

De novo and inherited dominant variants in U4 and U6 snRNA genes cause retinitis pigmentosa

Received: 2 June 2025

Accepted: 14 November 2025

Published online: 9 January 2026

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Small nuclear RNAs (snRNAs) combine with specific proteins to generate small nuclear ribonucleoproteins (snRNPs), the building blocks of the spliceosome. U4 snRNA forms a duplex with U6 and, together with U5, contributes to the tri-snRNP spliceosomal complex. Variants in *RNU4-2*, which encodes U4, have recently been implicated in neurodevelopmental disorders. Here we show that heterozygous inherited and de novo variants in *RNU4-2* and in four *RNU6* paralogs (*RNU6-1*, *RNU6-2*, *RNU6-8* and *RNU6-9*), which encode U6, recur in individuals with nonsyndromic retinitis pigmentosa (RP), a genetic disorder causing progressive blindness. These variants cluster within the three-way junction of the U4/U6 duplex, a site that interacts with tri-snRNP splicing factors also known to cause RP (*PRPF3*, *PRPF8*, *PRPF31*), and seem to affect snRNP biogenesis. Based on our cohort, deleterious variants in *RNU4-2* and *RNU6* paralogs may explain up to ~1.4% of otherwise undiagnosed RP cases. This study highlights the contribution of noncoding RNA genes to Mendelian disease and reveals pleiotropy in *RNU4-2*, where distinct variants underlie neurodevelopmental disorder and retinal degeneration.

While approximately 2 million individuals worldwide are affected by retinitis pigmentosa (RP), it is estimated that 30% to 50% remain without a conclusive genetic diagnosis, even after exome or genome sequencing is performed^{1–4}. This reflects high genetic heterogeneity, limited testing access and as-yet-unidentified disease genes, which in general carry pathogenic variants that are exceedingly rare in the control population^{5–7}.

Noncoding RNAs are essential to many cellular processes, including pre-messenger RNA (pre-mRNA) splicing, which is ensured by the spliceosome, a macromolecular complex that in its major form is composed of five small nuclear RNAs (snRNAs), U1, U2, U4, U5 and U6, and ~300 proteins⁸. Each snRNA associates with a specific set of proteins to form a small nuclear ribonucleoprotein (snRNP), the functional unit of the spliceosome. Variants in *RNU4-2*, one of the two paralogs encoding U4, have been linked to a common neurodevelopmental disorder (NDD) known as ReNU syndrome (OMIM: 620851). These variants account for up to 0.4% of all NDD cases and lead to systematic misrecognition of

donor splice sites by the spliceosome^{9–11}. Likewise, *RNU2-2* and *RNU5B-1* have been recently associated with NDDs^{11–13}.

Several spliceosomal proteins are also known to be involved in a wide range of hereditary diseases, including RP, as first noted by McKie and colleagues¹⁴. Specifically, of the ~100 genes that are currently associated with nonsyndromic RP⁵, the tri-snRNP splicing factor genes *PRPF3*, *PRPF4*, *PRPF8*, *PRPF31* and *SNRNP200* underlie the autosomal dominant form of the condition (adRP), with variants in *PRPF31* accounting for 10–20% of all adRP cases^{3,15}.

Here, we identify both inherited and de novo variants in *RNU4-2* and four paralogs of *RNU6*, encoding the U6 snRNA, as the molecular cause of adRP in 153 individuals across 67 families. We demonstrate that all identified variants cluster within the U4/U6 duplex, in a region that binds directly to *PRPF31* and *PRPF3* and indirectly to *PRPF6* and *PRPF8*^{16,17}. Furthermore, we show that such variants increase the association of U4 and U6 snRNAs with the splicing factors *SART3* and *PRPF31*, suggesting impaired snRNP biogenesis.

✉ e-mail: susanne.roosing@radboudumc.nl

Results

RNU4-2 variants underlie adRP

We initially examined a nonconsanguineous family with adRP (Family M1-A; Supplementary Fig. 1), in which seven of eight siblings (II:1–II:7) and their father (I:1) displayed classical RP features (Supplementary Fig. 2 and Supplementary Data 1). Genome sequencing was negative for pathogenic variants in known retinal disease-associated genes, but selective DNA variant filtering and shared haplotype analysis revealed a total of 55 variants that were absent from gnomAD v.4.1^{7,18} and co-segregated with RP. Of these, none was predicted to impact splicing (SpliceAI > 0.2)¹⁹ and only one was evolutionarily conserved (GERP = 4.03 and phyloP-vertebrate = 3.18)^{20,21}, a single-nucleotide insertion in the gene *RNU4-2* (NR_003137.2:n.18_19insA; Fig. 1a, Supplementary Fig. 1 and Supplementary Tables 1 and 2). This DNA change was present in one individual from the All of Us database²².

To find additional families, we first screened by Sanger sequencing a cohort of 1,891 individuals from the European Retinal Disease Consortium (www.erdc.info) with RP or Leber congenital amaurosis who remained undiagnosed after a large high-throughput screening using single molecule Molecular Inversion Probes²³. This analysis led to the identification of three additional families comprising 15 affected individuals segregating the same pathogenic variant (Supplementary Fig. 1 and Supplementary Tables 1 and 2). The n.18_19insA allele was significantly enriched in the RP cohort compared with both the gnomAD and the All of Us databases (analyzed control genomes: 76,215 and 414,000, respectively; Bonferroni-corrected *P* values = 2.6×10^{-3} and 6.9×10^{-5} , respectively, by two-sided Fisher's test; Supplementary Table 3). Additional screening of the *RNU4-2* sequence in the same cohort led to the identification of 28 other variants, one of which (n.56T>C) recurred in eight individuals from four families (Fig. 1a, Supplementary Fig. 1 and Supplementary Tables 1 and 2), was absent in controls and was significantly enriched in patients versus controls (Bonferroni-corrected *P* values = 6.4×10^{-5} (gnomAD) and 7.9×10^{-8} (All of Us); Supplementary Table 3).

Additional screening of 2,830 RP cases without previous genetic diagnosis from our respective institutions' cohorts, the UK National Genomic Research Library (hosting data from the Genomics England 100,000 Genomes Project²⁴ and from the NHS Genomic Medicine Service) uncovered an additional patient harboring n.18_19insA (for whom the variant was de novo) and six families (nine affected individuals) carrying the n.56T>C variant (Supplementary Fig. 1 and Supplementary Tables 1 and 2). Altogether, recurrent variants in *RNU4-2* were identified in 41 affected individuals from 15 families (Supplementary Fig. 3 and Supplementary Tables 1 and 2). Of note, incomplete penetrance was observed for nine obligate carriers, without visual symptoms (Supplementary Fig. 1). One carrier of n.56T>C was asymptomatic, with subnormal electroretinogram, diffuse atrophic changes in the periphery and attenuated vessels. Another individual with the same variant showed no clinical signs of disease upon examination, and seven (among whom four were deceased) were not clinically evaluated to determine their disease status. Our combined screening of *RNU4-2* also revealed 24 other unique rare DNA changes in 27 families, which were classified as variants of uncertain significance (VUS), as well as three benign changes (Supplementary Table 3).

Because U4 snRNA can also be transcribed from its paralog *RNU4-1*, which differs from *RNU4-2* at only four positions (n.37, n.88, n.99 and n.113; Supplementary Table 4), we next examined its sequence in our initial cohort and identified 63 variants, none of which were significantly enriched in cases compared with controls; also, these changes did not include variants at sites corresponding to n.18_19 and n.56 of *RNU4-2* (Fig. 1b and Supplementary Table 3). Notably, *RNU4-1* appears to be more tolerant to variation compared with *RNU4-2*, as evidenced by the numerous and frequent variants that are present in genomes from the general population (cumulative allele frequency

of 20.4% in *RNU4-1* versus 1.2% in *RNU4-2*; gnomAD v.4.1) (Fig. 1b,c and Supplementary Fig. 4), as already noted previously⁹.

Variants in U6 paralogues also cause RP

In the di-snRNP and the tri-snRNP complexes of the major spliceosome, U4 binds to U6 to form the U4/U6 RNA duplex. We therefore hypothesized that variants in U6 could also underlie adRP and extended our analysis to all five identical paralogous genes producing the U6 snRNA, scattered across the genome (*RNU6-1*, *RNU6-2*, *RNU6-7*, *RNU6-8* and *RNU6-9*; Supplementary Table 4). A screening of these genes by Sanger sequencing in our initial cohort of 1,891 RP families revealed 94 DNA changes in total. The n.55_56insG insertion recurred at the exact relative position in *RNU6-2*, *RNU6-8* and *RNU6-9* (four families per gene, 34 cases in total; Supplementary Fig. 1 and Supplementary Tables 1 and 2) and was significantly enriched in cases versus controls, who were all negative for this change (Bonferroni-corrected *P* value = 2.6×10^{-18} (gnomAD) and 5.1×10^{-27} (All of Us); Supplementary Table 3). Since this variant was identical in three U6 genes, we reasoned that the specific DNA change, rather than any particular paralog, was relevant to the etiology of the disease. We therefore repeated our analysis by collapsing the five *RNU6* genes and detected 66 unique variants. Another insertion, n.56_57insG, was identified in two unrelated families (once in *RNU6-2* and once in *RNU6-9*, four cases in total; Supplementary Fig. 1 and Supplementary Table 2) and found to be significantly enriched in cases versus controls (Bonferroni-corrected *P* value = 1.8×10^{-3} (gnomAD, a single *RNU6-2* positive individual of unknown status) and 2.1×10^{-5} (All of Us, no positive individuals); Supplementary Table 3). We then extended our analysis to the same international cohorts of patients that were previously analyzed (*n* = 2,830) and identified 74 additional cases from 38 families who were positive for either n.55_56insG or n.56_57insG (Supplementary Table 2).

In total, these two variants were detected in 112 affected individuals from 52 families, involving all *RNU6* paralogs except *RNU6-7*. The n.55_56insG insertion was present in most cases (102 individuals from 47 families), occurring in four of the five *RNU6* paralogs: *RNU6-1*, *RNU6-2*, *RNU6-8* and *RNU6-9*, while n.56_57insG was present in ten individuals from five families, in *RNU6-1*, *RNU6-2* and *RNU6-9* (Supplementary Tables 1 and 2 and Supplementary Figs. 1 and 3). Notably, n.55_56insG was confirmed to be a de novo event in eight individuals, clinically identified as sporadic cases. In 14 additional pedigrees, it was also observed in individuals born to unaffected parents, for which de novo inheritance was suspected but could not be confirmed, due to the lack of parental DNA. In contrast, no de novo events could be detected for n.56_57insG, which was identified exclusively in families with adRP (Supplementary Fig. 1). Similar to the screening of the *RNU4* paralogs, our analysis of *RNU6* paralogs revealed 66 VUSs and 23 benign variants, validated by Sanger sequencing (Supplementary Table 3).

In summary, we identified variants in *RNU4-2* or *RNU6* paralogs that underlie de novo or inherited dominant RP in 67 families. The overall phenotype across all cases was consistent with classical RP, based on clinical examination and electrophysiological testing, with symptomatic onset predominantly in adolescence (Supplementary Table 5). In addition, other concurrent ocular disease features were noted across individuals in the cohort: cystoid macular edema (55.9%), non-age-related lens opacities (23.6%) and various vitreomacular complications (30.6%) (Supplementary Table 5). Based on our data from these 4,722 RP cases, mostly of European descent and lacking a genetic diagnosis, we estimate that *RNU4-* and *RNU6*-associated RP could be responsible for ~1.4% of all molecularly undiagnosed individuals with this disease. Furthermore, considering that approximately 30% of RP diagnoses correspond to adRP^{25,26} and that our positive families include 24 isolated individuals, we can further infer that these variants may account for approximately 3.0% of undiagnosed adRP families.

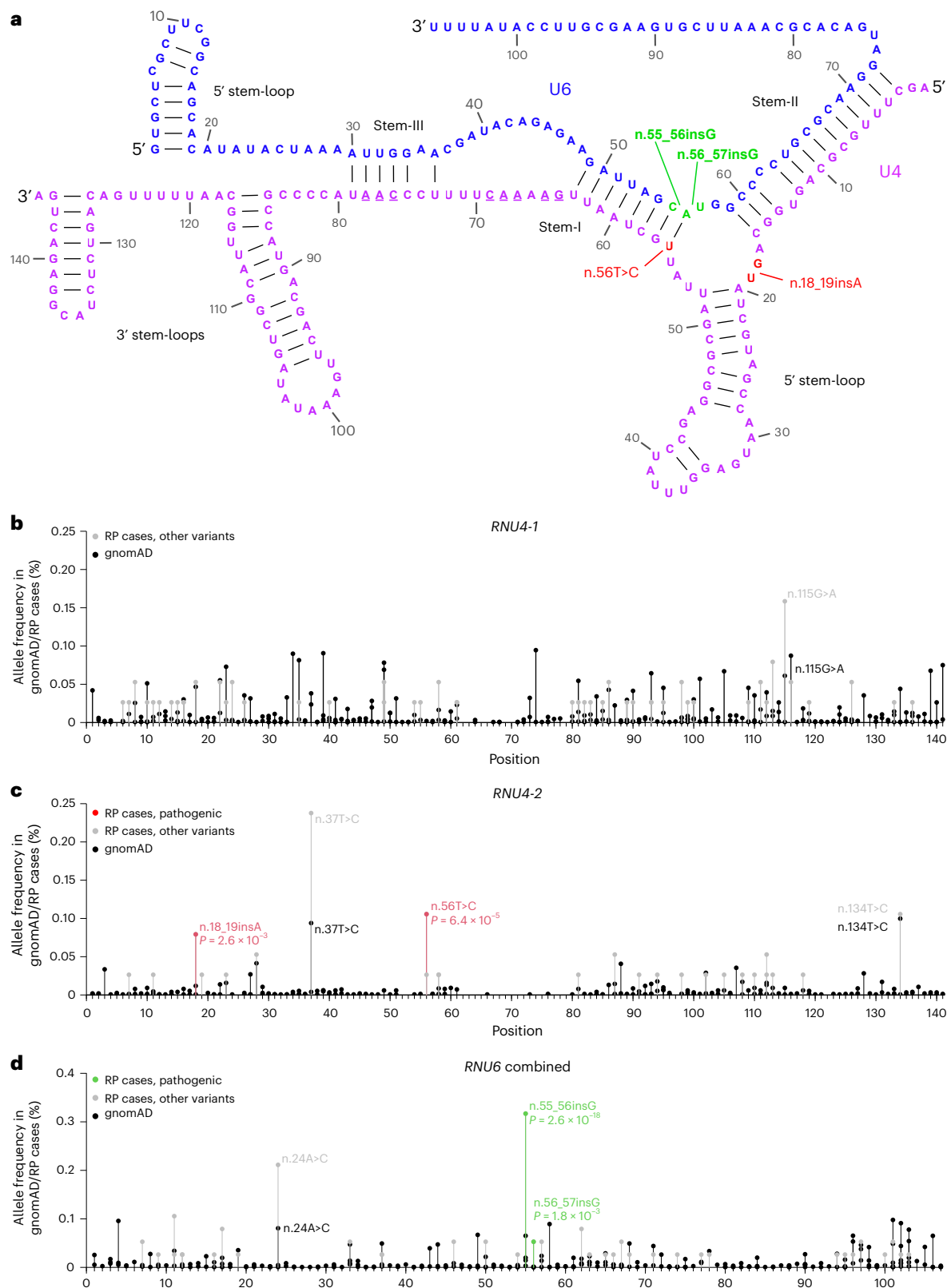


Fig. 1 | Structure of the U4/U6 duplex and rare variants found in RP cases and controls (gnomAD). **a**, Two-dimensional structure of the U4/U6 duplex, with recurrent variants identified in RP cases (in red for U4 and in green for U6), all clustering within the three-way junction. Nucleotides affected by variants previously observed in NDD cases are underlined. **b**, Rare variants affecting *RNU4-1*, defined as AF < 0.1% in gnomAD v.4.1, identified in RP cases and in

controls. **c**, Same as in **b** for *RNU4-2*, with recurrent pathogenic variants displayed in red. **d**, Same as **b** for all five *RNU6* paralogs combined, with recurrent causative variants displayed in green. Significant *P* values for variants enriched in RP cases versus controls from gnomAD are indicated (two-sided Fisher's test with Bonferroni correction).

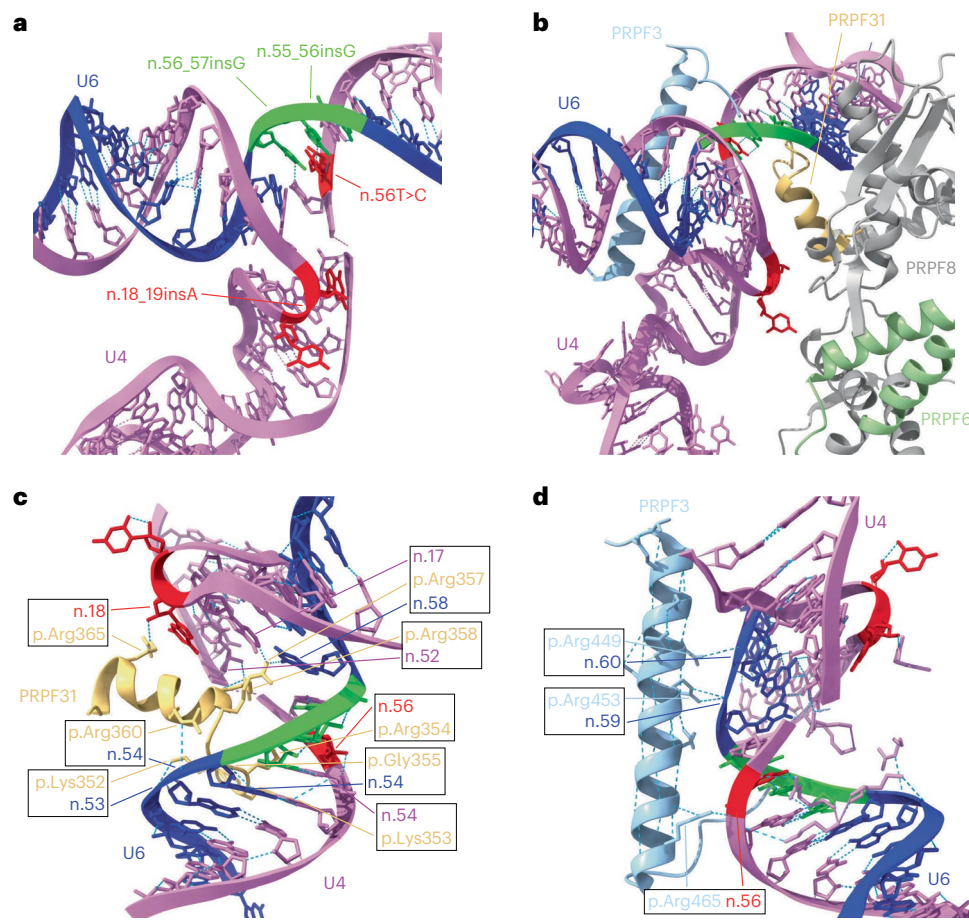


Fig. 2 | Three-dimensional structure, from cryo-electron microscopy (PDB 6QW6), of the U4/U6 duplex and its interactions with neighboring splicing factors. a, Naked U4/U6 pairing, showing the proximity of the causative variants

identified (red and green). **b**, Same as in **a** with interacting PRPF proteins. **c**, Direct interactions of nucleotides of the U4/U6 duplex with PRPF31, via hydrogen bonds. **d**, Same as **c** but for PRPF3.

Predicted effects of variants on the U4/U6 duplex

All RP variants are predicted to map in spatial proximity with each other, within the three-way junction delimited by stem-I and stem-II of the U4/U6 duplex and the 5' stem-loop of U4 (Figs. 1a and 2a). In particular, they are located in a different region compared with those underlying NDD (Fig. 1a). In silico two-dimensional modeling of RNA secondary structure predicted as well that the *RNU4-2* variant n.18_19insA inserts a nucleotide between stem-II and the U4 5' stem-loop (Supplementary Fig. 5a,b), while n.56T>C disrupts the first base-pairing of the U4/U6 duplex within stem-I (Supplementary Fig. 5a,c). Both changes lead to the extension of the internal loop, an event that is predicted to impact the overall stability of the duplex. In addition, n.18_19insA slightly modifies the orientation of the 5' stem-loop relative to stem-I and stem-II (Supplementary Fig. 5a,b).

In contrast, both n.55_56insG and n.56_57insG in *RNU6* paralogs are predicted to extend the length of stem-I by three additional base pairs, reduce the size of the internal loop and drastically change the orientation of the 5' stem-loop (Supplementary Fig. 5a,d,e). Interestingly, we observed that a benign insertion at the same position, n.55_56insT, was present in gnomAD v.4.1 in all five *RNU6* paralogs with a cumulative frequency of 0.12% ($n = 181$) (Supplementary Fig. 5f). While these models provide a coherent structural rationale for the observed clustering, the precise effects of the variants on U4/U6 architecture remain to be experimentally verified.

Analysis of cryo-electron microscopy data (PDB 6QW6)²⁷ confirmed that all RP variants identified reside in a region critical for binding of the U4/U6 duplex to the splicing factors PRPF31, PRPF3 and PRPF6, all previously associated with adRP^{16,17} (Fig. 2b). Specifically,

this region first engages PRPF31 or the PRPF3/PRPF4 complex, initiating the assembly interface, and is subsequently stabilized in its native orientation upon the coordinated binding of additional tri-snRNP components, including PRPF6 and PRPF8²⁸. The mutated and neighboring U4 and U6 nucleotides detected in RP cases directly participate in the binding of PRPF31 and PRPF3 (Fig. 2c,d), via hydrogen bonds with eight and three residues of these proteins, respectively. Notably, by querying the ClinVar database²⁹, we detected a missense variant affecting one of these residues, p.(Arg449Gly) of PRPF3, identified in a three-generation family with seven affected individuals having clinical features similar to those observed in most cases from our study³⁰.

Expression of *RNU4* and *RNU6* genes

Since the human genome contains several *RNU4* and *RNU6* pseudogenes³¹, we investigated whether any of these might be incorrectly annotated and could instead produce functional RNA, potentially contributing to the disease. In addition, we sought to understand why the various U4 and U6 paralogs appear to be differentially mutated, with *RNU4-1* and *RNU6-7* displaying none of the recurrent pathogenic variants. We used RNA sequencing (RNA-seq) data from human neurosensory retina (NSR), retinal pigment epithelium (RPE) and choroid that were enriched for small RNAs, applying stringent and paralog-aware bioinformatics analyses designed to mitigate the complexities associated with reads aligning against multiple paralogs and/or pseudogenes (Methods). *RNU4-2* was more highly expressed than *RNU4-1* in all tissues (average ratio: 1.63; Fig. 3a). Conversely, individual expression of *RNU6* genes and pseudogenes in the retina could not be reliably quantified by RNA-seq, since their sequences are identical, except for the last

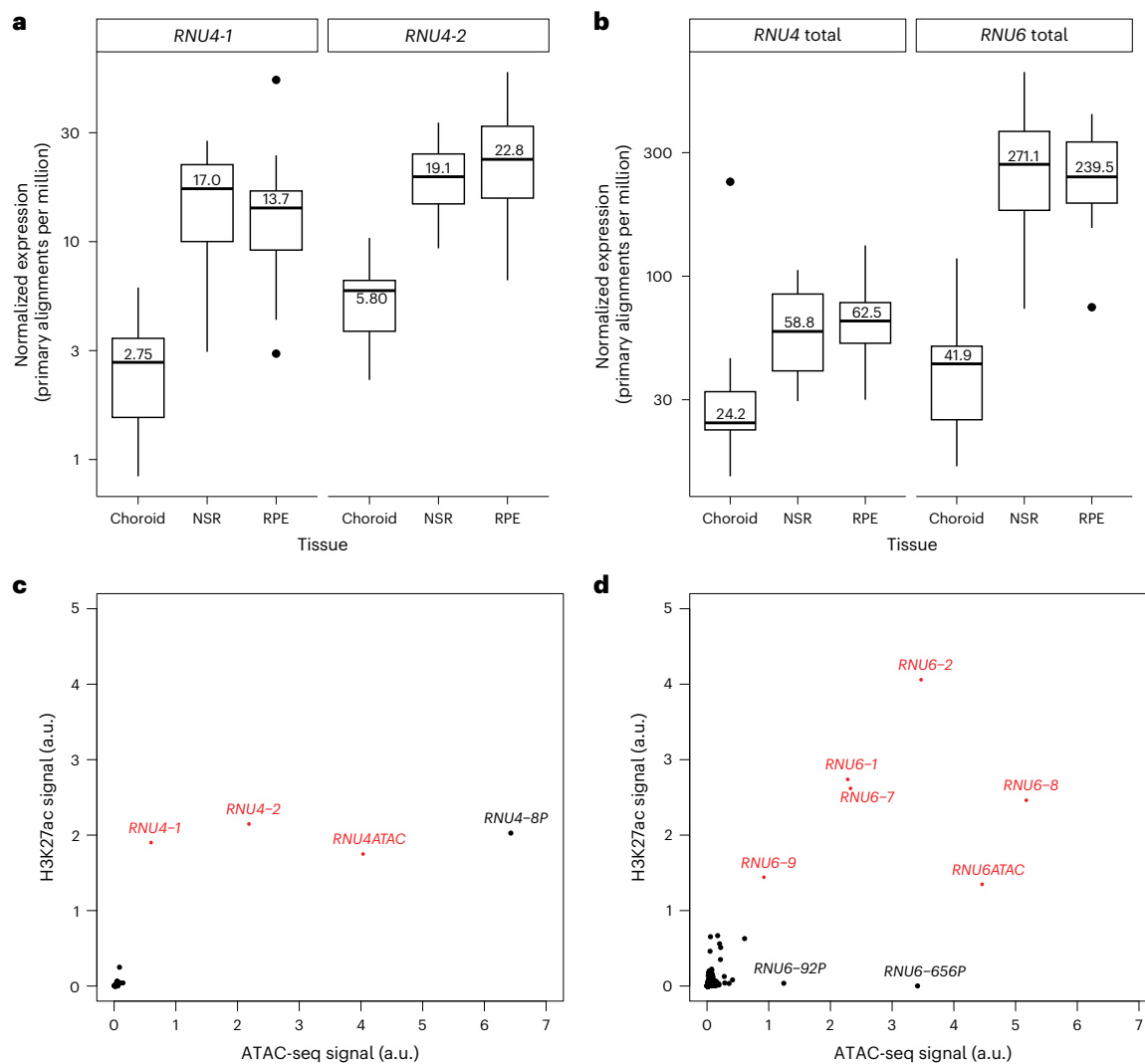


Fig. 3 | Expression and markers of transcriptional activity of *RNU4* and *RNU6* genes. **a**, Expression of *RNU4-1* and *RNU4-2* from RNA-seq of human donor choroid ($n = 13$), NSR ($n = 4$) and RPE ($n = 16$). For these boxplots, the tick line within boxes indicates the median (also expressed numerically), boxes represent the first and the third quartiles and whiskers indicate the largest observation smaller than or equal to the first quartile $- 1.5 \times \text{IQR}$ and the smallest observation greater than or equal to the third quartile $+ 1.5 \times \text{IQR}$. **b**, Same as in **a** for *RNU4* genes (*RNU4-1*, *RNU4-2* and pseudogenes) and for all *RNU6* genes (*RNU6-1*, *RNU6-2*, *RNU6-7*, *RNU6-8*, *RNU6-9* and pseudogenes). **c**, ATAC-seq and H3K27ac signals for *RNU4-1*, *RNU4-2*, *RNU4ATAC* (red) and 105 *RNU4* pseudogenes (black). **d**, Same as in **c** for five *RNU6* genes and *RNU6ATAC* (red), as well as for 1,312 *RNU6* pseudogenes (black). IQR, interquartile range.

nucleotide. Therefore, we compared the total expression of *RNU4* and *RNU6*, regardless of their respective paralogs and pseudogenes. *RNU6* expression was on average $3.39\times$ higher across the three tissues, compared with *RNU4* (Fig. 3b). Of note, NSR and RPE had higher expression of *RNU4* ($2.51\times$) and *RNU6* ($6.09\times$) with respect to the choroid, an ocular tissue not directly involved in vision, used as a control (Fig. 3b). This observation is in agreement with previous data showing that snRNA expression in the retina is approximately sixfold higher compared with muscle, testis, heart and brain³², indicating a high demand for snRNAs in these two retinal layers.

In addition, we analyzed the Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) and H3K27ac chromatin immunoprecipitation followed by sequencing (ChIP-seq) data from retinal tissues³³ in genomic regions spanning all *RNU4* and *RNU6* sequences. ATAC-seq assesses chromatin accessibility across the genome, while H3K27ac ChIP-seq reveals the presence of active enhancers. These data, combined, indicate potential active transcription at promoter regions. Our analysis showed clear transcription marks in all paralogous *RNU4* and *RNU6* genes in the retina (Fig. 3c,d).

greater than or equal to the third quartile $+ 1.5 \times \text{IQR}$. **b**, Same as in **a** for *RNU4* genes (*RNU4-1*, *RNU4-2* and pseudogenes) and for all *RNU6* genes (*RNU6-1*, *RNU6-2*, *RNU6-7*, *RNU6-8*, *RNU6-9* and pseudogenes). **c**, ATAC-seq and H3K27ac signals for *RNU4-1*, *RNU4-2*, *RNU4ATAC* (red) and 105 *RNU4* pseudogenes (black). **d**, Same as in **c** for five *RNU6* genes and *RNU6ATAC* (red), as well as for 1,312 *RNU6* pseudogenes (black). IQR, interquartile range.

Conversely, these signatures were absent from the 105 U4 pseudogenes and the 1,312 U6 pseudogenes, except for *RNU4-8P*, which displayed strong signals, but probably by virtue of its close proximity to the *ACTR1B* promoter. Of note, *RNU6-92P* and *RNU6-656P* had high ATAC-seq signals but very low H3K27ac signals at their respective promoters (Fig. 3d).

We performed the same analysis for other snRNA genes present in the human genome, which revealed a similar trend: all *RNU* genes, with the exception of *RNUSF-1*, had marks of active transcription and only a few among the thousands of *RNU* pseudogenes displayed signals compatible with potential expression, therefore representing plausible candidate genes for retinal disease (Supplementary Fig. 6). In addition, *RNU2-2* was recently implicated in NDD, yet without evidence of ocular involvement¹³. Interestingly, the same type of analysis, based on conservation and expression data from GTEx, was recently performed by others, showing similar results³⁴.

For *RNU6-7*, both ATAC-seq and H3K27ac signals were within the same range as those observed for other *RNU6* genes (Fig. 3d), and, therefore, the absence of pathogenic variants could not be explained

by a potential differential expression. We thus analyzed the genetic landscape of variations in healthy individuals in all five U6 paralogs and observed that *RNU6-7* displayed a lower number of variants, compared with the others (Supplementary Fig. 7). We also identified the recurrent variant n.55_56insG in *RNU6-7* in six control individuals of African or African American ancestry in gnomAD v.4.1 (allele frequency (AF) = 0.014%) and in 14 individuals of African origin in the All of Us database (AF = 0.013%). These seemingly contradictory observations merit further investigations in future studies.

Transcriptome analysis in patients

We performed transcriptome analysis following the collection of RNA from circulating leukocytes in nine affected individuals carrying variants in *RNU4-2*, *RNU6-1* and *RNU6-9* (three individuals per gene), as well as from 14 healthy controls (Supplementary Table 6). To avoid systematic errors linked to the use of different collection kits³⁵ across our cohort (Methods), we performed independent case–control tests for samples collected with PAXgene kits (three *RNU4-2* cases, six controls) or Tempus kits (six *RNU6* cases, eight controls). We identified 27 and eight differentially expressed genes in the two datasets, respectively, with no gene overlap (fold-change > 2 and false discovery rate (FDR) P value < 0.05; Supplementary Table 7), indicating no major differences in global gene expression in leukocytes in cases versus controls. We then further explored the data by investigating potential bias in pre-mRNA splicing, as performed in ref. 11. This analysis led to the identification of 107 upregulated and 67 downregulated 5' splice sites in the PAXgene set, and 37 upregulated and 13 downregulated 5' splice sites in the Tempus set, with only two sites in common between the datasets (in the genes *CLEC2D* and *SNHG29*; Supplementary Table 8). At these two sites the expression differences, although statistically significant, were below 10%.

We then examined the DNA sequences of the 224 differentially expressed splice sites, focusing on the occurrence of the 'AG' dinucleotide at positions –2/–1, which was previously reported to be enriched in sites with increased splicing in *RNU4-2* in patients with NDD¹¹. No differences in 'AG' frequency were observed between sites with increased versus decreased usage in either the PAXgene or Tempus samples (two-tailed Fisher's test, P = 0.63 and P = 0.52, respectively). We also compared nucleotide frequencies at each position between upregulated and downregulated splice sites, separately for the two groups, and found no significant differences (two-tailed Fisher's test with FDR correction, P < 0.05). A similar analysis of overlapping dinucleotides across the region flanking the 5' splice site (for example, positions –4/–3, –3/–2, up to +7/+8) revealed no significant differences.

Functional effects of RP variants

Since *PRPF* variants associated with RP affect primarily spliceosomal assembly³², we investigated whether the same phenomenon could be driven by the variants detected in this work. Specifically, we immunopurified ectopically expressed U4 and U6 snRNAs containing the RP variants and analyzed their association with specific markers for the U6 snRNP (SART3), the U4/U6 di-snRNP (SART3 and *PRPF31*), the U4/U6.U5 tri-snRNP (*PRPF31* and *SNRNP200*) and the U5 snRNP (*SNRNP200*). The combined results showed an increased association of snRNA constructs with RP variants with SART3 and partially with *PRPF31*, while the interaction with *SNRNP200* was unchanged or reduced (Fig. 4). For comparison, we included in our assays the U4 n.64_65insT variant, which causes NDD, and observed no significant alteration in the association with any of the proteins tested, compared with wild type (Fig. 4a). Additionally, no significant differences were detected between NDD and RP variants, pointing to the need for targeted functional studies to delineate their respective impacts on spliceosome dynamics. Similarly, U6 RNA bearing the n.55_56insT and n.57T>G variants, observed in healthy control individuals, presumably did not affect spliceosome formation, since the low amount of protein associated with them implies that they

entered the spliceosome assembly process only minimally (Fig. 4b). Taken together, the results indicate that RP pathogenic variants have potentially a specific dominant effect on snRNP biogenesis and delay the assembly process at the di-snRNP stage.

Discussion

The numerous genes associated with RP and allied diseases belong to a wide range of functional classes, from retina-specific biochemical pathways to ubiquitous cellular processes⁵. Yet, how these defects ultimately lead to retinal degeneration often remains unclear. The link between pathogenic variants in splicing factors of the tri-snRNP complex (RP-PRPFs), essential for survival in all eukaryotes, and RP, a phenotype limited to the eye, represents perhaps the most intriguing of these biological enigmas.

In this study, we identified recurrent heterozygous variants in *RNU4-2*, encoding U4 RNA, and in multiple paralogs of the U6 RNA as a cause of RP. Interestingly, these snRNAs are also an integral part of the di- and tri-snRNP and directly interact with some RP-PRPF proteins. In addition, similar to RP-PRPFs, they are also associated with the same specific phenotype: de novo or inherited adRP, with reduced penetrance for *RNU4-2* variants. Importantly, the clinical presentation of patients with *RNU4-2* and *RNU6* variants overlaps with that of other spliceosome-related forms of adRP, particularly showing an earlier onset—contrasting with the generally milder prognosis observed in most other adRP types^{36,37}—and a relatively high co-occurrence of features such as cataracts and cystoid macular edema, found in cases with *PRPF31*^{38,39}, *PRPF8*⁴⁰ and *SNRNP200*⁴¹ variants. Prevalence estimations indicate that these snRNA pathogenic changes may account for an elevated number of undiagnosed cases, and it is therefore surprising that the *RNU4* and *RNU6* genes have escaped disease association until now. A partial explanation for this phenomenon is that mainstream sequencing approaches are biased towards DNA-capturing procedures that do not include snRNA genes. Furthermore, although genome sequencing is increasingly being adopted in routine diagnostics, variants in snRNA genes may have remained undetected because they affect noncoding transcripts, which are more challenging to interpret and are often overlooked or deprioritized by standard analytical pipelines.

An intriguing feature of pathogenic changes in *RNU4-2* is their pleiotropy with respect to NDD (ReNU syndrome) and RP. Chen et al.⁹ described that more than half of the patients with ReNU also display some visual abnormalities, although only three were documented as having retinal phenotypes (one had an abnormal electroretinogram response, one had Leber congenital amaurosis and one presented with macular dysfunction). However, most cases were too young to display the typical symptoms of RP, which usually manifest during adolescence or early adulthood⁴². Although the exact mechanism for this phenotypic selectivity is unknown, *RNU4-2* variants represent a clear and new allelic series involving noncoding RNA genes. A recent preprint highlighted a strong effect of ReNU variants in an *RNU4-2* saturation genome editing experiment, while the RP variant n.56T>C in the same gene did not show any statistically significant effect⁴³. In addition, the region of *RNU4-2* containing RP variants showed function scores within the neutral range of the saturation genome editing assay⁴³, suggesting a potentially milder pathogenic effect compared with ReNU changes. It is therefore plausible that the RP variants identified in *RNU4-2* and *RNU6* paralogs could lead to photoreceptor death and subsequent visual loss, while having no influence on the development of the brain. Additionally, ReNU variants are located in the stem-III and the T-loop of the U4/U6 duplex and interfere with the proper recognition of intronic 5' splice signals, likely because these regions are involved in pairing pre-mRNA with U6 (ref. 9). In contrast, RP variants cluster in spatial proximity to the three-way junction, in regions not directly engaged in interactions with pre-mRNA but that are involved in the binding of various proteins, including RP-associated splicing factors.

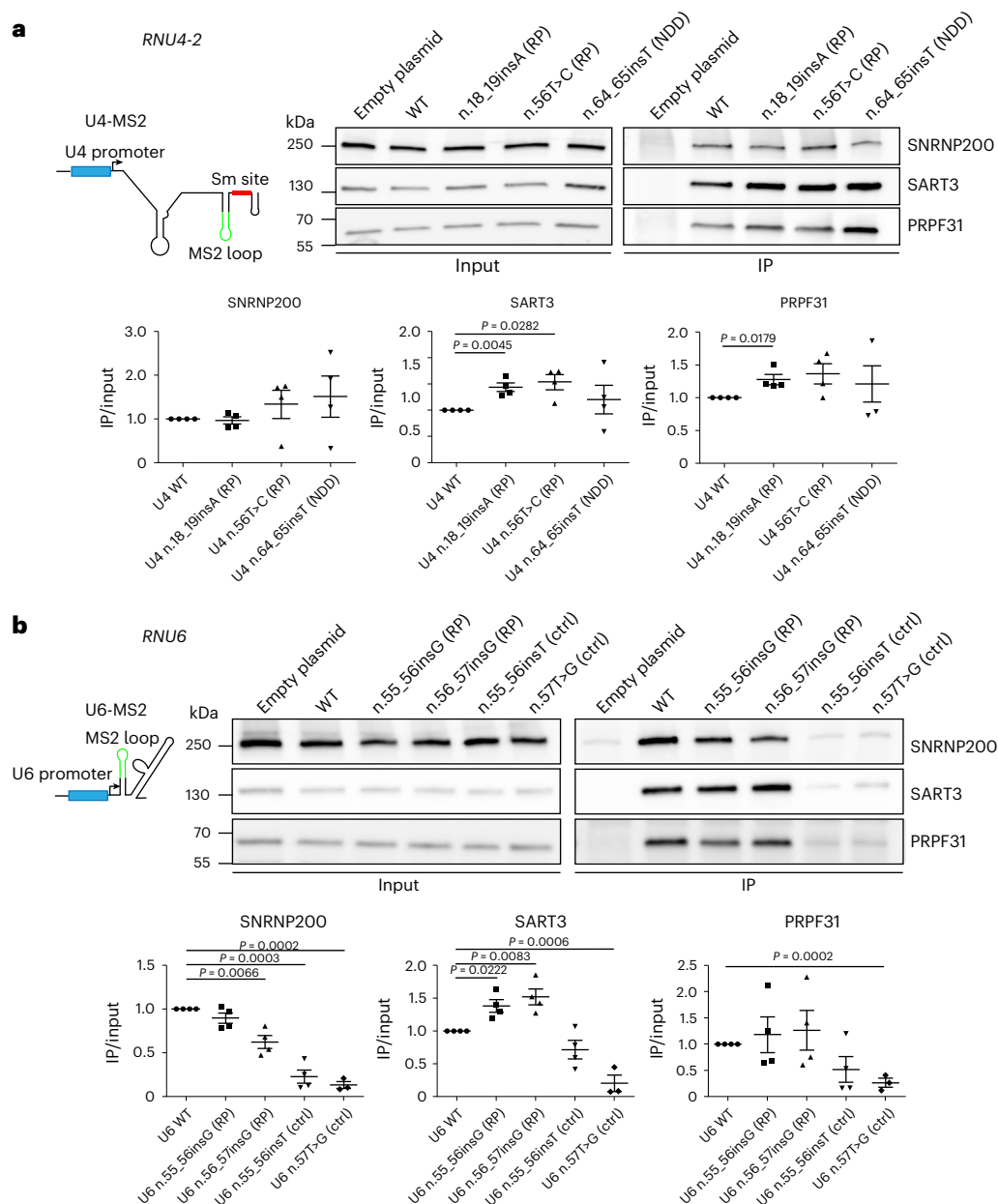


Fig. 4 | Effects of RP variants in *RNU4-2* and *RNU6* on snRNP maturation.

a,b, Immunoprecipitation of U4-MS2 (WT and variants) (**a**) and U6-MS2 (WT and variants) (**b**). snRNPs were immunoprecipitated via MS2-YFP by anti-GFP antibodies and co-precipitated proteins were detected by western blotting. The position of the MS2 loop (green) in snRNAs is indicated. Four independent experiments were quantified. Immunoprecipitated proteins are normalized to

input and U4 or U6 WT controls. Middle bars indicate average values and error bars the s.e.m. Statistical significance was analyzed by the two-tailed unpaired *t*-test and the *P* values were adjusted using the Benjamini–Hochberg FDR method to control for false discoveries. *P* values ≤ 0.05 are indicated. Full-length blots and antibody validation are provided as Source Data. Ctrl, control; IP, immunoprecipitation; WT, wild type.

Consistent with this evidence, we did not observe any major splicing anomalies in transcripts from patients with RP, with the only two significant events showing differences below 10% in expression. Conversely, our biochemical assays support a role for RP-associated variants in altering spliceosomal assembly. As the magnitude of the observed changes was in all instances rather moderate (less than 1.5-fold with respect to controls), we interpret these data as indicating that snRNA variants associated with RP are unlikely to prevent the assembly of spliceosomal complexes. Rather, they may cause a subtle alteration in snRNP dynamics, possibly affecting the efficiency of their biogenesis or recycling steps. In particular, the increased association with SART3 and, to a lesser extent, with PRPF31, together with unchanged or slightly reduced interaction with SNRNP200, may

indicate a modest delay in the transition from the di-snRNP to the tri-snRNP form. Moreover, the pathogenic variants identified in this study lie within regions of the U4/U6 duplex that directly contact the PRPF3 and PRPF31 proteins, two splicing factors linked to adRP whose mutations also delay spliceosomal complex assembly³².

In terms of specific molecular effect, our functional data show that snRNAs bearing RP variants display enhanced interaction with di-snRNP protein markers, suggesting that pathogenesis could result from a gain-of-function or dominant-negative mechanism, rather than from haploinsufficiency. This hypothesis is strengthened by the evidence that molecularly similar but benign variants, commonly observed in the general population, seem not to bind efficiently to di-snRNP markers and potentially not to be incorporated into the spliceosome,

supporting the idea that spliceosomal functions could be haploinsufficient with respect to heterozygous and snRNA-depleting variants.

Although DNA changes associating *RNU4-2* to ReNU syndrome have been primarily reported as de novo events^{9,10}, in our study most families with RP (61%) bore *RNU4-2* and *RNU6* changes as inherited variants. In part, this difference can be explained by the reduced reproductive fitness associated with NDD versus RP. Unlike ReNU syndrome, symptomatic onset (night-blindness and peripheral vision loss) in nonsyndromic adRP begins later in life, with severe central vision loss usually occurring after the onset of reproductive age. Another difference involves the inheritance of dominant variants, which in ReNU seems to be almost exclusively of maternal origin⁹. We did not observe the same trend for RP, with variants being inherited from either of the parents, possibly indicating the absence of any sex-specific negative selection during gametogenesis or at the embryonic stage.

The human genome contains two *RNU4* paralogs and five *RNU6* paralogs. This indicates that, assuming equal expression within paralogs, the presence of only ~25% of mutant U4 (heterozygous genotype, over two copies) or ~10% of mutant U6 (heterozygous genotype, over five copies) is sufficient to lead to a disease phenotype, again in support of a gain-of-function or dominant-negative molecular mechanism. This could be a crucial consideration for the development of potential gene-based therapies, as gene-augmentation strategies may be sub-optimal compared with gene correction or antisense oligonucleotide approaches. Our data also highlight the existence of mutational hotspots outside the coding regions of the human genome, emphasizing the need for further research into these parts of our genetic material, and show that the clustering of de novo pathogenic variants is not restricted to severe diseases with childhood onset⁴⁴, but may extend to milder pathologies, such as RP.

In conclusion, we identified four recurrent pathogenic variants in *RNU4-2* and in four of the five paralogs of the U6 snRNA as a frequent cause of de novo or inherited adRP. The immediate impact of these findings involves improved diagnosis and genetic counseling for patients with hereditary visual loss, especially for isolated cases who could potentially bear heterozygous de novo events. More fundamentally, this work substantially broadens our understanding of the genetic landscape of human disease, paving the way for the development of new molecular therapeutic approaches.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-025-02451-4>.

References

- Verbakel, S. K. et al. Non-syndromic retinitis pigmentosa. *Prog. Retin. Eye Res.* **66**, 157–186 (2018).
- Peter, V. G. et al. The first genetic landscape of inherited retinal dystrophies in Portuguese patients identifies recurrent homozygous mutations as a frequent cause of pathogenesis. *PNAS Nexus* **2**, pgad043 (2023).
- Perea-Romero, I. et al. Genetic landscape of 6089 inherited retinal dystrophies affected cases in Spain and their therapeutic and extended epidemiological implications. *Sci. Rep.* **11**, 1526 (2021).
- Conti, G. M. et al. Genetics of retinitis pigmentosa and other hereditary retinal disorders in western Switzerland. *Ophthalmic Res.* **67**, 172–182 (2024).
- Rivolta, C. et al. RetiGene, a comprehensive gene atlas for inherited retinal diseases. *Am. J. Hum. Genet.* **112**, 2253–2265 (2025).
- Hanany, M., Rivolta, C. & Sharon, D. Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases. *Proc. Natl Acad. Sci. USA* **117**, 2710–2716 (2020).
- Karczewski, K. J. et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
- Rogalska, M. E. et al. Transcriptome-wide splicing network reveals specialized regulatory functions of the core spliceosome. *Science* **386**, 551–560 (2024).
- Chen, Y. et al. De novo variants in the *RNU4-2* snRNA cause a frequent neurodevelopmental syndrome. *Nature* **632**, 832–840 (2024).
- Greene, D. et al. Mutations in the U4 snRNA gene *RNU4-2* cause one of the most prevalent monogenic neurodevelopmental disorders. *Nat. Med.* **30**, 2165–2169 (2024).
- Nava, C. et al. Dominant variants in major spliceosome U4 and U5 small nuclear RNA genes cause neurodevelopmental disorders through splicing disruption. *Nat. Genet.* **57**, 1374–1388 (2025).
- Jackson, A. et al. Analysis of R-loop forming regions identifies *RNU2-2* and *RNU5B-1* as neurodevelopmental disorder genes. *Nat. Genet.* **57**, 1362–1366 (2025).
- Greene, D. et al. Mutations in the small nuclear RNA gene *RNU2-2* cause a severe neurodevelopmental disorder with prominent epilepsy. *Nat. Genet.* **57**, 1367–1373 (2025).
- McKie, A. B. et al. Mutations in the pre-mRNA splicing factor gene *PRPC8* in autosomal dominant retinitis pigmentosa (RP13). *Hum. Mol. Genet.* **10**, 1555–1562 (2001).
- Weisschuh, N. et al. Genetic architecture of inherited retinal degeneration in Germany: a large cohort study from a single diagnostic center over a 9-year period. *Hum. Mutat.* **41**, 1514–1527 (2020).
- Mozaffari-Jovin, S. et al. The Prp8 RNase H-like domain inhibits Brr2-mediated U4/U6 snRNA unwinding by blocking Brr2 loading onto the U4 snRNA. *Genes Dev.* **26**, 2422–2434 (2012).
- Liu, S. et al. Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science* **316**, 115–120 (2007).
- Chen, S. et al. A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).
- Jaganathan, K. et al. Predicting splicing from primary sequence with deep learning. *Cell* **176**, 535–548 (2019).
- Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
- Davydov, E. V. et al. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* **6**, e1001025 (2010).
- Denny, J. C. et al. The “All of Us” Research Program. *N. Engl. J. Med.* **381**, 668–676 (2019).
- Panneman, D. M. et al. Cost-effective sequence analysis of 113 genes in 1,192 probands with retinitis pigmentosa and Leber congenital amaurosis. *Front. Cell Dev. Biol.* **11**, 1112270 (2023).
- Caulfield, M. et al. National Genomic Research Library. *Figshare* <https://doi.org/10.6084/m9.figshare.4530893.v7> (2024).
- Daiger, S. P., Bowne, S. J. & Sullivan, L. S. Genes and mutations causing autosomal dominant retinitis pigmentosa. *Cold Spring Harb. Perspect. Med.* **5**, a017129 (2014).
- Sullivan, L. S. et al. Prevalence of disease-causing mutations in families with autosomal dominant retinitis pigmentosa: a screen of known genes in 200 families. *Invest. Ophthalmol. Vis. Sci.* **47**, 3052–3064 (2006).
- Charenton, C., Wilkinson, M. E. & Nagai, K. Mechanism of 5' splice site transfer for human spliceosome activation. *Science* **364**, 362–367 (2019).
- Hardin, J. W., Warnasooriya, C., Kondo, Y., Nagai, K. & Rueda, D. Assembly and dynamics of the U4/U6 di-snRNP by single-molecule FRET. *Nucleic Acids Res.* **43**, 10963–10974 (2015).

29. Landrum, M. J. et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980–D985 (2014).
 30. Zhong, Z. et al. Two novel mutations in *PRPF3* causing autosomal dominant retinitis pigmentosa. *Sci. Rep.* **6**, 37840 (2016).
 31. Denison, R. A., Van Arsdell, S. W., Bernstein, L. B. & Weiner, A. M. Abundant pseudogenes for small nuclear RNAs are dispersed in the human genome. *Proc. Natl Acad. Sci. USA* **78**, 810–814 (1981).
 32. Tanackovic, G. et al. *PRPF* mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum. Mol. Genet.* **20**, 2116–2130 (2011).
 33. D'Haene, E. et al. Comparative 3D genome analysis between neural retina and retinal pigment epithelium reveals differential cis-regulatory interactions at retinal disease loci. *Genome Biol.* **25**, 123 (2024).
 34. Prasetyo, N. K. & Gardner, P. P. Assessing the robustness of human ncRNA notation at HGNC. Preprint at *bioRxiv* <https://doi.org/10.1101/2024.12.08.627405> (2024).
 35. Skogholt, A. H. et al. Gene expression differences between PAXgene and Tempus blood RNA tubes are highly reproducible between independent samples and biobanks. *BMC Res. Notes* **10**, 136 (2017).
 36. Hamel, C. Retinitis pigmentosa. *Orphanet J. Rare Dis.* **1**, 40 (2006).
 37. Grover, S. et al. Visual acuity impairment in patients with retinitis pigmentosa at age 45 years or older. *Ophthalmology* **106**, 1780–1785 (1999).
 38. Bodenbender, J. P. et al. Clinical and genetic findings in a cohort of patients with *PRPF31*-associated retinal dystrophy. *Am. J. Ophthalmol.* **267**, 213–229 (2024).
 39. Waseem, N. H. et al. Mutations in the gene coding for the pre-mRNA splicing factor, *PRPF31*, in patients with autosomal dominant retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.* **48**, 1330–1334 (2007).
 40. Maubaret, C. G. et al. Autosomal dominant retinitis pigmentosa with intrafamilial variability and incomplete penetrance in two families carrying mutations in *PRPF8*. *Invest. Ophthalmol. Vis. Sci.* **52**, 9304–9309 (2011).
 41. Yusuf, I. H. et al. Clinical characterization of retinitis pigmentosa associated with variants in *SNRNP200*. *JAMA Ophthalmol.* **137**, 1295–1300 (2019).
 42. Berson, E. L. Retinitis pigmentosa. The Friedenwald Lecture. *Invest. Ophthalmol. Vis. Sci.* **34**, 1659–1676 (1993).
 43. De Jonghe, J. et al. Saturation genome editing of *RNU4-2* reveals distinct dominant and recessive neurodevelopmental disorders. Preprint at *medRxiv* <https://doi.org/10.1101/2025.04.08.25325442> (2025).
 44. Veltman, J. A. & Brunner, H. G. De novo mutations in human genetic disease. *Nat. Rev. Genet.* **13**, 565–575 (2012).
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.
- © The Author(s) 2026

Mathieu Quinodoz^{1,2,3,135}, Kim Rodenburg^{4,135}, Zuzana Cvackova⁵, Karolina Kaminska^{1,2}, Suzanne E. de Bruijn^{4,6}, Ana Belén Iglesias-Romero^{1,2}, Erica G. M. Boonen^{4,7}, Mukhtar Ullah¹, Nick Zomer⁴, Marc Folcher^{1,2}, Jacques Bijon^{8,9}, Lara K. Holtes⁴, Stephen H. Tsang^{10,11}, Zelia Corradi⁴, K. Bailey Freund^{8,12}, Stefanida Shliaga⁴, Daan M. Panneman⁴, Rebekkah J. Hitti-Malin⁴, Manir Ali¹³, Ala'a AlTalibishi¹⁴, Sten Andréasson¹⁵, Georg Ansari², Gavin Arno^{16,17,18}, Galuh D. N. Astuti⁴, Carmen Ayuso^{19,20}, Radha Ayyagari²¹, Sandro Banfi^{22,23}, Eyal Banin²⁴, Tahsin Stefan Barakat²⁵, Mirella T. S. Barboni²⁶, Miriam Bauwens^{27,28}, Tamar Ben-Yosef²⁹, Virginie Bernard³⁰, David G. Birch³¹, Pooja Biswas²¹, Fiona Blanco-Kelly^{19,20}, Beatrice Bocquet^{32,33}, Camiel J. F. Boon^{34,35}, Kari Branham³⁶, Dominique Bremond-Gignac^{37,38}, Alexis Ceecee Britten-Jones³⁹, Kinga M. Bujakowska⁴⁰, Cyril Burin des Roziers^{38,41,42}, Elizabeth L. Cadena⁴³, Giacomo Calzetti^{44,45}, Francesca Cancellieri^{1,2}, Luca Cattaneo⁴⁶, Naomi Chadderton⁴⁷, Peter Charbel Issa^{48,49}, Luisa Coutinho-Santos⁵⁰, Stephen P. Daiger⁴³, Elfride De Baere^{27,28}, Marieke De Bruyne^{27,28}, Berta de la Cerdá⁵¹, John N. De Roach⁵², Julie De Zaeytjij^{53,54}, Ronny Derks⁴, Claire-Marie Dhaenens⁵⁵, Lubica Dudakova⁵⁶, Jacque L. Duncan⁵⁷, G. Jane Farrar⁴⁷, Nicolas Feltgen², Beau J. Fenner^{58,59,60}, Lidia Fernández-Caballero^{19,20}, Juliana M. Ferraz Sallum⁶¹, Simone Gana⁴⁶, Alejandro Garanto^{4,62}, Jessica C. Gardner¹⁷, Christian Gilissen⁴, Roser González-Duarte⁶³, Kensuke Goto⁶⁴, Sam Griffiths-Jones⁶⁵, Tobias B. Haack^{66,67,68}, Lonkeke Haer-Wigman⁴, Alison J. Hardcastle¹⁷, Takaaki Hayashi⁶⁹, Elise Héon⁷⁰, Lies H. Hoefsloot²⁵, Alexander Hoischen^{4,71}, Josephine P. Holtan⁷², Carel B. Hoyng⁷³, Manuel Benjamin B. Ibanez IV^{74,75}, Chris F. Inglehearn¹³, Takeshi Iwata⁷⁶, Brynjar O. Jensson⁷⁷, Kaylie Jones³¹, Vasiliki Kalatzis^{32,33}, Smaragda Kamakari⁷⁸, Marianthi Karali^{22,79}, Ulrich Kellner⁸⁰, Caroline C. W. Klaver^{73,81}, Krisztina Knézy²⁶, Robert K. Koenekekoop⁸², Susanne Kohl⁸³, Taro Kominami⁶⁴, Laura Kühlewein⁸⁴, Tina M. Lamey⁵², Rina Leibu⁸⁵, Bart P. Leroy^{27,53,54}, Petra Liskova^{56,86}, Irma Lopez⁸², Víctor R. de J. López-Rodríguez⁸⁷, Quinten Mahieu^{27,28}, Omar A. Mahroo^{17,18,88,89}, Gaël Manes³², Luke Mansard^{30,32,90}, M. Pilar Martín-Gutiérrez⁹¹, Nelson Martins^{27,28}, Laura Mauring^{92,93,94}, Martin McKibbin^{13,95}, Terri L. McLaren⁵², Isabelle Meunier^{32,33}, Michel Michaelides^{17,18}, José M. Millán⁹⁶, Kei Mizobuchi⁶⁹, Rajarshi Mukherjee⁹⁵,

Zoltán Zsolt Nagy²⁶, Kornelia Neveling⁴, Monika Otdak⁹⁷, Michiel Oorsprong⁴, Yang Pan⁷⁶, Anastasia Papachristou⁹⁸, Antonio Percesepe⁹⁹, Maximilian Pfau², Eric A. Pierce⁴⁰, Emily Place⁴⁰, Raj Ramesar¹⁰⁰, Francis Ramond^{30,101}, Florence Andrée Rasquin¹⁰², Gillian I. Rice⁶⁵, Lisa Roberts¹⁰⁰, María Rodríguez-Hidalgo^{103,104}, Javier Ruiz-Ederra^{103,105}, Ataf H. Sabir^{106,107}, Ai Fujita Sajiki⁶⁴, Ana Isabel Sánchez-Barbero^{19,20}, Asodu Sandeep Sarma²⁴, Riccardo Sangermano⁴⁰, Cristina M. Santos^{50,108}, Margherita Scarpato²², Hendrik P. N. Scholl^{109,110,111}, Dror Sharon²⁴, Sabrina G. Signorini¹¹², Francesca Simonelli⁷⁹, Ana Berta Sousa^{113,114}, Maria Stefanidou¹¹⁵, Kari Stefansson^{77,116}, Katarina Stingl⁸⁴, Akiko Suga⁷⁶, Patrick Sulem⁷⁷, Lori S. Sullivan⁴³, Viktória Szabó²⁶, Jacek P. Szaflik^{117,118}, Gita Taurina¹¹⁹, Alberta A. H. J. Thiadens⁸¹, Carmel Toomes¹³, Viet H. Tran^{120,121}, Miltiadis K. Tsilimbaris⁹⁸, Pavlina Tsoka⁹⁸, Veronika Vaclavik¹²⁰, Marie Vajter^{56,86}, Sandra Valeina¹²², Enza Maria Valente^{46,123}, Casey Valentine¹⁰⁰, Rebeca Valero⁶³, Sophie Valleix^{38,41,42}, Joseph van Aerschot¹²⁴, L. Ingeborgh van den Born⁷, Mattias Van Heetvelde^{27,28}, Virginie J. M. Verhoeven^{25,81}, Andrea L. Vincent¹²⁵, Andrew R. Webster^{17,18}, Laura Whelan^{47,126,127}, Bernd Wissinger⁸³, Georgia G. Yioti¹¹⁵, Kazutoshi Yoshitake⁷⁶, Juan C. Zenteno^{87,128}, Roberta Zeuli²², Theresia Zuleger⁶⁶, Chaim Landau¹²⁹, Allan I. Jacob¹³⁰, Siying Lin^{18,131,132}, Frans P. M. Cremers⁴, Winston Lee¹³³, Jamie M. Ellingford^{65,134}, David Stanek⁵, Susanne Roosing^{4,136} ✉ & Carlo Rivolta^{1,2,3,136}

¹Ophthalmic Genetics Group, Institute of Molecular and Clinical Ophthalmology Basel (IOB), Basel, Switzerland. ²Department of Ophthalmology, University of Basel, Basel, Switzerland. ³Department of Genetics, Genomics and Cancer Sciences, University of Leicester, Leicester, UK. ⁴Department of Human Genetics, Radboud University Medical Center, Nijmegen, the Netherlands. ⁵Laboratory of RNA Biology, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic. ⁶Department of Otorhinolaryngology, Hearing & Genes, Radboud University Medical Center, Nijmegen, the Netherlands. ⁷The Rotterdam Eye Hospital, Rotterdam Ophthalmic Institute, Rotterdam, the Netherlands. ⁸Vitreous Retina Macula Consultants of New York, New York, NY, USA. ⁹Department of Ophthalmology, Rothschild Foundation Hospital, Paris, France. ¹⁰Departments of Ophthalmology, Pathology & Cell Biology, Columbia Stem Cell Initiative, Vagelos College of Physicians and Surgeons Columbia University Irving Medical Center, New York, NY, USA. ¹¹Edward S. Harkness Eye Institute, Jonas Children's Vision Care (JCVC), Columbia University Irving Medical Center, New York-Presbyterian Hospital, New York, NY, USA. ¹²Department of Ophthalmology, NYU Grossman School of Medicine, New York, NY, USA. ¹³Division of Molecular Medicine, Leeds Institute of Medical Research, School of Medicine, University of Leeds, Leeds, UK. ¹⁴Molecular Biology Research Unit, St John Eye Hospital Group, Jerusalem, Palestine. ¹⁵Department of Ophthalmology, University Hospital of Lund, Lund, Sweden. ¹⁶JC Self Research Institute, Greenwood Genetic Center, Greenwood, SC, USA. ¹⁷UCL Institute of Ophthalmology, University College London, London, UK. ¹⁸National Institute of Health Research Biomedical Research Centre, Moorfields Eye Hospital, London, UK. ¹⁹Department of Genetics & Genomics, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain. ²⁰Center for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain. ²¹Department of Ophthalmology, Shiley Eye Institute, University of California San Diego, La Jolla, CA, USA. ²²Medical Genetics, Department of Precision Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy. ²³Genomic Medicine, Telethon Institute of Genetics and Medicine, Pozzuoli, Italy. ²⁴Hadassah Medical Center, Division of Ophthalmology, The Hebrew University of Jerusalem, Jerusalem, Israel. ²⁵Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands. ²⁶Department of Ophthalmology, Semmelweis University, Budapest, Hungary. ²⁷Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium. ²⁸Department of Biomolecular Medicine, Ghent University, Ghent, Belgium. ²⁹Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel. ³⁰GCS Auragen, Plan France Medecine Genomique, Lyon, France. ³¹Retina Foundation of the Southwest, Dallas, TX, USA. ³²Institute for Neurosciences of Montpellier (INM), Montpellier University, INSERM, Montpellier, France. ³³National Reference Centre for Inherited Sensory Diseases, University of Montpellier, Montpellier University Hospital, Montpellier, France. ³⁴Department of Ophthalmology, Leiden University Medical Center (LUMC), Leiden, the Netherlands. ³⁵Department of Ophthalmology, Amsterdam University Medical Center, Amsterdam, the Netherlands. ³⁶Department of Ophthalmology and Human Genetics, University of Michigan, Ann Arbor, MI, USA. ³⁷Department of Ophthalmology, University Hospital Necker Enfants Malades, APHP, Paris Cité University, Paris, France. ³⁸UMRS1138, Centre de Recherche des Cordeliers, Université Paris Cité, INSERM, Paris, France. ³⁹Department of Optometry and Vision Sciences, The University of Melbourne, Melbourne, Victoria, Australia. ⁴⁰Ocular Genomics Institute, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA. ⁴¹Service de Médecine Génomique des Maladies de Système et d'Organe, Hôpital Cochin, AP-HP Centre Université Paris Cité, Paris, France. ⁴²Laboratoire de Biologie Médicale SeqOIA, Site Broussais, Paris, France. ⁴³Department of Epidemiology and Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA. ⁴⁴Vista Vision Eye Clinic, Brescia, Italy. ⁴⁵Clinical Translation Group, Institute of Molecular and Clinical Ophthalmology Basel (IOB), Basel, Switzerland. ⁴⁶Neurogenetics Research Center, IRCCS Mondino Foundation, Pavia, Italy. ⁴⁷Institute of Genetics, School of Genetics and Microbiology, Trinity College Dublin, The University of Dublin, Dublin, Ireland. ⁴⁸Department of Ophthalmology, TUM University Hospital, School of Medicine and Health, Technical University of Munich, Munich, Germany. ⁴⁹Oxford Eye Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK. ⁵⁰Serviço de Oftalmologia, Instituto de Oftalmologia Dr. Gama Pinto, Lisboa, Portugal. ⁵¹Department of Cell Therapy and Regenerative Medicine, Andalusian Molecular Biology and Regenerative Medicine Centre, CABIMER, Seville, Spain. ⁵²Australian Inherited Retinal Disease Registry and DNA Bank, Department of Medical Technology and Physics, Sir Charles Gairdner Hospital, Perth, Western Australia, Australia. ⁵³Department of Head and Skin, Ghent University, Ghent, Belgium. ⁵⁴Department of Ophthalmology, Ghent University Hospital, Ghent, Belgium. ⁵⁵Université de Lille, INSERM U1172 - LiNCog - Lille Neuroscience & Cognition, Lille, France. ⁵⁶Department of Paediatrics and Inherited Metabolic Disorders, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic. ⁵⁷Department of Ophthalmology, University of California, San Francisco, Wayne and Gladys Valley Center for Vision, San Francisco, CA, USA. ⁵⁸Department of Medical Retina, Singapore National Eye Centre, Singapore, Singapore. ⁵⁹Ophthalmology and Visual Sciences Academic Clinical Program, Duke-NUS Graduate Medical School, Singapore, Singapore. ⁶⁰Singapore Eye Research Institute, Singapore, Singapore. ⁶¹Department of Ophthalmology, Federal University of São Paulo, UNIFESP, São Paulo, Brazil. ⁶²Department of Pediatrics, Amalia Children's Hospital, Radboud University Medical Center, Nijmegen, the Netherlands. ⁶³Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain. ⁶⁴Department of Ophthalmology, Nagoya University Graduate School of Medicine, Nagoya, Japan. ⁶⁵School of Biological Sciences, Division of Evolution, Infection, and Genomics, The University of Manchester, Manchester, UK. ⁶⁶Institute for Medical Genetics and

Applied Genomics, University of Tübingen, Institute for Ophthalmic Research, University Hospital Tübingen, Tübingen, Germany. ⁶⁷Center for Rare Disease, University of Tübingen, Tübingen, Germany. ⁶⁸Genomics for Health in Africa (GHA), Africa-Europe Cluster of Research Excellence (CoRE), Ghana, South Africa. ⁶⁹Department of Ophthalmology, The Jikei University School of Medicine, Minato-ku, Japan. ⁷⁰Department of Ophthalmology and Vision Sciences, Ocular Genetics Program, The Hospital for Sick Children, Toronto, Ontario, Canada. ⁷¹Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands. ⁷²Department of Ophthalmology, Oslo University Hospital, Oslo, Norway. ⁷³Department of Ophthalmology, Radboud University Medical Center, Nijmegen, the Netherlands. ⁷⁴Department of Ophthalmology, Section of Pediatric Ophthalmology, Strabismus, and Ocular Genetics, DOH Eye Center, East Avenue Medical Center, Quezon City, Philippines. ⁷⁵Section of Pediatric Ophthalmology, Strabismus, and Ocular Genetics, Makati Medical Center, Makati City, Philippines. ⁷⁶Division of Molecular and Cellular Biology, National Institute of Sensory Organs, NHO Tokyo Medical Center, Tokyo, Japan. ⁷⁷deCODE genetics, Amgen Inc., Reykjavik, Iceland. ⁷⁸Department of Inherited Retinal Dystrophies, Ophthalmic Genetics Group, OMMA, Ophthalmological Institute of Athens, Athens, Greece. ⁷⁹Eye Clinic, Multidisciplinary Department of Medical, Surgical and Dental Sciences, University of Campania “Luigi Vanvitelli”, Naples, Italy. ⁸⁰Center for Rare Retinal Diseases, Augenzentrum Siegburg, MVZ Augenärztliches Diagnostik- und Therapiezentrum Siegburg GmbH, Siegburg, Germany. ⁸¹Department of Ophthalmology, Erasmus MC University Medical Center, Rotterdam, the Netherlands. ⁸²Department of Pediatric Surgery, Division of Pediatric Ophthalmology, Montreal Children’s Hospital, McGill University Health Center (MUHC), Montreal, Quebec, Canada. ⁸³Institute for Ophthalmic Research, Centre for Ophthalmology, University Hospital Tübingen, Tübingen, Germany. ⁸⁴University Eye Hospital, Centre for Ophthalmology, University Hospital Tübingen, Tübingen, Germany. ⁸⁵Department of Ophthalmology, Rambam Health Care Campus, Haifa, Israel. ⁸⁶Department of Ophthalmology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic. ⁸⁷Department of Genetics, Institute of Ophthalmology “Conde de Valenciana”, Mexico City, Mexico. ⁸⁸Section of Ophthalmology, King’s College London, St. Thomas’ Hospital Campus, London, UK. ⁸⁹Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. ⁹⁰Molecular Genetics Laboratory, Montpellier University, CHU Montpellier, Montpellier, France. ⁹¹Department of Ophthalmology, Fundación Jiménez Díaz University Hospital, Madrid, Spain. ⁹²Department of Clinical Genetics, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia. ⁹³Eye Clinic, Tartu University Hospital, Tartu, Estonia. ⁹⁴Genetics and Personalized Medicine Clinic, Tartu University Hospital, Tartu, Estonia. ⁹⁵Department of Ophthalmology, Leeds Teaching Hospitals NHS Trust, St James’s University Hospital, Leeds, UK. ⁹⁶Instituto de Investigación Sanitaria La Fe (IIS La Fe) and CIBERER, Valencia, Spain. ⁹⁷Department of Histology and Embryology, Medical University of Warsaw, Warsaw, Poland. ⁹⁸Department of Ophthalmology, School of Medicine, University of Crete, Heraklion, Crete, Greece. ⁹⁹Department of Medicine and Surgery, Medical Genetics, University of Parma, Parma, Italy. ¹⁰⁰UCT/MRC Precision and Genomic Medicine Research Unit, Division of Human Genetics, Department of Pathology, Institute of Infectious Disease and Molecular Medicine (IDM), Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa. ¹⁰¹Department of Medical Genetics, Saint-Etienne University Hospital, Saint-Etienne, France. ¹⁰²Department of Ophthalmology, HUB-Erasme Hospital, Brussels, Belgium. ¹⁰³Department of Neuroscience, Biodonostia Health Research Institute, Donostia-San Sebastián, Spain. ¹⁰⁴Department of Genetic, Physical Anthropology and Animal Physiology, University of the Basque Country (UPV/EHU), Leioa, Spain. ¹⁰⁵Department of Ophthalmology, University of the Basque Country (UPV/EHU), San Sebastián, Spain. ¹⁰⁶West Midlands Clinical Genetics Unit, Birmingham Women’s and Children’s NHS Foundation Trust, Birmingham, UK. ¹⁰⁷College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK. ¹⁰⁸NOVA4Health, NOVA Medical School, Faculdade de Ciências Médicas, NMS, FCM, Universidade NOVA de Lisboa, Lisboa, Portugal. ¹⁰⁹Department of Clinical Pharmacology, Medical University of Vienna, Vienna, Austria. ¹¹⁰Pallas Kliniken AG, Pallas Klinik Zürich, Zürich, Switzerland. ¹¹¹European Vision Institute, Basel, Switzerland. ¹¹²Department of Child Neurology and Psychiatry, IRCCS Mondino Foundation, Pavia, Italy. ¹¹³Department of Medical Genetics, ULS Santa Maria, Lisboa, Portugal. ¹¹⁴Laboratory of Basic Immunology, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal. ¹¹⁵Department of Ophthalmology, School of Medicine, University of Ioannina, Ioannina, Greece. ¹¹⁶Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. ¹¹⁷Department of Ophthalmology, Medical University of Warsaw, Warsaw, Poland. ¹¹⁸SPKSO Ophthalmic University Hospital in Warsaw, Warsaw, Poland. ¹¹⁹Medical Genetics and Prenatal Diagnostics Clinic, Children’s Clinical University Hospital, Riga, Latvia. ¹²⁰Department of Ophthalmology; Oculogenetics Unit, Jules Gonin University Hospital; University of Lausanne, Lausanne, Switzerland. ¹²¹Centre for Gene Therapy & Regenerative Medicine, King’s College London, London, UK. ¹²²Eye Disease Clinic, Children’s University Hospital Riga, Riga, Latvia. ¹²³Department of Molecular Medicine, University of Pavia, Pavia, Italy. ¹²⁴Department of Ophthalmology, Pediatric Ophthalmology and Ophthalmogenetics, University Hospitals Leuven, Leuven, Belgium. ¹²⁵Department of Ophthalmology, New Zealand National Eye Centre, Faculty of Medical and Health Science, University of Auckland, Auckland, New Zealand. ¹²⁶School of Pharmacy and Biomolecular Sciences (PBS), RCSI University of Medicine and Health Sciences, Dublin, Ireland. ¹²⁷FutureNeuro Research Ireland Centre, RCSI University of Medicine and Health Sciences, Dublin, Ireland. ¹²⁸Department of Biochemistry, Faculty of Medicine, UNAM, Mexico City, Mexico. ¹²⁹Bonei Olam - Center for Rare Jewish Genetic Diseases, Brooklyn, NY, USA. ¹³⁰Physicians Dialysis, Miami, FL, USA. ¹³¹Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust, Saint Mary’s Hospital, Manchester, UK. ¹³²Faculty of Biology, Medicine and Health, School of Biological Sciences, Division of Evolution, Infection and Genomics, University of Manchester, Manchester, UK. ¹³³Department of Ophthalmology, Columbia University Irving Medical Center, New York, NY, USA. ¹³⁴Genomics England Ltd, London, UK. ¹³⁵These authors contributed equally: Mathieu Quinodoz, Kim Rodenburg. ¹³⁶These authors jointly supervised this work: Susanne Roosing, Carlo Rivolta. ✉e-mail: susanne.roosing@radboudumc.nl

Methods

Patients and DNA samples

This study adhered to the tenets of the Declaration of Helsinki, and signed, informed consent was obtained from all participants. All procedures were conducted in accordance with Institutional Review Board-approved human research protocols and were approved by the ethics committees of the Radboud University Medical Center (Nijmegen, the Netherlands) and the Rotterdam Eye Hospital (Rotterdam, the Netherlands) (MEC-2010-359; OZR protocol no. 2009-32), and the local ethics committees of all other participating institutions.

Clinical characterization and analysis

Complete ophthalmic examinations were performed by a retinal specialist, which included measurement of best-corrected visual acuity and intraocular pressures, and examination of anterior segment and fundus (dilated). Color fundus photographs and montages were captured using the FF450plus Fundus Camera (Carl Zeiss Meditec) and Optos 200 Tx (Optos). Fundus autofluorescence images (488-nm excitation) and high-resolution spectral-domain optical coherence tomography (SD-OCT) scans were acquired using the Spectralis HRA+OCT module (Heidelberg Engineering). Hyper-autofluorescent ring contours were analyzed using a custom program in FIJI software (National Institute of Mental Health)⁴⁵. Progression rates were calculated using linear mixed-effects regression in R (v.4.0.4) with time (years) since baseline as the primary independent variable, baseline ring size as a covariate and inter-ocular differences as a random effect. Photoreceptor+ thickness was assessed on horizontal SD-OCT scans through the fovea using a semi-automated procedure⁴⁶. Photoreceptor+ was defined as the distance between the Bruch's membrane/choroid interface and the inner nuclear layer/outer plexiform layer boundary. Layer segmentation was performed in a semi-automated manner using a custom software in MatLab (MathWorks). Full-field electroretinogram recordings were conducted using the Espion Visual Electrophysiology System (Diagnosys) according to International Society for Clinical Electrophysiology of Vision (ISCEV) standards⁴⁷.

Genome sequencing and annotation

Genomic DNA from probands was isolated from peripheral blood lymphocytes according to standard procedures. Sequencing was performed by BGI Tech Solutions using the DNBseq Sequencing Technology, with a minimal median coverage per genome of 30×. The processing of the sequencing data was performed by using BWA mem (v.0.7.17)⁴⁸, Picard (v.2.14.0-SNAPSHOT) (<http://broadinstitute.github.io/picard>) and GATK (v.4.1.4.1)⁴⁹ for mapping to the human genome reference sequence (build hg19/GRCh37) and variant calling². For variant annotation, we used ANNOVAR⁵⁰ with the addition of splicing predictions by MaxEntScan⁵¹ and SpliceAI¹⁹.

Assessment of variants

Human Genome Variation Society (HGVS) notations of the variants were retrieved using VariantValidator⁵² and American College of Medical Genetics and Genomics (ACMG) classification⁵³ was applied according to the ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2023⁵⁴. In particular, we used the PS4_strong criterion for variants significantly enriched in cases versus controls (gnomAD v.4.1 and All of Us), as assessed for each variant by two-tailed Fisher's exact test in R (fisher.test function), in agreement with the ACMG recommendations⁵³ (odds ratio > 5.0, lower bound of the confidence interval > 1.0, corrected *P* value < 0.05), but only for variants present in at least three probands to avoid any bias from imbalanced case-control sets. This assessment was made using probands only (*n* = 1,891) for all variants, except for those in Supplementary Table 1, which were assessed in 4,722 individuals. PM6 and PP1 were applied according to ClinGen Sequence Variant Interpretation (SVI) recommendations. Specifically, PM6_sup was applied when two unrelated families had de novo variants without

parental confirmation, given that RP is a 'phenotype consistent with gene but not highly specific and high genetic heterogeneity'. PP1_sup, PP1_mod and PP1_strong were assigned when the variant segregated with disease in ≥1, ≥2 and ≥5 informative meioses, respectively⁵⁵. We defined thresholds for PM2_sup and BS2 based on the frequency of *RHO* p.(Pro23His), the most prevalent variant causing adRP, which was detected once in gnomAD v.4.1 and 13 times in All of Us. Specifically, PM2_sup was assigned to variants that were present fewer than two times in gnomAD v.4.1 and fewer than 14 times in All of Us. BS2 was applied to variants that were observed more than four times in gnomAD v.4.1 or more than 28 times in All of Us, that is, twice the values of p.(Pro23His). PM2_sup was not applied to variants for which the PS4 criterion had already been used, to avoid double-counting evidence related to their low frequency in gnomAD. BA1 was considered for variants with allele frequency >5% in gnomAD v.4.1 or the All of Us databases, whereas BS1 was assigned to variants with allele frequencies greater than expected for disease (1/2,000 = 0.05%).

Screening by Sanger sequencing

Genomic DNA was collected, and *RNU4-1*, *RNU4-2*, *RNU6-1*, *RNU6-2*, *RNU6-7*, *RNU6-8* and *RNU6-9* genes were amplified using standard PCR procedures. *RNU4-1*, *RNU4-2*, *RNU6-1*, *RNU6-2*, *RNU6-7*, *RNU6-8* and *RNU6-9* PCR fragments were sequenced using Sanger sequencing and screened for novel variants in these genes.

Two-dimensional modeling of the effect of variants and three-dimensional representation

We utilized RNAfold WebServer to model the effect of variants with default parameters⁵⁶ and RNAcanvas was used for drawing the structure⁵⁷. We used ChimeraX with PDB file, using PDB file 6QW6 to draw three-dimensional representation of the U4/U6 duplex with and without surrounding PRPF proteins.

RNA-seq experiments and analysis

RNA was isolated from human donor eye tissue, which was collected and dissected according to a reported procedure⁵⁸ from an ethically approved Research Tissue Bank (UK NHS Health Research Authority reference no. 15/NW/0932). Total RNA was isolated from four NSR samples, 16 pelleted RPE samples and 13 choroid samples that had been stored in RNeasy Lysis Buffer (Qiagen), using an Animal Tissue RNA Purification kit (Norgen Biotek), as per manufacturer's instructions. Sequencing libraries were prepared using the NEBnext multiplex small RNA library preparation kit, as per manufacturer's protocols, with size selection performed using Ampure beads. Paired-end sequencing (2 × 75 base pairs (bp)) was performed on an Illumina HiSeq 4000.

NEBnext adapters were removed from sequencing reads using trimmomatic (v.0.39) before alignment against the GRCh38 reference genome with bowtie⁵⁹ (v.1.3). No mismatches between sequencing reads and the reference genome were allowed, and no restriction was set on multi-mapping reads. Sequence read counts were restricted to primary alignments using samtools (v.1.21)⁶⁰, and therefore only counted once if they aligned to multiple *RNU4* (*n* = 90) or *RNU6* (*n* = 1,277) genes or pseudogenes. Calculations were drawn from read 1 datasets and normalized for the total read count achieved for the sample. Total *RNU4* and *RNU6* expression was based on all annotated genes and pseudogenes in GENCODE v.38.

ATAC-seq and H3K27ac ChIP-seq data

ATAC-seq data from ref. 61 (eight different experiments) and H3K27ac ChIP-seq data from ref. 62 (five different experiments) were downloaded as bigwig files from the RegRet database (<http://genome.ucsc.edu/s/stvdsomp/RegRet>)⁶³. For both data types, the signal (the genes and 500 bp on each side) was extracted using bedtools (v.2.27.1) after conversion using bigWigToWig (v.4.69). We quantified the signal for all *RNU* genes and pseudogenes first by normalizing the signal of each

experiment to the maximum and then summing them. For *RNU4*, we quantified two genes and 105 pseudogenes, while for *RNU6* we assessed five genes and 1,312 pseudogenes, in addition to *RNU4ATAC* and *RNU6ATAC*.

RNA-seq from blood RNA

Peripheral blood samples were collected from affected individuals and controls using either Tempus Blood RNA tubes (Applied Biosystems) or PAXgene Blood RNA tubes (Qiagen). Total leukocyte RNA was extracted with the Tempus Spin RNA Isolation Kit (Applied Biosystems) or the Preserved Blood RNA Purification Kit II (Norgen Biotek), respectively, following the manufacturers' protocols. Following the quality assessment of RNA integrity and concentration, 100 ng of input RNA per sample was subsequently processed for library preparation using the KAPA RNA HyperPrep Kit with RiboErase (HMR) and KAPA Globin Depletion Hybridization Oligos (Roche). Sequencing was performed on an Illumina NovaSeq 6000 platform with 2 × 101-bp paired-end reads. To improve quality score calculations for the final base, one additional base was sequenced in both read 1 and read 2. The Q30 value for all RNA-seq data was ≥91.1%. Adapters were trimmed with Skewer (v.0.2.2)⁶⁴.

Reads were aligned to reference transcripts from Ensembl (v.110, GRCh38) using STAR (v.2.7.11a) with the option `--twopassMode Basic`. DESeq2 (v.1.46.0) with default options was used for differential expression analysis between different groups according to sample origin (Tempus or PAXgene tubes) and presence/absence of the pathogenic *RNU* genotypes, with fold-change > 2 and FDR *P* value < 0.05. We used rMATS⁶⁵ to assess differential alternative splicing, separately for the Tempus and PAXgene sets and with specific options (`--allow-clipping --variable-read-length --anchorLength 1 --novelSS --task both --libType fr-unstranded -t paired --readLength 101`). We further used the Python scripts from ref. 11 to process the rMATS output and filter the data according to a mean coverage > 7, an FDR *P* value < 0.1 and a deltaPSI value > 0.05. The R function `fisher.test` with default parameters was used to assess differences in base compositions at splicing sites, at each position, as well as differences for 2-mers (for example, positions -4/-3 to +7/+8).

U4 and U6 snRNP analysis

U4 n.18_19insA, n.56T>C and n.64_65insT variants were introduced by site-directed mutagenesis into the plasmid expressing U4-MS2⁶⁶. The full-length U6 sequence, including 256 bp upstream and 93 bp downstream of the *RNU6-1* gene, was inserted into the pcDNA3 plasmid lacking the CMV promoter. The MS2 loop was inserted between nucleotides 10 and 11. U6 n.55_56insG, n.55_56insT, n.56_57insG and n.57G>T variants were introduced by site-directed mutagenesis. U4- and U6-expressing plasmids were transfected into HeLa cells stably expressing MS2-YFP protein. At 24 h after transfection, snRNAs were immunoprecipitated using anti-GFP antibodies and co-precipitated proteins were analyzed by western blotting⁶⁶.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Research on the de-identified patient data used in this publication from the Genomics England 100,000 Genomes Project and the NHS GMS dataset can be carried out in the Genomics England Research Environment subject to a collaborative agreement that adheres to patient-led governance. All interested readers will be able to access the data in the same manner that the authors accessed the data. For more information about accessing the data, interested readers may contact research-network@genomicsengland.co.uk or access the relevant information on the Genomics England website: <https://www.genomicsengland.co.uk/research>. Sharing of additional sequencing or

blood RNA-seq data is subject to the European General Data Protection Regulation (GDPR) applicable in the countries of residence of the tested individuals and may become available upon a data transfer agreement approved by local ethical committees. Patient sample identifiers from this study can be released upon reasonable request from 'M1-A to M9-B' to the corresponding local 'DNA-number'. Specific variant requests or other data are available from the corresponding author (S.R.) upon reasonable request. The data generated during this study (causative variants from Supplementary Table 1) are submitted to the Leiden Open (source) Variation Database (LOVD) (<http://www.lovd.nl>) and ClinVar (accession codes [SCV006562526](https://www.ncbi.nlm.nih.gov/sra/PRJNA1256119) to [SCV006562534](https://www.ncbi.nlm.nih.gov/sra/PRJNA1256119)). Sequences of primers used in this study are listed in Supplementary Table 9. Additional details regarding PCR conditions or primer design are available upon request. Detailed information on antibodies used are provided in Supplementary Table 10. Small RNA-seq datasets analyzed in this study are available at the NCBI Sequence Read Archive through accession PRJNA1256119 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA1256119>). The genes and pseudogenes analyzed are present in Supplementary Table 11 and the read counts are available in Supplementary Table 12. Source data are provided with this paper.

References

- Petersen-Jones, S. M. et al. Patients and animal models of CNGβ1-deficient retinitis pigmentosa support gene augmentation approach. *J. Clin. Invest.* **128**, 190–206 (2018).
- Lee, W. et al. *Cis*-acting modifiers in the ABCA4 locus contribute to the penetrance of the major disease-causing variant in Stargardt disease. *Hum. Mol. Genet.* **30**, 1293–1304 (2021).
- Robson, A. G. et al. ISCEV Standard for full-field clinical electroretinography (2022 update). *Doc. Ophthalmol.* **144**, 165–177 (2022).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
- Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **11**, 377–394 (2004).
- Freeman, P. J., Hart, R. K., Gretton, L. J., Brookes, A. J. & Dalgleish, R. VariantValidator: accurate validation, mapping, and formatting of sequence variation descriptions. *Hum. Mutat.* **39**, 61–68 (2018).
- Richards, S. 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).
- Ellard, S. et al. ACGS Best Practice Guidelines for Variant Classification 2019. ACGS Best Practice Guidelines for Variant Classification 2019 (ACGS, 2019); <https://www.acgs.uk.com/media/11285/uk-practice-guidelines-for-variant-classification-2019-v1-0-3.pdf>
- Biesecker, L. G. et al. ClinGen guidance for use of the PP1/BS4 co-segregation and PP4 phenotype specificity criteria for sequence variant pathogenicity classification. *Am. J. Hum. Genet.* **111**, 24–38 (2024).
- Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res.* **36**, W70–W74 (2008).
- Johnson, P. Z. & Simon, A. E. RNAcanvas: interactive drawing and exploration of nucleic acid structures. *Nucleic Acids Res.* **51**, W501–W508 (2023).

58. McHarg, S. et al. Mast cell infiltration of the choroid and protease release are early events in age-related macular degeneration associated with genetic risk at both chromosomes 1q32 and 10q26. *Proc. Natl Acad. Sci. USA* **119**, e2118510119 (2022).
59. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
60. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
61. Wang, J. et al. ATAC-Seq analysis reveals a widespread decrease of chromatin accessibility in age-related macular degeneration. *Nat. Commun.* **9**, 1364 (2018).
62. Cherry, T. J. et al. Mapping the *cis*-regulatory architecture of the human retina reveals noncoding genetic variation in disease. *Proc. Natl Acad. Sci. USA* **117**, 9001–9012 (2020).
63. Van de Sompele, S. et al. Multi-omics approach dissects *cis*-regulatory mechanisms underlying North Carolina macular dystrophy, a retinal enhanceropathy. *Am. J. Hum. Genet.* **109**, 2029–2048 (2022).
64. Jiang, H., Lei, R., Ding, S. W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
65. Shen, S. et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-seq data. *Proc. Natl Acad. Sci. USA* **111**, E5593–E5601 (2014).
66. Roithová, A. et al. The Sm-core mediates the retention of partially-assembled spliceosomal snRNPs in Cajal bodies until their full maturation. *Nucleic Acids Res.* **46**, 3774–3790 (2018).

Acknowledgements

We thank all patients and their family members for their help and participation in this study. We thank S. Föhr for technical and administrative support; S. van der Velde-Visser, E. Blokland and M. Jacobs-Camps for sample registration and administration; and T. Rosseel (computational) and S. Van de Sompele (clinical reporting) for their support. The title page was formatted using AuthorArranger, a tool developed at the National Cancer Institute (National Institutes of Health, NIH). C.R. was supported by the Swiss National Science Foundation (SNSF) Grant No. 310030_204285 entitled ‘Genomics of inherited retinal diseases’. S.R. was supported by the Foundation Fighting Blindness Career Development Award (grant no. CD-GE-0621-0809-RAD), Radboudumc Starter grant (no. OZI-23.009) and NWO Aspasia (grant no. 015.021.028). S.R., C.R., E.D.B., M.B., S.B. and D. Stanek were supported by HORIZON-MSCA-2022-DN (grant no. 101120562, ProgRET). E.D.B., S.R. and C.R. were supported by the EJPRD19-234 Solve-RET. This work has been funded by a Foundation Fighting Blindness Program Project Award (grant no. PPA-0622–0841-UCL) (to A.J.H., S.R. and S.E.d.B.). S.R. and F.P.M.C. were supported by the Gelderse Blindenstichting, the Algemene Nederlandse Vereniging ter voorkoming van Blindheid, Oogfonds, Landelijke Stichting voor Blinden en Slechtienden, Rotterdamse Stichting Blindenbelangen, Stichting Blindenhulp, Stichting tot Verbetering van het Lot der Blinden and Stichting Blinden-Penning. M.Q. was supported by the RetinAward 2021. S.H.T. – Jonas Children’s Vision Care (JCVC) is supported by the National Institute of Health grant nos. U01EY030580, U01EY034590 R24EY028758, R24EY027285, 5P30EY019007, R01EY033770 and R01EY018213, R01EY024698; the Foundation Fighting Blindness grant no. TA-GT-0321-0802-COLU-TRAP; Richard Jaffe; the NYEE Foundation; the Rosenbaum Family Foundation; the Gebroe Family Foundation; the Piyada Phanaphat Fund; the Research to Prevent Blindness (RPB) Physician-Scientist Award; and unrestricted funds from RPB, New York, NY, USA. C.A. was supported by Instituto de Salud Carlos III (ISCIII) of the Ministerio de Ciencia e Innovación and Unión Europea – European Regional Development Fund (FEDER) (grant nos. PI22/00321 and IMP/00009), Centro de Investigación Biomédica en Red Enfermedades Raras

(CIBERER, grant no. 06/07/0036), IIS-FJD BioBank (grant no. PT23/00114), the Organización Nacional de Ciegos Españoles (ONCE), the European Regional Development Fund (FEDER) and the University Chair UAM-IIS-FJD of Genomic Medicine. This work was performed by using the data contained in the ‘Programa Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología en Medicina Genómica (IMPACT-GENÓMICA)’, coordinated by the CIBERER and founded by ISCIII. L.F.-C. was supported by Centro de Investigación Biomédica en Red (CIBER). R.A. was supported by the National Eye Institute (NEI) (grant nos. RO1 EY030591, RO1 EY031663, T32 EY026590 and P30 EY22589). C.C.W.K. was supported by the Combined Ophthalmic Research Rotterdam grant no. 8.2.0. S.B. was supported by the Italian Telethon Foundation and by the European Union HORIZON-MSCA-2021-DN-01 (grant no. 101073316, RETORNA). T.S.B. was supported by ZonMw Vidi (grant no. 09150172110002), and acknowledges support from Stichting 12q. G.J.F. and N.C. were supported by Fighting Blindness Ireland (grant nos. FB22FAR, FB16FAR), Fighting Blindness Ireland – Health Research Charities Ireland (grant no. MRCG-2016-14) and the Science Foundation Ireland (grant nos. 16/IA/4452 and 22/FFP-A/10544). J.M.E. was supported by the Macular Society (United Kingdom), the National Institute for Health and Care Research (NIHR) Manchester Biomedical Research Centre (BRC) (grant no. NIHR203308) and the University of Manchester Core Genomics Technology Facility. T.B.-Y. was supported by the Israel Science Foundation (grant no. 331/24). A.C.B.-J. is supported by the University of Melbourne Research Fellowship. K.M.B. was supported by the National Eye Institute (NEI) (grant nos. RO1 EY035717 and P30 EY014104 (MEE core support)), the Iraty Award 2023, the Lions Foundation and RPB (Unrestricted Grant). L.S.S., E.L.C. and S.P.D. were supported by grants from the Foundation Fighting Blindness (grant no. EGI-GE1218-0753-UCSD) and the Brett & Jane Eberle Foundation. E.D.B. and B.P.L. were supported by Ghent University Special Research Fund (grant no. BOF20/GOA/023) and E.D.B. (grant no. 1802220N) and B.P.L. (grant no. 1803816N) are Senior Clinical Investigators of the Research Foundation-Flanders (FWO). N.M. and S.S. are Ph.D. fellows of HORIZON-MSCA-2022-DN ProgRET (grant no. 101120562). R.A. was supported by the Foundation Fighting Blindness. J.L.D. was supported by the UCSF Vision Core shared resource of the NIH/NEI grant no. P30 EY002162, the Foundation Fighting Blindness, an unrestricted grant from RPB and the All May See Foundation. T.I. was supported by research grants from the Japan Agency for Medical Research and Development (AMED) (grant nos. 20ek0109493h0001, 21ek0109493h0002, 22ek0109493h0003, 23ek0109617h0002, 24ek0109617h0003). R.K.K. was supported by The Montreal Children’s Hospital Foundation, The Vision Sciences Research Network (VSRN), The NIH (grant no. RO1 EY030499-01, Dr. Lentz), The Canadian Institutes for Health Research (CIHR), Fighting Blindness Canada (FBC) and Fonds de Recherche du Québec - Santé (FRQS). R.K.K. participates in the NAC Attack clinical trial, which is funded by the NIH via grant nos. UG1EY033286, UG1EY033293, UG1EY033286 and UG1EY033292. T.M.L., T.L.M. and J.N.D.R. were supported by Retina Australia (awarded to the Australian Inherited Retinal Disease Registry and DNA Bank). O.A.M. was supported by the Wellcome Trust (grant no. 206619/Z/17/Z). M.P. was supported by the BrightFocus Foundation (grant no. M2024009N). E.A.P. was supported by the National Eye Institute (NEI) (grant no. RO1 EY012910). R.R. was supported by Retina South Africa and the South African Medical Research Council (MRC). S.G.S. and E.M.V. were supported by the Italian Ministry of Health (grant no. PNRR-MR1-2023-12377314). D. Stanek was supported by the Project P JAC grant no. CZ.02.01.01/00/22_008/0004575 RNA for therapy, Co-Funded by the European Union. T.B.H. was supported by the European Commission (Recon4IMD – grant no. GAP-101080997) and the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG, grant nos. 418081722 and 433158657 to T.B.H.). P.L. and L.D. were supported by a research grant (no. NW24-06-00083)

from the Ministry of Health of the Czech Republic and grant no. UNCE/24/MED/022. V.R.d.J.L.-R. and J.C.Z. were supported by the Velux Stiftung Grant no. 1860. M.A., M.M., C.F.I., J.C.G., A.J.H. and C.T. are supported by Retina UK and Fight for Sight UK (RP Genome Project Grant no. GR586). A.J.H., J.C.G., M. Michaelides, O.A.M., A.R.W., G. Arno and S.L. were supported by the National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital and UCL Institute of Ophthalmology. S.L. was funded by a Medical Research Council (MRC) Clinician Scientist Fellowship (grant no. UKRI440). J.M.M. was supported by Instituto de Salud Carlos III (ISCIII) of the Spanish Ministry of Health (grant no. PI22/00213), CIBERER (grant no. 06/07/1030), grant no. CIPROM/2023/26 from the Generalitat Valenciana and IMPaCT-GENOMICA (grant no. IMP/00009) co-funded by ISCIII and FEDER. S.R., C.C.W.K., C.J.F.B. and A.G. were supported by the Dutch Ministry of Education, Culture and Sciences, Gravitation grant no. 024.006.034 Lifelong VISION. W.L. was supported by the National Institute of Health/National Eye Institute (grant no. 1K99EY036930-01). This work is supported by partners of the European Reference Network for Rare Eye Diseases ERN-EYE (Grant Agreement No. 101085439, C.J.F.B., E.D.B., C.B.H., S.K., B.P.L., P.L., L.H.-W., K. Stingl, L.I.v.d.B.). Novartis contributed funding for the preceding RP-LCA smMIPs panel design and subsequent sequencing (to F.P.M.C., S.R. and D.M.P.). Novartis was not involved in the study design; collection, analysis or interpretation of data; the writing of this article; or the decision to submit it for publication. This research was made possible through access to data generated by the 2025 French Genomic Medicine initiative and present in the National Genomic Research Library, which is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The National Genomic Research Library holds data provided by patients and collected by the NHS as part of their care, and data collected as part of their participation in research. The National Genomic Research Library is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure.

Author contributions

M.Q., F.P.M.C., S.R. and C.R. conceived and designed the study. M.Q., K.R., K.K. and M.U. analyzed pedigrees and genotypes. J.B., S.H.T., C.L., A.I.J., K.B.F. and W.L. contributed to the collection of material and clinical data for family M1-A. F.P.M.C., S.E.d.B., D.M.P., R.J.H.-M. and S.R. coordinated the collection of DNA samples. K.R., S.E.d.B., E.G.M.B., N.Z., L.K.H., Z. Corradi, S.S., D.M.P. and R.J.H.-M. performed the majority of the Sanger sequencing screening. K.K., A.B.I.-R. and M.F. carried out the molecular biology experiments. M.Q. and C.R.

performed the statistical analyses. M.Q. and J.M.E. retrieved and analyzed large-scale data. K.R. and S.R. coordinated the genotyping of the cohorts. W.L. was responsible for clinical data analysis. Z. Cvackova and D. Stanek conducted all in vitro experiments. F.P.M.C., C.R. and S.R. supervised the project. M.Q., K.R., F.P.M.C., J.M.E., D. Stanek, S.R. and C.R. prepared the original draft of the manuscript. M.Q., K.R., S.R. and C.R. reviewed and edited the manuscript. M.A., A.A., S.A., G. Ansari, G. Arno, G.D.N.A., C.A., R.A., S.B., E.B., T.S.B., M.T.S.B., M.B., T.B.-Y., V.B., D.G.B., P.B., F.B.-K., B.B., C.J.F.B., K.B., D.B.-G., A.C.B.-J., K.M.B., C.B.d.R., E.L.C., G.C., F.C., L.C., N.C., P.C.I., L.C.-S., S.P.D., E.D.B., M.D.B., B.d.l.C., J.N.D.R., J.D.Z., R.D., C.-M.D., L.D., J.L.D., G.J.F., N.F., B.J.F., L.F.-C., J.M.F.S., S.G., A.G., J.C.G., C.G., R.G.-D., K.G., S.G.-J., T.B.H., L.H.-W., A.J.H., T.H., E.H., L.H.H., A.H., J.P.H., C.B.H., M.B.B.I., C.F.I., T.I., B.O.J., K.J., V.K., S. Kamakari, M.K., U.K., C.C.W.K., K.K., R.K.K., S. Kohl, T.K., L.K., T.M.L., R.L., B.P.L., S.L., P.L., I.L., V.R.d.J.L.-R., Q.M., O.A.M., G.M., L. Mansard, M.P.M.-G., N.M., L. Muring, M. McKibbin, T.L.M., I.M., M. Michaelides, J.M.M., K.M., R.M., Z.Z.N., K.N., M. Otdak, M. Oorsprong, Y.P., A. Papachristou, A. Percesepe, M.P., E.A.P., E.P., R.R., F.R., F.A.R., G.I.R., L.R., M.R.-H., J.R.-E., A.H.S., A.F.S., A.I.S.-B., A.S.S., R.S., C.M.S., M. Scarpato, H.P.N.S., D. Sharon, S.G.S., F.S., A.B.S., M. Stefanidou, K. Stefansson, K. Stingl, A.S., P.S., L.S.S., V.S., J.P.S., G.T., A.A.H.J.T., C.T., V.H.T., M.K.T., P.T., V.V., M.V., S. Valeina, E.M.V., C.V., R.V., S. Valleix, J.v.A., L.I.v.d.B., M.V.H., V.J.M.V., A.L.V., A.R.W., L.W., B.W., G.G.Y., K.Y., J.C.Z., R.Z. and T.Z. were instrumental in acquiring funding, clinical data, validation of genetic or molecular data, and reviewing and editing the final manuscript. All authors approved the final content of this work.

Competing interests

The authors affiliated with deCODE genetics/Amgen Inc. (B.O.J., K. Stefansson, P.S.) are employed by the company. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41588-025-02451-4>.

Correspondence and requests for materials should be addressed to Susanne Roosing.

Peer review information *Nature Genetics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒ ☐ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☐ ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection No software was used for the data collection.

Data analysis INSPIRE (<https://github.com/jiazhaoy97/INSPIRE>, <https://inspire-tutorial.readthedocs.io/en/latest>), python v3.8.15, pytorch v1.12.1, pandas v1.5.3, numpy v1.24.4 were used for designing and running the algorithm. SpiceMix v1.0.0 (<https://github.com/ma-compbio/SpiceMix>), NSFH v1.0.0 (<https://github.com/willtownes/nsf-paper>), PASTE v1.4.0 (<https://github.com/raphael-group/paste>), PRECAST v1.6.4 (<https://github.com/feiyong/PRECAST>), mofapy2 v0.7.2 (<https://github.com/bioFAM/mofapy2>), GraphST v1.0.0 (<https://github.com/JinmiaoChenLab/GraphST>), Harmony v0.1 (<https://github.com/immunogenomics/harmony>), Seurat v4.3.0 (<https://github.com/satijalab/seurat>), LIGER v1.0.1 (<https://github.com/welch-lab/liger>), STAGATE v1.0.0 (<https://github.com/zhanglabtools/STAGATE>), Banksy v0.1.6 (<https://github.com/prabhakarlab/Banksy>), SpaGCN v1.2.7 (<https://github.com/jianhuupenn/SpaGCN>), SpatialGlue v1.0.0 (<https://github.com/JinmiaoChenLab/SpatialGlue>), BayesSpace v1.5.1 (<https://github.com/edward130603/BayesSpace>), BASS v1.1.0 (<https://github.com/zhengli09/BASS>), SpatialLeiden v0.2.0 (<https://github.com/HiDiHlabs/SpatialLeiden>) were used for the benchmarking study. scanpy v1.8.2, anndata v0.9.2, louvain v0.8.0, umap-learn v0.5.5, matplotlib v3.5.3 were used for preprocessing and post-processing of the spatial transcriptomics data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All data used in this work are publicly available through online sources.

Human dorsolateral prefrontal cortex dataset profiled by Visium platform (<https://research.libd.org/spatialLIBD/>);

Mouse brain sagittal anterior, sagittal posterior, and coronal sections profiled by Visium (<https://www.10xgenomics.com/datasets>);

Mouse brain sections profiled by Slide-seq V2 (https://singlecell.broadinstitute.org/single_cell);

Mouse brain sections profiled by MERFISH (<https://doi.brainimaginglibrary.org/doi/10.35077/act-bag>);

Mouse whole-embryo section profiled by seqFISH (<https://crukci.shinyapps.io/SpatialMouseAtlas/>);

Mouse skin sections during wound healing profiled by the Visium platform (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178758>);

Human breast cancer sections profiled by Xenium (<https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast>);

Mouse whole-embryo datasets across different developmental time points profiled by Stereo-seq (<https://db.cngb.org/stomics/mosta/>);

Mouse hypothalamic preoptic region sections profiled by MERFISH (<https://doi.org/10.5061/dryad.8t8s248>);

Mouse hippocampus region sections profiled by SRARmap PLUS (<https://doi.org/10.5281/zenodo.7458952>);

Mouse brain sections profiled by the ST platform (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE147747>).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

N.A.

Reporting on race, ethnicity, or other socially relevant groupings

N.A.

Population characteristics

N.A.

Recruitment

N.A.

Ethics oversight

N.A.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences

☐ Behavioural & social sciences

☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

We used 11 publicly available data in the manuscript.

Data exclusions

We removed genes, cells, and spots by applying standard preprocessing steps such as removing lowly expressed genes, cells, and spots.

Replication

While we did not generate any data, all code is provided openly to enable others to replicate our results.

Randomization

Randomization does not apply to this work because there is no experimental data collected in this study.

Blinding

Blinding does not apply to this work because there is no experimental data collected in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.
Novel plant genotypes	Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.