Nucleic Acids Research, 2022 **1**
https://doi.org/10.1093/nar/gkac706

Domain specific mutations in dyskerin disrupt 3′ end processing of scaRNA13

Neha Nagpal[1], Albert K. Tai[2,3], Jayakrishnan Nandakumar[4] and Suneet Agarwal[1,*]

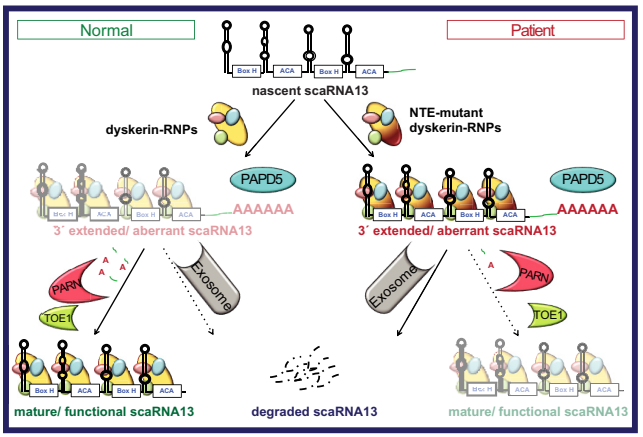
[1]Division of Hematology/Oncology and Stem Cell Program, Boston Children's Hospital; Pediatric Oncology, Dana-Farber Cancer Institute; Harvard Stem Cell Institute; Department of Pediatrics, Harvard Medical School; Manton Center for Orphan Disease Research; Harvard Initiative in RNA Medicine; Boston, MA, USA, [2]Department of Immunology, Tufts University School of Medicine, Boston, MA, USA, [3]Data Intensive Studies Center, Tufts University, Medford, MA, USA and [4]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

Received December 06, 2021; Revised July 27, 2022; Editorial Decision July 28, 2022; Accepted August 10, 2022

ABSTRACT

Mutations in *DKC1* (encoding dyskerin) cause telomere diseases including dyskeratosis congenita (DC) by decreasing steady-state levels of TERC, the non-coding RNA component of telomerase. How *DKC1* mutations variably impact numerous other snoRNAs remains unclear, which is a barrier to understanding disease mechanisms in DC beyond impaired telomere maintenance. Here, using DC patient iPSCs, we show that mutations in the dyskerin N-terminal extension domain (NTE) dysregulate scaRNA13. In iPSCs carrying the del37L NTE mutation or engineered to carry NTE mutations via CRISPR/Cas9, but not in those with C-terminal mutations, we found scaRNA13 transcripts with aberrant 3′ extensions, as seen when the exoribonuclease *PARN* is mutated in DC. Biogenesis of scaRNA13 was rescued by repair of the del37L *DKC1* mutation by genome-editing, or genetic or pharmacological inactivation of the polymerase PAPD5, which counteracts PARN. Inspection of the human telomerase cryo-EM structure revealed that in addition to mediating intermolecular dyskerin interactions, the NTE interacts with terminal residues of the associated snoRNA, indicating a role for this domain in 3′ end definition. Our results provide mechanistic insights into the interplay of dyskerin and the PARN/PAPD5 axis in the biogenesis and accumulation of snoRNAs beyond TERC, broadening our understanding of ncRNA dysregulation in human diseases.

GRAPHICAL ABSTRACT



INTRODUCTION

Dysregulation of non-coding RNAs (ncRNAs) is increasingly linked to human genetic diseases, not only by lesions in the ncRNAs themselves but also protein factors that control their biogenesis, post-transcriptional base modifications, and incorporation into ribonucleoprotein (RNP) complexes (1–4). Mutations in the genes encoding these protein factors are typically hypomorphic, presumably because complete ablation and the resultant disruption of a broad span of RNAs would be inviable. Moreover, they show pleiotropy, likely reflecting the particular sensitivity of one or a few potential target ncRNAs in different cell types and at different developmental stages. However, in most cases, how mutations in general ncRNA factors impact subsets of potential targets to cause disease remains unclear.

This conundrum is exemplified by telomere diseases, a spectrum of Mendelian disorders caused by mutations in factors regulating telomere maintenance and frequently

*To whom correspondence should be addressed. Tel: +1 617 919 4610; Fax: +1 617 919 3359; Email: suneet.agarwal@childrens.harvard.edu

© The Author(s) 2022. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

impacting the telomerase RNP (5–9). Dyskeratosis congenita (DC) is a severe form of telomere disease associated with multiple degenerative defects including epithelial and mucosal abnormalities, hematopoietic failure, lung and liver fibrosis, and cancer (10,11). DC and telomere diseases show broad heterogeneity including age of onset and disease manifestations that remain incompletely unexplained by genotype (12,13). In stem cells, telomerase reverse transcriptase (TERT) restores telomere length by adding 5′-GGTTAG-3′ hexameric repeats using the non-coding telomerase RNA component (TERC) template (14–17). TERC is a 451 nucleotide box H/ACA containing scaRNA (small Cajal body specific RNA) that bridges TERT and the H/ACA RNP complex (dyskerin-NHP2-NOP10-GAR1-TCAB1) (17–21). Whereas TERT and TERC can form a minimal functional telomerase holoenzyme *in vitro*, the H/ACA complex is critical for the stability and trafficking of telomerase RNP in cells (18,22–24). Remarkably, mutations not only in telomerase-specific factors such as *TERT* and *TERC* but also in several 'general' ncRNA factors controlling TERC biogenesis present clinically as DC/telomere diseases (9,25). These include factors comprising the H/ACA RNP complex as well as those regulating nascent RNA end-processing, thus potentially affecting numerous ncRNAs. The finding of disrupted TERC levels in the setting of all of these various genetic mutations provides a central pathophysiologic mechanism for DC/telomere diseases (8,25). However, understanding the broader impact of these DC-associated lesions on ncRNAs promises to illuminate new molecular mechanisms underlying their differential regulation as well as the basis of phenotypic heterogeneity.

In terms of the H/ACA RNP complex, mutations in *DKC1* (encoding dyskerin), *NOP10*, *NHP2*, *TCAB1* and the RNP assembly factor *NAF1* (the product of which is exchanged with GAR1 in the mature RNP) are associated with DC/telomere diseases (26–30). Almost all box H/ACA small nucleolar RNAs (snoRNAs; of which scaRNAs are a subset) are intron-encoded and are bound by the H/ACA RNP co-transcriptionally or soon after processing of the lariat, which is critical for defining the 5′ and 3′ ends of the mature snoRNA by protection from exonucleases (19,31,32). Dyskerin is a highly-conserved, 514 amino acid pseudouridine synthase that is guided by one of over a hundred possible associated snoRNAs to various ribosomal and small nuclear RNAs targets for pseudouridylation (33,34). Pathogenic *DKC1* mutations in DC are typically missense or small in-frame deletions, and clustered in two regions, the N-terminal extension (NTE) and a distinct region comprising the pseudouridine synthase and archaeosine transglycosylase (PUA) domain and downstream residues towards the C-terminus of dyskerin (35,36). While impacts on H/ACA snoRNAs beyond TERC due to *DKC1* mutations might be expected and have been demonstrated variably in certain contexts (8,37–42), the mechanistic basis for these differences is not well understood.

With respect to general ncRNA processing factors, mutations in *PARN* (encoding poly(A) specific ribonuclease) and *ZCCHC8* (encoding a nuclear exosome targeting (NEXT) complex component) also cause DC/telomere diseases (43–47). In both cases, aberrant TERC forms with post-transcriptional, non-templated addition of nucleotides to the 3′ end have been found, and are associated with TERC destabilization and telomerase insufficiency (45,46). PARN is a 3′ exonuclease originally thought to regulate the turnover of mRNAs, but in humans appears to play a more prominent role in the post-transcriptional processing of not only nascent snoRNAs such as TERC but a variety of ncRNAs including microRNAs and Y-RNAs (45,48–51). With respect to TERC and certain other snoRNAs, PARN counteracts the non-canonical poly(A) polymerase, PAPD5, which oligoadenylates nascent transcripts and targets them for degradation by the RNA exosome (52–55). In the context of H/ACA RNPs, the exonucleolytic action of PARN removes unprotected ncRNA 3′ ends extending beyond contacts with dyskerin, eliminating their accessibility to PAPD5 and thereby promoting stabilization over degradation of nascent transcripts. Consistent with this, disrupting PAPD5 via genetic means or using small molecules reduces adenylation of nascent TERC transcripts and is sufficient to reverse quantitative defects in TERC, telomerase levels and telomere maintenance caused by *PARN* loss-of-function mutations in human cells (45,53,55–57). As with *DKC1*, *PARN* mutations would be expected to affect several other ncRNA substrates (48). However, transcriptomic analysis revealed a specific snoRNA, scaRNA13, as the only RNA affected to an equal or greater degree than TERC in the setting of PARN deficiency, and consistently differentially expressed after PAPD5 inactivation in *PARN*-mutant patient iPSCs (45,56). The mechanisms underlying the specific or predominant effects of *PARN* mutations on a subset of potential ncRNA substrates remain poorly understood.

Here, we studied the impact of *DKC1* mutations on H/ACA snoRNAs, and their relationship to 3′ end processing by the PARN/PAPD5 axis. Using induced pluripotent stem cells (iPSCs) from patients carrying lesions in different domains of dyskerin, we found defects specifically in the 3′ end maturation of scaRNA13 when the dyskerin NTE domain was disrupted. CRISPR-Cas9-mediated engineering of dyskerin NTE mutations in normal iPSCs recapitulated the 3′ end processing defects in scaRNA13. Pharmacological inhibition or genetic deletion of PAPD5 in NTE-mutant iPSCs led to reversal of the aberrant scaRNA13 phenotype. Genome editing of scaRNA13 to disrupt its unique tandem structure or homology-directed repair of the mutated dyskerin NTE promoted proper end maturation of scaRNA13. Structural analysis utilizing recently published cryo-EM studies places the specific residues of the NTE in question in proximity to the 3′ end of bound RNA (35). Together, our findings identify a novel function for dyskerin in regulating post-transcriptional processing of scaRNA13 by PAPD5, which may be explained both by the unusual tandem architecture of scaRNA13 and distinct functions of the dual dyskerin molecules comprising the H/ACA RNP.

MATERIALS AND METHODS

Patient samples

Institutional Review Boards approved the protocols for procurement of biological samples, which were obtained under written informed consent in accordance with the Declaration of Helsinki.

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

Cell lines and iPSCs

Derivation, characterization, and culture conditions of iPSCs from fibroblasts for a normal female subject (NHSF2), and *DKC1*- and *PARN*-mutant patients were as described (45,56,58,59). iPSCs from normal male fibroblasts and *DKC1* p.T49M patient fibroblasts were generated using the 4-in-1-dTomato lentiviral reprogramming vector (kind gift of A. Schambach (60)). For feeder-free culture, iPSCs were maintained in Essential 8 medium (Life Technologies) on hES-qualified Matrigel matrix (BD Biosciences) and sub-cultured using Accutase (Stem Cell Technologies). HEK 293 cells were obtained from ATCC, maintained in DMEM with 10% FCS, and sub-cultured using trypsin 0.05%.

RNA interference

shRNA constructs targeting human *DKC1* were as described (58).

Expression constructs

The coding region of human U93 cloned in pCMV-globin expression vector was used for transient overexpression of scaRNA13 (kind gift of B. Jady and T. Kiss) (61). The expression cassette containing U93 was amplified from pCMV-globin-U93 using primers given in Supplementary Table S1 and cloned into Sleeping Beauty vector pSB482-GFP-blast vector (pSBbi-Bla; kind gift of R. Marschalek) (62).

Viral vector production and transduction

Lentiviruses were produced using pCMV-dR8.91 and pCMV-VSV-G in HEK293 cells as described (45,58). Cells were transduced with viral vectors in the presence of protamine sulfate for 12–16 h. Cells were selected (blasticidin, 5 μg/ml or puromycin, 0.2 μg/ml) and analyzed after confirming stable expression.

Transient transfection

Transfection was performed using the X-TremeGene 9 transfection kit (Millipore Sigma, 6365787001) following the manufacturer's protocol. Briefly, 2.5 μg of overexpression or downregulation (shRNA) plasmid DNA constructs were used for each transfection (2×10^5 cells) and samples were collected and analyzed 96 h after transfection.

Stable or transposon transfection

The expression cassette containing the coding region of U93 was amplified from pCMV-globin-U93 (61) using primers given in Supplementary Table S1 and cloned into the Sleeping Beauty vector pSB482-GFP-blast vector (pSBbi-Bla; kind gift of R Marschalek) (62). Transfer of scaRNA13-beta-globin expression cassette for long-term overexpression was achieved using Sleeping Beauty vector containing transposase (62). Delivery of both the vectors containing transposase and transposon (harboring the U93 cassette) was performed via electroporation (Lonza, VCA-1003) following the manufacturer's protocol. Cells were selected (blasticidin, 5 μg/ml) for 3 weeks and analyzed for overexpression and downstream effects.

CRISPR-Cas9 knockout generation

Knockout generation was done by RNP electroporation as described (56). Briefly, sgRNAs were obtained from Synthego Corporation and gRNA sequences are given in Supplementary Table S1. Electroporation was done using Lonza 4D nucleocuvettes (V4XP-3024, 100 μl) as described (56). Modified synthetic gRNAs (with first and last three nucleotides having 2′-*O*-methyl-3′-phosphorothrioate modification) and 3xNLS-SpCas9 were mixed in equimolar concentrations and incubated for 15 min at RT. iPSCs were derived to single cell suspension using Accutase (Stem Cell Technologies) and resuspended in P3 solution immediately before electroporation. 50 μM RNPs were mixed with 1 M iPSCs and electroporation was done with program CA-137. Cells were maintained in E8 media supplemented with Y-27632 ROCK inhibitor (Abcam) for *in vitro* culture on Matrigel (Corning) matrix.

CRISPR-Cas9 editing validation

CRISPR-Cas9 efficiency was determined by Sanger sequencing followed by western or northern blotting for change in protein or RNA levels, respectively.

Sanger sequencing

Genomic DNA was isolated from iPSCs using GenJET genomic DNA purification kit (Thermo Fisher Scientific). Primers used for *AAVS1*, *DKC1, PAPD5* and *SCARNA13* gene amplification and sequencing are provided in Supplementary Table S1. Sanger sequencing was performed by Genewiz and the efficiency of editing was analyzed by ICE software (Synthego Corporation).

Western blotting

DC patient iPSCs were lysed in 2X Laemmli sample buffer (Bio-Rad, 1610737) and total cellular lysates were then subjected to SDS–PAGE followed by transfer to PVDF membranes. Detection of PAPD5 and DKC1 was done using anti-PAPD5 antibodies (Atlas antibodies, HPA042968, 1:1000) and anti-DKC1 antibodies (SantaCruz Biotechnology, sc-48794), respectively. HRP-conjugated goat anti-rabbit secondary antibody (Bio-Rad 170-5046, 1:10000) was used for HRP-conjugation and chemiluminescent detection using Clarity Western ECL substrate (Bio-Rad, 1705060). The same membrane was further probed with HRP-conjugated anti-actin antibodies (Santa Cruz sc-1615, dilution 1:1000) to control for protein loading. Imaging and quantification of chemiluminescent signals were performed using the Biorad ChemiDoc Touch imaging system.

Northern blotting

Total RNA (3.5–5 μg) was electrophoresed on a 1.5% agarose/formaldehyde gel followed by capillary transfer to Hybond N+ membranes (Amersham, GE Healthcare; RPN303B) in 10× SSC. Blot was hybridized for 12hrs with α-^{32}P-dCTP-labeled full-length RNA probe for TERC, or

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

full-length, 5′- or 3′-half probes for scaRNA13, in ULTRAhyb buffer (Life Technologies, AM8669M). 18S rRNA signal obtained from ethidium bromide staining was used for normalization, and signal quantifications were performed using ImageJ software. Primers used for probe amplification are given in Supplementary Table S1.

Telomere length measurement

To analyze telomere length in cell cultures, terminal restriction fragment (TRF) length analysis was performed (TeloTAGGG telomere length assay kit, Roche Life Science; 12209136001). Briefly, 2.5–3 μg of genomic DNA was electrophoresed and probed using kit reagents following manufacturer's protocol. Images were taken using Chemidoc Touch Imaging System (Bio-Rad).

3′ Rapid amplification of cDNA ends (RACE)

The RACE protocol was performed as described (45). Briefly, total RNA was extracted using TRIzol (Ambion), followed by ligation of 5′-adenylated-3′-blocked adapter (New England Biolabs) using T4 RNA ligase truncated KQ (New England Biolabs). Ligated RNAs were purified and cDNAs were synthesized using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific). PCR amplification was performed using gene specific primers and PCR amplicons were resolved on 2.5% agarose gels to visualize mature and extended transcripts. Primers used are given in Supplementary Table S1.

3′ RACE amplicon deep sequencing

Linker-ligated cDNAs obtained using the above-described RACE protocol were used for the generation of TERC and scaRNA13 amplicons, which after purification using Qiaquick PCR purification kit (Qiagen, 28106) were used for library preparation for deep sequencing. Library preparation, sequencing, and analysis was performed as described (45), with the addition of size selection (BluePippin) of the library prior to sequencing. Cumulative percent oligo-adenylated species were calculated as the fraction of all oligo-adenylated RNA species divided by total RNA species within eight genomically-encoded nucleotides of the mature 3′ end. The script used for analysis is given at https://github.com/alberttai/tailAcount.

Structural analysis

Views of the dyskerin-TERC interface in Figure 5A and Supplementary Figure S5A are derived from the cryo-EM structure of the H/ACA RNP lobe of human telomerase (PDB accession code: 7BGB) (35). Other protein subunits of the H/ACA RNP are not shown for clarity.

Transcriptome or RNA-Seq analysis

Total RNA was isolated and RNA quality was checked by both fragment analyzer (Agilent Fragment Analyzer) and electrophoresis on 1.5% agarose/formaldehyde gel. 0.1–1 μg of total RNA was used as input for library preparation using Illumina Stranded Total RNA with Ribo-Zero Plus (Illumina 20040529). The pooled libraries were sequenced on one lane of NextSeq 550 High Output V2.5 chemistry using single read 75 bases format. Resulting sequencing data was demultiplexed into fastq format using Illumina bcl2fastq. Fastqc reports were generated for the demultiplexed data to ensure data quality, and the reads from each sample was subsequently mapped to human reference genome (UCSC hg38) using HISAT(v2.1.0). Normalized expression value (FPKM) was then generated with Cuffflinks (v2.1.1). Differential expression analyses were then performed using either Cuffkinks or Qlucore Omics Explorer. An adjust *P*-value of 0.05 was used as cutoff the significantly differentially expressed genes during pair-wise comparison. Expression profiles of *DKC1* NTE-mutant iPSCs treated with PAPD5 inhibitor RG7834 were compared to those of their isogenic "no indel" control clones and parent (bulk) WT iPSCs.

Quantification and statistical analysis

Error bars presented mean with standard error. *P* values were calculated based on two-tailed Student's *t*-test, and one-way and two-way ANOVA and $P > 0.05$ was defined as not significant. All the statistical analysis was done using GraphPad Prism 8 software. Statistical and replicate details can be found in the figure legends.

RESULTS

Disruption of scaRNA13 maturation by *PARN* mutations is rescued by specific inactivation of PAPD5

To assess the global effects of alterations in the PARN/PAPD5 axis on the transcriptome, we inactivated PAPD5 in *PARN*-mutant patient iPSCs carrying different compound heterozygous mutations (p.N7H/gene deletion; p.S87L/non-functional gene; p.K58X/p.K59R) using two different approaches: CRISPR-Cas9 mediated *PAPD5* gene disruption and pharmacological inhibition of PAPD5 using small molecules (56). By RNA-Seq analysis, we identified scaRNA13 as the most significantly differentially expressed RNA in PAPD5-inactivated iPSCs (Supplementary Figure S1A). Low levels of scaRNA13, which depends on PARN for its stability and 3′ end processing (45,48,56), were rescued by specific inactivation of PAPD5 in *PARN*-mutant patient iPSCs (Figure 1A, B). Next, we performed RNA ligation mediated 3′ rapid amplification of cDNA ends (RLM-RACE) in PAPD5 inactivated *PARN*-mutant patient iPSCs, and observed a decrease in scaRNA13 extended forms, with a concomitant increase in mature forms (Figure 1C). To investigate these findings more closely, we profiled the 3′ ends of scaRNA13 by deep sequencing, and found a significant decrease in the proportion of transcripts extended beyond the canonical 3′ end and/or post-transcriptionally adenylated (Figure 1D, E). Similar results were obtained using the small molecule PAPD5 inhibitor RG7834 (Supplementary Figure S1). These results confirm that PAPD5 is responsible for 3′ adenylation of nascent scaRNA13 transcripts, and that

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

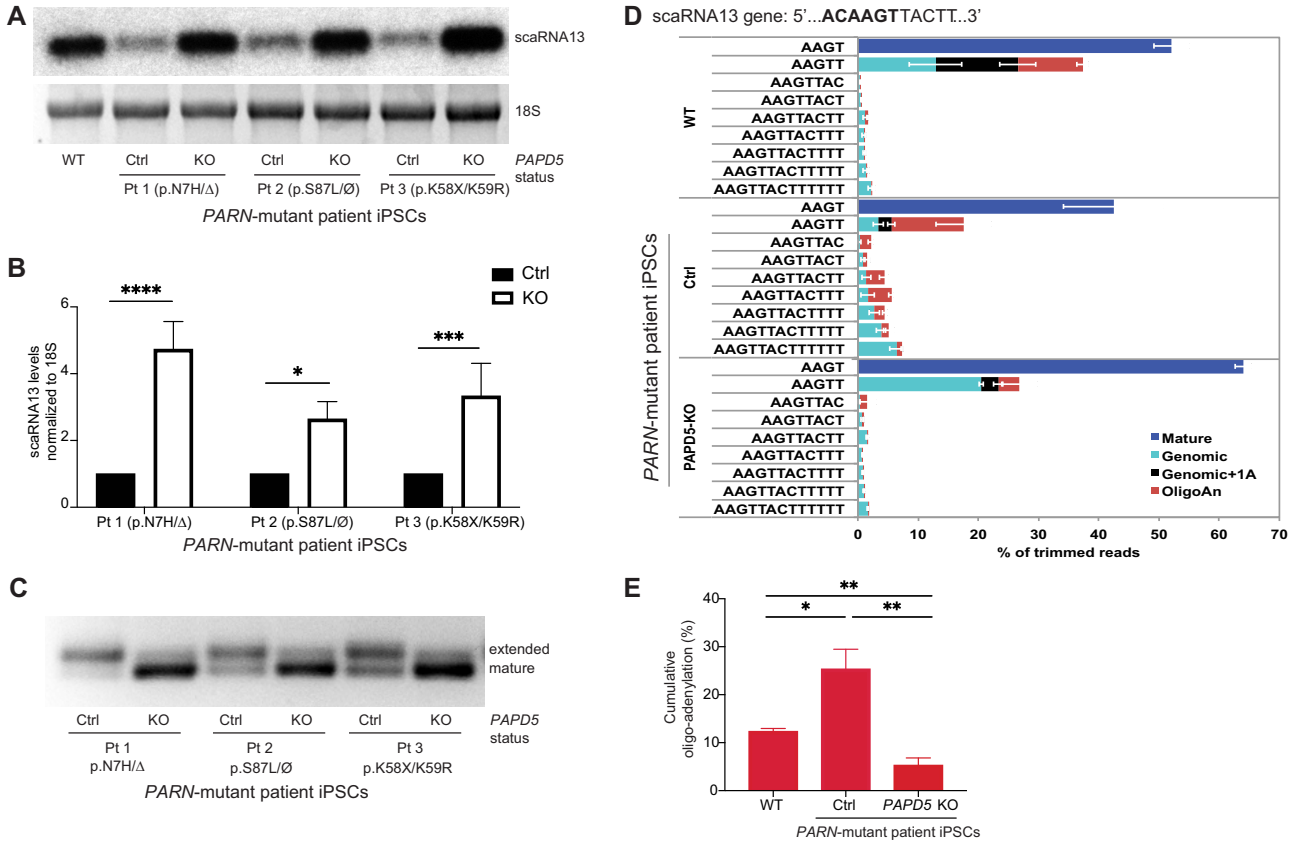


Figure 1. PAPD5 inactivation rescues scaRNA13 3′ end maturation in *PARN*-mutant patient iPSCs. (**A**) scaRNA13 RNA levels by northern blot in iPSCs from a normal subject (WT) versus three *PARN*-mutant DC patients (genotypes indicated). Pairwise comparisons to their respective *PAPD5* CRISPR knockout clone (KO) are shown. Inverted image of ethidium bromide staining of 18S rRNA is shown as a loading control. (**B**) Quantification of relative scaRNA13 levels normalized to 18S rRNA after *PAPD5* KO in patient iPSCs (from A). Mean ± S.D. is shown for $n = 3$ biological replicates. Statistics: $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$, two-way ANOVA. (**C**) Pairwise comparisons of scaRNA13 3′ end profiles using RNA-ligation mediated 3′ rapid ligation of complementary DNA ends (3′RACE) and agarose gel electrophoresis of iPSC clones from *PARN*-mutant DC patients before (Ctrl) and after (KO) *PAPD5* gene disruption by CRISPR/Cas9. $n = 3$ independent experiments, one representative image is shown. (**D**) scaRNA13 3′ RACE PCR products from from two different WT iPSCs and three *PARN*-mutant patient iPSCs as in (C) were subjected to deep sequencing and aligned to the scaRNA13 gene (canonical 3′ end sequence in bold). The relative proportions of mature scaRNA13 and extended scaRNA13 RNA species are depicted. Transcripts with oligo(A) additions of any length (*n*) are in solid red. Biological replicates for two WT and three patient iPSCs are shown with means ± S.D. (**E**) Quantitation of cumulative scaRNA13 oligo-adenylated species determined by deep sequencing as in (D). Graph shows mean ± S.D. Statistics: two-tailed *t*-test, $*P < 0.05$, $**P < 0.005$.

inactivation of PAPD5 rescues maturation and steady-state levels of scaRNA13 in a PARN deficient background.

The N-terminal extension (NTE) domain of dyskerin differentially regulates scaRNA13 maturation

Because box H/ACA snoRNAs including scaRNA13 require the H/ACA RNP complex for their productive maturation, we studied their end processing in cells from three male DC patients with pathogenic hemizygous *DKC1* mutations in different domains of dyskerin. One patient carries deletion of a highly conserved leucine residue at position 37 (del37L) while the other two patients carry missense variants, A386T and A353V in the C-terminal extension (CTE) and PUA domains, respectively (Figure 2A). To determine the impact of specific *DKC1* mutations on scaRNA biogenesis, we performed RLM-RACE on 27 annotated scaRNAs. We found scaRNA13 3′ end processing defects specifically in del37L patient iPSCs with the accumulation of an extended form (Figure 2B, Supplementary Figure S2A), a phenotype reminiscent of *PARN*-mutant patient iPSCs (Figure 1C). In keeping with these findings, scaRNA13 steady state levels were more severely reduced in del37L patient iPSCs than in iPSCs from patients carrying A353V or A386T mutations (Figure 2C,D). By Sanger sequencing, we determined that no coding mutations in the *PARN* gene were detectable in del37L patient cells. Deep sequencing of scaRNA13 ends confirmed an accumulation of oligo-adenylated forms in the setting of the del37L mutation (Figure 2E, F), but distinct from those found with *PARN* mutations. Oligo-adenylated scaRNA13 transcripts in *DKC1* del37L iPSCs had a median total tail length of 7 nucleotides, with a majority of species (85%) extended by only one genomically-encoded nucleotide beyond the canonical end, and carrying median A-tails 6 nucleotides in length (Figure 2G, H). In contrast, those in *PARN*-mutant iPSCs had a median total tail length of 10 nucleotides, with genomic extensions at the +1 position representing only

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

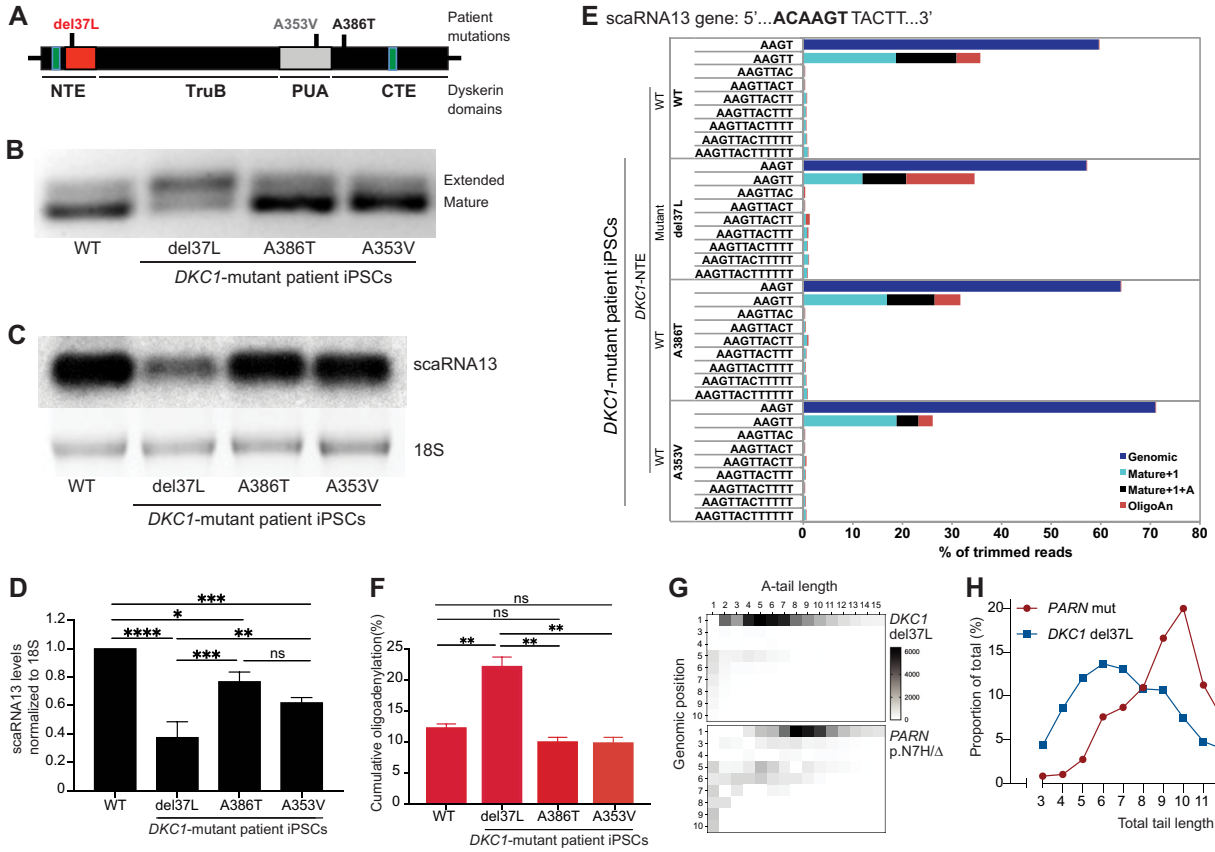


Figure 2. Domain-specific effects of *DKC1* mutations on scaRNA13. (**A**) Schematic showing domains of dyskerin including NTE (N-terminal extension), TruB (tRNA pseudouridine synthase B-like), PUA (pseudouridine synthase and archaeosine transglycosylase) and CTE (C-terminal extension). Mutations in the three DC patients under study are indicated. (**B**) scaRNA13 3′ end profiles in three *DKC1*-mutant DC patients compared to that of normal (WT), using RNA-ligation mediated 3′ RACE. $n = 3$ independent experiments, one representative image is shown. (**C**) scaRNA13 RNA levels by northern blot in three *DKC1*-mutant DC patients compared to normal (WT) control. Inverted image of ethidium bromide staining of 18S rRNA is shown as a loading control. (**D**) Quantification of relative scaRNA13 levels normalized to 18S rRNA in patient iPSCs in (C) compared to WT iPSCs. Mean ± S.D. is shown for $n = 3$ biological replicates. Statistics: ns- not significant, $*P < 0.05$, $**P < 0.005$, $***P < 0.001$, $****P < 0.0001$, one-way ANOVA. (**E**) 3′ scaRNA13 RACE PCR products from two WT and three *DKC1*-mutant patient iPSCs as in (B) were subjected to deep sequencing and aligned to the scaRNA13 gene. Relative proportions of mature scaRNA13 and extended scaRNA13 RNA species are depicted. Oligo(A) additions of any length (n) are in solid red. (**F**) Quantitation of cumulative scaRNA13 oligo-adenylated species determined by deep sequencing as in (D). Statistics: two-tailed t-test, ns = not significant, $*P < 0.05$, $**P < 0.005$. (**G**) Heat map of oligo-adenylated forms of *DKC1* del37L iPSCs ($n = 55\,888$ reads) versus *PARN* mutant iPSCs (p.N7H/Δ; $n = 64\,224$ reads), depicting number of reads mapped for a given A-tail length and genomic position. Legend depicts read number. (**H**) Graph of proportion of all extended transcript forms ≥3 nucleotides in length, as a function of total tail length (genomic + adenylated) in *DKC1* del37L versus *PARN*-mutant (mut) iPSCs.

56% of oligo-adenylated transcripts, and with median A-tails of seven nucleotides that varied in length in a relationship inversely proportional to genomic position (Figure 2G, H). Thus in *DKC1* del37L iPSCs cells, the dominant scaRNA13 extended transcript forms comprising the upper RACE amplicon (Figure 2B) bear similarities to, but can be distinguished from, those found in *PARN*-mutant cells (Figure 1C). In contrast to the specific effects of del37L *DKC1* mutation on scaRNA13 3′ end processing, we did not find a domain-specific effect of dyskerin dysfunction on TERC biogenesis, as 3′ end processing of TERC was unchanged in all three patients (Supplementary Figure S2B). As previously observed (56), TERC levels were uniformly low in all three patient iPSCs (Supplementary Figure S2C, D). Taken together, these results indicate that in addition to PARN, dyskerin is also required for the proper 3′ end processing of scaRNA13.

CRISPR-Cas9 mediated engineering of the dyskerin NTE disrupts 3′ end maturation of scaRNA13

To rigorously test the involvement of the dyskerin NTE in scaRNA13 biogenesis, we used CRISPR-Cas9 to mutagenize the endogenous hemizygous *DKC1* locus in normal male iPSCs. After transient electroporation of a CRISPR RNP with a guide RNA targeting *DKC1* exon 3 near amino acid 44, we isolated several mutant iPSC clonal lines as well as genetically-unmodified clones. As expected, given the essential role of dyskerin, only iPSCs carrying in-frame *DKC1* mutations were able to be cloned and propagated, and both the mutant and unmodified clones had detectable dyskerin protein levels (Figure 3A, Supplementary Figure S3A). Amongst the iPSCs generated to carry a variety of N-terminal dyskerin mutations, we found significantly diminished TERC and scaRNA13 steady state levels compared to clones without modifications (Figure 3B,C, Supplementary

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

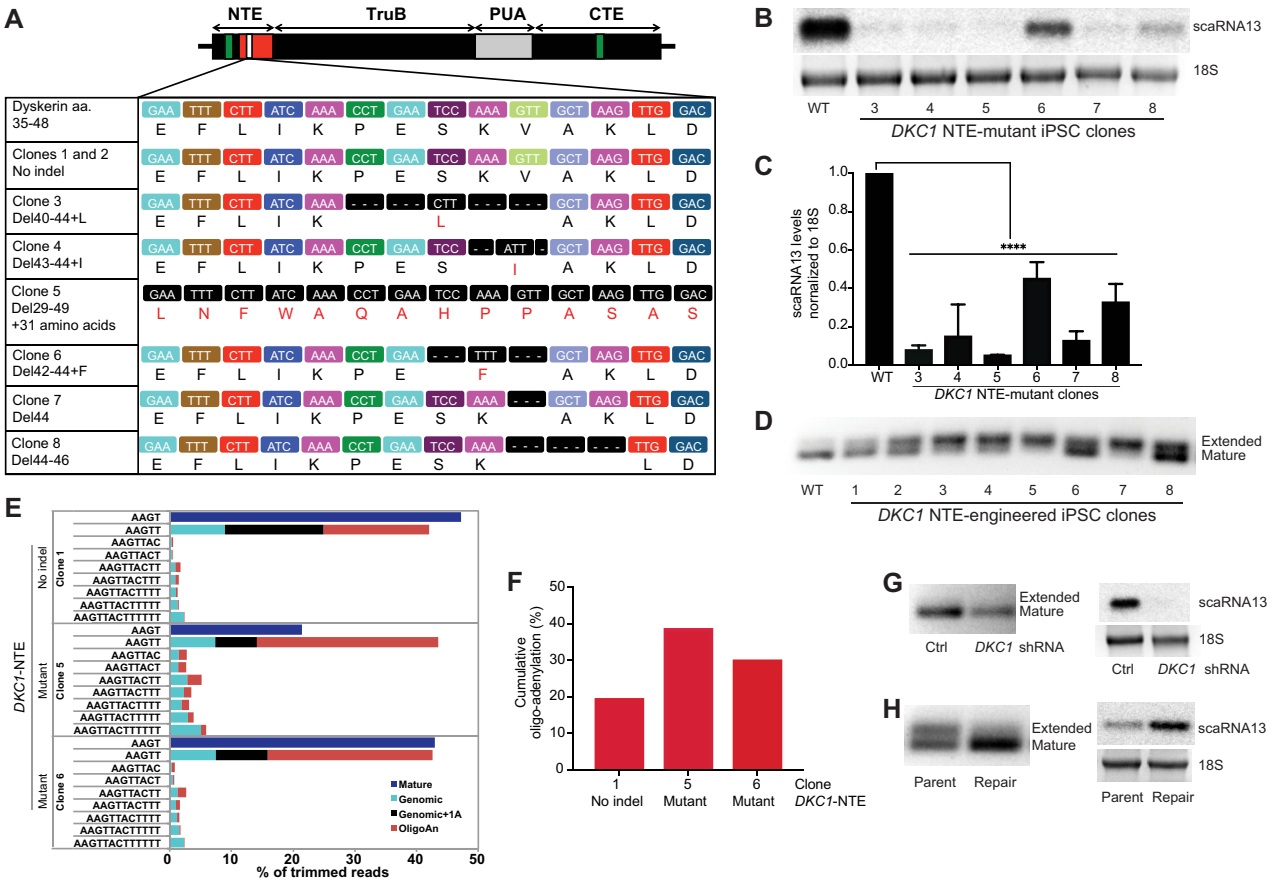


Figure 3. CRISPR-Cas9-mediated mutagenesis of the dyskerin N-terminal extension (NTE) disrupts 3′ end maturation of scaRNA13. (**A**) Schematic depicting mutagenesis of the NTE domain (amino acids 35–48) of dyskerin via CRISPR-Cas9 RNP targeting of *DKC1* in normal male iPSCs, and the isolated clones under study: unmodified *DKC1* locus (no indel; clones 1 and 2), in-frame deletion / mutation of one or more amino acids (clones 3–4 and 6–8), or loss of exon 3 with a large heterologous in-frame insertion (clone 5). (**B**) scaRNA13 RNA levels in *DKC1* NTE-mutant iPSC clones as in (A) by northern blot. 18S rRNA is shown as loading control.$n = 3$ independent experiments, one representative image is shown. (**C**) Quantification of relative scaRNA13 levels in *DKC1* NTE-mutant iPSC clones as in (B) compared to WT iPSCs, normalized to 18S rRNA. Mean ± S.D. is shown for $n = 3$ biological replicates. Statistics: $^{****}P < 0.0001$, one-way ANOVA. (**D**) scaRNA13 3′ end profiles in *DKC1* NTE-mutant iPSC clones as in (A) using 3′ RACE. $n = 3$ independent experiments, one representative image is shown. (**E**) scaRNA13 3′ RACE PCR products from three iPSC clones [1 (*DKC1* NTE-no indel), 5 and 6 (*DKC1* NTE-mutant)] were subjected to deep sequencing and aligned to the scaRNA13 gene. The relative proportions of mature scaRNA13 and extended scaRNA13 RNA species are depicted. Oligo(A) additions of any length (n) are in solid red. (**F**) Quantitation of cumulative scaRNA13 oligo-adenylated species of clone 5 and clone 6 as representative of mutant-NTE dyskerin versus clone 1 (no indel) as control. (**G**) scaRNA13 3′ end profiles by RACE (*left*) and RNA levels by northern blot (*right*) in WT iPSCs with transient shRNA-mediated knock-down of *DKC1* versus control. $n = 3$ independent experiments, one representative image is shown. (**H**) scaRNA13 3′ end profiles by RACE (*left*) and RNA levels by northern blot (*right*) in del37L patient iPSCs, comparing original iPSCs (Parent) to those in which the *DKC1* locus underwent CRISPR-Cas9-mediated homology-directed repair (Repair). $n = 3$ independent experiments, one representative image is shown.

Figure S3B). Next, we performed RLM-RACE to profile scaRNA 3′ ends and observed extended forms specifically of scaRNA13 in multiple NTE-mutant clones (Figure 3D), but not in 26 other scaRNAs, as seen in del37L patient iPSCs. Deep sequencing of 3′ RLM-RACE products showed that extended scaRNA13 forms were composed predominantly of increased post-transcriptionally oligo-adenylated species (Figure 3E, F). Although dyskerin NTE domain mutations led to a decrease in TERC steady levels, unlike scaRNA13, we found no changes in the end maturation of TERC (Supplementary Figure S3C-E). To further establish the specific role of the dyskerin NTE in scaRNA13 biogenesis, we knocked down *DKC1* using short hairpin RNA (shRNA) and found no change in scaRNA13 3′ end processing (Figure 3G). As expected, we observed a decrease in TERC and scaRNA13 steady state levels indicative of dyskerin's known role in sno/scaRNA binding and stabilization (Figure 3G, Supplementary Figure S3F). We next performed CRISPR-Cas9-mediated homology directed repair (HDR) of the del37L locus in patient iPSCs. Repaired clones showed a restoration of scaRNA13 3′ end processing (Figure 3H), and scaRNA13 and TERC levels (Figure 3H, Supplementary Figure S3G). Taken together, these results demonstrate that the NTE domain of dyskerin is selectively involved in scaRNA13 3′ end maturation.

PAPD5 inhibition restores scaRNA13 3′end processing defects in dyskerin mutant iPSCs

PAPD5 inactivation counteracts 3′ end processing defects in TERC, scaRNA13 and other ncRNAs in the setting of mutations or disruption of the *PARN* gene (48,55–57).

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

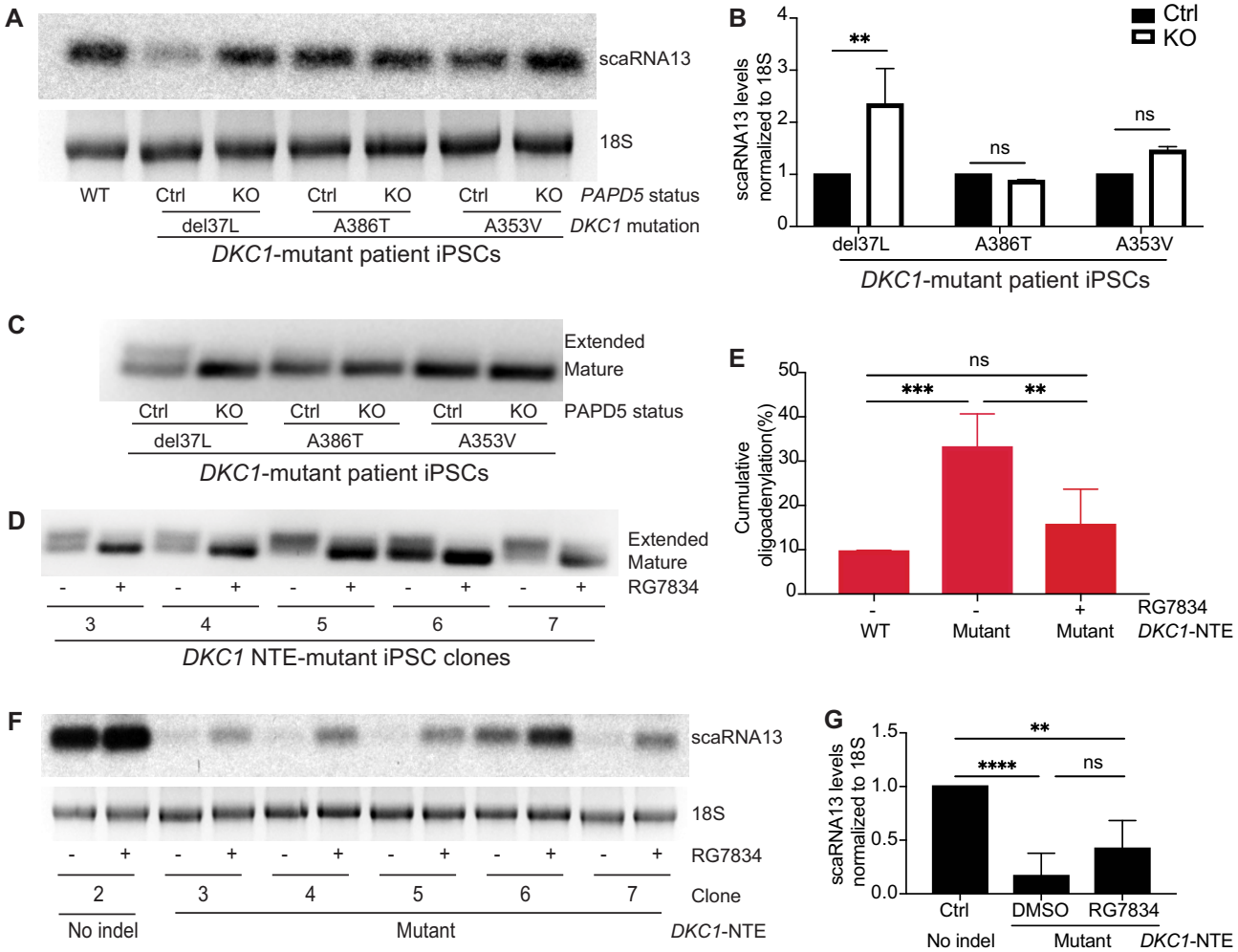


Figure 4. PAPD5 inactivation augments scaRNA13 3′ end processing in *DKC1* NTE-mutant iPSCs. (**A**) scaRNA13 RNA levels by northern blot in three *DKC1*-mutant DC patients and their respective CRISPR *PAPD5* knockout clones (KO). Normal iPSCs (WT) shown as a baseline. Inverted image of ethidium bromide staining of 18S rRNA is shown as a loading control. (**B**) Quantification of relative scaRNA13 levels in response to *PAPD5* KO in patient iPSCs in (A), normalized to 18S rRNA. Mean ± S.D. is shown for n = 3 biological replicates. Statistics: ns- not significant, $^{**}P < 0.005$, 2way ANOVA. (**C**) Pairwise comparisons of scaRNA13 3′ end profiles in control vs CRISPR *PAPD5* knockout clones of three DC patients with *DKC1* mutations in different domains of dyskerin (*see* Figure 2A), using 3′ RACE. n = 3 independent experiments, one representative image is shown. (**D**) scaRNA13 3′ RACE profiles in five *DKC1* NTE-engineered iPSC clones (*see* Figure 3A), after seven days of RG7834 (+) versus DMSO (–) treatment. *n* = 3 independent experiments, one representative image is shown. (**E**) scaRNA13 3′ RACE PCR products from iPSCs as in (C) and (D) were subjected to deep sequencing and aligned to the scaRNA13 gene. The relative proportions of mature scaRNA13 and extended scaRNA13 RNA species were determined. Graph shows quantitation of cumulative scaRNA13 oligo-adenylated species in seven different *DKC1* NTE-mutant iPSCs treated with and without RG7834, compared to three iPSCs with intact *DKC1* NTE as controls. Statistics: one-way ANOVA ns: not significant, $^{**}P < 0.05$, $^{***}P < 0.001$. (**F**) scaRNA13 RNA levels by northern blot in *DKC1* NTE-mutant iPSC clones after seven days of RG7834 (+) versus DMSO (–) treatment. Clone 2 with unmodified (no indel) *DKC1* NTE represents a baseline (Ctrl). (**G**) Quantification of relative scaRNA13 levels normalized to 18S rRNA in different *DKC1* NTE-mutant iPSC clones in response to PAPD5 inhibitor treatment, compared to control clone 2 (*DKC1* NTE no indel (Ctrl)) as shown in (F). Mean ± S.D. is shown for five independent *DKC1* NTE-mutant clones. Statistics: ns = not significant, $^{**}P < 0.005$, $^{****}P < 0.0001$, one-way ANOVA.

We therefore asked whether PAPD5 inhibition could rescue scaRNA13 3′ end maturation in dyskerin NTE mutant iPSCs. We inactivated *PAPD5* in *DKC1*-mutant patient iPSCs using CRISPR-Cas9 genome editing and isolated clonal lines, and found that scaRNA13 3′ end processing and scaRNA13 steady state levels were restored in the setting of the del37L mutation (Figure 4A–C). Similar effects were seen via PAPD5 inhibition using the small molecule RG7834 (56) (Supplementary Figure S4A). We next inhibited PAPD5 using RG7834 in engineered dyskerin NTE-mutant iPSC lines, and again found that scaRNA13 3′ end processing defects were reversed (Figure 4D, E). However, scaRNA13 steady state levels could not be fully restored by PAPD5 inhibition alone (Figure 4F,G), suggesting a residual defect in dyskerin beyond its function in scaRNA13 biogenesis in the setting of the engineered mutations. In keeping with this, despite restoring 3′ end processing, we observed only a partial rescue in TERC steady state levels in response to PAPD5 inhibition in the engineered dyskerin NTE-mutant lines (Supplementary Figure S4B–D). These observations are in contrast to the full restoration of both scaRNA13 and TERC steady state levels in *DKC1*

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

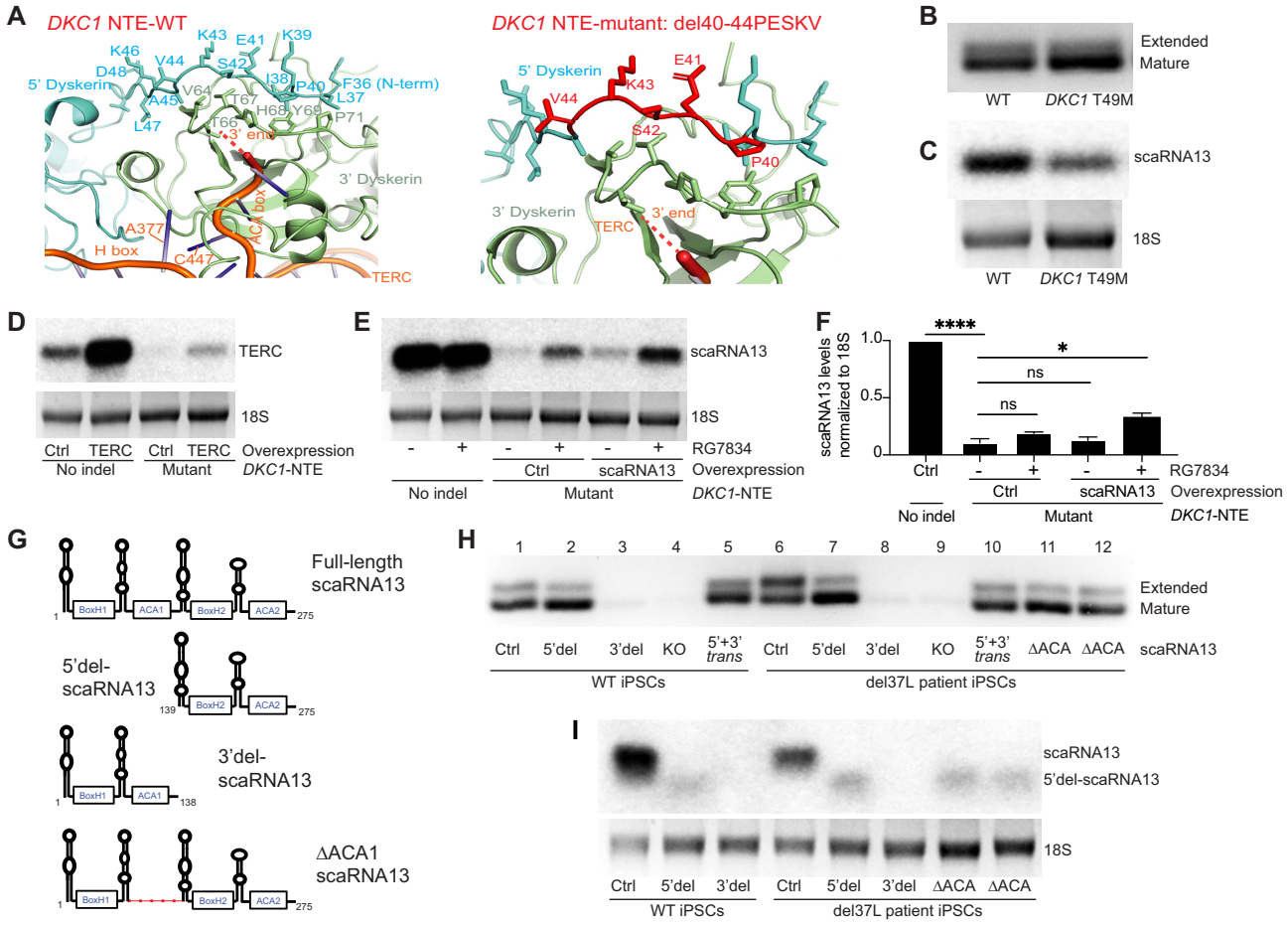


Figure 5. The NTE domain is implicated in interactions between dyskerin molecules and the 3′ end of bound RNAs in H/ACA RNPs. (**A**) Dyskerin-TERC interface within the H/ACA RNP lobe of the human telomerase cryo-EM structure shown in a cartoon representation. 3′ terminal nucleotides 450 and 451 of TERC, which are not resolved in the cryo-EM structure, are represented with a dashed line (red). *Left:* Anti-parallel like interaction surface formed between the N-terminal regions of 5′ dyskerin (cyan; F35-T48 shown in sticks) and 3′ dyskerin (green; subset of residues within V64-P71 shown in sticks) that protects the 3′ end of TERC is highlighted. *Right:* Magnified views of the interaction between the N-termini of the two dyskerin molecules where residues mutated in *DKC1* NTE-mutant iPSC clone 3 (*see* Figure 3A) are shown as red sticks. (**B**) scaRNA13 3′ end profiles in *DKC1* p.T49M mutant patient iPSCs compared to normal (WT). n = 3 independent experiments, one representative image is shown. (**C**) scaRNA13 RNA levels by northern blot of samples in (B). Inverted image of ethidium bromide staining of 18S rRNA is shown as a loading control. n = 3 independent experiments, one representative image is shown. (**D**) TERC levels by northern blot in CRISPR *DKC1* iPSC clone 2 (NTE-no indel) and clone 5 (NTE-mutant), comparing cells in which TERC has been stably overexpressed versus empty vector (Ctrl). $n = 2$ independent experiments, one representative image is shown. (**E**) scaRNA13 levels by northern blot in *DKC1* NTE-mutant iPSC clone 5, where scaRNA13 has been stably overexpressed using a transposon-based expression vector (scaRNA13) versus control (Ctrl), followed by seven days of RG7834 (+) versus DMSO (–) treatment. *DKC1DKC1*iPSC clone 2 (NTE-no indel) is shown as a control. (**F**) Quantification and comparison of relative normalized scaRNA13 levels in (E). Mean ± S.D. is shown for two *DKC1* NTE-mutant clones. Statistics: ns-not significant, $*P < 0.05$, $****P < 0.0001$, one-way ANOVA. (**G**) Schematics of scaRNA13 predicted secondary structures after genome-engineering. *First*: Full-length scaRNA13 secondary structure showing tandem, duplex hinge-hairpin-hinge box H and ACA domains. *Second*: Schematic of predicted mature 3′ half of scaRNA13 (dubbed 5′del-scaRNA13). *Third*: Schematic of the predicted mature 5′ half of scaRNA13 (dubbed 3′del-scaRNA13). *Fourth*: Schematic of box ACA1 mutant scaRNA13 (dubbed ΔACA1-scaRNA13). (**H**) scaRNA13 3′ end profiles by RACE in WT and del37L patient iPSCs clones carrying CRISPR-Cas9 engineered mutations in the endogenous scaRNA13 locus (predicted product in parentheses): *Ctrl* (full-length scaRNA13), *5′del* (3′ half scaRNA13); *3′del* (5′ half scaRNA13), *KO* (null; biallelic deletion of scaRNA13 residues 21–273), *5′+3′trans* (compound heterozygous; 5′del- and 3′del-scaRNA13), *ΔACA1* (3′ half scaRNA13, due to disruption of 5′ box ACA1). n = 3 independent experiments, one representative image is shown. (**I**) Northern blot showing transcript levels and sizes using a probe targeting the 3′ half of scaRNA13 (5′del-scaRNA13), in engineered clones in (H). $n = 2$ independent experiments, one representative image is shown.

del37L patient iPSCs (Figure 4A, Supplementary Figure S4E), and are possibly due to a milder impact on dyskerin function in this single amino acid deletion compared to the CRISPR-engineered mutations. Taken together, these observations provide a novel link between the dyskerin NTE and PARN/PAPD5 axis in scaRNA13 biogenesis, and provide a mechanism for the effects of PAPD5 inhibition in the setting of *DKC1* mutations (56,63).

Dual function of the NTE domain of dyskerin in H/ACA RNPs

To gain mechanistic insights into how mutations in the NTE domain of dyskerin impact snoRNP formation, we analyzed the recently published cryo-EM structure of the human telomerase RNP (35). The dyskerin-TERC interaction lies in the H/ACA RNP lobe of the human telomerase structure, showing how two juxtaposed dyskerin molecules

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

bring together the H and the ACA box of TERC and protect the 3′ end of TERC from deregulated processing (Figure 5A, Supplementary Figure S5A). Dyskerin dimerization within the RNP includes an anti-parallel interaction between extended regions at the N-termini of the two dyskerin molecules. The 3′ dyskerin-half of this interface interacts with the 3′ end of TERC. We envision that mutations that either delete or otherwise mutate residues in the L37-T48 region of dyskerin will disrupt the structure of the 5′ dyskerin N-terminus, destabilize its interaction with 3′ dyskerin N-terminus, and ultimately result in a reduced ability of dyskerin to bind the 3′ end of TERC/snoRNAs (Figure 5A). This may result in decreased protection of the snoRNA 3′ end, thereby, providing greater access to PAPD5, the RNA exosome and other ribonucleases.

The impact of L37-T48 mutations on 3′ end processing is proposed to be in addition to a general destabilizing effect of pathogenic *DKC1* mutations caused by impaired binding of sno/scaRNAs. In keeping with this dual function and demarcation of the region involved in 3′ end protection, cells from a patient carrying a pathogenic *DKC1* mutation of residue T49M did not show aberrant scaRNA13 3′ end processing, but showed lower TERC and scaRNA13 levels (Figure 5B,C, Supplementary Figure S5B-E). Furthermore, ectopic expression of TERC in the NTE-domain mutant cells yielded only partial restoration of TERC levels and telomere length in iPSCs (Figure 5D, Supplementary Figure S5F,G). Similarly, ectopic expression of scaRNA13 could not restore scaRNA13 levels in the settings of mutant NTE domain of dyskerin, and partial restoration of scaRNA13 steady state levels was observed only when scaRNA13 overexpression was combined with PAPD5 inhibition (Figure 5E, F, Supplementary Figure S5H). Taken together, our results suggest a dual function of dyskerin, not only in general sno/scaRNA stabilization through binding and nucleating the H/ACA RNP, but also 3′ end protection of certain species like scaRNA13 mediated by NTE residues.

Specific effects of dyskerin NTE mutations on scaRNA13 are governed by its tandem structure

The effects of dyskerin NTE mutations on 3′ end-processing appeared to be most pronounced for scaRNA13, but for unclear reasons. ScaRNA13 is unusual amongst sno/scaRNAs in possessing a 275 nucleotide tandem structure, in effect a duplex snoRNA with two box H and two ACA domains (Figure 5G) (61). We hypothesized that the dual structure of scaRNA13 may underlie its processing sensitivity in the setting of dyskerin NTE mutations. To test this, we used CRISPR-Cas9 RNP with two gRNAs in del37L patient iPSCs to precisely delete nucleotides 21–138 of the 5′ hairpin-hinge-hairpin biallelically in the endogenous scaRNA13 loci, leaving only the sequence coding for the 3′ half of scaRNA13 intact (dubbed *5′del-scaRNA13*) (Figure 5G). Strikingly, RLM-RACE analysis showed that the 3′ end processing defect seen with scaRNA13 was eliminated in the setting of the monomeric form, 5′del-scaRNA13, despite harboring the same 3′ sequence context in the endogenous locus (Figure 5H, compare lanes 6 and 7). We confirmed that an intact monomeric 5′del-scaRNA13 form was detected by northern blot, albeit with reduced steady state levels (Figure 5I). To explore this further, we deleted nucleotides 139–273 encompassing the scaRNA13 3′ hairpin-hinge-hairpin on both endogenous loci in WT and *DKC1* del37L iPSCs, leaving only the sequence coding for the 5′ half of scaRNA13 intact (*3′del-scaRNA13*; see Figure 5G). No PCR product could be obtained with the conventional RACE strategy due to the position of the RACE primer near the 3′ hairpin (Figure 5H). However, using a primer from the 5′ end of scaRNA13 in RLM-RACE, we detected transcripts corresponding to the 5′ hairpin-hinge-hairpin (*3′del-scaRNA13*), further confirmed by northern blot (Figure 5SI–K). By deep sequencing of these RACE amplicons, we found that the 3′del-scaRNA13 fragment (5′ half) terminated three nucleotides after the ACA1 motif, irrespective of variations in the 3′ sequence context, thus undergoing maturation like a *bona fide* scaRNA. We also found no increase in oligoadenylation of the the 3′del-scaRNA13 fragment in the *DKC1* del37L cells compared to WT iPSCs (Supplementary Figure S5L). These results confirm the dual H/ACA RNP structure of scaRNA13, and are consistent with our hypothesis that its tandem structure is responsible for its sensitivity to dyskerin NTE mutations. Moreover, our results predict that the dyskerin NTE-mutant acts through an intact 5′ H/ACA RNP in *cis* to disrupt scaRNA13 3′ end processing. To investigate this further, we created compound heterozygous 5′del- and 3′del-scaRNA13 cells (dubbed *5′+3′-trans*) and found that although the 5′ H/ACA RNP is present in the cell, it is not capable of altering 3′ end processing of scaRNA13 in *trans* (Figure 5H, lane 10). We next created homozygous mutants of the linker region deleting the ACA1 motif (ΔACA1; Figure 5G), which is predicted to be essential for formation of the 5′ H/ACA RNP. Indeed we found the ΔACA1 mutation abrogated formation of full-length scaRNA13 (Figure 5I), but the resulting 5′del-scaRNA13 fragment (3′ half) had intact 3′ end processing (Figure 5H, lanes 11 and 12). Taken together, these results suggest that dyskerin NTE-mutants disrupt an essential intramolecular interaction in scaRNA13 via 5′ sequences that must form an H/ACA RNP, resulting in 3′ end processing defects. More broadly, these results demonstrate that the structure of a scaRNA also underlies its specific sensitivity to particular *DKC1* mutations.

DISCUSSION

In this study, we provide a mechanistic explanation for differential snoRNA perturbations due to disease-causing variants in a mutational hotspot in *DKC1*. By examining an allelic series of DC patient mutations coupled with CRISPR-Cas9 engineered mutations in human iPSCs, we defined a domain-dependent role of *DKC1*-mutations on ncRNA 3′ end maturation. Specifically, we identified scaRNA13 as the only one of 27 scaRNAs analyzed that is differentially modified post-transcriptionally in patient iPSCs carrying a *DKC1* NTE domain mutation. Consistent with the findings in patient cells, we observed selective effects on scaRNA13 when we engineered NTE mutations in iPSCs using CRISPR-Cas9. Pathogenic mutations across different domains of dyskerin are sufficient to impair the stability of TERC (8,64). However, we did not observe

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

defects in 3′ end maturation of TERC due to *DKC1* mutations, nor domain-specific effects on other members of the scaRNA family. Importantly, quantitative depletion of dyskerin was sufficient to dramatically decrease steady state levels of H/ACA RNAs including TERC and scaRNA13, yet did not phenocopy scaRNA13 3′ end processing abnormalities found in the setting of dyskerin NTE mutations. These results indicate that domain-specific, qualitative defects can result from pathogenic *DKC1* mutations to perturb specific ncRNAs, beyond quantitative or more global effects on H/ACA snoRNA binding and stabilization (32,64).

Our study was motivated by our observations that diminished TERC levels in the setting of *DKC1* mutations show responsiveness to PAPD5 inhibitors (56), suggesting that dyskerin plays a role in TERC 3′ end protection from PAPD5, which is important to elucidate for therapeutic development in DC/telomere diseases. However, we found no direct evidence of this with respect to proximal TERC 3′ end processing by PARN. One potential explanation lies in the observation that the dyskerin-H/ACA RNP complex stimulates productive maturation of longer nascent TERC transcript forms by the exosome-associated nuclease RRP6, which may be in competition with PAPD5-mediated degradation (65). Alternatively, a shift in the balance of PARN/PAPD5 activities, independent of a direct role for dyskerin, may be sufficient to promote proximal 3′ end maturation and augment TERC levels in the setting of *DKC1* mutations. In this context, a distinguishing feature of the work presented here lies in revealing a direct impact of the integrity of the dyskerin-H/ACA RNP on proximal 3′ end maturation and protection of an associated snoRNA against PAPD5-mediated oligo-adenylation and destabilization.

One of the striking findings of our study is that NTE mutations and even complete disruption of the NTE domain was compatible with cell viability, as opposed to shRNA-mediated knockdown or genetic ablation of dyskerin (66). In particular, we isolated and propagated several mutant clones with an in-frame replacement of residues 29–49 of dyskerin with 31 heterologous amino acids (exemplified by clone 5, see Figure 3A). This was a result of recurrent recombination with the chr1:28606460–28606551 locus, likely due to homology in the flanking regions or fixed proximity in nuclear chromatin. Of note, all the iPSCs we engineered carried deletions of multiple amino acids in the NTE, unlike *DKC1* NTE mutations in patients that generally affect a single amino acid (36,67). The molecular phenotypes in NTE-engineered iPSCs were more pronounced than in the patient-derived (del37L) NTE-mutant iPSCs, in that we were unable to completely rescue TERC or scaRNA13 levels by overexpression in CRISPR-Cas9 generated NTE-mutant dyskerin iPSC clones, nor by inhibition of PAPD5. These observations support the notion that the functional defects caused by documented pathogenic mutations in dyskerin are relatively modest, and loss-of-function beyond what is observed in DC patients is likely incompatible with life (66).

It was unclear why scaRNA13 3′-end processing was specifically compromised by mutations in NTE domain of dyskerin, in contrast to TERC and other H/ACA RNAs. We noted that scaRNA13 is unique in possessing a tandem structure composed of two H/ACA RNAs (61), and found that its composite architecture contributes towards differences in its 3′ end processing and sensitivity to dyskerin NTE mutations. In the context of TERC, several pathogenic *DKC1* mutations map to intermolecular interactions of the two dyskerin molecules involving the '5′ dyskerin' NTE and a hydrophobic pocket created by the '3′ dyskerin' NTE and PUA regions (35). It is possible that the higher order complex of four copies of dyskerin-NOP10-NHP2-GAR1 with four hairpins alters dyskerin NTE-3′ RNA end interactions uniquely in the scaRNA13 RNP, diminishing PARN engagement and/or making the scaRNA13 3′ end more susceptible to oligo-adenylation by PAPD5. Consistent with this, our RACE amplicon sequencing data showed that oligo-adenylation occurred predominantly +1 nucleotide downstream of the canonical scaRNA13 end in the context of the *DKC1* del37L NTE mutation, a phenotype distinct from the adenylation of longer genomic extensions seen in *PARN*-mutant cells. Taken together, we speculate that dyskerin NTE mutations impair completion of 3′ end maturation of scaRNA13 in the RNP, enabling PAPD5-mediated oligo-adenylation and further destabilization. An alternative but less likely explanation is that the tandem architecture of scaRNA13 stabilizes it compared to other H/ACA RNAs in the setting of NTE mutations, permitting isolation and observation of the intermediate 3′ extended form. Further structural and biochemical studies will be required to address these possibilities.

Our ability to associate the specific defect in scaRNA13 shown here with DC disease phenotypes is restricted both by a lack of sufficiently robust clinical annotation across relevant genotypes and, as for most snoRNAs, by a limited understanding of scaRNA13 functions. With respect to the latter, scaRNA13 upregulation has been associated with hepatocellular carcinoma progression in one study (68), and *in silico* predictions suggest scaRNA13 may serve as a guide for U2 and/or U5 snRNA pseudouridylation (61,69,70). How these functions might relate to DC clinical manifestations when scaRNA13 is decreased by mutations in *DKC1* or *PARN* remains to be elucidated.

Overall, our study reveals disease-causing mutations in the N-terminal domain of dyskerin can impact specific snoRNAs such as scaRNA13, expanding the repertoire of dyskerin functions to include 3′ end maturation of H/ACA RNAs. We propose that further defining the composition and determinants of the dyskerin-dependent snoRNAome in the setting of specific mutations will uncover novel disease mechanisms.

DATA AVAILABILITY

The data that support the findings of this study are available within the manuscript and associated extended data files. Unprocessed data will be provided on request. There are no restrictions on data availability from this study. The RNA-seq data have been deposited in NCBI's Gene Expression Omnibus and can be accessed through GEO series accession number GEO: GSE190173. The RLM-RACE-seq data have been deposited to NCBI's Sequence Read Archive and can be accessed through the BioProject PRJNA785578 (Biosamples SUB10749232). The

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

code used in this study can be accessed at https://github.com/alberttai/tailAcount. Further requests for information should be directed to the lead contact, Suneet Agarwal, at suneet.agarwal@childrens.harvard.edu.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank patients and families for participation in research; Dr Yick Fong for critical reading of the manuscript; Dr Akiko Shimamura and the BCH Bone Marrow Failure/Myelodysplastic Syndrome registry, supported by grants from the NIH (RC2DK122533).

FUNDING

Boston Children's Hospital (BCH) Manton Center for Orphan Disease Research and Uplifting Athletes (to N.N.); National Institutes of Health [R01DK107716, R33HL154133 to S.A., R01GM120094 to J.N.]; Team Telomere/Million Dollar Bike Ride/Penn Orphan Disease Center, Harvard Stem Cell Institute, and philanthropic gifts (to S.A.); American Cancer Society [RSG-17-037-01-DMC to J.N.]. Funding for open access charge: Boston Children's Hospital.

Conflict of interest statement. N.N. and S.A. are listed as co-inventors on provisional patent applications that includes PAPD5 inhibitors used in this manuscript.

REFERENCES

1. de Almeida,R.A., Fraczek,M.G., Parker,S., Delneri,D. and O'Keefe,R.T. (2016) Non-coding RNAs and disease: the classical ncRNAs make a comeback. *Biochem. Soc. Trans.*, **44**, 1073–1078.
2. Esteller,M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.*, **12**, 861–874.
3. Gebauer,F., Schwarzl,T., Valcarcel,J. and Hentze,M.W. (2021) RNA-binding proteins in human genetic disease. *Nat. Rev. Genet.*, **22**, 185–198.
4. Jonkhout,N., Tran,J., Smith,M.A., Schonrock,N., Mattick,J.S. and Novoa,E.M. (2017) The RNA modification landscape in human disease. *RNA*, **23**, 1754–1769.
5. Armanios,M. and Blackburn,E.H. (2012) The telomere syndromes. *Nat. Rev. Genet.*, **13**, 693–704.
6. Calado,R.T. and Young,N.S. (2009) Telomere diseases. *N. Engl. J. Med.*, **361**, 2353–2365.
7. Bertuch,A.A. (2016) The molecular genetics of the telomere biology disorders. *RNA Biol*, **13**, 696–706.
8. Mitchell,J.R., Wood,E. and Collins,K. (1999) A telomerase component is defective in the human disease dyskeratosis congenita. *Nature*, **402**, 551–555.
9. Niewisch,M.R. and Savage,S.A. (2019) An update on the biology and management of dyskeratosis congenita and related telomere biology disorders. *Expert Rev Hematol*, **12**, 1037–1052.
10. Agarwal,S. (2018) Evaluation and management of hematopoietic failure in dyskeratosis congenita. *Hematol. Oncol. Clin. North Am.*, **32**, 669–685.
11. Kirwan,M. and Dokal,I. (2008) Dyskeratosis congenita: a genetic disorder of many faces. *Clin. Genet.*, **73**, 103–112.
12. Vulliamy,T. and Dokal,I. (2006) Dyskeratosis congenita. *Semin. Hematol.*, **43**, 157–166.
13. Niewisch,M.R., Giri,N., McReynolds,L.J., Alsaggaf,R., Bhala,S., Alter,B.P. and Savage,S.A. (2022) Disease progression and clinical outcomes in telomere biology disorders. *Blood*, **139**, 1807–1819.
14. Bodnar,A.G., Ouellette,M., Frolkis,M., Holt,S.E., Chiu,C.P., Morin,G.B., Harley,C.B., Shay,J.W., Lichtsteiner,S. and Wright,W.E. (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science*, **279**, 349–352.
15. Vaziri,H., Dragowska,W., Allsopp,R.C., Thomas,T.E., Harley,C.B. and Lansdorp,P.M. (1994) Evidence for a mitotic clock in human hematopoietic stem cells: loss of telomeric DNA with age. *Proc. Nat. Acad. Sci. U.S.A.*, **91**, 9857–9860.
16. Bryan,T.M. and Cech,T.R. (1999) Telomerase and the maintenance of chromosome ends. *Curr. Opin. Cell Biol.*, **11**, 318–324.
17. Feng,J., Funk,W.D., Wang,S.S., Weinrich,S.L., Avilion,A.A., Chiu,C.P., Adams,R.R., Chang,E., Allsopp,R.C., Yu,J. *et al.* (1995) The RNA component of human telomerase. *Science*, **269**, 1236–1241.
18. Mitchell,J.R., Cheng,J. and Collins,K. (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3′ end. *Mol. Cell. Biol.*, **19**, 567–576.
19. Kiss,T., Fayet-Lebaron,E. and Jady,B.E. (2010) Box H/ACA small ribonucleoproteins. *Mol. Cell*, **37**, 597–606.
20. Venteicher,A.S., Abreu,E.B., Meng,Z., McCann,K.E., Terns,R.M., Veenstra,T.D., Terns,M.P. and Artandi,S.E. (2009) A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science*, **323**, 644–648.
21. Jady,B.E., Bertrand,E. and Kiss,T. (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.*, **164**, 647–652.
22. Ganot,P., Caizergues-Ferrer,M. and Kiss,T. (1997) The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.*, **11**, 941–956.
23. Fu,D. and Collins,K. (2003) Distinct biogenesis pathways for human telomerase RNA and H/ACA small nucleolar RNAs. *Mol. Cell*, **11**, 1361–1372.
24. Weinrich,S.L., Pruzan,R., Ma,L., Ouellette,M., Tesmer,V.M., Holt,S.E., Bodnar,A.G., Lichtsteiner,S., Kim,N.W., Trager,J.B. *et al.* (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat. Genet.*, **17**, 498–502.
25. Nagpal,N. and Agarwal,S. (2020) Telomerase RNA processing: implications for human health and disease. *Stem Cells*, **28**, 1532–1543.
26. Heiss,N.S., Knight,S.W., Vulliamy,T.J., Klauck,S.M., Wiemann,S., Mason,P.J., Poustka,A. and Dokal,I. (1998) X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat. Genet.*, **19**, 32–38.
27. Walne,A.J., Vulliamy,T., Marrone,A., Beswick,R., Kirwan,M., Masunari,Y., Al-Qurashi,F.H., Aljurf,M. and Dokal,I. (2007) 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.
28. Vulliamy,T., Beswick,R., Kirwan,M., Marrone,A., Digweed,M., Walne,A. and Dokal,I. (2008) Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. *Proc. Nat. Acad. Sci. U.S.A.*, **105**, 8073–8078.
29. Zhong,F., Savage,S.A., Shkreli,M., Giri,N., Jessop,L., Myers,T., Chen,R., Alter,B.P. and Artandi,S.E. (2011) Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev.*, **25**, 11–16.
30. Stanley,S.E., Gable,D.L., Wagner,C.L., Carlile,T.M., Hanumanthu,V.S., Podlevsky,J.D., Khalil,S.E., DeZern,A.E., Rojas-Duran,M.F., Applegate,C.D. *et al.* (2016) Loss-of-function mutations in the RNA biogenesis factor NAF1 predispose to pulmonary fibrosis-emphysema. *Sci. Transl. Med.*, **8**, 351ra107.
31. Hamma,T. and Ferre-D'Amare,A.R. (2010) The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J. Biol. Chem.*, **285**, 805–809.
32. Egan,E.D. and Collins,K. (2012) An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol. Cell. Biol.*, **32**, 2428–2439.
33. Garus,A. and Autexier,C. (2021) Dyskerin: an essential pseudouridine synthase with multifaceted roles in ribosome biogenesis, splicing, and telomere maintenance. *RNA*, **27**, 1441–1458.
34. Watkins,N.J., Gottschalk,A., Neubauer,G., Kastner,B., Fabrizio,P., Mann,M. and Luhrmann,R. (1998) Cbf5p, a potential pseudouridine

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022

synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA*, **4**, 1549–1568.
35. Ghanim,G.E., Fountain,A.J., van Roon,A.M., Rangan,R., Das,R., Collins,K. and Nguyen,T.H.D. (2021) Structure of human telomerase holoenzyme with bound telomeric DNA. *Nature*, **593**, 449–453.
36. Podlevsky,J.D., Bley,C.J., Omana,R.V., Qi,X. and Chen,J.J. (2008) The telomerase database. *Nucleic Acids Res.*, **36**, D339–D343.
37. Bellodi,C., McMahon,M., Contreras,A., Juliano,D., Kopmar,N., Nakamura,T., Maltby,D., Burlingame,A., Savage,S.A., Shimamura,A. *et al.* (2013) H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation. *Cell Rep.*, **3**, 1493–1502.
38. Balogh,E., Chandler,J.C., Varga,M., Tahoun,M., Menyhard,D.K., Schay,G., Goncalves,T., Hamar,R., Legradi,R., Szekeres,A. *et al.* (2020) Pseudouridylation defect due to DKC1 and NOP10 mutations causes nephrotic syndrome with cataracts, hearing impairment, and enterocolitis. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 15137–15147.
39. Yoon,A., Peng,G., Brandenburger,Y., Zollo,O., Xu,W., Rego,E. and Ruggero,D. (2006) Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science*, **312**, 902–906.
40. Schwartz,S., Bernstein,D.A., Mumbach,M.R., Jovanovic,M., Herbst,R.H., Leon-Ricardo,B.X., Engreitz,J.M., Guttman,M., Satija,R., Lander,E.S. *et al.* (2014) Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell*, **159**, 148–162.
41. Penzo,M., Rocchi,L., Brugiere,S., Carnicelli,D., Onofrillo,C., Coute,Y., Brigotti,M. and Montanaro,L. (2015) Human ribosomes from cells with reduced dyskerin levels are intrinsically altered in translation. *FASEB J.*, **29**, 3472–3482.
42. Thumati,N.R., Zeng,X.L., Au,H.H., Jang,C.J., Jan,E. and Wong,J.M. (2013) Severity of X-linked dyskeratosis congenita (DKCX) cellular defects is not directly related to dyskerin (DKC1) activity in ribosomal RNA biogenesis or mRNA translation. *Hum. Mutat.*, **34**, 1698–1707.
43. Tummala,H., Walne,A., Collopy,L., Cardoso,S., de la Fuente,J., Lawson,S., Powell,J., Cooper,N., Foster,A., Mohammed,S. *et al.* (2015) Poly(A)-specific ribonuclease deficiency impacts telomere biology and causes dyskeratosis congenita. *J. Clin. Invest.*, **125**, 2151–2160.
44. Stuart,B.D., Choi,J., Zaidi,S., Xing,C., Holohan,B., Chen,R., Choi,M., Dharwadkar,P., Torres,F., Girod,C.E. *et al.* (2015) Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat. Genet.*, **47**, 512–517.
45. Moon,D.H., Segal,M., Boyraz,B., Guinan,E., Hofmann,I., Cahan,P., Tai,A.K. and Agarwal,S. (2015) Poly(A)-specific ribonuclease (PARN) mediates 3′-end maturation of the telomerase RNA component. *Nat. Genet.*, **47**, 1482–1488.
46. Gable,D.L., Gaysinskaya,V., Atik,C.C., Talbot,C.C., Kang,B., Stanley,S.E., Pugh,E.W., Amat-Codina,N., Schenk,K.M., Arcasoy,M.O *et al.* (2019) ZCCHC8, the nuclear exosome targeting component, is mutated in familial pulmonary fibrosis and is required for telomerase RNA maturation. *Genes Dev.*, **33**, 1381–1396.
47. Dhanraj,S., Gunja,S.M., Deveau,A.P., Nissbeck,M., Boonyawat,B., Coombs,A.J., Renieri,A., Mucciolo,M., Marozza,A., Buoni,S. *et al.* (2015) Bone marrow failure and developmental delay caused by mutations in poly(A)-specific ribonuclease (PARN). *J. Med. Genet.*, **52**, 738–748.
48. Son,A., Park,J.E. and Kim,V.N. (2018) PARN and TOE1 constitute a 3′ end maturation module for nuclear Non-coding RNAs. *Cell Rep.*, **23**, 888–898.
49. Shukla,S., Bjerke,G.A., Muhlrad,D., Yi,R. and Parker,R. (2019) The RNase PARN controls the levels of specific miRNAs that contribute to p53 regulation. *Mol. Cell*, **73**, 1204–1216.
50. Shukla,S. and Parker,R. (2017) PARN modulates y RNA stability and its 3′-End formation. *Mol. Cell. Biol.*, **37**, e00264-17.
51. Berndt,H., Harnisch,C., Rammelt,C., Stohr,N., Zirkel,A., Dohm,J.C., Himmelbauer,H., Tavanez,J.P., Huttelmaier,S. and Wahle,E. (2012) Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA*, **18**, 958–972.
52. Nguyen,D., Grenier St-Sauveur,V., Bergeron,D., Dupuis-Sandoval,F., Scott,M.S. and Bachand,F. (2015) A polyadenylation-dependent 3′ end maturation pathway is required for the synthesis of the human telomerase RNA. *Cell Rep.*, **13**, 2244–2257.
53. Shukla,S., Schmidt,J.C., Goldfarb,K.C., Cech,T.R. and Parker,R. (2016) Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects. *Nat. Struct. Mol. Biol.*, **23**, 286–292.
54. Tseng,C.K., Wang,H.F., Burns,A.M., Schroeder,M.R., Gaspari,M. and Baumann,P. (2015) Human telomerase RNA processing and quality control. *Cell Rep.*, **13**, 2232–2243.
55. Boyraz,B., Moon,D.H., Segal,M., Muosieyiri,M., Aykanat,A., Tai,A.K., Cahan,P. and Agarwal,S. (2016) Post-transcriptional manipulation of TERC reverses molecular hallmarks of telomere disease. *J. Clin. Invest.*, **126**, 3377–3382.
56. Nagpal,N., Wang,J., Zeng,J., Lo,E., Moon,D.H., Luk,K., Braun,R.O., Burroughs,L.M., Keel,S.B., Reilly,C. *et al.* (2020) Small-Molecule PAPD5 inhibitors restore telomerase activity in patient stem cells. *Cell Stem Cell*, **26**, 896–909.
57. Roake,C.M., Chen,L., Chakravarthy,A.L., Ferrell,J.E., Raffa,G.D. and Artandi,S.E. (2019) Disruption of telomerase RNA maturation kinetics precipitates disease. *Mol. Cell*, **74**, 688–700.
58. Agarwal,S., Loh,Y.H., McLoughlin,E.M., Huang,J., Park,I.H., Miller,J.D., Huo,H., Okuka,M., Dos Reis,R.M., Loewer,S. *et al.* (2010) Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. *Nature*, **464**, 292–296.
59. Paulsen,B.S., Mandal,P.K., Frock,R.L., Boyraz,B., Yadav,R., Upadhyayula,S., Gutierrez-Martinez,P., Ebina,W., Fasth,A., Kirchhausen,T. *et al.* (2017) Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR-Cas9 genome editing. *Nat. Biomed. Eng.*, **1**, 878–888.
60. Warlich,E., Kuehle,J., Cantz,T., Brugman,M.H., Maetzig,T., Galla,M., Filipczyk,A.A., Halle,S., Klump,H., Scholer,H.R. *et al.* (2011) Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming. *Mol. Ther.*, **19**, 782–789.
61. Kiss,A.M., Jady,B.E., Darzacq,X., Verheggen,C., Bertrand,E. and Kiss,T. (2002) A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains. *Nucleic Acids Res.*, **30**, 4643–4649.
62. Kowarz,E., Loscher,D. and Marschalek,R. (2015) Optimized sleeping beauty transposons rapidly generate stable transgenic cell lines. *Biotechnol. J.*, **10**, 647–653.
63. Fok,W.C., Shukla,S., Vessoni,A.T., Brenner,K.A., Parker,R., Sturgeon,C.M. and Batista,L.F.Z. (2019) Posttranscriptional modulation of TERC by PAPD5 inhibition rescues hematopoietic development in dyskeratosis congenita. *Blood*, **133**, 1308–1312.
64. MacNeil,D.E., Lambert-Lanteigne,P. and Autexier,C. (2019) N-terminal residues of human dyskerin are required for interactions with telomerase RNA that prevent RNA degradation. *Nucleic Acids Res.*, **47**, 5368–5380.
65. Tseng,C.K., Wang,H.F., Schroeder,M.R. and Baumann,P. (2018) The H/ACA complex disrupts triplex in hTR precursor to permit processing by RRP6 and PARN. *Nat. Commun.*, **9**, 5430.
66. He,J., Navarrete,S., Jasinski,M., Vulliamy,T., Dokal,I., Bessler,M. and Mason,P.J. (2002) Targeted disruption of Dkc1, the gene mutated in X-linked dyskeratosis congenita, causes embryonic lethality in mice. *Oncogene*, **21**, 7740–7744.
67. Knight,S.W., Heiss,N.S., Vulliamy,T.J., Greschner,S., Stavrides,G., Pai,G.S., Lestringant,G., Varma,N., Mason,P.J., Dokal,I. *et al.* (1999) X-linked dyskeratosis congenita is predominantly caused by missense mutations in the DKC1 gene. *Am. J. Hum. Genet.*, **65**, 50–58.
68. Lan,T., Yuan,K., Yan,X., Xu,L., Liao,H., Hao,X., Wang,J., Liu,H., Chen,X., Xie,K. *et al.* (2019) LncRNA SNHG10 facilitates hepatocarcinogenesis and metastasis by modulating its homolog SCARNA13 via a positive feedback loop. *Cancer Res.*, **79**, 3220–3234.
69. Schattner,P., Barberan-Soler,S. and Lowe,T.M. (2006) A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *RNA*, **12**, 15–25.
70. Lestrade,L. and Weber,M.J. (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.*, **34**, D158–D162.

Downloaded from https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkac706/6677329 by guest on 30 August 2022