91	1415
	clusters accounting for 90% of snapshots	20	27	613
dyskerin/Nop10 interaction surface	total	50	78	997
	clusters accounting for 90% of snapshots	14	25	438

d

WT	
dyskerin / Nop10 contact site:	
E206 (DKC1) -	- T16 (Nop10): 23.2%
	- L17 (Nop10): 22.4%
	- K18 (Nop10): 54.2%
	min 1 contact to res(16-18): 63.5%
dyskerin active site:	
D125(CG)-$U_{substrate}$(N1) distance: 4.9 $\pm$ 0.2 Å	
D125 (OD1, OD2)-	- $U_{substrate}$ (C2'): 99.5% (< 5Å)
	- $U_{substrate}$ (N3): 99.8%
DKC1(E206K)	
dyskerin / Nop10 contact site:	
K206 (DKC1) -	min 1 contact to res(16-18): 0.1%
	- Ile205 (DKC1): 1.3%
	- Tyr207 (DKC1): 10.4%
dyskerin active site:	
D125(CG)-$U_{substrate}$(N1) distance: 5.5 $\pm$ 0.5 Å	
D125 (OD1, OD2)-	- $U_{substrate}$ (C2'): 5.8% (< 5Å)
	- $U_{substrate}$ (N3): 29.0%
Nop10(T16M)	
dyskerin / Nop10 contact site:	
E206 (DKC1) -	- M16 (Nop10): 0%
	- L17 (Nop10): 1.5%
	- K18 (Nop10): 41.3%
	min 1 contact to res(16-18): 42.6%
dyskerin active site:	
D125(CG)-$U_{substrate}$(N1) distance: 6.9 $\pm$ 0.5 Å	
D125 (OD1, OD2)-	- $U_{substrate}$ (C2'): 7.3% (< 5Å)
	- $U_{substrate}$ (N3): 1.8%

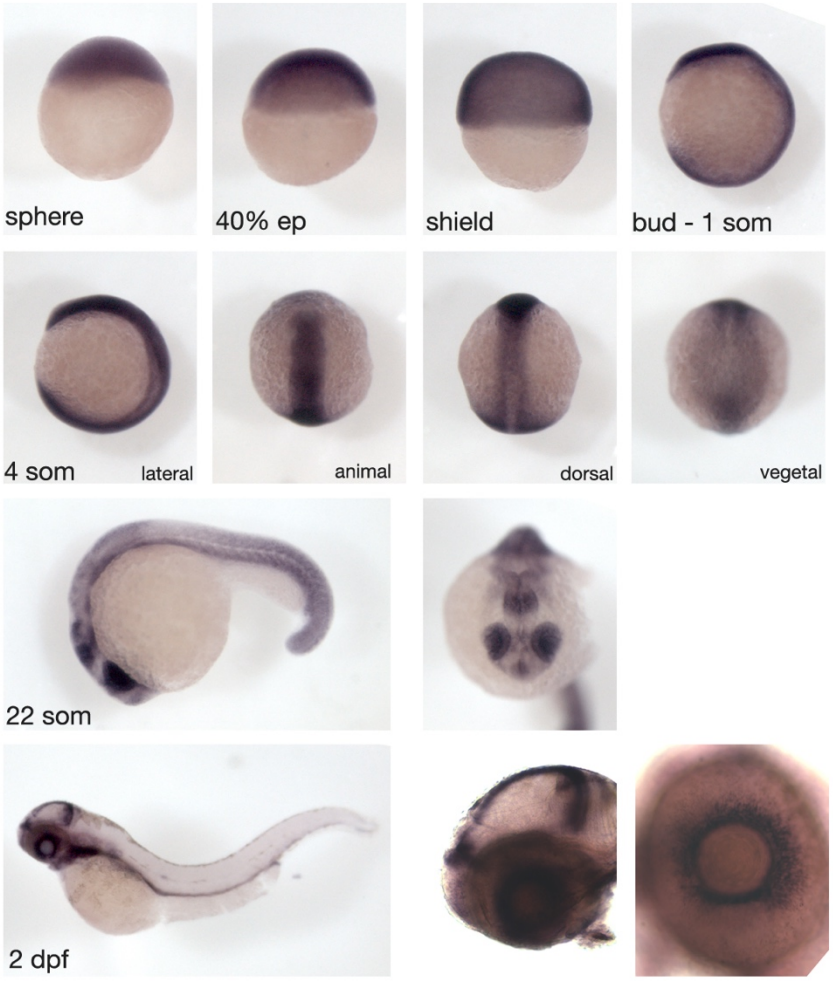
(a) Comparison of the most populated clusters sampled during the last 300 ns of the three simulations. (b) Root-mean-square deviation of backbone atoms of the protein components from the starting structure along the trajectories. (c) Results of clustering (using a 1 Å cutoff) of the equilibrium trajectories (300-600 ns), based on the backbone atoms of the protein components of the H/ACA snoRNPs and the interaction surface between the dyskerin and NOP10 proteins (containing the entire NOP10 model and residues of dyskerin that have an atom reaching within 6 Å of NOP10). (d) H-bonds and protein contacts between the 206 residue of dyskerin and the facing residues of NOP10 (% of snapshots in which the donor-acceptor distance is <3.3 Å, if not otherwise indicated) and some characteristic contacts between Asp125 of dyskerin and the uridine (U) of the coordinated substrate at the active site.

Figure S6. Characterization of *dkc1* loss-of-function.



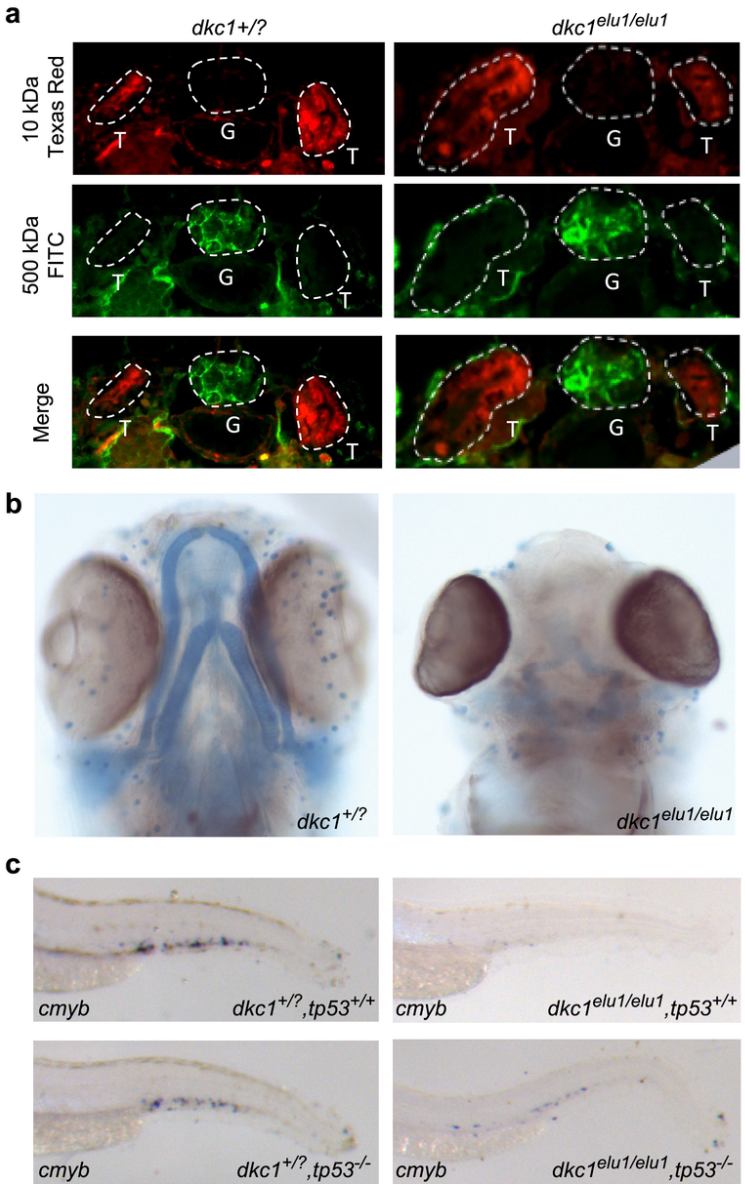
(a) Domain structure of the *dkc1* CDS, highlighting the position of the generated mutations (elu1 = c.566_567insTCATGGT, p.Arg190Hisfs*14; elu2 = c.567_568insGTG, p.Leu189_Arg190insVal; elu8 = c.1117_20delCGCA, p.Arg373Asnfs*13). (b) The Dkc1 deficient fish develop smaller eyes and head, bloated fourth ventricle and pericardial edema. The phenotype of the *elu1* mutant is identical to the one observed in the *elu8* mutants and the morphants. (c) A phenocopy of the $dkc1^{elu1/elu1}$ phenotype can also be achieved by blocking the translation in 36 hpf wild-type (WT) embryos with cycloheximide (CHX). (d) Representative embryos, phenotypically rescued with WT or E206K *DKC1* mRNA at the indicated ratios are shown. (e) The level of *dkc1* transcripts is downregulated in 36 hpf $dkc1^{elu1/elu1}$ embryos ($p = 0.0073$) according to whole transcriptome analysis, suggesting the activation of nonsense mediated decay. (f) Western blot analysis indicates that dyskerin levels are reduced in 4 dpf $dkc1^{elu1/elu1}$ larvae. (g) No major cell death can be detected by either acridine orange or TUNEL staining in mutants.

Figure S7. The expression of *dkc1* in wild-type zebrafish larvae.



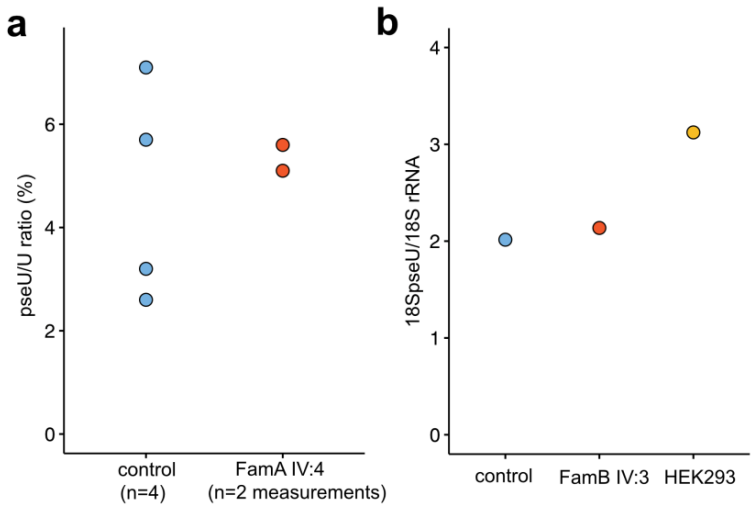
During early development at the somite stage *dkc1* is ubiquitously expressed, whereas by 2 dpf expression was limited to the proliferative areas of the eyes and optic tecta.

Figure S8. Additional phenotyping of *dkc1*$^{elu1/elu1}$ larvae



(a) No altered glomerular leakage was observed in 5 dpf mutant larvae upon the co-injection of 500kDa FITC-labeled dextran and 10kDa Texas Red-labeled dextran into the cardinal vein, as the 500kDa dextran was not visible in the tubules. T: tubules, G: glomeruli (b) Severe microphthalmia and absence of jaw cartilage as shown by alcian blue staining at 5 dpf. (c) Hematopoietic defects observed in *dkc1*$^{elu1/elu1}$ larvae were partially rescued on a *tp53*$^{-/-}$ background as shown by *cmyb* staining.

Figure S9. The pseudoU/U ratio in the skin and the pseudouridylation of fibroblast 18S rRNA remain unaltered.



(a) The pseudoU/U ratio in the skin of the index girl with skewed X-chromosome inactivation (FamA, IV:4) is similar to that of four age-matched controls. (b) The pseudouridylation of 18S rRNA in the fibroblasts of patient FamB, IV:3 is similar to that of an age-matched control.

Supplementary Methods

Telomere length analysis

a) Southern blot

For each sample, 2.5 µg of gDNA was digested with HinfI and RsaI. The digested gDNA fragments were separated on a 0.8% agarose gel and digestion was confirmed by ethidium bromide staining. The gDNA was then transferred to a nitrocellulose membrane via capillary action and probed overnight with an α-32P-labelled oligonucleotide for human telomere DNA (from an EcoRI-digested pKazu-hTelo plasmid (55xTTAGGG repeats, cloned into pCR4 TOPO-blunt vector)). Similarly, a DNA marker (Hyperladder™ 1kb) was probed with a *Schizosaccharomyces pombe* telomere DNA, delivered from EcoRI & BamHI-digested p1742 plasmid. The estimated median length of telomere restriction fragments was calculated using Image J with Excel and TeloMetrix (2). Median telomere length of the four severely affected children (FamA, IV:4, IV:14, FamB IV:3, V:2) and six age-matched controls (FamA IV:8, IV:10, IV:17, IV:18 and two patients with *NPHS2*-related nephrotic syndrome) was compared by Mann-Whitney U test.

b) Monochrome multiplex quantitative PCR (MM-qPCR)

The MM-qPCR analysis was performed as previously described (3). with some alterations. In brief, 20ng of gDNA, SYBR green (0.75x) and primers (300 nM) for human telomerase (hTelo) and the single copy gene (Albumin) was run in triplicate; see Table S1 for primer sequences. In addition, five concentrations of reference gDNA from HEK293T (ATCC® CRL-3216™), were used to generate a standard curve for relative quantitation. T/S ratio was then calculated as T (the number of nanograms of the standard DNA that matches the experimental sample for copy number of the telomere template) divided by S (the number of nanograms of the standard DNA that matches the experimental sample for copy number of albumin). Telomere copy number is presented as fold-enrichment against the reference HEK293T standard. If T/S >1.0, the sample has an average telomere length greater than that of the reference HEK293T cells.

c) Flow-Fluorescence *in-situ* hybridization (FISH)

For FamA, peripheral blood mononuclear cells (PBMCs) were isolated from peripheral blood samples and washed as described earlier (4). Cells were stained using the 'Telomere PNA Kit/FITC for Flow Cytometry' (Code K5327, Dako, Glostrup, Denmark) following the manufacturer's instructions (5). Cells were analyzed on a BD FACSAria flowcytometer (BD Biosciences, San Jose, CA, USA). Following the gating of viable

lymphocytes in the G0/1-phase, mean fluorescence intensity was established in samples hybridized with the FITC-conjugated telomere specific DNA probe. Telomere length was determined by the subtraction of cell autofluorescence from the fluorescence of telomere specific stained samples. For FamB, V:2, clinical assessment of telomere length was conducted by Flow-FISH on lymphocyte subset cells in the diagnostic lab Repeat Diagnostics Inc. (Vancouver, Canada).

Generation of dyskerin and NOP10-coding plasmids

The pRSF_DKC1 construct encoding wild-type (WT) human dyskerin has been described previously (6). A Flag-tag was inserted to produce an N-terminal fusion protein. The Flag-dyskerin encoding cDNA was cloned into the EcoRI-NotI sites of pcDNA 3.1 Zeo (+) vector. A GFP-tagged human NOP10 encoding vector (Origene, #NM_018648) was used for co-immunoprecipitation analysis. To produce NOP10 for pressure tuning fluorescence spectroscopy measurements, the NOP10-encoding cDNA was cloned into pcDNA3.1/V5-His-TOPO (Thermo Fisher Sc.) at the HindIII-XhoI sites to produce NOP10-V5-His. The *DKC1* c.616 G>A and the *NOP10* c.47C>T mutations were generated by site-directed mutagenesis (QuikChange II Site-Directed Mutagenesis Kit (Agilent, #200518)). All primers are listed in Table S1. All constructs were verified by Sanger sequencing.

Expression of dyskerin and NOP10 variants

Human embryonic kidney cells (HEK293) were grown at 37°C, 5% CO_2 in DMEM media and 10% FBS (Gibco, Thermo Fisher) and 1% penicillin/streptomycin (Gibco, 15140122) as previously described (7). For pressure tuning fluorescence spectroscopy measurements, cells were transiently transfected using CalPhos Mammalian Transfection Kit (Clontech) with vectors encoding Flag-tagged dyskerin or V5-tagged NOP10 variants. For co-immunoprecipitation analysis, cells were transfected with Flag-tagged dyskerin or GFP-tagged NOP10 variants using polyethylenimine (PEI). A control plate was sham-transfected with PEI and DPBS. Cells were incubated for 48h post transfection in both experiments.

Analysis of the subcellular localization of dyskerin

A human immortalized podocyte cell line (AB8/13) provided by M. Saleem (University of Bristol, UK) was cultured on type I collagen–coated coverslips and transfected with plasmids encoding Flag-tagged WT, p.Glu206Lys, dyskerin using FuGENE HD according to the manufacturer's instructions (Promega FUGENE

HD Transfection reagent, E2311, Thermo Fisher Sc.) and incubated for 48 hrs. Transfected podocytes were fixed in ice cold 4% PFA/100% ethanol for 5 minutes, blocked with PBS-BSA 1% for 15 minutes, incubated with rabbit anti-flag primary antibodies (F2555, Sigma-Aldrich, 1:500) for 1 hour, followed by Alexa Fluor 647 conjugated goat anti-rabbit IgG secondary antibodies (A-21235, Life Technologies, 1:200) and Alexa Fluor 555 phalloidin (A34055, Life Technologies, 1:200) both for 1 hour; all conducted at room temperature. Nuclei were stained with Hoechst. Confocal optical slices were captured using a 40× oil objective lens (Leica Microsystems), an optical slice thickness of 800 nm and an x-y pixel size of 86 nm (Leica SP8 confocal microscope, Necker Imaging Facility).

Zebrafish histology

In situ hybridization experiments were done as previously described (8) using probes for *ifabp (9)*, *try* (10), *rag1* (11), *gata1* (12) and *cmyb* (13). Embryos stained with *ccnD1* were mounted in 80% glycerol and imaged, or embedded in OCT (Sakura Finetek, 4583), sectioned and imaged. For Neutral Red staining, 4 dpf larvae (n = 10) fixed in 4% paraformaldehyde (PFA) were embedded into JB-4 resin (Sigma-Aldrich, EM0100) according to manufacturer's protocol. Series of 10 μm sections were prepared using a Leica microtome, following which sections were stained in 0.001% Neutral Red solution for 30 minutes, washed in dH_2O, dried and mounted with DPX Mountant (Sigma-Aldrich, 06522). For proteinuria screening, embryos were injected with FITC-labeled 500kDa (Sigma-Aldrich, FD500S) and Texas-Red-labeled 10kDa dextrans (Thermo Fisher, D1828) (both 20mg/ml, 4nl) at 5 dpf and fixed six hours later as previously described (14). For immunofluorescence, fixed zebrafish larvae (n = 30) were prepared for paraffin embedding and sectioned at 5 μm. A sample of embryos were taken for hematoxylin and eosin staining, while the remaining sections underwent immunostaining with antibodies against WT1 (1:50, 6F-H2 #05-753, Merck Millipore;) and an Alexa-594 conjugated secondary (1:300, Invitrogen) as previously described (15). Imaging was performed on Zeiss Axioimager Z1.

To visualize cells undergoing apoptosis 0.023 g acridine orange dye powder (Sigma-Aldrich, A6014) was diluted in 9.5 ml MilliQ water and added to E3 embryo medium in a 10x dilution. Embryos were incubated for 30 minutes at 28.5°C, then rinsed in E3 medium, embedded in 3% methylcellulose (Sigma-Aldrich, M0387) and imaged under a fluorescent microscope.

Western blot of dyskerin and p53 in zebrafish

For the Western blotting of both dyskerin and p53 n = 12 *dkc1*$^{+/?}$ and n = 12 *dkc1*$^{elu1/elu1}$ 5 dpf embryos were used. Protein extraction was performed following standardised protocols (16). Whole lysates were loaded on 4–20% MiniPROTEAN® TGX StainFree™ Protein Gel and transferred to a Trans-Blot®Turbo™ Mini Nitrocellulose membrane (Bio-Rad). Membranes were blocked with 5% non-fat milk in PBS and were incubated with dyskerin antibody (H-300, SCBT) or anti-p53 antibody (ab77813, Abcam) as primary antibodies in PBS - 0.1% Tween-20 and goat anti-rabbit IgG-HRP as secondary antibody (sc-2004, SCBT). Gamma tubulin antibody (T5326, Sigma-Aldrich) was used as a loading control. Proteins were visualized by chemiluminescence (Western Blotting Luminol Reagent, sc-2048, SCBT) and densitometry was performed on a Molecular Imager VersaDoc MP 5000 System (Bio-Rad).

Cycloheximide (CHX) treatments

CHX was used to impair translation in developing zebrafish embryos as previously described (17). In brief, WT zebrafish embryos were treated in 10 μM CHX (Sigma-Aldrich, C7698) from 36 hpf, the first time point when the mutant phenotype of *dkc1*$^{elu1/elu1}$ embryos becomes apparent.

Multiple sequence alignments, data analysis and visualization

Multiple sequence alignments and visualizations were performed in R (R Core Team, 2019) using the msa and ggplot2 packages (18, 19).

Table S1. Primer sequences

Gene	Primer sequence
MM-qPCR Telomere length analysis	
hTelo F hTelo R	5'-ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT-3' 5'-TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTAACA-3'
Albumin F Albumin R	5'-CGGCGGCGGGCGGCGCGGGCTGGGCGGAAATGCTGCACAGAATCCTTG-3' 5'-GCCCGGCCCGCCGCGCCCGTCCCGCCGGAAAAGCATGGTCGCCTGTT-3'
X-inactivation measurements	
K206 WT F K206 WT R	5'-GACCATCTACGAGCAAAATGATTG-3' 5'-CTGCATCTGACCAACTCC-3'
E206 F E206K TM	5'-GGACCATCTACGAGAGCAAAATGATTA-3' 5'-FAM-AGGAATCTTTTGGGTGAGTTGTGAGGC-BBQ-3'
Dyskerin and NOP10 plasmid mutagenesis for pressure tuning fluorescence spectroscopy	
EcoRI-Flag-DKC1 F	5'-CCGAATTCCCACCATGGACTACAAAGACGATGACGACAAGGCGGATGCGGAAGTAAT-3'
NotI-DKC1 R	5'-TAGAAGCTTGCGGCCGCTTTA-3'
HindIII-NOP10 F	5'-CCCAAGCTTCCACCATGTTTCTCCAGTATTACCTCAA-3'
NOP10-XhoI R	5'-CCGCTCGAGTCGAGGACAGGGCGCGGTT-3'
Dyskerin and NOP10 plasmid mutagenesis for co-immunoprecipitation	
E206K DKC1 F E206K DKC1 R	5'-CGAGAGCAAAATGATTAAATACGATCCTGAAAGAAGATTAGG-3' 5'- CGAGAGCAAAATGATTAAATACGATCCTGAAAGAAGATTAGG-3'
T16M NOP10 F T16M NOP10 R	5'-CGAGCAGGGAGATCGAGTCTATATGCTGAAGAAATTTGACCCGATGG-3' 5'-CCATCGGGTCAAATTTCTTCAGCATATAGACTCGATCTCCCTGCTCG-3'
Zebrafish mutagenesis and genotyping (exon 7)	
Zebrafish exon 7 gRNA target sequence (PAM sequence bracketed)	5'-AGATGGTGCGCACTCGCAGT(TGG)-3'
Zebrafish exon 7 sequencing forward	5' GGGGGTTTAGGGTTCGGTTT 3'
dkc1 exon 7 reverse	5' ACAAAAGTGAGGGAAAAGGTTCA 3'
Zebrafish genotyping *elu1* (exon 7)	5' TGAAGCGCCAACTTCATGGT 3'
Zebrafish genotyping *elu2* (exon 7)	5' TGTGAAGCGCCAACTGGTG 3'
Zebrafish genotyping *wild-type* (exon 7)	5' GTGAAGCGCCAACTGCGA 3'
Zebrafish mutagenesis and genotyping (exon 11)	
Zebrafish exon 11 gRNA target sequence (PAM sequence bracketed)	5'-GGTCCAAGACCCCATTTGCG(TGG)-3'
Zebrafish genotyping *elu8* (exon 11)	5' AGAGACACATATCCAAAT 3'
Zebrafish genotyping *wild-type* (exon 11)	5' AGAGACACATATCCACGC 3'
Zebrafish exon 11 sequencing forward	5' ACTGATGGCTCTGCAAGACTTT 3'
dkc1 exon 11 reverse	5' ACAACAAAGCCTGATTGCACTA 3'
Anti-sense, synthetic morpholino (MO)	5' CTTTCTTACTTTCAGTGTCCGCCAT 3'

SI Web Resource

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/

SI References

1. I. Bekerman, P. Gottlieb, M. Vaiman, Variations in eyeball diameters of the healthy adults. *J Ophthalmol* **2014**, 503645 (2014).
2. J. D. Grant *et al.*, Telometric: a tool providing simplified, reproducible measurements of telomeric DNA from constant field agarose gels. *BioTechniques* **31**, 1314-1316, 1318 (2001).
3. R. M. Cawthon, Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res* **37**, e21 (2009).
4. G. Toldi *et al.*, The frequency of peripheral blood CD4+ CD25high FoxP3+ and CD4+ CD25- FoxP3+ regulatory T cells in normal pregnancy and pre-eclampsia. *Am J Reprod Immunol* **68**, 175-180 (2012).
5. M. Hultdin *et al.*, Telomere analysis by fluorescence in situ hybridization and flow cytometry. *Nucleic Acids Res* **26**, 3651-3656 (1998).
6. C. Trahan, C. Martel, F. Dragon, Effects of dyskeratosis congenita mutations in dyskerin, NHP2 and NOP10 on assembly of H/ACA pre-RNPs. *Hum Mol Genet* **19**, 825-836 (2010).
7. A. M. Waters *et al.*, The kinetochore protein, CENPF, is mutated in human ciliopathy and microcephaly phenotypes. *J Med Genet* **52**, 147-156 (2015).
8. S. Maegawa, M. Varga, E. S. Weinberg, FGF signaling is required for {beta}-catenin-mediated induction of the zebrafish organizer. *Development* **133**, 3265-3276 (2006).
9. M. Andre *et al.*, Intestinal fatty acid binding protein gene expression reveals the cephalocaudal patterning during zebrafish gut morphogenesis. *Int J Dev Biol* **44**, 249-252 (2000).
10. F. Biemar *et al.*, Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev Biol* **230**, 189-203 (2001).
11. S. Sawamiphak, Z. Kontarakis, D. Y. Stainier, Interferon gamma signaling positively regulates hematopoietic stem cell emergence. *Dev Cell* **31**, 640-653 (2014).
12. Q. Long *et al.*, GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene. *Development* **124**, 4105-4111 (1997).
13. A. C. Morris, M. A. Forbes-Osborne, L. S. Pillai, J. M. Fadool, Microarray analysis of XOPS-mCFP zebrafish retina identifies genes associated with rod photoreceptor degeneration and regeneration. *Invest Ophthalmol Vis Sci* **52**, 2255-2266 (2011).
14. G. Dorval *et al.*, TBC1D8B Loss-of-Function Mutations Lead to X-Linked Nephrotic Syndrome via Defective Trafficking Pathways. *Am J Hum Genet* **104**, 348-355 (2019).
15. J. L. Huang *et al.*, Vascular Endothelial Growth Factor C for Polycystic Kidney Diseases. *J Am Soc Nephrol* **27**, 69-77 (2016).
16. V. Link, A. Shevchenko, C. P. Heisenberg, Proteomics of early zebrafish embryos. *BMC Dev Biol* **6**, 1 (2006).
17. M. Poulain, T. Lepage, Mezzo, a paired-like homeobox protein is an immediate target of Nodal signalling and regulates endoderm specification in zebrafish. *Development* **129**, 4901-4914 (2002).
18. U. Bodenhofer, E. Bonatesta, C. Horejs-Kainrath, S. Hochreiter, msa: an R package for multiple sequence alignment. *Bioinformatics* **31**, 3997-3999 (2015).
19. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (2016).