GPATCH11 variants cause mis-splicing and earlyuncle-onset retinal dystrophy with neurological impairment

Andrea Zanetti[1], Lucas Fares-Taie[1], Jeanne Amiel[2], Pierre David[3], Stéphanie Moriceau[4], Nicolas Goudin[5], Jérôme Roger[6], Ida Chiara Guerrera[7], Vincent Jung[7], Cyril Gitiaux[8], Isabelle Audo[9], Matthieu Robert[10], Nathalie Boddaert[11], Sabine Sigaudy[12], Nicole Philip-Sarles[12], Sylvain Briault[13], Carlo Rivolta[14-15], Karolina Kaminska[14-15], Danielle Amana[16], Ange-Line Bruel[17], Christel Thauvin[18], Josseline Kaplan[1], Jean-Michel Rozet[1], Isabelle Perrault[1*]

[1]Laboratory of Genetics in Ophthalmology (LGO), INSERM UMR1163, Institute of Genetic Diseases, Imagine and Paris Cité University, 75015 Paris, France

[2]Laboratory of Embryology and Genetics of Malformations, INSERM UMR1163, Institute of Genetic Diseases, Imagine and Paris Cité University, 75015 Paris, France.

[3]Transgenesis platform, Laboratory of Animal Experimentation and Transgenesis (LEAT) of the Structure Fédérative de Recherche Necker, INSERM US24/CNRS UMSS3633, Institute of Genetic Diseases, Imagine, 75015 Paris, France

[4]Platform for Neurobehavioral and metabolism, Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR 3633, Institute of Genetic Diseases, Imagine, 75015 Paris, France

[5]Necker Bioimage Analysis Core Facility of the Structure Fédérative de Recherche Necker, INSERM US24/CNRS UAR 3633, 75015 Paris, France

[6]Paris-Saclay Institut of Neuroscience, CERTO-Retina France, CNRS, Paris-Saclay University, 91400 Saclay, France

[7]Proteomic Platform Necker, INSERM US24/CNRS UAR 3633, Structure Fédérative de Recherche Necker, Paris, France.

[8]Reference center for neuromuscular pathologies Paris Nord Est, Department of Clinical Neurophysiology, University Hospital Necker-Enfants Malades, Paris Cité University, 75015 Paris, France

[9]Sorbonne Université, INSERM, CNRS, Institut de la Vision, Paris, France; Centre Hospitalier National d'Ophtalmologie des Quinze-Vingts, National Rare Disease Center REFERET F-75012 Paris, France.

[10]Ophthalmology Department, University Hospital Necker-Enfants Malades, APHP, 75015 Paris, France

[11]Pediatric-Radiology Deparment, University Hospital Necker-Enfants Malades, APHP, 75015 Paris, France

[12]Medical Genetics Department, Hospital Timone Enfant, 13000 Marseille, France

[13]Genetics Department, Regional Hospital Orleans (CHRO), 45000 Orleans, France

[14]Institute of Molecular and Clinical Ophthalmology Basel (IOB), Basel, Switzerland

[15]Department of Ophthalmology, University of Basel, Basel, Switzerland.

[16]Ophthalmology Department, Hospital Center of Orleans, 45000 Orleans, France

[17]INSERM UMR1231, GAD, University of Burgundy, FHU-TRANSLAD, CHU Dijon-Bourgogne, 21000 Dijon, France

[18]Reference Center for Rare Diseases "Developmental Abnormalities and Malformation Syndromes" of the East, Genetic center, Hospital d'Enfants, FHU TRANSLAD, CHU Dijon Bourgogne, 21000 Dijon, France

[*]Correspondence: isabelle.perrault@inserm.fr

ABSTRACT

G-patch domain-containing proteins have emerged as pivotal regulators of RNA metabolism, and the dysfunction of these proteins has been linked to a large range of phenotypic traits. This study provides compelling evidence that biallelic mutations in *GPATCH11* lead to severe early-onset photoreceptor degeneration, neurological issues, and skeletal abnormalities in humans. Analysis of exome datasets from affected individuals of four independent families identified variants in a new gene, *GPATCH11*. The aim of the study was to determine the genetic source of the patients' symptoms and characterize the role of *GPATCH11*. Using engineered disease variants in mice, we showed abnormal levels of spliceosome components in mutant retinas and substantial splicing deregulation, indicating a role for GPATCH11 in spliceosome homeostasis and splicing activity. Additionally, we identified deregulated gene expression, partly independent of splicing abnormalities, suggesting the involvement of GPATCH11 in transcriptional regulation. The affected genes were found to be related to phototransduction and the transmission of visual messages, protein homeostasis, RNA homeostasis, and cilia metabolism. The presence of GPATCH11 in ciliary basal bodies further suggests a potential role in ciliary functions. Studying GPATCH11 dysfunction elucidates the complex roles of this protein in RNA metabolism, highlighting the significance of GPATCH11 in maintaining proper gene expression and its contributions to various aspects of retinal, neurological, and skeletal function or development. As demonstrated here, studying rare genetic disorders can help to define broader functions beyond the specific disease context, and study provides new insight into the broader roles of GPATCH11.

INTRODUCTION

By analysis of exome sequencing in four affected families, we present compelling evidence that biallelic mutations in *GPATCH11* are responsible for a severe and early-onset neurodegenerative retinal disease, affecting both rod and cone photoreceptors, associated with neurological problems and skeletal abnormalities (including facial dysmorphia). *GPATCH11*, also known as coiled-coil domain containing 75 (*CCDC75*) and centromere protein Y (*CENP-Y*), belongs to the less explored group of GPATCH domain-containing proteins. While *GPATCH11* has been shown to be localized in the nucleus[1], its specific role in RNA metabolism is not fully understood.

The spliceosome is a large ribonucleoprotein complex that plays a crucial role in pre-mRNA splicing, intron excision, and exon ligation to generate functional mRNA. Comprising five core small nuclear ribonucleoprotein particles (snRNPs) and numerous other protein factors, the spliceosome requires precise coordination for proper assembly, activation, and regulation to ensure accurate RNA splicing and cellular function[2].

Among the contributors to the spliceosome, G-patch domain-containing (GPATCH) proteins form a distinct group that is characterized by a glycine-rich motif and multiple RNA-binding motifs[3,4]. Despite sharing common features, GPATCH proteins exhibit considerable variations in size, domain composition, and cellular localization[3]. Through their protein–protein and protein-nucleic acid interactions, they contribute significantly to RNA metabolism by regulating the remodelling of RNAs and ribonucleoprotein complexes by the DEAH-box RNA helicases with which they associate[3,5]. In humans, 23 GPATCH-domain-containing proteins have been identified[3]. Of these, a dozen have been reported to be involved in pre-mRNA splicing and/or transcription regulation, and others participate in ribosome biogenesis, RNA export, rRNA and snoRNA maturation, and telomere maintenance[6,7].

Consistent with the roles of GPATCH proteins in these various processes, various phenotypic traits have been associated with GPATCH variants, but only a few have been confirmed as disease-causing mutations. Notably, *RBM10*[8], *SON*[9], and possibly *GPKOW*[10], as well as *GPATCH3*[11], have

been reported as disease-causing mutations. Considering the diverse roles of GPATCH proteins in RNA metabolism, dysfunctions and mutations in these proteins have been linked to a wide spectrum of anomalies, which prominently include skeletal abnormalities and craniofacial abnormalities[11,12], neurological disorders[13], cardiovascular conditions and angiogenesis anomalies[14], and immune system dysfunctions[15]. These craniofacial abnormalities and neurological signs overlap with symptoms seen in spliceosomopathies resulting from mutations in specific genes, such as *TXNL4A*[16], *RBM8A*[17], *SNRPB*[18], *EIF4A3*[19], *EFTUD2*[20] and *SF3B4*[21]-associated, but not in *PRPF3*[22], *PRPF31*[23], *PRPF4*[24], *PRPF6*[25], *PRPF8*[26], and *SNRNP200*[27]. The latter group of genes is associated with non-syndromic retinitis pigmentosa (RP), a condition primarily affecting rod photoreceptors[28–31]. The identification of mutations in pre-mRNA splicing genes involving defects in the splicing process as a mechanism of photoreceptor degeneration is known but most certainly underestimated.

Our findings reveal the accumulation of several spliceosome components in the retinas of the mutant mice mimicking one of the observed human variants, while the abundance of others is decreased, suggesting a role for GPATCH11 in spliceosome homeostasis, linking GPATCH11 to GPATCH-associated and spliceosome-associated diseases[2]. Accordingly, we observed substantial deregulation of splicing in the retinas of mutant mice. Furthermore, we describe dysregulated expression of genes involved in phototransduction and synaptic transmission of the visual message. Photoreceptor-connecting cilium functions were also deregulated, suggesting a potential role for GPATCH11 in cilia functions. This is further supported by the presence of GPATCH11 in ciliary basal bodies, in addition to its nuclear localization.

Our study highlights the significance of GPATCH11 in maintaining proper cellular function and its contribution to various aspects of retinal, neurological and bone function and development. However, in a broader sense, understanding the implications of GPATCH11 dysfunction and mutations elucidates the intricate roles of GPATCH11 protein in RNA metabolism and its impact on diverse clinical manifestations.

RESULTS

Biallelic *GPATCH11* mutations cause retinal dystrophy, neurodevelopmental delay, and behavioural problems with or without seizures and skeletal anomalies.

The families were initially referred to a genetic consultation for distinctive presenting symptoms, namely, severe visual dysfunction from birth (Family 1), intellectual disability (Family 2), intellectual disability with encephalopathy (Family 3) and intellectual disability with retinitis pigmentosa (Family 4). However, a review of clinical files, re-examination and/or disease progression revealed common features of progressive retinal degeneration, intellectual disability, and dysmorphic features (**Fig.1**).

The index case in Family 1 (F1:IV-2) and an elder affected sibling (F1:IV-1) manifested severe visual deficiency with nystagmus and altered electroretinography (ERG) results in the first months of life (**Table 1**). Ophthalmological evaluation in childhood revealed poor vision (visual acuity of 1/20 and hand movements during their childhood, for F1:IV-1 and F1:IV-2 respectively; **Table 1**), macular atrophy and peripheral pigmentary deposits at the fundus (**Fig. 2a**), and outer retinal thinning affecting predominantly the perimacular region at SD-OCT (**Fig. 2b**). As they grew up, the two affected individuals manifested agitation, hyperactivity, frustration, and delay in language acquisition (**Table 1**). Neuro-pediatric workup and neuroimaging revealed neurodevelopmental anomalies with normal cerebral MRI. Recently, the elder affected child manifested seizures. Their affected paternal relatives (F1:III-3 and F1:III-6; **Fig. 1**), who were born to a first-cousin marriage, were seen in a north African hospital for severe visual dysfunction and marked neurodevelopmental delay manifesting as poor language skills, mild ataxia and general movement disorganization. The eldest affected relative passed away in the second decade of life during an episode of fever (**Table 1**).

The three affected siblings of Family 2 (**Fig. 1**) were initially addressed for intellectual disability, speech and walking delay and dyspraxia. The eldest affected subject (F2:II-2) was reported to have displayed normal electroencephalography (EEG) and MRI. Before 5 years old, she manifested

visual problems, and ophthalmological examination revealed reduced visual acuity (4/10 left and right eyes, LRE), hypermetropia and flat ERG (**Table 1**). The eldest of the two affected siblings in Family 3 (F3:II-1; **Fig. 1**) presented with short stature, psychomotor retardation, encephalopathy and seizures, lack of muscular coordination for voluntary movement in early childhood, and diabetes in the second decade of life, and she died in the third decade of life of thalamic stroke. The younger affected chld (F3:II-2) presented the same clinical features (**Table 1**). She and the affected subjects from Families 1 (F1:IV-1, F1:IV-2) and 2 (F2:II-2) were examined by dysmorphologists who noted similar facial dysmorphic features consisting of enophthalmos, a short philtrum, large incisors and diastema. The affected children in Family 4 (F4:VI-2, **Fig. 1**) were reported to display night blindness since childhood but were diagnosed of retinitis pigmentosa (RP) with peripheral deposit of pigments in the second decade of life (**Fig. 2a-b**). The cognitive deficits with global developmental delay were noticed in the first years of life and dysmorphic features were reported (**Table 1**).

Analysis of exome datasets from affected individuals and their unaffected relatives identified the *GPATCH11* splice-site variant c.328+1G>T (p.?) (NM_174931.4) in homozygosity and in compound heterozygosity with the nonsense variant c.454C>T (p.Arg152*) in affected subjects from consanguineous Families 1 and 3 and nonconsanguineous Family 2, respectively (**Fig. 1**). Haplotype reconstruction using exome datasets from the three families, all of whom originate from Maghreb, identified a short (1.27 Mb) common haplotype encompassing the splice-site variant c.328+1G>T (**Supplementary Table 1**), supporting an ancient founder effect. In Family 4, originating from Europe, exome analysis identified homozygosity for another splice-site variant, c.449+1G>C (p.?) (**Fig. 1**). Consistent with autosomal recessive disease transmission, the parents of affected individuals in the four families carried these variants in single heterozygosity (**Fig. 1**).

The three variants were absent from controls in the Exome Sequencing Project, 1000 Genomes Project, and Exome Aggregation Consortium. *In silico* predictions suggest that the consensus donor

splice-site variants c.328+1G>T and c.449+1G>C cause skipping of in-frame exon 4 and out-frame exon 5, respectively.

Exome sequencing in four independent families allowed the identification of 3 different variants in *GPATCH11*.

Cells from individuals carrying the c.328+1G>T variant produce a mutant GPATCH11 protein lacking the internal portion of the GPATCH domain

To validate the *in silico* predictions described above, we next performed RT-PCR analysis of RNA isolated from the fibroblasts of the affected siblings in Family 1 carrying the c.328+1G>T variant in homozygosity F1:IV-1 (P1) and F1:IV-2 (P2). This analysis showed skipping of the 42 bp-long exon 4, encoding a 14 amino acid-long sequence (aa 96–109) that links the helix and loop braces of the GPATCH domain encoded by the adjacent exons[3] (**Fig. 3a, c, d**).

Western blot analysis using an antibody specific to a sequence downstream of the GPATCH domain detected a unique band in patient cells. The molecular weight corresponding to this band was lower than that of the wild-type counterpart in control fibroblasts and consistent with a 14-residue deletion (**Fig. 3e, f**). Analysis of blood RNA from the affected subject of Family 2 (F2:II-2; P5) showed apparent homozygosity for the variant c.328+1G>T, suggesting nonsense-mediated mRNA decay (NMD) of the mRNA transcribed for the trans allele carrying the c.454C>T (p.Arg152*) nonsense variant (**Fig. 3b**).

We showed that the variant c.328+1G>T produces a mutant GPATCH11 protein and the variant c.454C>T alters the formation of the GPATCH11 protein due to NMD.

GPATCH11 is a nuclear and basal body protein, and alterations in the GPATCH domain and CCDC do not affect its subcellular localization

To analyse the expression of GPATCH11 in cellular models, confocal microscopy analysis of immune-stained control fibroblasts fixed by methanol was performed and confirmed the supposed nuclear localization of the GPATCH11 protein[1] and revealed its basal body localization, suggesting

a role in cilia homeostasis. Of note, basal body staining was decreased in PFA-fixed cells (**Supplementary Fig. 1a**). To refine the subcellular localization of the protein, fibroblasts were immunostained for GPATCH11 together with the cilium axoneme, centrosome and centrosome linker using acetylated-α-tubulin, centrin-3 and rootletin antibodies, respectively. Stimulated-emission-depletion (STED) microscopy detected GPATCH11 in the centrosome linker region anchoring numerous coiled-coil proteins to the proximal end of centrioles, including notable colocalization with rootletin (**Fig. 4**). Immunocytochemistry of the fibroblasts from the siblings F1:IV-1 and F1:IV-2 showed that the c.328+1G>T mutant protein displayed a normal nuclear distribution and abundance, despite alteration of the GPATCH domain, and a normal basal body localization (**Supplementary Fig. 1b, Fig. 4**). To assess whether the coiled-coil domain-containing (CCDC) domain of GPATCH11 participates in its basal body localization, a homozygous in-frame deletion of the coding sequence in exon 3 was generated by the CRISPR–Cas9 technique in the hTERT-RPE1 line, and the presence of shortened mRNA and protein was confirmed by RT–PCR and Western blot analysis, respectively (**Supplementary Fig. 2a, b**). The subcellular localization of the CCDC-deleted GPATCH11 isoform in these cells was comparable to that of its wild-type counterpart in control hTERT-RPE1 cells.

Even if GPATCH11 contains a CCDC domain, which can act in the ciliary function[32,33], and the immunostaining of GPATCH11 showed a specific localization in the fibres between the two centrioles at the base of the primary cilia, our experiments did not show any absence nor mis-localization of GPATCH11 protein in affected patients' fibroblasts nor CCDC-deleted hTERT-RPE1.

Cilia biogenesis and function are not affected by the c.328+1G>T variant

To investigate the impact of c. 328+1G>T variant we studied cilia formation in F1:IV-1 and F1:IV-2 fibroblasts and in CCDC-deleted hTERT-RPE1 cells by labelling the primary cilium basal body and axoneme with pericentrin and acetylated α-tubulin antibodies, respectively.

The abundance of ciliated cells and mean axonemal length were not significantly different from those in controls, indicating that neither alteration of the GPATCH domain nor deletion of the CCDC domain affected ciliation (**Supplementary Fig. 2c, d; Supplementary Fig. 3a, b**). The Sonic Hedgehog (SHH) pathway, a critical developmental signalling pathway functionally linked to primary cilia in vertebrates[34], was analysed in the F1:IV-1, F1:IV-2 and controls fibroblasts, as determined by RT–qPCR analysis of *SMO*, *GLI1*, *GLI2* and *PTCH1* expression upon smoothened agonist (SAG)-mediated activation, which revealed no significant difference between groups (**Supplementary Fig. 4**).

The ciliary structure and SHH pathway are not impacted by the c. 328+1G>T variant or the CCDC-deleted domain.

Deletion of the central part of the GPATCH domain from *Gpatch11* in mice causes rapid retinal degeneration, memory dysfunction and infertility

As described previously, the c.328+1G>T variant is responsible of the skipping of exon 4 in human and to model it from the mRNA *in vivo*, the corresponding exon in the mouse genome (exon 5) was ablated by the 9-mediated double-cut exon deletion technique using intronic guides located on either side of the target exon (**Fig. 5a**). Mice carrying exon 5 deletion in homozygosity (*Gpatch11*$^{\Delta 5/\Delta 5}$) were viable and developed normally (the aspect, the weight and their daily comportment are normal). However, males were completely infertile and exhibited smaller than normal and empty testes (**Supplementary Fig. 5**). In contrast, females exhibited normal fertility.

The expected exclusion of exon 5 from the mRNA and presence of a shortened protein product in the retina of *Gpatch11*$^{\Delta 5/\Delta 5}$ mice were verified by RT–PCR and Sanger sequencing and Western blot analysis, respectively (**Fig. 5b, c**). The abundance of the mutant protein in *Gpatch11*$^{\Delta 5/\Delta 5}$ mouse retina and the wild-type counterpart in the retinas of the *Gpatch11*$^{WT/WT}$ control littermates were not significantly different according to Western blot analyses (**Fig. 5c, d**).

Considering the retinal defects found in human patients, we characterized the retinal phenotype of *Gpatch11*$^{\Delta 5/\Delta 5}$ mice through longitudinal functional and structural studies based on

electroretinography (ERG), histology and immunohistochemistry analysis using *Gpatch11*$^{WT/WT}$ littermates as controls. *Gpatch11*$^{\Delta5/\Delta5}$ pups displayed normally developed and layered retinas at the age of 15 days (1 day from eye opening), with normal rod-specific but moderately reduced cone-specific responses to light stimulation (**Fig. 6a, b**). Rod and cone responses declined progressively as the mice aged, the photoreceptor nuclear layer became slimmer, and the rhodopsin and S and M cone opsin immunostaining weakened, supporting the gradual loss of rod and cone photoreceptors, respectively. By the age of 3 months, approximately half of the photoreceptor nuclei were lost, and the depletion of the opsin content was even more severe in rods and, more importantly, in cone outer segments, where light is captured and transduced (**Supplementary Fig. 6**). Consistently, ERG responses were highly altered. At the age of 6 months, ERG responses were flat, and photoreceptor nuclei were almost completely lost (**Fig. 6a, b**).

Presence of neurological impairment in the four families of this study encouraged us to investigate the presence of neurological deficits in *Gpatch11*$^{\Delta5/\Delta5}$ mice in 1-month-old *Gpatch11*$^{\Delta5/\Delta5}$ mice (the opsin content in photoreceptors and ERG responses consistent with vision) and *Gpatch11*$^{WT/WT}$ littermates. The mice were subjected to novel object recognition (NOR), contextual fear conditioning (CFC), and Morris water maze (MWM) tests and the open field test (OFT) to assess episodic memory, associative memory, spatial memory and anxiety-like behaviour, respectively. Both male and female *Gpatch11*$^{\Delta5/\Delta5}$ mice demonstrated an inability to recognize novel objects in NOR experiment (**Fig. 7a**) and significantly shorter context-elicited freezing times during the training and testing phases in CFC experiment (**Fig. 7b**), supporting defective episodic and associative memory, respectively. In contrast, spatial memory and anxiety-like behaviour were not significantly different from those of controls (**Supplementary Fig. 7a, b**).

Exploration of the retina and the behaviour of our mice model confirmed presence of the major symptoms described in the human syndrome.

Wild-type and mutant *Gpatch11* mRNA and protein products are detected in all retinal cells and the GPATCH11 Δ5 protein accumulates in the nucleus of mouse retinal cells

The retinal expression of wild-type and mutant *Gpatch11* mRNA isoforms was examined in eye sections by RNAscope using a probe targeting the two isoforms. Wild-type mRNA in *Gpatch11*$^{WT/WT}$ mice and the mutant counterpart in *Gpatch11*$^{\Delta 5/\Delta 5}$ littermates were detected and quantified with no significant difference in all retinal layers (**Fig. 6c, Supplementary Fig. 8**).

Immunostaining of the wild-type and mutant protein isoforms in retinal sections using the GPATCH11 antibody showed the presence of the protein in all retinal nuclei (**Fig. 6d**). Of note, we show that in photoreceptor nuclei the protein displayed a perinuclear distribution, contrasting with the nuclear localization observed in other retinal cells. This inverted pattern in photoreceptor nuclei has been previously ascribed to a peculiar nuclear architecture in nocturnal mammalian photoreceptors[35].

Interestingly, quantification of wild-type and mutant mRNA and protein isoforms over a 6-month period revealed comparable mRNA abundance but increased GPATCH11 immunostaining in *Gpatch11*$^{\Delta 5/\Delta 5}$mice compared to *Gpatch11*$^{WT/WT}$ littermates, suggesting accumulation of the mutant protein (**Fig. 6d**, **e**). Because protein degradation is often modulated by ubiquitination or other post-translational modifications, we investigated posttranslational modifications through proteomic Nextprot database (https://www.nextprot.org/entry/NX_Q8N954/proteomics?isoform=NX_Q8N954-1), which did not identify any modification of the 96-109 amino-acid deleted sequence whose loss could account for altered protein degradation.

Loss of the internal portion of the GPATCH domain of GPATCH11 affects the splicing and expression of genes related to visual perception in the retina.

Whether the loss of the internal portion of the GPATCH domain of GPATCH11 affected gene expression and/or splicing was investigated by comparing the transcriptomes of retina from 15-day-old *Gpatch11*$^{\Delta 5/\Delta 5}$ mice and *Gpatch11*$^{WT/WT}$ littermates. This resulted in the identification of 160 differentially expressed genes (DEGs, **Fig. 8a**). GO enrichment analysis indicated the enrichment of genes related to photoreceptors, nonmotile cilium and response to light stimuli (phototransduction

cascade and visual signal propagation) in the *Gpatch11*$^{\Delta 5/\Delta 5}$ mice compared to wild-type littermates (**Fig. 8b-d**). Circular visualization based on the calculated z scores indicated that most of the pathways, for which Heatmaps are shown, were enriched in downregulated DEGs in particular (**Fig. 8b, d)**. Neither inflammation nor cell death pathways were deregulated upon loss of the internal portion of the GPATCH domain, supporting the view that the transcriptome changes at day 15 were primary consequences of alterations in *Gpatch11*.

Differentially expressed isoform analysis identified 299 splicing events, mainly reduced skipped exons (SEs) and increased mutually exclusive exons (MXEs), in 178 genes (**Fig. 9a**). Differentially spliced genes (DSGs) included genes related to the photoreceptor response to light (phototransduction and synaptic transmission of the visual message), photoreceptor connecting the cilium assembly, protein homeostasis, mitochondria, nucleic acid (chromatin, DNA and RNA) binding and the regulation of RNA splicing (**Fig. 9b**). Most of these genes showed no differences in expression in the DEG analysis, suggesting that splicing events were not responsible for the changes in expression (**Fig. 9c**). Specifically, the intersection of DEGs and DSGs identified only 12 genes: *Arr3, Asap3, Cabp4, Ccnl2, Dhrs3, Lbhd1, Mpp4, Pex5l, Pitpnm3, Tulp1, Unc13b,* and *Vtn,* all of which, except *Asap3,* were downregulated upon loss of the internal portion of the GPATCH domain (**Fig. 9c, d**). The splicing events in *Asap3, Cabp4, Lbhd1, Pitpnm3, Tulp1,* and *Vtn* are expected to introduce premature termination codons in the mRNA and trigger nonsense-mediated mRNA decay. In contrast, analysis of the splicing events in *Arr3,* which is reported as an example, suggests that mis-splicing is not the main driver of expression deregulation in this gene (**Fig. 9e-g**) as well as in *Ccnl2, Dhrs3, Mpp4, Pex5l,* and *Unc13b* (**Supplementary Table 9**).

Loss of the internal portion of the GPATCH domain of GPATCH11 affects the expression of proteins related to visual perception and to spliceosome in the retina.

Mass spectrometry (MS) analysis of total-retina lysates from 21-day-old *Gpatch11*$^{\Delta 5/\Delta 5}$ mice and *Gpatch11*$^{WT/WT}$ littermates detected 150 proteins in the *Gpatch11*$^{\Delta 5/\Delta 5}$ mice with either decreased (n = 63/150) or increased (n = 87/150) abundance (**Fig. 10a**). The downregulated proteins were

involved in several of the pathways identified by transcriptome analysis, namely, the response to light stimulation, synaptic transmission of the visual message, mitochondria, cilia, protein homeostasis and RNA binding and splicing (**Fig. 10a**). Nine of these proteins were encoded by downregulated DEGs, namely, *Arr3, Cabp4, Eml3, Lhbd1, Nxnl2, Mpp4, Pex5l, Smad11*, and *Tulp1* (**Fig. 10b**). Among them, *Tulp1* and *Cabp4* have been found to be associated with LCA15[36] and cone-rod synaptic disorder (CRSD)[37], respectively, and *Arr3* encodes cone arrestin, a key player in the light-dark adaptation of cones[38]. Among the 54 remaining proteins in this group, 12/54 are involved in RNA processing, and 8/12 are involved in splicing and present in spliceosomal complexes: Snrpa (in U1), Ddx17 and Fus (in Complex A), Ddx42 and Dnajc8 (in U2), Fam32a (in Complex C), Hnrnpm and Hnrnph3 (in hnRNP) (**Fig. 10a**). Among the 87 proteins upregulated in *Gpatch11*$^{\Delta 5/\Delta 5}$ mice, 3 were encoded by genes with increased mRNA abundance (e.g., *Fgf2, Apobec2, Txnl4a;* **Fig. 10c, d**). The remaining 84 proteins are encoded by genes whose RNA expression was similar in *Gpatch11*$^{\Delta 5/\Delta 5}$ mice and *Gpatch11*$^{WT/WT}$ littermates. GPATCH11 was included in this list of these 84 upregulated proteins, consistent with immunostaining studies showing accumulation of the mutant product in the retinas of *Gpatch11*$^{\Delta 5/\Delta 5}$ mice (**Fig. 10e**).

GO enrichment analysis showed that RNA-binding proteins represented more than half of the overrepresented proteins, including Prpf39 (in U1), Cherp, a G-patch domain-containing protein that is known to play a role in pre-mRNA splicing; Sf3b2; Sf3b3 and Sf3b4 (in U2); Prpf3 (in U4/U6); Ddx23; Prpf6 and Txnl4 (in U5); Prpf40a and Rbm39 (in Complex A); Dhx16 (in 1st Step); SrfS6 (in SR); and Lsm14a (in LSm) (**Fig. 10a**). Several down- and upregulated spliceosomal proteins were reported to be involved in nonretinal (Txnl4, Sf3b4 and Fus) and retinal-specific (Prpf3 and PrPf6) spliceosomopathies, respectively[2].

Together, these results are consistent with a role for GPATCH11 in splicing and the transcriptional regulation of genes important to retinal function and possibly other functions through the deregulation of nonretinal spliceosomopathy-associated genes (Txnl4, Sf3b1, Sf3b4 and Fus).

Furthermore, these results support a role for GPATCH11 in primary cilium assembly and/or function despite the lack of cilia homeostasis defects in fibroblasts from patients.

DISCUSSION

In recent years, G-patch domain-containing proteins have emerged as pivotal regulators of RNA metabolism, assuming diverse roles that encompass pre-mRNA splicing and transcriptional regulation. GPATCH11 is one of the lesser-known members of this family[3]. Here, we report the identification of biallelic *GPATCH11* mutations in four families; three of these share a donor/acceptor splice site variant, supporting a founder effect. The families were initially referred to a genetic consultation for distinctive presenting symptoms. The fact that the four families were seen for distinct major symptoms reflects the variable severity of retinal disease, which represented the primary symptom in Family 1. Various clinical features, including seizure, short stature, and diabetes, were associated with this retinal disease. These symptoms were reported in the eldest individuals available for examination in the series and could not yet be evaluated in the youngest affected subjects. Additional symptoms associated with *GPATCH11* mutations are possible considering the sudden death in teenage and young adulthood of two unrelated individuals from a fever episode, which could indicate defective responses to pathogens, and thalamic stroke. Therefore, specific explorations of immune responses in individuals carrying *GPATCH11* disease-causing variations certainly deserves consideration. Furthermore, considering the severe phenotype affecting male gonads in the mice, the fertility of affected male individuals in Families 1, 2 and 4 would be interesting to examine.

Interestingly, genotype–phenotype correlation analysis revealed no major difference in disease presentation between individuals homozygous for the splice site variant (Families 1 and 3) and subjects carrying a nonsense mutation in trans with the splice site variant (Family 2) or the individual carrying a frameshifting splice site mutation (Family 4). Although the frameshift is predicted to lead to absence of the protein, partial bypassing of protein truncation through self-correcting mechanisms cannot be excluded. In order to validate this hypothesis, a skin biopsy-derived fibroblasts of the patient of the Family 4 would be interesting to analyse.

By analysing the loss of integrity of the G-patch domain of GPATCH11 in mice, we show that it recapitulates the retinal and neurological abnormalities observed in the affected individuals. In contrast, we found no remarkable skeletal, metabolic, or immune involvement, and male infertility was observed, which might be explored in the future.

In our study using mutant mice, we observed that the retina developed and layered normally and responded appropriately to light at eye opening. These findings, combined with the variability in disease onset and severity in affected individuals from our cohort, suggest that GPATCH11 may play a crucial role in the function and maintenance of photoreceptor cells but not in the initial development of the retina. Nevertheless, the origin of the neurological anomalies remains ambiguous, as it is unclear whether they result from developmental issues or arise later in life, particularly considering that neurological problems in the affected siblings of Family 1 manifested rather late in childhood. Conversely, the presence of dysmorphic features provides evidence that GPATCH11 serves essential functions both during early tissue development and in the subsequent maintenance of proper function. The mutations in GPATCH proteins have been linked to a wide spectrum of anomalies, which prominently include skeletal abnormalities and craniofacial abnormalities as shown in *GPATCH3* and *GPKOW*[10,11].

Interestingly, both gene expression and splicing were significantly deregulated in the retinas of the mutant mice. Reduced mRNA abundance could be a direct outcome of splicing aberrations leading to the degradation of mRNA isoforms containing premature termination codons or an indirect consequence of splicing dysregulation. However, we identified certain genes, such as *Arr3*, that exhibited reduced mRNA levels and abnormal splicing without alteration of the reading frame. This finding suggests that GPATCH11, as other G-patch domain-containing proteins such as SON and ZGPAT[3], is involved in both pre-mRNA splicing and transcriptional regulation.

Broader functions of GPATCH11 in RNA homeostasis, extending beyond pre-mRNA splicing, may elucidate its association with a more diverse spectrum of clinical phenotypes than are present in traditional spliceosomopathies. Additionally, we observed GPATCH11 accumulation, which might

result from the loss of residues essential for its degradation or be an indirect consequence of broader deregulation[39]. This accumulation led to the massive deregulation of RNA-binding protein complexes, most of which also showed increased accumulation. Many of these proteins are known to be involved in the spliceosome, and mutations in their components have been linked to nonretinal and retinal-specific spliceosomopathies, collectively presenting many, although not all, of the clinical features observed in patients carrying mutations in *GPATCH11*.

Through our subcellular localization analysis of GPATCH11, we found that GPATCH11 is not only present in the nucleus[1] but also, for the first time, localizes to the linker region that anchors various coiled-coil proteins to the proximal end of centrioles. The integrity of this region is known to be vital for maintaining centrosome cohesion, and mutations in linker proteins have been implicated in ciliopathies, such as primary microcephaly (MCPH), which is characterized by small brain size and intellectual disability, as well as other conditions, such as microphthalmia, Usher syndrome, Alström syndrome, myopic maculopathy, Alzheimer's disease, microdeletion syndrome, and neuroblastoma[40]. However, it has not been demonstrated whether *GPATCH11* mutations compromise the integrity of the centriolar linker region, as our experiment showed that fibroblasts from patients carrying an exon 4-skipping site mutation displayed unremarkable cilia formation. Nevertheless, in mice, we did observe significant deregulation of gene expression and splicing of non motile cilium-genes encoding which are components of the photoreceptor connecting cilium and outer segments, which could be related to the retinal phenotype. The potential impact of *GPATCH11* mutations on cilia and their contribution to various symptoms, including retinal dystrophy, bone anomalies, male sterility, and neurological problems, which are all common features of ciliopathies, is currently unknown. Further research is warranted to better understand the intricate roles of GPATCH11 in ciliary function and its implications for the diverse clinical manifestations associated with *GPATCH11* mutations, focusing on hTERT-RPE1 cell line and the CRISPR-Cas9 technology in order to delete the gene.

In conclusion, our findings reveal that *GPATCH11* mutations give rise to a complex syndrome that is characterized by clinical features that overlap with both nonretinal and retinal-specific spliceosomopathies, as well as other GPATCH-associated phenotypes. Notably, the loss of GPATCH11 integrity exerts a substantial impact on pre-mRNA splicing, affecting several core components of the spliceosome, several of which have been associated with retinal, haematopoietic lineage, craniofacial skeleton, spinal cord, and limb abnormalities. Additionally, our analysis in a mouse model indicates that GPATCH11 is involved in the regulation of gene transcription, specifically for genes essential in the function and maintenance of rod and cone photoreceptor cells and brain development. These findings collectively elucidated the diverse and intricate roles of GPATCH11 in RNA metabolism and possibly cilia metabolism, highlighting its significance in various physiological processes and providing valuable insights into the molecular basis of the observed clinical features associated with *GPATCH11* mutations.

MATERIAL AND METHODS

Families

Subjects

The study involved ten affected subjects (six females and four males) with retinal dystrophy from four families, three of which are multiplex and consanguineous (Family 1 and 3), one is a multiplex and non-consanguineous (Family 2) and one is non-multiplex and consanguineous (Family 4). Family 1 was referred for early-onset retinal dystrophy with neurodevelopmental delay and consists in an inbred north African pedigree with several loops of consanguinity and four affected individuals (two affected children and two affected relatives, one of whom deceased). Family 2 was referred for intellectual disability and comprises three affected individuals, one sibling and their two unrelated parents from North Africa. Family 3 includes two affected subjects born to consanguineous parents from North Africa and had been reported previously[41]. Family 4 includes one affected individual from Europe. All individuals or legal representatives consented with the study, which received approval from the institutional review boards Comité de Protection des Personnes Ile de France II (Necker), Reference Center for Congenital Abnormalities and Malformative Syndromes in Dijon (France) and *Orphonomix units* for genetic testing, located in several hospitals in France. All participating individuals of Family 4 signed a consent form, in agreement with the Declaration of Helsinki and the ARVO statement on human subjects. Genomic DNA was extracted from peripheral blood by standard procedures.

Gene-panel testing and exome sequencing (ES)

The *GPATCH11* variants were identified by IRD panel followed by whole-exome sequencing (WES) in research settings (Family 1) or in diagnostic settings in clinical laboratories and direct WES (Families 2, 3 and 4). Genomic DNA libraries were generated from DNA (F1:III-6; F1:IV-1; F1:IV-2; F2:I-1, F2:I-2; F2:II-2) sheared with a Covaris S2 Ultrasonicator via SureSelect^{XT} Library Prep Kit (Agilent). Regions of interest (ROIs) were captured with the SureSelect All Exon V5 kit (Agilent) and sequenced on an Illumina HiSeq2500 HT system (Illumina). Data analysis was

performed with a homemade pipeline (POLYWEB)13 created by the *Imagine* Institute Bioinformatics core facilities of Paris Descartes University. The following algorithms were used to predict the consequences of variants identified with WES: PolyPhen-2, SIFT and MutationTaster. Allele frequencies were evaluated via the gnomAD population database. The Segregation analysis was performed in parents of affected individuals via Sanger sequencing. It demonstrated biparental transmission in all the affected individuals. WES and variant filtration in Family 3 has been previously described[41]. WES analysis of individuals from Family 4 was performed at Otogenetics (Norcross, GA) using Agilent V4 (Santa Clara, CA, USA) and Illumina HiSeq 2000 with 30 × coverage. Variants in *GPATCH11* were annotated using GRCh38 reference genome and GenBanck transcript (NM_174931.4).

Culture of skin-derived fibroblasts and hTERT-RPE1

Skin biopsies were obtained from two probands of Family 1, namely F1:IV-1 (P1) and F1:IV-2 (P2) and four control individuals (C1, C2, C3 and C4) from dermal fibroblast culture. Fibroblasts and hTERT-RPE1 were maintained at 37°C, 5% CO_2 in Opti-MEM GlutaMAX I medium and Dulbecco modified eagle's medium (DMEM F-12), respectively, supplemented with 10% FBS (Fetal Bovine Serum), 1% streptomycin/penicillin (Thermo Fisher Scientific).

In silico analysis of variants on splicing

The effect on the splicing of the *GPATCH11* variants identified in the F1:IV-1 and F1:IV-2 affected individuals was analyzed *in silico* using Alamut visual, a software which examines variations of human genome and combines data with high-quality missense and splicing predictors.

gDNA preparation

Fibroblasts and hTERT-RPE1 were trypsinated with Trypsine-EDTA 0,05% (Thermo Fisher Scientific). Cell pellets were then incubated for 7 min at 65°C followed by 2 min at 98°C with Quickextract (Lucigen).

RNA preparation and cDNA synthesis

Cell lysis and RNA extraction from affected individuals and control fibroblasts were performed using the RNAeasy Mini kit, according to the manufacturer's protocol (Qiagen). mRNA (500 ng) was converted and synthetized in cDNA using Verso cDNA kit, according to the manufacturer's instructions (Thermo Fisher Scientific).

RNA for the patient (F2:II2; P5) was extracted from peripheral blood samples following PAXgene Blood RNA Kit (50) v2 protocol (Qiagen).

PCR and RT-PCR

gDNA and cDNA were PCR amplified in a buffer containing 10 µM of specific primers to *GPATCH11* (**Supplementary Table 2, 3**), dNTP (100 µM), $MgCl_2$ (25 mM) and 0,5 units of DNA polymerase (GoTaq, Promega).

Sanger Sequencing

PCR products were subjected to a sequencing reaction using BigDye Terminator v3.1 Cycle Sequencing Kit, according to the recommendation of manufacturer (Applied Biosystem). Purified products were sequenced on a 350xL Genetic Analyzer ABI and analyzed using the Sequencing Analysis v5.2 sequence software.

Western blot analysis

P1, P2 and controls fibroblasts and hTERT-RPE1 lines were treated with radioimmunoprecipitation assay (RIPA) lysis buffer (Thermo Fisher Scientific) for 1 hour in ice and sonicated (1 pulse of 20 seconds at 70% of amplitude). The concentration of proteins was quantified with the Bradford assay. 50 µg of proteins were prepared in a solution containing LDS 1X and 10% of 2-Mercaptoethanol. The lysates were heated at 95°C for 8 minutes and denatured proteins were separated by electrophoresis in a Mini-PROTEAN TGX Stain-Free Any kD gel, according to the recommendations of the supplier (BioRad). Separated proteins were transferred onto a PVDF membrane using the TransBlot Turbo Mini-size PVDF membrane system (BioRad). The membrane was incubated overnight at 4°C with the polyclonal rabbit Anti-GPATCH11 primary antibody (also named Anti-CCDC75; Abcam) (1/4000). The antigenic sequence of the antibody corresponds to

amino acids 111-192 of human GPATCH11. The goat anti-rabbit IgG-HRP was used as secondary antibody (Invitrogen; 1/4000). Blots were revealed using Clarity Western ECL Substrate (BioRad) and the ChemiDoc XRS+ Imagin System (BioRad). Western blot images were acquired and analyzed with Image Lab software 3.0.1 build 18 (Bio-Rad). The abundance of GPATCH11 relative to ß-actin was estimated in each cell line by densitometry using Image Lab software (BioRad). A Student's test was performed.

Immunocytochemistry analysis

Cells were seeded at $2x10^5$ cells/well on glass coverslip in 12-well or in 6-well plates for 24 hours. After 24 hours, the cells were fixed with PFA or Methanol and immunostained with polyclonal rabbit anti-gpatch11 (1:200; Abcam) and nuclei were stained using DAPI (Invitrogen). The images were scanned with Spinning Disk Zeiss microscope (Zeiss) using a 40x/1.3 Oil objective. For the analysis, we first used a machine learning with Ilastik[42] (v1.3.3post3) for GPATCH11 and Cellpose2[43] for nucleus detection. Then we measured the mean intensity of nuclear and cytoplasmic GPATCH11 with Fiji[44] macro. Image figures were made through FigureJ[45].

To analyze the cilia structure, the cells were incubated in serum free medium for other 24 hours, to promote ciliation. To analyze ciliation, the cells were fixed in Methanol and immunostained as using mouse monoclonal anti-α-tubulin acetylated (1:2000, Sigma-Aldrich), rabbit anti-pericentrin (1:5000, Abcam), polyclonal rabbit anti-gpatch11 (1:200; Abcam), anti-rootletin (1:200, Santa Cruz Biotechnology), anti-centrine3 (1:200, Abnova), primary antibodies, and Alexa-Flour 488-conjugated donkey anti-mouse, 555-conjugated donkey anti-rabbit (1:1000; Life Technologies), 514-conjugated goat anti-mouse and ATTO 550-conjugated goat anti-rabbit secondary antibodies (1:700; Life Technologies and Sigma-Aldrich). Nuclei were stained using DAPI (Invitrogen). The images were scanned with Spinning Disk Zeiss microscope (Zeiss) using a 40x/1.3 Oil objective and with Confocal Leica SP8 gSTED (Leica). The final images were analyzed with Fiji[44], through which the number of ciliated cells and the length of cilia were counted. The mean of ciliated cells was calculated from three independent experiments (>100 cells for each cell line). One-way

ANOVA on GraphPad was used for statistical analysis of ciliation. The significance was given by Sidak's test and the comparison of values is considered highly significant when p-value<0,05.

RT-qPCR analysis

P1, P2 and controls fibroblasts were serum-starved for 48 h, and either exposed to a smoothened agonist (SAG, 100 nM) or negative control for 24 h. RNA was extracted separately for each condition and converted into cDNA. RT-qPCR was performed with the cDNA of control and patient fibroblasts. cDNAs (5 µl of a 1:25 dilution in nuclease-free water) were subjected to real-time PCR amplification in a Mastercycle (Eppendorf) using the Sybr Green PCR Master Mix kit (Applied Biosystem), following the supplier protocol. For each cDNA sample, the mean of cycle threshold (Ct) values was calculated from triplicates (SD <0.5 Ct). SMO, GLI1, GLI2, PTCH1 expression levels were normalized by using the relation ΔCt, thanks to the housekeeping genes GUSB and HPRT1 (**Supplementary Table 4**).

Generation of genome-edited hTERT-RPE1

hTERT-RPE1 lines was generated via a CRISPR-Cas9 genome editing strategy, as previously described[46]. Briefly, two sgRNAs targeting the exon 3 of GPATCH11 was designed using CRISPOR software (http://crispor.tefor.net/) (**Supplementary Table 5**) and selected by on-target and off-target score. The CRISPR-Cas9/sgRNA RNP complex was formed using a crRNA XT recognition domain annealed with ATTO550 tagged tracrRNA transactivator domain (IDT, #1075928) and then complexed with the S. pyogenes HiFi Cas9 Nuclease V3 nuclease (IDT, #1081061), following IDT recommendations. The RNP complexes were mixed with hTERT-RPE1 cells and transferred into a 16-well reaction cuvette of the 4D-Nucleofector System (Lonza). Cells were nucleofected using program CA137 on the 4D-Nucleofector system. After 24 h, Atto550 positive cells were sorted using flow cytometry (BD FACS ARIA II SORP, BD Biosciences) and transferred into 96-well plates for single-cell selection.

Animals

Generation of Gpatch11$^{\Delta 5/\Delta 5}$ mice model

CRISPR-Cas9 mutagenesis has been performed to generate a mutant mouse model encoding a *Gpatch11* mRNA missing the exon 5, with the help of the Transgenesis platform of the LEAT Facility at the Imagine Institut. The two guide RNAs (sgRNAs) targeting the introns encompassing the exon 5 of the gene *Gpacth11* were designed and selected thanks to CRISPOR program (http://crispor.tefor.net/) (**Supplementary Table 6**). C57BL/6J female mice (4 weeks old) were super ovulated by intraperitoneal injection of 5 IU PMSG (SYNCRO-PART® PMSG 600 UI, Ceva) followed by 5 IU hCG (Chorulon 1500 UI, Intervet) at an interval of 46–48 h and mated with C57BL/6J male mice. The next day, zygotes were collected from the oviducts and exposed to hyaluronidase (H3884, Sigma-Aldrich) to remove the cumulus cells and then placed in M2 medium (M7167, Sigma-Aldrich) into a CO2 incubator (5% CO2, 37 °C). Recombinant Cas9 protein, tracrRNA and crRNA were purchased from Integrated DNA technologies. 200ng/uL tracrRNA:crRNA duplex and 2 µM Cas9 protein were mixed and allowed to form an active ribonucleoprotein (RNP) complex for 10 minutes at room temperature were added in Opti-MEM buffer (31985062, ThermoFisher Scientific). A glass chamber with 1 mm gap platinum plate electrodes (CUY5001P1–1.5, NEPA GENECo. Ltd) was filled with 7 µL of RNPs containing medium. Several batches of zygotes were aligned between the electrodes, and repeated pulses of electroporations were delivered by the NEPA21 electroporator, allowing the RNPs/ssODN to enter the zygotes. Surviving zygotes were placed in KSOM medium (MR-106-D, Merck-Millipore) and cultured overnight to two-cell stage and then transferred into the oviduct of B6CBAF1 pseudo-pregnant females. The newborn mice were genotyped from genomic and exonic DNA, by PCR amplification followed by Sanger sequencing using appropriate primers (**Supplementary Table 7**), combined with TIDE analysis (https://tide-calculator.nki.nl/; data not shown). F0 founders' mice carrying the deletion have been crossed with C57BL/6J wildtype animals to remove potential off targets. Backcrossed heterozygous mice were then intercrossezd to obtain homozygous mice carrying the GRCm39 chr17:1779147497-1779147632 deletion. Wildtypes (*Gpatch11*$^{WT/WT}$) C57BL/6J mice were used as reference in all analyses. Animal procedures were performed with

approval from the French Ministry of Research, in compliance with the French Animal Care and Use Committee from the Paris Descartes University (APAFIS#31460) and in accordance with ethical principles in the LEAT Facility of Imagine Institute.

Western Blot analysis

Retina was extracted from 15-day-old *Gpatch11*$^{\Delta 5/\Delta 5}$ and *Gpatch11*$^{WT/WT}$ mice and the tissues were lysed with RIPA-PIC (Sigma-Aldrich) and proceed further for proteomic analysis as described previously.

Electroretinographic analysis of mice model

Gpatch11$^{WT/WT}$ and *Gpatch11*$^{\Delta 5/\Delta 5}$ mice of age of 15-days to 6-months were analyzed by electroretinography. They were left in the dark the day before ERG recordings, which were performed using Celeris (Diagnosys LLC). For the experiment, the animals were weighted and lightly anesthetized by intramuscular injection of Ketamine 100 mg/mL (120 mg/kg) (Vibrac France) and Rompun 2% xylazine (16 mg/kg) (Bayer). During the procedure, the mice were kept at 37°C over the platform. The pupils were dilated using tropicamide (Mydriaticum 2 mg/0.4 mL) and phenylephrine (Néosynéphrine Faure 5%, Europhta) and the eyes were wet by a drop of contact sterile gel (LacriGel, Lacryvisc gel ophthalmic, Novartis) to ensure corneal protection and to maintain electrical contact. One electrode was placed onto each cornea and miniature subdermal electrodes were inserted at the base of mice tail to record. To assess rod function, dark-adapted mice were subjected to 0.01, 0.1, 1 and 3 cd.s/m2 light stimulations. To record cone function, mice were light adapted (*i.e.* rods were saturated/bleached for 8 minutes) and subsequently subjected to 3 and 10 cd.s/m2 light stimulation. The *a* and *B* waves (which measure the photoreceptors and bipolar cells activates respectively) amplitudes were analyzed with the software Diagnosys. The statistical analysis was ran using GraphPad software according to the significance of the post hoc Sidak's test (two-way ANOVA): the significance of the difference in *a*-wave and *B*-wave amplitudes between *Gpatch11*$^{\Delta 5/\Delta 5}$ mice models and same age *Gpatch11*$^{WT/WT}$ mice was determined.

Histological sections of mice organs

Mice tissues were obtained after euthanasia by cervical dislocation performed according to protocol guidelines for animal safety. Eyes of 15-days to 6-months old *Gpatch11*$^{WT/WT}$ and *Gpatch11*$^{\Delta 5/\Delta 5}$ were sampled, fixed in Paraformaldehyde 4% (PFA 4%) and kept at 4°C, overnight. The samples were washed four times with PBS 1X, dehydrated in ethanol gradient and embedded in paraffin. Tissue sections (5 µm) were stained using Hematoxylin-Eosin (HE; Histological platform, Institut Necker Enfants Malades) and imaged using a NanoZoomer S210 microscope (Hamamatsu). Images were analyzed using the software NDP.view software (Hamamatsu). The thickness of the retina outer (ONL) and inner (INL) nuclear layers of the retina were plotted with respect to the distance (0.25, 0.50, 0.75, 1, 1.25, 1.50, 1.75, 2 mm) from the optic nerve. Three mice from each group were included in the analysis. The ONL thickness of *Gpatch11*$^{\Delta 5/\Delta 5}$ mice models was compared with that of the same age *Gpatch11*$^{WT/WT}$ mice by two-way ANOVA with post hoc Sidak's test.

RNAScope analysis

RNA Scope RNAscope was performed on mice eyes, following RNAs using Advanced Cell Diagnostics (ACD) probes specific for *Gpatch11*, excluding the region of the exon 5 (Nucleotide Accession# (e.g.NM_000942.4)/miRbase(e.g.MIMAT0009197)/genome co-ordinates(e.g.hg38_chr9:84669778-85027070strand=+): https://www.ncbi.nlm.nih.gov/gene/53951). RNAscope probes were designed by ACD. Images were recorded on Spinning Disk Zeiss microscope (Zeiss) using a 63x/1.4 Oil objective. For analysis we first used Fiji[44] to prepare images for RNAScope quantification which have been done on Icy[47] (v2.4.3.0) with the Spot detector plugin. Image figures were made through FigureJ[45].

Immunohistochemistry of mice model

Eyes, previously paraffin embedded, were deparaffinized and rehydrated by Histoclear and decreasing ethanol gradients. Incubation in 10 mM Trisodium citrate pH 6, 0.05% Tween for 30 minutes at 95°C in bain-marie was performed for the Antigen retrival. Sections were blocked for 1 hour with blocking solution containing 5% BSA in PBS. Primary antibodies were prepared in blocking solution and slides were further incubated overnight at 4°C in humidifying chamber. Mice

eyes were immunostained using rabbit anti-GPATCH11 (1:100, Abcam) primary antibody. Sections were washed three times with PBS and incubated for 1 hour with Alexa-Flour 555-conjugated donkey anti-rabbit (IgG, 1:200, Life Technologies ThermoFisher Scientific). DAPI (Roche, Mannheim, Germany) was diluted with PBS to final 1.25 μg/mL and used to label nuclei and sections were washed with PBS and then mounted with Fluoromount medium (Sigma) under glass coverslip. Images were recorded on Spinning Disk Zeiss microscope (Zeiss) using a 63x/1.4 Oil objective. For the analysis, we first used machine learning for the detection of GPATCH11 and tissue using Ilastik[42] (v1.3.3post3). Then we measured the mean intensity of GPATCH11 with a Fiji[44] macro. Image figures were made through FigureJ[45]. Mice eyes were also immunostained using rabbit anti-Blue S-Opsin (1:100, Invitrogen), anti-M-Opsin (1:100, Invitrogen) and anti-Rhodopsin (1:100, Novus) primary antibody and Alexa-Flour 555-conjugated donkey anti-rabbit and 633-conjugated goat anti-mouse (1:200, Life Technologies ThermoFisher Scientific).

Behavioural characterization of mice model

Novel object recognition paradigm (NOR)

We used a modified version of the NOR test described by [48]. The test sessions were recorded with a camera placed above the test arena (a grey plastic box, 60 × 40 × 32 cm). The light intensity was equal in all parts of the arena (approximately 20 lx). Two different objects were used: a blue ceramic pot (diameter: 6.5 cm; maximum height: 7.5 cm) and a clear glass (diameter: 8.5 cm; maximum height: 7 cm). The object that serves as the novel object and the left or right placement of the objects were counterbalanced within each group. The NOR paradigm consists of three phases (over 3 days): a habituation phase, a training phase, and a testing phase. Mice were always placed in the center of the arena at the start of each exposure. The NOR paradigm consists of three phases over 3 days. On day 1 (habituation phase): mice were given 5 min to explore the arena, without any objects and were then taken back to their home cage. On day 2 (training phase): mice were allowed to explore, for 10 min, two identical objects arranged in a symmetric opposite position from the center of the arena and were then transported to their home cage. On day 3 (testing phase): mice were given 15 minutes to

explore two objects: a familiar object and a novel one, in the same arena. The following behaviors were scored as exploration: sniffing, licking, touching the object with the nose or with the front legs, or proximity (≤ 1 cm) between the nose and the object. If the mouse was on top of the object or completely immobile, exploration was not scored. The preference index for the novel object was calculated as (time spent exploring the new object/the total time spent exploring both objects), and the discrimination index was calculated as (time spent exploring the new object -time spent exploring the familiar object) / (total time spent exploring both objects). Behavior was scored on video by observer blinded to the mice.

3-foot shock contextual fear conditioning (CFC)

Mice were tested individually inside the conditioning chambers (Bioseb (France); 25 × 25 × 25 cm) located inside a larger, insulated plastic cabinet that provided protection from outside light and noise (67 × 55 × 50 cm, Bioseb, France). Floors of the chamber consisted of 27 stainless steel bars wired to a shock generator with scrambler for the delivery of foot shock. Signal generated by the mice movements was recorded and analyzed through a high sensitivity weight transducer system and analysis of time active/time immobile (Freezing) was performed. The CFC procedure took place over two consecutive days. On day 1, mice were placed in the conditioning chamber, and received 3 foot-shocks (1.5 s, 0.5 mA), which were administrated at 60, 120, and 180 s after the animals were placed in the chamber. They were returned to their home cages 60 s after the final shock. Contextual fear memory was assessed 24 h after conditioning by returning the mice to the chamber and measuring freezing behavior during a 4-min retention test. Freezing was scored and analyzed automatically using Packwin 2.0 software (Bioseb, France). Freezing behavior was considered to occur if the animals froze for a period of at least two seconds.

Open field test (OFT)

This test takes advantage of the aversion of rodents to brightly lit areas. Each mouse is placed in the center of the OFT chamber (43 × 43 cm chamber) and allowed to explore for 30 min. Mice were monitored throughout each test session by infrared light beam activity monitor using actiMot2 Software (PhenoMaster Software, TSE). The overall motor activity was quantified as the total

distance traveled (ambulation). Anxiety was quantified by measuring the time and distance spent in the center versus periphery of the open-field chamber.

Morris Water Maze (MWM)

The Morris water maze (MWM) is a white circular swimming pool (diameter: 200 cm, walls: 60 cm high), which was in a room with various distal cues. The pool was filled with water (depth: 50 cm) maintained at 22C±1C, which was made opaque by the addition of a nontoxic white paint. A 12 cm round platform was hidden 1.0 cm below the water surface. The pool was located in a brightly lit room. Extra maze geometric and high-contrast cues were mounted on the walls of the swimming pool with the ceiling providing illumination. Each daily trial consisted of four swimming trials, in which each mouse was placed in the pool facing the wall of the tank and allowing the animal to swim to the platform before 120 s had elapsed. A trial terminated when the animal reached the platform, where it remained for 5 s. Mice were removed and placed back in their home cages for a 5 min inter-trial interval. The starting point differed at each trial, and different sequences of release points were used from day to day. Swimming time to the platform was calculated as an evaluation of performance of the mice to locate the target.

Transcriptomic analysis

Total RNAs of retina were extracted using the RNAeasy Mini kit, according to the manufacturer's protocol (Qiagen). The concentration and the purity of total RNA were determined using an Agilent 2100 Bioanalyzer. RNA-Seq was performed for *Gpatch11*$^{WT/WT}$ (n=5) and *Gpatch11*$^{\Delta 5/\Delta 5}$ (n= 5) mice retina at 15 days. RNA-Seq libraries were constructed from 1ug of total RNAs using the TrueSeq Stranded mRNA Sample Prep (Illumina) and paired-end sequencing of 100 bases length was performed on a Novaseq 6000 Illumina (IlluminaPass-filtered reads were mapped using STAR 2.7.10b and aligned to Ensembl genome assembly GRCm39.108. For differential expression analysis a count table of the genes features was obtained using FeatureCounts[49]. For gene level analysis, EdgeR R package was used for normalization, differential expression analysis and to compute TPM (Transcripts Per Million) values[50]. Filtering of the differentially expressed genes

(DEGs) with a fold change (FC) greater or <2, an FDR (False Discovery Rate) lower than 0.05 and a minimum expression of five TPMwas applied.

Alternate splicing analysis was performed using rMATS[51] allowing the analysis of differential exon usage for distinct types of events (skipped exons, alternative 5' and 3' splice sites, retained introns, and mutually exclusive exons). For each comparison being made, sorted BAM files produced by STAR (two-pass) were used to run rMATS with default settings. The identified splicing changes were filtered out using a minimum count of 2 for at least one sample, a false discovery rate (FDR) of <0.05 and a change in inclusion-level difference of more than 20%. Comprehensive gene list analysis enriched biological pathways, gene annotation from differential expression and alternate splicing analysis were based on the Gene Ontology classification system using Metascape[52]. R packages were used for data mining including GOPlot for pathway data graphical representation[53]. Transcripts candidates were tested on *Gpatch11*$^{\Delta 5/\Delta 5}$ retina and *Gpatch11*$^{WT/WT}$ *(***Supplementary Table 8**). Raw data were submitted to BioStudies. Data are available via BioStudies with identifier S-BSST1157.

S-Trap digestion / TTP MS/MS

S-TrapTM micro spin column (Protifi, Hutington, USA) digestion was performed on 50 µg of cell lysates according to manufacturer's instructions. Briefly, samples were supplemented with 20% SDS to a final concentration of 5%, reduced with 20mM TCEP (Tris-2-carboxyethyl-phosphine hydrochloride) and alkylated with 50mM CAA (chloracetamide) for 15min in the dark. Aqueous phosphoric acid was then added to a final concentration of 2.75% following by the addition of S-Trap binding buffer (90% aqueous methanol, 100mM TEAB, pH7.1). Mixtures were then loaded on S-Trap columns. Six washes were performed for thorough SDS elimination. Samples were digested with 5µg of trypsin (Promega) for immunoprecipitation samples and 12.5µg for input samples at 47°C for 90min. After elution, peptides were vacuum dried and resuspended in 2% ACN, 0.1% formic acid in HPLC-grade water prior to MS analysis. The input samples were resuspended in 125 µL and a volume of 1 µL was injected on a nanoelute (Bruker Daltonics, Germany) HPLC (high-

performance liquid chromatography) system coupled to a timsTOF Pro (Bruker Daltonics, Germany) mass spectrometer. HPLC separation (Solvent A : 0.1% formic acid in water, 2% acetonitrile; Solvent B : 0.1% formic acid in acetonitrile) was carried out at 250nL/min using a packed emitter column (C18, 25 cm×75µm 1.6µm) (Ion Optics, Australia) using a 70min gradient elution (2 to 13% solvent B during 41min; 13 to 20% during 23min; 20% to 30% during 5min; 30% to 85% for 5min and finally 85% for 5min to wash the column). Mass-spectrometric data were acquired using the parallel accumulation serial fragmentation (PASEF) acquisition method in DDA mode. The measurements were carried out over the m/z range from 100 to 1700 Th. The range of ion mobilities values from 0.75 to 1.25 V s/cm2(1/k0). The total cycle time was set to 1.17s and the number of PASEF MS/MS scans was set to 10.

Data Processing Following LC-MS/MS acquisition

The MS files were processed with the MaxQuant software version 2.1.3.0 and searched with Andromeda search engine against the database of Mus musculus from swissprot and trembl (release 11/2021, 88132 entries). To search parent mass and fragment ions, we set an initial mass deviation of 4.5 ppm and 20 ppm respectively. The minimum peptide length was set to 7 aminoacids and strict specificity for trypsin cleavage was required, allowing up to two missed cleavage sites. Carbamidomethylation (Cys) was set as fixed modification, whereas oxidation (Met) and N-term acetylation were set as variable modifications. Match between runs was allowed. LFQ minimum ratio count was set to 1. The false discovery rates (FDRs) at the protein and peptide level were set to 1%. Scores were calculated in MaxQuant as described previously[54]. The reverse and common contaminants hits were removed from MaxQuant output. Proteins were quantified according to the MaxQuant label-free algorithm using LFQ intensities[54,55]. Match between runs was allowed for the input samples. Four independent lysate replicates of gpatch11 *Gpatch11*$^{WT/WT}$ and gpatch11 homozygote gpatch11 mutant were analyzed with Perseus software (version 1.6.15.0) freely available at www.perseus-framework.org[56]. For statistical comparison, we set two groups, each containing up to 4 biological replicates. We then filtered the data to keep only proteins with at least

3 valid values in at least one group. Next, the data were imputed to fill missing data points by creating a Gaussian distribution of random numbers with a standard deviation of 33% relative to the standard deviation of the measured values and 1.8 standard deviation downshift of the mean to simulate the distribution of low signal values. We performed a t-test, FDR<0.05, S0=0.01. Volcano plot of proteins was performed in Perseus on logarithmised LFQ intensities at FDR<0.05, S0=0.01. Raw data from the lysates were submitted to ProteomeXchange via the PRIDE database. Data are available via ProteomeXchange with identifier PXD041849.

REFERENCES

1. Ohta, S. *et al.* The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**, 810–821 (2010).
2. Griffin, C. & Saint-Jeannet, J.-P. Spliceosomopathies: Diseases and mechanisms. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **249**, 1038–1046 (2020).
3. Bohnsack, K. E., Ficner, R., Bohnsack, M. T. & Jonas, S. Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol. Chem.* **402**, 561–579 (2021).
4. Robert-Paganin, J., Réty, S. & Leulliot, N. Regulation of DEAH/RHA helicases by G-patch proteins. *BioMed Res. Int.* **2015**, 931857 (2015).
5. Studer, M. K., Ivanović, L., Weber, M. E., Marti, S. & Jonas, S. Structural basis for DEAH-helicase activation by G-patch proteins. *Proc. Natl. Acad. Sci.* **117**, 7159–7170 (2020).
6. Abdelhaleem, M. Helicases: An Overview. in *Helicases* (ed. Abdelhaleem, M. M.) vol. 587 1–12 (Humana Press, 2009).
7. Sloan, K. E. & Bohnsack, M. T. Unravelling the Mechanisms of RNA Helicase Regulation. *Trends Biochem. Sci.* **43**, 237–250 (2018).
8. Johnston, J. J. *et al.* Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. *Am. J. Hum. Genet.* **86**, 743–748 (2010).
9. Wynn, S. L. *et al.* Organization and conservation of the GART/SON/DONSON locus in mouse and human genomes. *Genomics* **68**, 57–62 (2000).
10. Zang, S. *et al.* GPKOW is essential for pre-mRNA splicing in vitro and suppresses splicing defect caused by dominant-negative DHX16 mutation in vivo. *Biosci. Rep.* **34**, e00163 (2014).
11. Ferre-Fernández, J.-J. *et al.* Whole-Exome Sequencing of Congenital Glaucoma Patients Reveals Hypermorphic Variants in GPATCH3, a New Gene Involved in Ocular and Craniofacial Development. *Sci. Rep.* **7**, 46175 (2017).
12. Huang, T., Yan, G. & Guan, M. Zinc Homeostasis in Bone: Zinc Transporters and Bone Diseases. *Int. J. Mol. Sci.* **21**, 1236 (2020).

13. Carroll, R. *et al.* Variant in the X-chromosome spliceosomal gene GPKOW causes male-lethal microcephaly with intrauterine growth restriction. *Eur. J. Hum. Genet. EJHG* **25**, 1078–1082 (2017).
14. Si, W. *et al.* Angiotensin II increases angiogenesis by NF-κB-mediated transcriptional activation of angiogenic factor AGGF1. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **32**, 5051–5062 (2018).
15. Li, M. *et al.* Grass carp (Ctenopharyngodon idella) GPATCH3 initiates IFN 1 expression via the activation of STING-IRF7 signal axis. *Dev. Comp. Immunol.* **112**, 103781 (2020).
16. Wieczorek, D. *et al.* Compound heterozygosity of low-frequency promoter deletions and rare loss-of-function mutations in TXNL4A causes Burn-McKeown syndrome. *Am. J. Hum. Genet.* **95**, 698–707 (2014).
17. Albers, C. A. *et al.* Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat. Genet.* **44**, 435–439, S1-2 (2012).
18. Bacrot, S. *et al.* Mutations in SNRPB, encoding components of the core splicing machinery, cause cerebro-costo-mandibular syndrome. *Hum. Mutat.* **36**, 187–190 (2015).
19. Favaro, F. P. *et al.* A noncoding expansion in EIF4A3 causes Richieri-Costa-Pereira syndrome, a craniofacial disorder associated with limb defects. *Am. J. Hum. Genet.* **94**, 120–128 (2014).
20. Lines, M. A. *et al.* Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes mandibulofacial dysostosis with microcephaly. *Am. J. Hum. Genet.* **90**, 369–377 (2012).
21. Bernier, F. P. *et al.* Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. *Am. J. Hum. Genet.* **90**, 925–933 (2012).
22. Chakarova, C. F. *et al.* Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum. Mol. Genet.* **11**, 87–92 (2002).
23. Deery, E. C. *et al.* Disease mechanism for retinitis pigmentosa (RP11) caused by mutations in the splicing factor gene PRPF31. *Hum. Mol. Genet.* **11**, 3209–3219 (2002).
24. Chen, X. *et al.* PRPF4 mutations cause autosomal dominant retinitis pigmentosa. *Hum. Mol. Genet.* **23**, 2926–2939 (2014).
25. Tanackovic, G. *et al.* A missense mutation in PRPF6 causes impairment of pre-mRNA splicing and autosomal-dominant retinitis pigmentosa. *Am. J. Hum. Genet.* **88**, 643–649 (2011).

26. 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).
27. Cvačková, Z., Matějů, D. & Staněk, D. Retinitis pigmentosa mutations of SNRNP200 enhance cryptic splice-site recognition. *Hum. Mutat.* **35**, 308–317 (2014).
28. Tsang, S. H. & Sharma, T. Retinitis Pigmentosa (Non-syndromic). *Adv. Exp. Med. Biol.* **1085**, 125–130 (2018).
29. Růžičková, Š. & Staněk, D. Mutations in spliceosomal proteins and retina degeneration. *RNA Biol.* **14**, 544–552 (2016).
30. 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).
31. Yang, C. *et al.* Pre-mRNA Processing Factors and Retinitis Pigmentosa: RNA Splicing and Beyond. *Front. Cell Dev. Biol.* **9**, 700276 (2021).
32. Jerber, J. *et al.* The coiled-coil domain containing protein CCDC151 is required for the function of IFT-dependent motile cilia in animals. *Hum. Mol. Genet.* **23**, 563–577 (2014).
33. Becker-Heck, A. *et al.* The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat. Genet.* **43**, 79–84 (2011).
34. Gigante, E. D. & Caspary, T. Signaling in the primary cilium through the lens of the Hedgehog pathway. *Wiley Interdiscip. Rev. Dev. Biol.* **9**, e377 (2020).
35. Solovei, I. *et al.* Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* **137**, 356–368 (2009).
36. Hanein, S. *et al.* Leber congenital amaurosis: comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis. *Hum. Mutat.* **23**, 306–317 (2004).
37. Zeitz, C. *et al.* Mutations in CABP4, the gene encoding the Ca2+-binding protein 4, cause autosomal recessive night blindness. *Am. J. Hum. Genet.* **79**, 657–667 (2006).

38. Zhang, H. *et al.* Light-dependent redistribution of visual arrestins and transducin subunits in mice with defective phototransduction. *Mol. Vis.* **9**, 231–237 (2003).

39. Awwad, S. W., Darawshe, M. M., Machour, F. E., Arman, I. & Ayoub, N. Recruitment of RBM6 to DNA Double-Strand Breaks Fosters Homologous Recombination Repair. *Mol. Cell. Biol.* **43**, 130–142 (2023).

40. Remo, A., Li, X., Schiebel, E. & Pancione, M. The Centrosome Linker and Its Role in Cancer and Genetic Disorders. *Trends Mol. Med.* **26**, 380–393 (2020).

41. Bruel, A.-L. *et al.* Increased diagnostic and new genes identification outcome using research reanalysis of singleton exome sequencing. *Eur. J. Hum. Genet. EJHG* **27**, 1519–1531 (2019).

42. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).

43. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).

44. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

45. Mutterer, J. & Zinck, E. Quick-and-clean article figures with FigureJ. *J. Microsc.* **252**, 89–91 (2013).

46. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

47. de Chaumont, F. *et al.* Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* **9**, 690–696 (2012).

48. Glatigny, M. *et al.* Autophagy Is Required for Memory Formation and Reverses Age-Related Memory Decline. *Curr. Biol. CB* **29**, 435-448.e8 (2019).

49. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

50. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* **26**, 139–140 (2010).

51. Shen, S. *et al.* MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res.* **40**, e61 (2012).

52. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

53. Walter, W., Sánchez-Cabo, F. & Ricote, M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinforma. Oxf. Engl.* **31**, 2912–2914 (2015).

54. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

55. Luber, C. A. *et al.* Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity* **32**, 279–289 (2010).

56. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

ACKNOWLEDGMENTS

This work was supported by grants from Retina France; Fondation JED Belgique; Fondation Visio to JMR and IP. Fondation des Aveugles de France, PhD international Institut Imagine to AZ. JMR is member of the European Reference Network for Rare Eye Diseases (ERN-EYE), which is co-funded by the Health Program of the European Union under the Framework Partnership Agreement n°739534. We thank the Transgenesis, Neurobehavioral and metabolism, Necker Bioimage Analysis and Proteomic platforms for valuable discussions and help.

AUTHOR CONTRIBUTIONS

A.Z. performed *in vitro and in vivo* experiments, analyzed the data and wrote the paper

P.D., L.F-T. performed creation of animal models

J.A C.G I.A. M.R. N.B. AL. B. N.P-S. D.A. J.K. K.K. collected clinical data

S.M. performed behavioral analysis of mice

N.G. created analytic tools

J.R analyzed data from RNASeq

C.G analyzed data from mass spectrometry

S.B. collected skin biopsy from patients

I.P analyzed exome data from Family 1

J.A analyzed exome data from Family 2

C.T. analyzed exome data from Family 3

C.R. analyzed exome data from Family 4

J-M.R. I.P. supervised the research, designed the experiments and wrote the paper

All authors discussed the results and participated in manuscript preparation and editing

COMPETING INTERESTS

The authors declare no competing interests.

Table 1: **Summary of clinical features of ten affected individuals from four families with *GPATCH11* variants.** Ophthalmological, neurological and dysmorphological investigation in all affected individuals. Abbreviation used: NA, not available; NP, not performed.

Individual	**F1:III-3**	**F1:III-6**	**F1:IV-1**	**F1:IV-2**	**F2:II-2**	**F2:II-3**	**F2:II-4**	**F3:II-1**	**F3:II-2**	**F4:II-2**
Age range	Dead in the second decade of life because of fever	In the second decade of life	In the second decade of life	In the first decade of life	In the first decade of life	In the first decade of life	In the first decade of life	Dead in the third decade of life because of stroke	In the third decade of life	In the fourth decade of life
Ophthalmological Features										
Age at diagnosis	N.A.	Birth	First moths of life	Birth	First years of life	Birth	Birth	N.A.	N.P.	Second decade of life
Nystagmus	N.A.	+	+	+	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Photophobia	N.A.	N.A.	+	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	+
Night blindness	N.A.	N.A.	N.A.	N.A.	N.A.	+	N.A.	N.A.	N.A.	+ since early childhood; color vision trouble since the second decade of life
Visual acuity (RE) (LE)	N.A.	N.A.	1/20	Hand movements	4/10; hypermetropia	8/10	N.A.	Myopia	Myopia	0,4 OD; 0,1 OS
Refractive error	N.A.	N.A.	+8.75 (-1.5) 5° +8.5 (-1.25) 170°	+1.25 (-0.25) 45° +0.00 (-0.75) 170°	+1 (-0.50) 140° -0.50 (-0.75) 75°	+1 +1.25	N.A.	N.A.	N.A.	N.A.
ERG electronegativity	N.P.	N.A.	Flat	Flat	Flat	N.A.	N.A.	N.A.	N.A.	N.A.
Fundus	N.A.	N.A.	Macular atrophy; pigmentary deposits in periphery	Macular atrophy; pigmentary deposits in periphery	Pigmentary deposits in periphery; light foveolar hypoplasia	N.A.	N.A.	N.A.	Optic atrophy; pigmentary deposits; bruised vessels	Retinis pigmentosa signs: moderately pale optic nerve, narrowed retinal vessels, peripheral pigmentation (bone spicules)
Macular SD-OCT	N.P.	N.P.	Global alteration of the retina	N.P.	N.A.	N.A.	N.A.	N.A.	N.A.	Decreased retinal thickness (mainly external layers with relative foveal sparing OD)
Autofluorescence	N.A.	N.A.	Hyper autofluorescence of the macula surrounded by hypo autofluorescence	Hyper autofluorescence of the macula surrounded by hypo autofluorescence	Hyper autofluorescence of the macula	N.A.	N.A.	N.A.	N.A.	N.A.
Neurological features										
Neurological examination	Psychomotor retardation; intellectual disability; no language	Psychomotor retardation; intellectual disability; no language	IME Psychomotor retardation; language disorder;	IME Psychomotor retardation; language disorder	ULIS Psychomotor retardation; intellectual disability;	Language: do sentences	N.A.	Psychomotor retardation; discrete ataxia; encephalopat	Psychomotor retardation; ataxia ; encephalopat hy; autism	Cognitive deficit, global developmenta l delay; mainstream

			hyperactivity; epilepsy: continuous spikes and waves during sleep		language disorder			hy; retardation; language: says few words; epilepsy at few months old	spectrum disorder; self- and hetero-aggression; retardation; no language; epilepsy	schooling until the second decade of life; currently autonomous for daily activities.
MRI	N.A.	N.A.	Normal	Normal	Normal	Normal	N.A.	Hypoplasia of frontal lobes	N.A.	Normal
Dysmorphological features										
Morphological examination	Facial dysmorphia	Facial dysmorphia	Facial dysmorphia; enophtalmia; short philtrum; large incisors; diastema; oligodontia; flat feet; hyperlaxity	Facial dysmorphia; enophtalmia; short philtrum; large incisors; diastema; enamel dysplasia; flat feet; hyperlaxity	Facial dysmorphia; enophtalmia; short philtrum; large incisors; broad ovula; flat feet	Facial dysmorphia	N.A.	Facial dysmorphia; enophtalmia; thick lips; prominent columella; short philtrum; large incisors; hyperlaxity; short stature: 143 cm	Facial dysmorphia; enophtalmia; thick lips; prominent columella; short philtrum; large incisors; abdominal obesity; short stature: 146 cm	Some dysmorphic features noticed (not suggestive of a specific syndrome), macrocephaly, large external ears, small pre-auricular tubercle on the right side, small and deep-set eyes, long and prominent nose with long columella, hypoplastic nasal alae, short 5th toes
Other										
Metabolism examination	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	Diabetes discovered in the second decade of life	Diabetes discovered in the second decade of life	N.A.

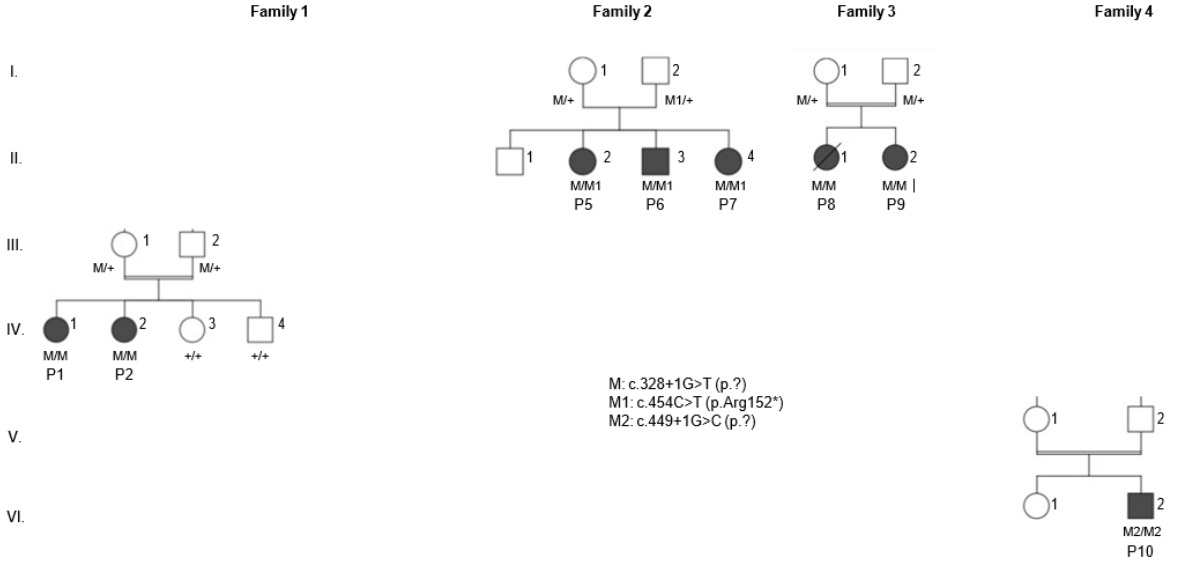


Fig.1: **GPATCH11 variants are responsible for syndromic inherited retinal diseases.** Pedigree and segregation analysis of the four families carrying *GPATCH11* variants. M: c.328+1G>T (p.?); M1: c.454C>T (p.Arg152*); M2: c.449+1G>C (p.?); +, wild-type allele; P, patient. F1:IV-1 (P1) and F1:IV-2 (P2) are affected individuals whose fibroblasts were analyzed.

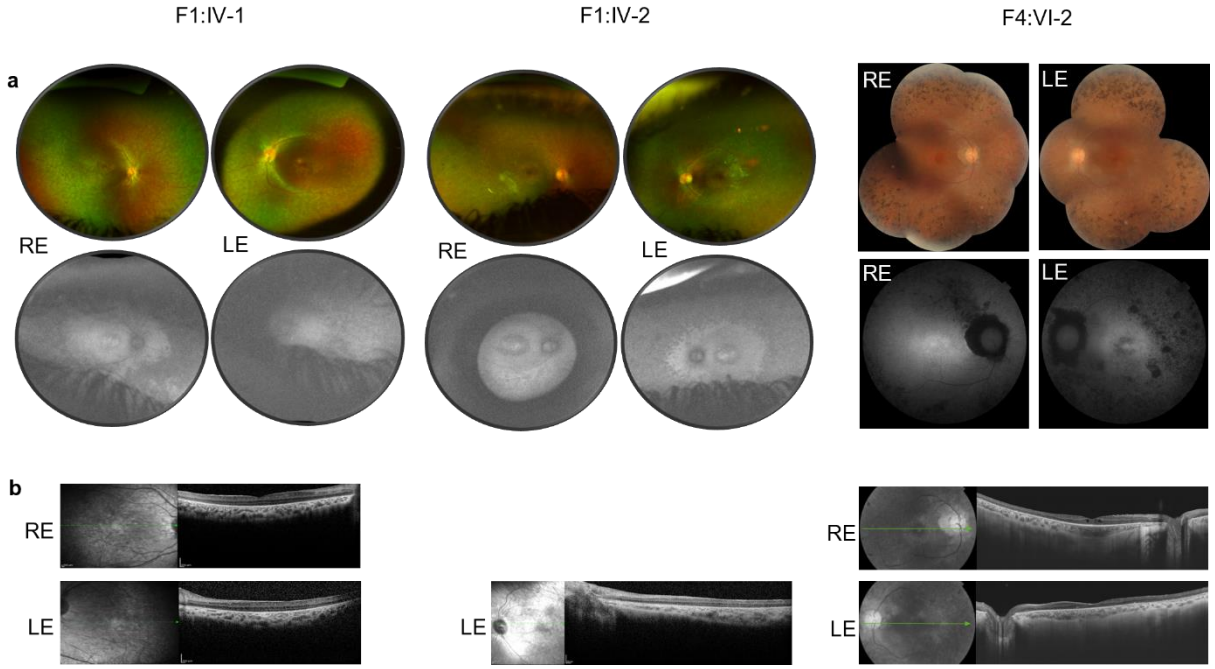


Fig.2: **Clinical analysis of affected individuals** (**a**) Color fundus photographs of the affected individuals F1:IV-1, F1:IV-2 and F4:VI-2. RE: right eye; LE: left eye. (**b**) Spectral-domain optical coherence tomography (SD-OCT) of the macular region of the retina of the affected individuals F1:IV-1, F1:IV-2 and F4:VI-2.

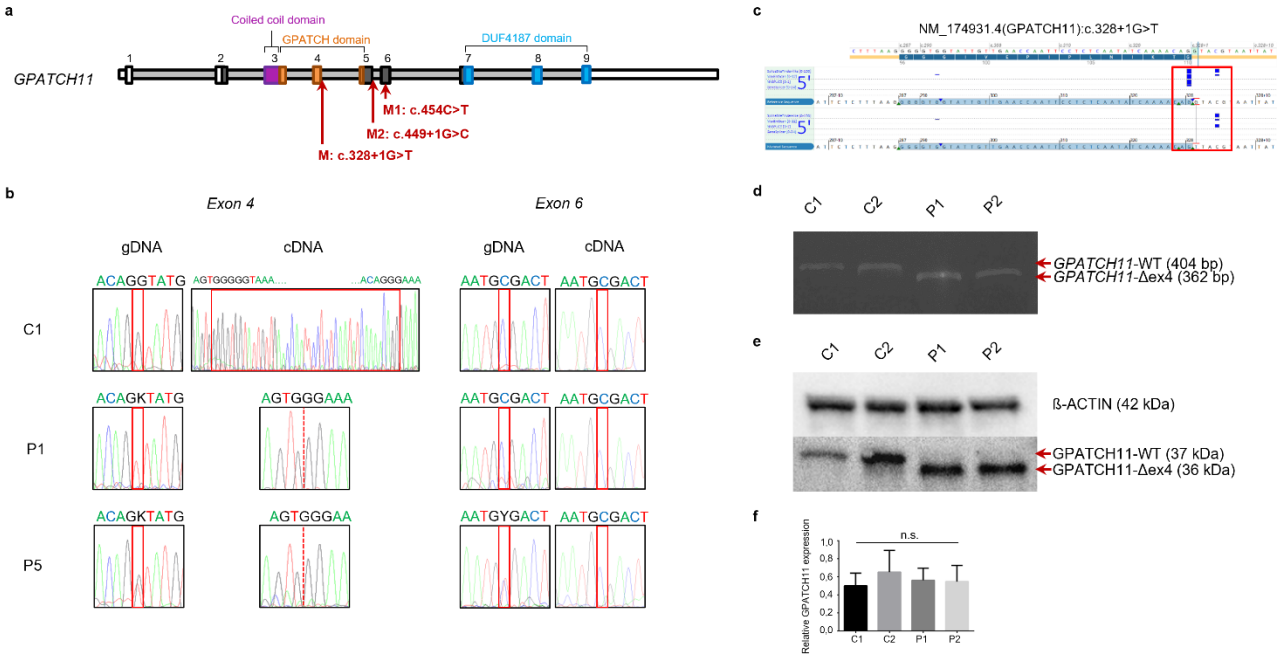


Fig.3: **Analysis of the variants of *GPATCH11*.** (**a**) Diagram of the human *GPATCH11* gene showing position of the c.328+1G>T, c.454C>T and c.449+1G>C variants. (**b**) Chromatogram of gDNA and the cDNA of the exon 4 and exon 6 of controls and affected individuals' fibroblasts. (**c**) Predicted effect of c.328+1G>T (p.?) on donor splice-site (in blue) identified in Family 1, analyzed by NNSPLICE, SpliceSiteFinder-like, MaxEntScan and GeneSplicer (Alamut Visual 2.0). (**d**) Electrophoresis of human *GPATCH11* cDNA of control and affected individuals' fibroblasts after the PCR. One band is clearly visible in each line. Two bands correspond to *GPATCH11*: *GPATCH11-Δex4*: 362 bp and *GPATCH11-WT*: 404 bp. (**e**) Western Blot analysis of controls (C1, C2) and affected individuals (P1, P2) fibroblasts. The specific epitope for the anti-GPATCH11 antibody used is encoded for the amino acids 111 to 192. (**f**) The abundance of the GPATCH11 products in affected individual fibroblasts is similar to that of the control fibroblasts. Graphic bars represent the mean±SEM derived from four experimental replicates. n.s., not significant.

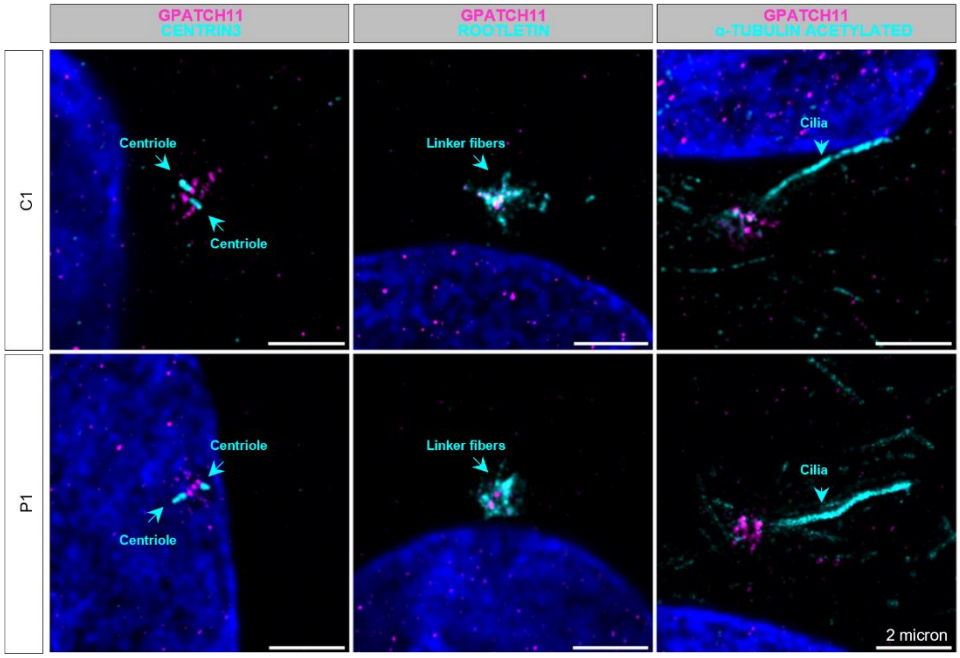


Fig.4: Expression of GPATCH11 in control and affected individuals' fibroblasts, with c.328+1G>T variant. Immunostaining of GPATCH11 (magenta), CENTRIN3, ROOTLETIN and ACETYLATED α-TUBULIN (cyan) proteins in control and affected individual fibroblasts, after 24 hours of starvation. DAPI is used to label the nucleus. C1: control fibroblasts; P1: affected individual fibroblasts affected by the c.328+1G>T variant. Scale bars, 2µm

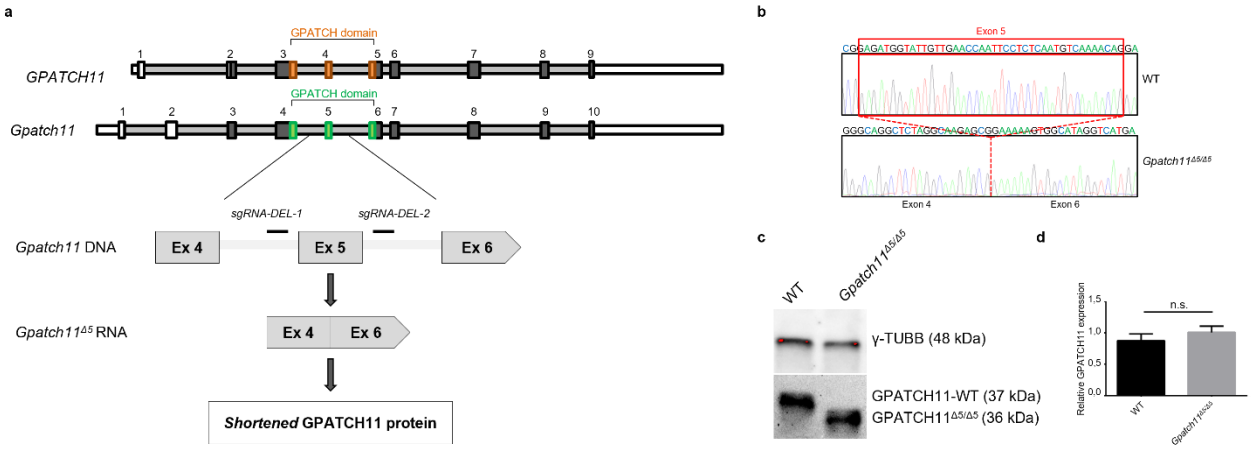


Fig.5: Creation of *Gpatch11*$^{\Delta5/\Delta5}$ mice model. (a) Diagram of the human *GPATCH11* gene and the murine *Gpatch11* gene. Human exon 4 corresponds to murine exon 5 and schematic illustration of CRISPR/Cas9-mediated exon 5 deletion from the mouse *Gpatch11* locus. (**b**) Chromatogram of the cDNA of the exon 4, 5 and 6 of WT: *Gpatch11*$^{WT/WT}$ and *Gpatch11*$^{\Delta5/\Delta5}$ retina samples. The exon 5 is missing in mutated mice. (**c**) Western blot of protein of mice retina from WT: *Gpatch11*$^{WT/WT}$ and *Gpatch11*$^{\Delta5/\Delta5}$. (**d**) The abundance of the GPATCH11 products in *Gpatch11*$^{\Delta5/\Delta5}$ is similar to that of the WT: *Gpatch11*$^{WT/WT}$. Graphic bars represent the mean±SEM derived from three experimental replicates. n.s., not significant.

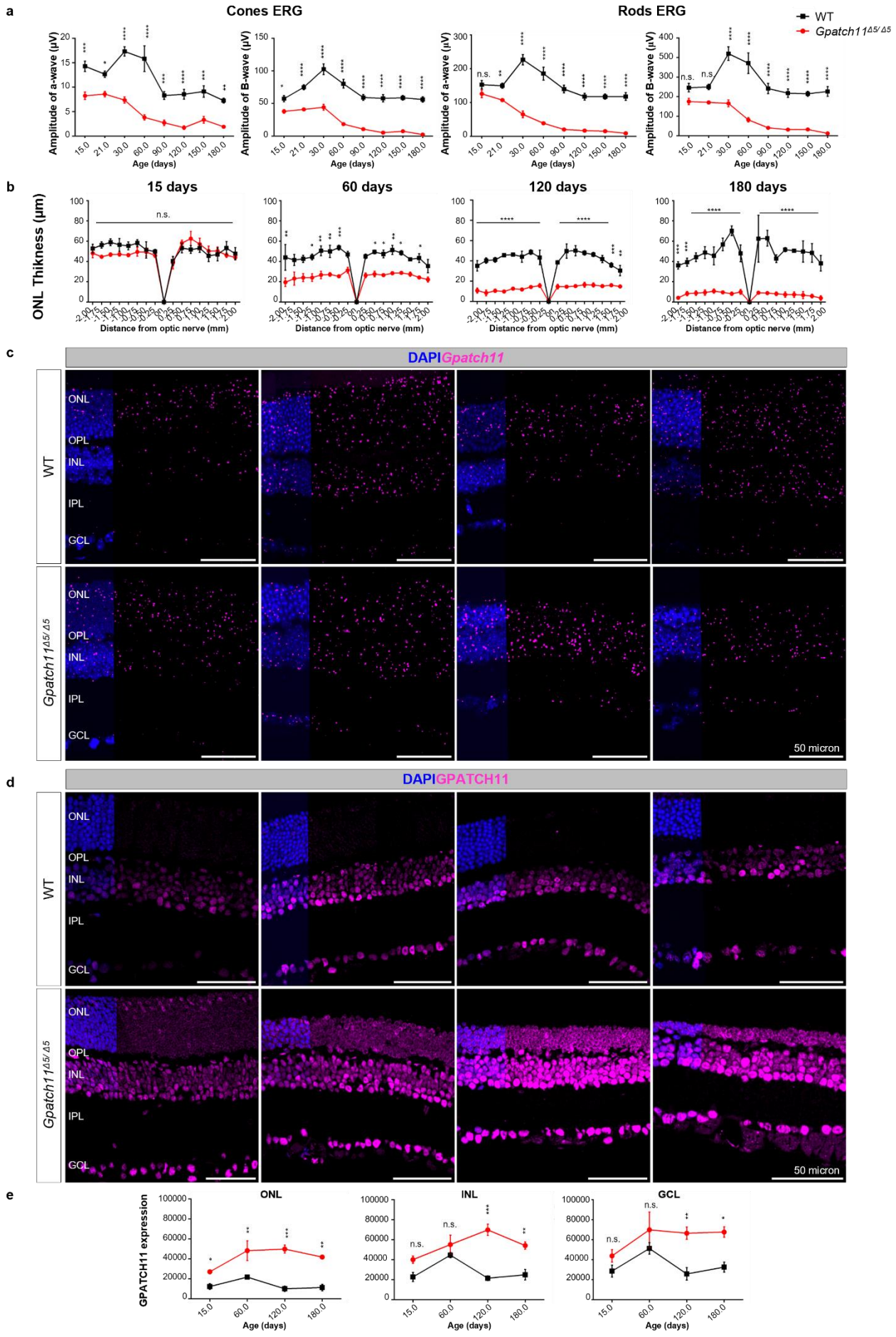

a
Cones ERG
Rods ERG
WT
Gpatch11Δ5/Δ5
Amplitude of a-wave (µV)
Amplitude of B-wave (µV)
Age (days)
b
15 days
60 days
120 days
180 days
ONL Thikness (µm)
Distance from optic nerve (mm)
c
DAPIGpatch11
WT
Gpatch11Δ5/Δ5
ONL
OPL
INL
IPL
GCL
50 micron
d
DAPIGPATCH11
e
ONL
INL
GCL
GPATCH11 expression

Fig.6: **Homozygous deletion of exon 5 in *Gpatch11* leads to retinal degeneration mimicking early-onset retinal dystrophy.** (**a**) Light adapted (photopic) and dark-adapted (scotopic) ERG recordings in WT: *Gpatch11*$^{WT/WT}$ (in black) and *Gpatch11*$^{\Delta 5/\Delta 5}$ (in red) mice at 15, 21, 30, 60, 90, 120, 150, 180 days. Bars represent the mean±SEM derived from nine biological replicates. (**b**) Spider graph demonstrating the outer nuclear layer (ONL) thickness measured in WT: *Gpatch11*$^{WT/WT}$ (in black) and *Gpatch11*$^{\Delta 5/\Delta 5}$ (in red) mice at 15, 60, 120 and 180 days. Bars represent the mean±SEM derived from three biological replicates. (**c**) RNAScope of *Gpatch11* (magenta) RNA in retina sections. DAPI is used to label the cell nucleus. Scale bars, 50 µm. (**d**) Immunostaining of retina sections with anti-GPATCH11 (magenta) antibody. Nuclei are stained with DAPI (blue). Scale bars, 50 µm. (**e**) Quantification of GPATCH11 protein expression in the ONL, in the inner nuclear layer (INL) and in the ganglion cells layer (GCL), in WT: *Gpatch11*$^{WT/WT}$ (in black) and *Gpatch11*$^{\Delta 5/\Delta 5}$ (in red). Bars represent the mean±SEM derived from three biological replicates.

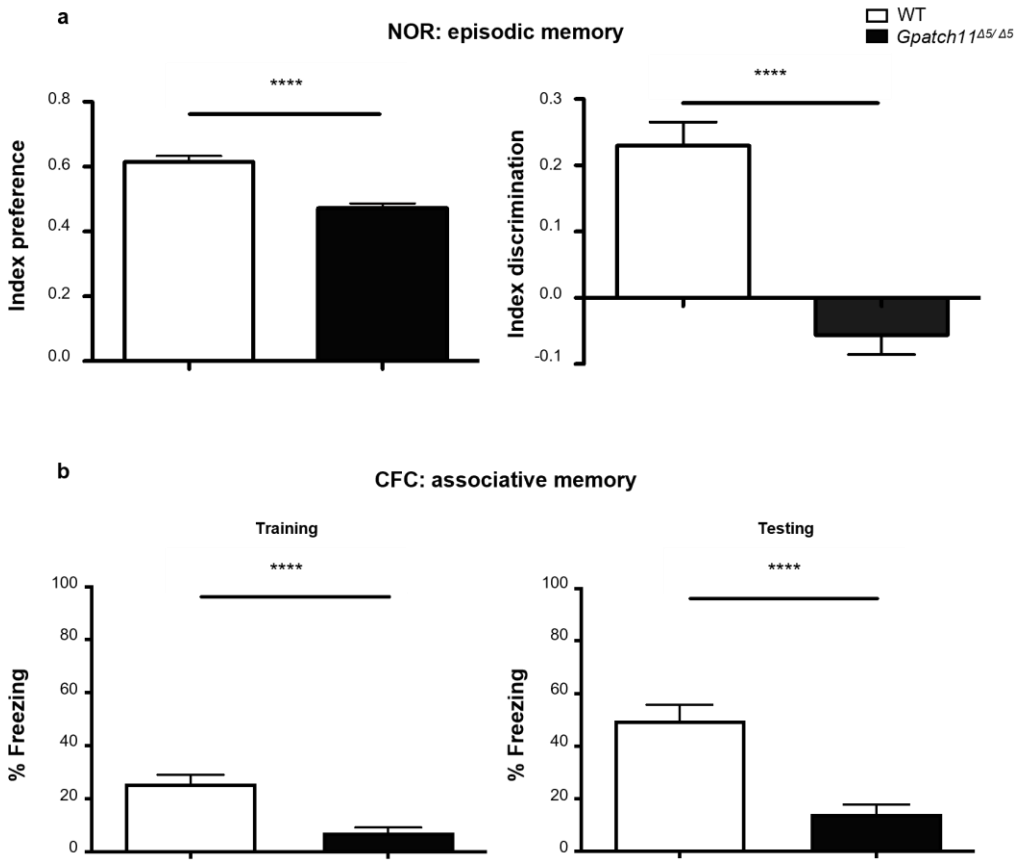
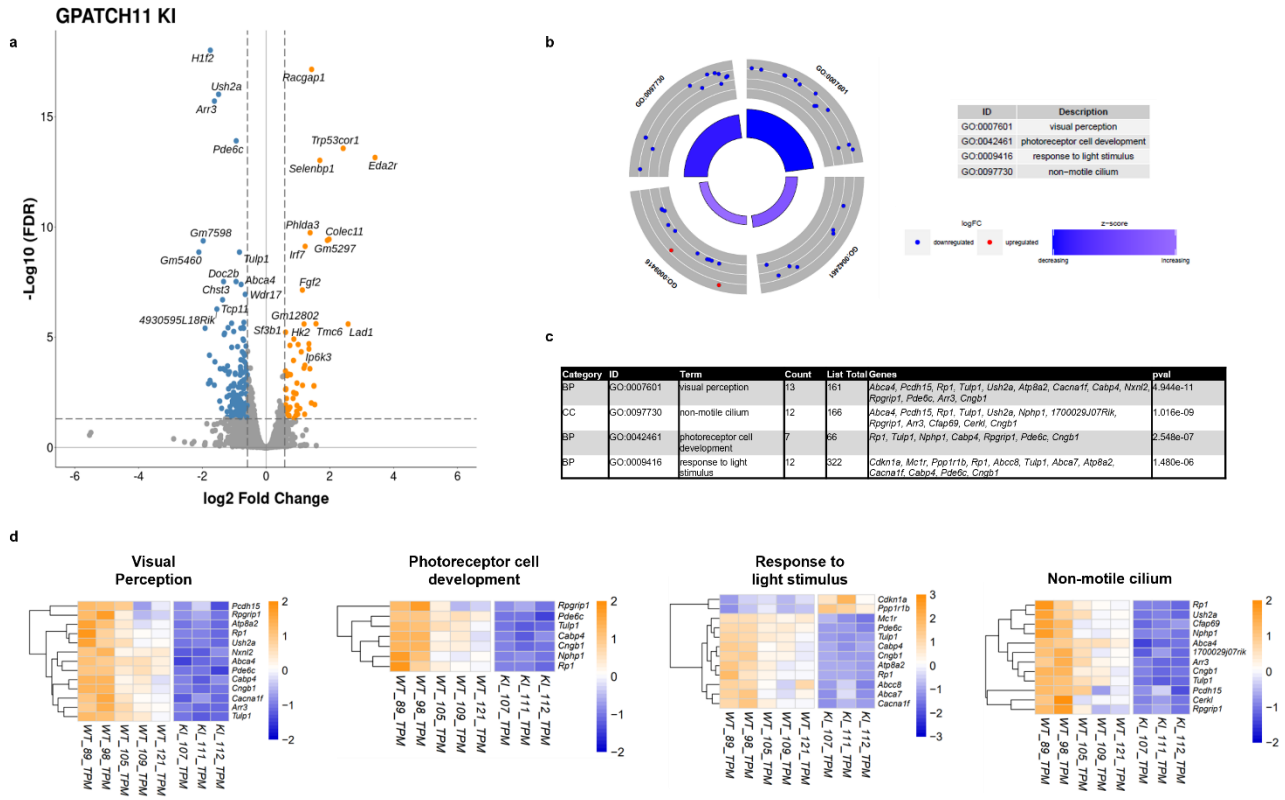
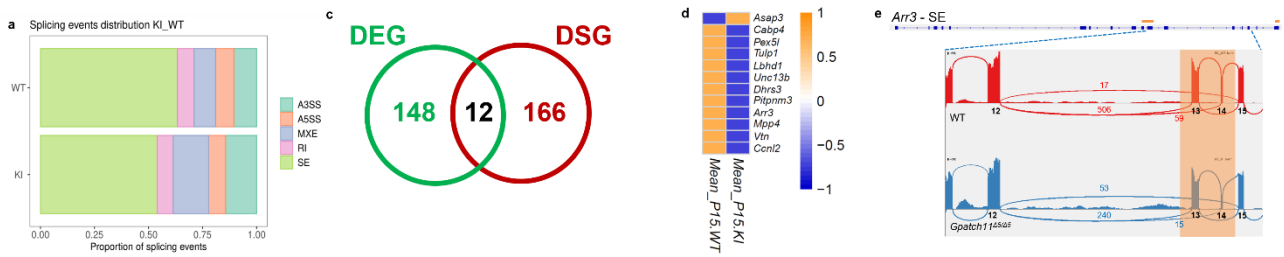


Fig.7: **Behavioural tests on mouse model performed in 1-month old *Gpatch11*$^{WT/WT}$ and *Gpatch11*$^{\Delta 5/\Delta 5}$ mice.** (**a**) Novel object recognition (NOR). Discrimination index and Preference index was measured. (**b**) Contextual fear conditioning (CFC). Percentage freezing after 24 h testing and after 24 h of training was measured (Student's t test).



Category	ID	Term	Count	List Total	Genes	pval
BP	GO:0007601	visual perception	13	161	*Abca4, Pcdh15, Rp1, Tulp1, Ush2a, Atp8a2, Cacna1f, Cabp4, Nxnl2, Rpgrip1, Pde6c, Arr3, Cngb1*	4.944e-11
CC	GO:0097730	non-motile cilium	12	166	*Abca4, Pcdh15, Rp1, Tulp1, Ush2a, Nphp1, 1700029J07Rik, Rpgrip1, Arr3, Cfap69, Cerkl, Cngb1*	1.016e-09
BP	GO:0042461	photoreceptor cell development	7	66	*Rp1, Tulp1, Nphp1, Cabp4, Rpgrip1, Pde6c, Cngb1*	2.548e-07
BP	GO:0009416	response to light stimulus	12	322	*Cdkn1a, Mc1r, Ppp1r1b, Rp1, Abcc8, Tulp1, Abca7, Atp8a2, Cacna1f, Cabp4, Pde6c, Cngb1*	1.480e-06

Fig.8: Identification of genes and pathways deregulated in retina from *Gpatch11*$^{\Delta 5/\Delta 5}$ mice using whole transcriptome analysis. (**a**) Volcano plot of differentially expressed genes for *Gpatch11*$^{\Delta 5/\Delta 5}$ *vs* WT: *Gpatch11*$^{WT/WT}$. Difference in gene expression Fold Change (FC) is plotted on the x-axis (log2 scale), and False Discovery Rate (FDR) adjusted significance is plotted on the y-axis (log10 scale). Genes significantly up- or downregulated are indicated in orange and blue, respectively. (**b**) Circular visualization of selected GO enriched pathways. Down- (blue dots) and up-regulated genes (red dots) within each GO pathway are plotted based on logFC. Z score bars indicate if an entire biological process is more likely to be increased or decreased based on the genes it comprises. (**c**) Table of the over-represented GO pathways of interest identified with Metascape using the DEGs from the selected clusters. The different categories of pathways are BP, Biological Process; CC, Cell Compartment. (**d**) Heatmaps from Metascape represented GO pathways of interest. Blue: low expression; orange: high expression. Raw data were submitted to BioStudies. Data are available via BioStudies with identifier S-BSST1157.



b

Term	ID	Genes
Visual perception	GO:0007601, GO:0001654, mmu04744, GO:0001750, R-MMU-	*Cnga1, Gnat1, Rgs9, Rs1, Tulp1, Unc119, Arr3, Rho, Slc24a1, Prcd, Prom1, Rax, Slc25a25, Tmem237*
Synpatic function	GO:0016188, GO:0098984, GO:0030658, GO:0051342	*Atp6v1b2, Unc13b, Slc17a7, rims2, Dmd, Mpdz, Ptprt, Homer3, Cpeb4, Egln1, Ablim1, Ap1b1, Calm2, Calm3,Sec31a, Cabp4, Egln1*
Cilium Assembly	R-MMU-5617833	*Dctn1, Arl6, Ift74, Bbs7, Atat1, Ocrl*
Protein homeostasis	GO:0036503, GO:2000058	*Dnajb2, Edem3, Brsk2, Ube4a, Dda1, Agtpbp1, Glmn*
Mitochondrial	GO:0031966	*Aifm1, Rhot1, Cox20, Slc25a23, Atpsckmt, Ndufaf5, Parl, Gpx4, Pet100*
Chromatin binding	GO:0003682	*Hmgn1, Nap1l4, Ncoa1, Mcm8, Pbrm1, Smarca2,Dhx30, Ccnt2, Kdm1a, Camta2, Kmt5b, Uimc1, Nrip1*
Regulation of RNA splicing	GO:0043484	*Clk4, Ccnl2, Pcbp4, Tra2a, Tmbim6, Rbm39*
Signal transduction	GO:0051020, GO:1901797	*Gdi2, Gnb1, Pex5l, Ranbp3, Cyrib, Arhgap1,Mdm2, Mdm4, Pcbp4, Kdm1a*

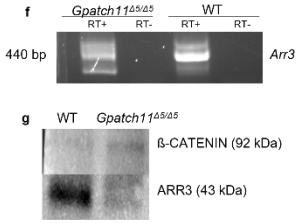
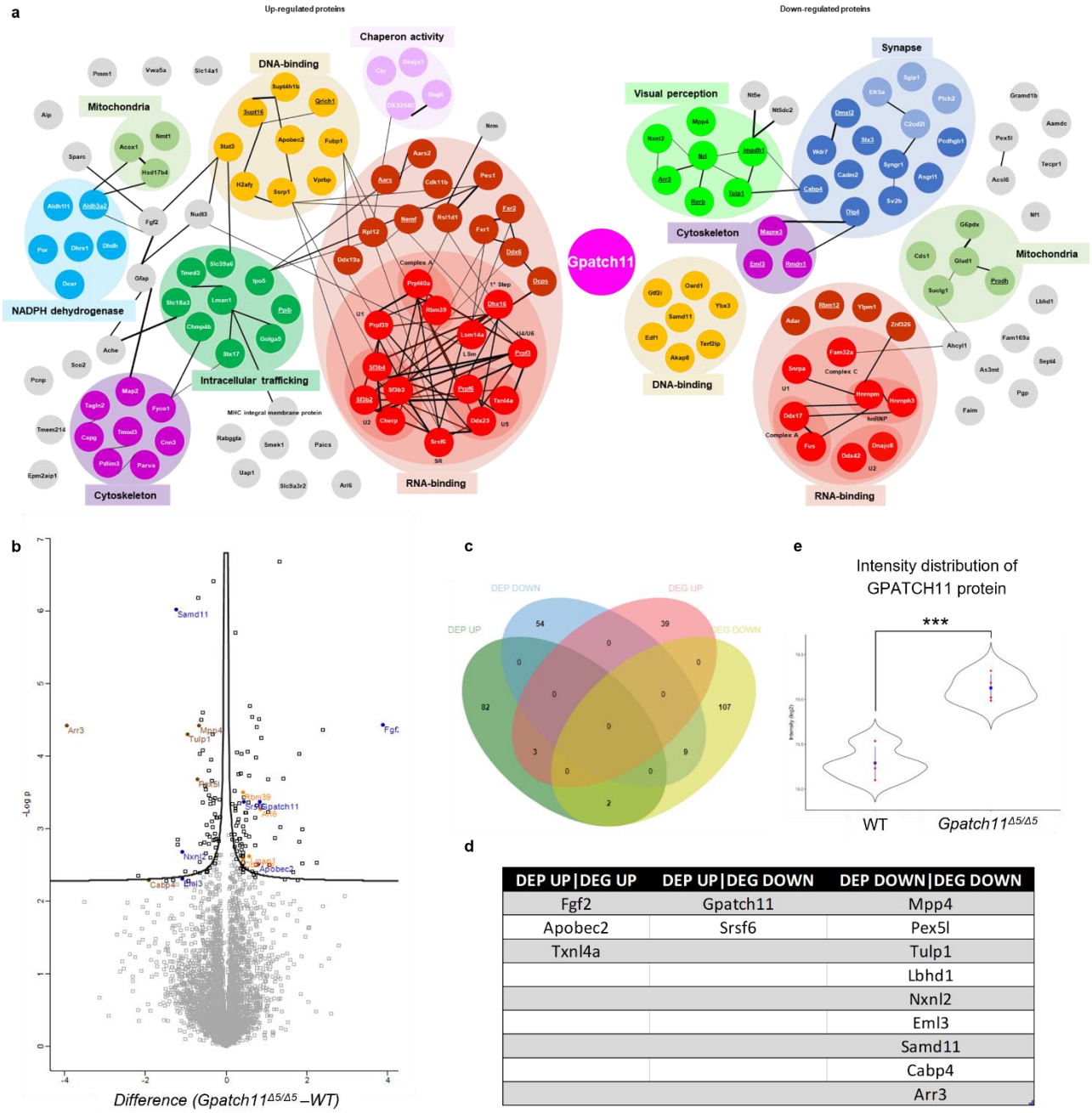


Fig.9: Identification of alternate splicing in retina from *Gpatch11$^{\Delta 5/\Delta 5}$* mice. (**a**) Chart from rMATs analysis showing that a majority of splicing events corresponded to skipped exons. A5SS and A3SS, alternative 5′ and 3′ splice sites; SE, skipped exons; RI, retained introns; MXE, mutually exclusive exons. (**b**) Table of the over-represented GO pathways of interest identified with Metascape using the DSGs from the selected clusters. (**c**) Venn diagrams comparing DEG and DSG in WT: *Gpatch11$^{WT/WT}$* and *Gpatch11$^{\Delta 5/\Delta 5}$* retina samples. The number of deregulated and mis-spliced transcripts is 12. (**d**) Heatmaps of the 12 deregulated and mis-spliced transcripts. Blue: low expression; orange: high expression. (**e**) Sashimi plots for the *Arr3* representing the alternate splicing event in retina from WT: *Gpatch11$^{WT/WT}$* (blue) and *Gpatch11$^{\Delta 5/\Delta 5}$* (red). Data are representative of the replicates for each experimental group. Orange highlights in the sashimi plots indicate the alternative splicing events and numbers correspond to the number of junctions reads for each event. (**f**) Electrophoresis of *Arr3* cDNA of WT: *Gpatch11$^{WT/WT}$* and *Gpatch11$^{\Delta 5/\Delta 5}$* retina samples after the PCR. (**g**) Western blot of ARR3 protein of mice retina from WT: *Gpatch11$^{WT/WT}$* and *Gpatch11$^{\Delta 5/\Delta 5}$*. Raw data were submitted to BioStudies. Data are available via BioStudies with identifier S-BSST1157.



d

DEP UP|DEG UP	DEP UP|DEG DOWN	DEP DOWN|DEG DOWN
Fgf2	Gpatch11	Mpp4
Apobec2	Srsf6	Pex5l
Txnl4a		Tulp1
		Lbhd1
		Nxnl2
		Eml3
		Samd11
		Cabp4
		Arr3

Fig.10: Mass Spectrometry analysis of murine retina lysates. (**a**) **S**chematic representation of protein interaction found in total lysate of *Gpatch11*$^{\Delta 5/\Delta 5}$ mice retina. (**b**) Volcano plots representation of differentially expressed proteins for *Gpatch11*$^{\Delta 5/\Delta 5}$ *vs* WT: *Gpatch11*$^{WT/WT}$ **(c)** Venn diagrams comparing DEG up-regulated and down-regulated and DEP up-regulated and down-regulated *Gpatch11*$^{\Delta 5/\Delta 5}$. DEP: differentially expressed proteins; DEG: differentially expressed genes. **(d)** Genes identified in the intersection of Venn diagram **(e)** Violin plot representing the intensity distribution of the GPATCH11 protein. Raw data are available via ProteomeXchange with identifier PXD041849. SR: Serine and arginine-rich; LSm: Like Sm.