Raquel Silva   ORCID iD: 0000-0002-1455-3082

Unique non-coding variants upstream of *PRDM13* are associated with a spectrum of developmental retinal dystrophies including Progressive Bifocal Chorioretinal Atrophy

Raquel S Silva[1,2,§], Gavin Arno[1,2,§] [,*], Valentina Cipriani[1-4], Nikolas Pontikos[1,2,4], Sabine Defoort-Dhellemmes[5], Ambreen Kalhoro[2], Keren J Carss[6,7], F Lucy Raymond[7,8], Claire Marie Dhaenens[9], Hanne Jensen[10], Thomas Rosenberg[10], Veronica van Heyningen[1,2], Anthony T Moore[1,2,11], Bernard Puech[5], Andrew R Webster[1,2,*]

[1]UCL Institute of Ophthalmology, University College London, London, UK

[2]Moorfields Eye Hospital, London, UK

[3] William Harvey Heart Centre, Queen Mary University, London, UK

[4]UCL Genetics Institute, University College London, London, UK

[5]Exploration de la Vision et Neuro-Ophtalmologie, Centre Hospitalier Universitaire, Lille, France

[6] Department of Haematology, University of Cambridge, Cambridge, UK

[7]NIHR BioResource - Rare Diseases, Cambridge University Hospitals, Cambridge Biomedical Campus, Cambridge, UK

[8]Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK

[9] Univ Lille, Inserm UMR-S 1172, CHU Lille, Biochemistry sand Molecular Biology Department - UF Génopathies, F-59000 Lille, France [10] Department of Ophthalmology, The Kennedy Eye Clinic, Rigshospitalet Glostrup, Denmark

[11]Ophthalmology Department, UCSF School of Medicine, San Francisco, CA, USA

[§]These authors contributed equally to this work

[*]To whom correspondence should be addressed.

Corresponding authors:

Andrew R. Webster, MD (Res), FRCOphth, UCL Institute of Ophthalmology, 11-43 Bath Street, London, EC1V 9EL, United Kingdom; email: andrew.webster@ucl.ac.uk

Gavin Arno, PhD, UCL Institute of Ophthalmology, 11-43 Bath Street, London, EC1V 9EL, United Kingdom; email: g.arno@ucl.ac.uk

Abstract

The autosomal dominant progressive bifocal chorioretinal atrophy (PBCRA) disease locus has been mapped to chromosome 6q14-16.2 which overlaps the North Carolina Macular Dystrophy (NCMD) locus MCDR1. NCMD is a non-progressive developmental macular dystrophy, in which variants upstream of *PRDM13* have been implicated. Whole genome sequencing was performed to interrogate structural variants (SVs) and single nucleotide variants (SNVs) in eight individuals, six affected individuals from two families with PBCRA and two individuals from an additional family with a related developmental macular dystrophy. A SNV (chr6:100,046,804 T>C), located 7.8kb upstream of the *PRDM13* gene, was shared by all PBCRA-affected individuals in the disease locus. Haplotype analysis suggested that the variant arose independently in the two families. The two affected individuals from family 3, were screened for rare variants in the PBCRA and NCMD loci. This revealed a *de novo* variant in the proband, 21 bp from the first SNV (chr6:100,046,783 A>C). This study expands the non-coding variant spectrum upstream of *PRDM13* and suggests altered spatio-temporal expression of *PRDM13* as a candidate disease mechanism in the phenotypically distinct but related conditions, NCMD and PBCRA.

Graphical Abstract
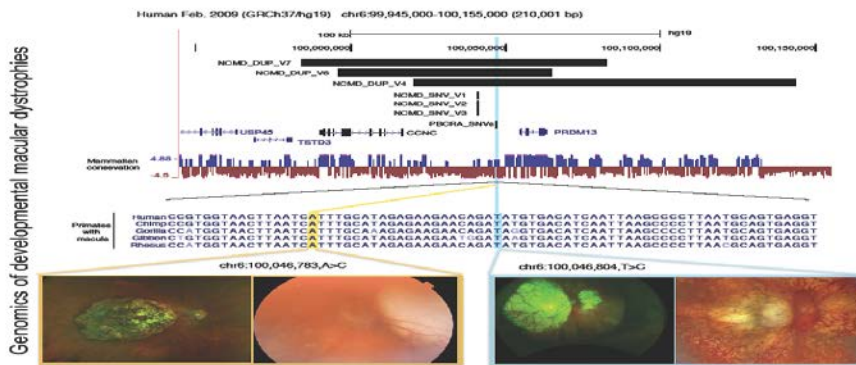
Two ultra-rare single nucleotide variants were identified in three unrelated families with a phenotypically related congenital macular dystrophies. This study expands on the non-coding variant spectrum upstream of PRDM13 and suggests altered spatio-temporal expression of PRDM13 as a candidate disease mechanism in the phenotypically distinct but related conditions.

1 Introduction

Progressive bifocal chorioretinal atrophy (PBCRA) (MIM# 600790) is a rare dominantly inherited condition originally described in a single British kindred by Douglas and colleagues in 1968 (Douglas, Waheed, & Wyse, 1968), corresponding to family 1 in this study. PBCRA is characterised by distinctive slowly progressing chorioretinal atrophic lesions at the macula evident in early life, with subsequent atrophy nasal to the optic nerve head. Affected individuals have life-long poor vision, nystagmus and myopia. Electroretinography reveals generalized rod and cone photoreceptor dysfunction (Douglas et al., 1968; Godley et al., 1996).

PBCRA maps to chromosome 6q14-16.2 (Kelsell et al., 1995) overlapping the Macular Dystrophy Region 1 (MCDR1, MIM# 136550) associated with North Carolina Macular Dystrophy (NCMD), another rare dominant developmental macular disorder. The close linkage assignment and some shared phenotypic features has led to speculation of a similar aetiology (Godley et al., 1996; Kelsell et al., 1995).However there are a number of key phenotypic differences between the disorders. Although both affect the macula from birth, PBCRA shows disease progression with electrophysiological evidence of widespread retinal dysfunction. In contrast, NCMD is a stationary disorder with restricted retinal involvement.

Initial investigations of the protein coding regions at this locus did not identify any definitive causative gene defects. Recently, the likely variants underlying NCMD at the MCDR1 locus have been elucidated: three novel non-coding SNVs located in a predicted DNase1 hypersensitivity site (DHS), in an intergenic region upstream of *PRDM13, TSTD3* and *CCNC* were identified in 12 NCMD families (Small et al., 2016). Additionally, three distinct tandem duplications encompassing *PRDM13* and its upstream region were also reported (Bowne et al., 2016; Manes et al., 2017; Small et al., 2016). Tandem duplications were also identified on a second locus at 5p21 (MCDR3, MIM# 608850), where the critical region for phenotype was narrowed down to a 39kb non-coding region (Cipriani, Silva, et al., 2017; Small et al., 2016). The disease mechanism at both loci is undetermined.

Here, we report the variant likely to be causative of PBCRA, a recurrent non-coding SNV located 7.8kb upstream of *PRDM13* in two families with clinically

defined PBCRA. In addition, we report a *de novo* variant in a proband and affected offspring affected by a similar developmental macular dystrophy.

2 Materials & Methods

2.1 Participants and retinal imaging

Affected individuals and unaffected family members were ascertained from the Inherited Eye Disease clinics (Moorfields Eye Hospital, London, United Kingdom), the Centre Hospitalier Regional Universitaire de Lille (France) and the Kennedy Eye Clinic, Department of Ophthalmology, Rigshospitalet Glostrup (Denmark). The study protocol adhered to the tenets of the Declaration of Helsinki and received approval from the local ethics committee. Written informed consent was obtained from all participants, or their parents, prior to their inclusion in the study. Where possible, patients underwent full clinical examination including slit lamp and biomicroscopic fundus examination. Color fundus photography was conducted using 35° (Topcon; Carl Zeiss) or 60° (Canon USA) retinal cameras or ultra–wide-field confocal scanning laser imaging (Optos). Fundus autofluorescence (FAF) was performed with 30° or 55° (Spectralis), or ultra–wide-field (Optos) imaging, with 488 nm (Spectralis) or 532nm (Optos) excitation wavelength. Spectral domain optical coherence tomography scans (SD-OCT, Spectralis; Cirrus) and kinetic visual fields (Goldmann or Octopus 900 Perimeter; Haag-Streit) were also performed.

2.2 Genomic analysis

Genomic DNA was purified from peripheral blood mononuclear cells of eight affected individuals from three unrelated families (Figure 1A and 2E) for whole-genome sequencing (WGS). WGS was performed independently or as part of the National Institute for Health Research BioResource – Rare Disease study (NIHR BR-RD) as previously described (Carss et al., 2017; Cipriani, Silva, et al., 2017), generating minimum coverage of 15X for 95% of the genome. Reads were aligned to the hg19 human reference sequence (build GRCh37) with novoalign (version 3.02.08). The aligned reads were sorted by base pair position and duplicates were marked using NovoSort. Variant calling from the aligned reads was performed using the GATK HaplotypeCaller. Variant Effect Predictor (VEP) was used for variant annotation. Following genome alignment, variant calling and annotation, only variants that passed standard quality filters were considered in subsequent analyses.

Variant filtering and prioritisation was performed based on rarity in publicly available datasets (Kaviar (Glusman, Caballero, Mauldin, Hood, & Roach, 2011), the Genome Aggregation Database (gnomAD, http://gnomad.broadinstitute.org/)), and the NIHR BR-RD dataset of 6,688 individuals.

In the first instance, likely pathogenic variants in a panel of 192 genes previously associated with inherited retinal disease were excluded as the cause (Carss et al., 2017). In addition, all known genetic causes of NCMD at the MCDR1 and MCDR3 loci were interrogated.

Subsequent analysis was focused on the previously defined PBCRA linkage region (Kelsell et al., 1995) (D6S249;D6S301: GRCh37 [hg19] chr6:98,117,898-103,695,199). Shared rare (MAF<0.01 control datasets) variants in the linkage region were ascertained from the 4 affected individuals from family 1. Assuming a genetic cause affecting the same gene as family 1, the remaining affected individuals were analysed in the same way. For this extremely rare dominantly-inherited severe congenital visual disorder, a variant within the linked region was considered a likely candidate if absent from gnomAD and the NIHR BR-RD control cohort. Ultra-rare candidate structural rearrangements were also investigated as previously described (Carss et al., 2017). The disease associated variants have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) under accession number SUB5050319.

Six affected individuals (2 from family 1 and 4 from family 2) were selected for SNP-array genotyping using the Illumina HumanOmniExpress-24 v1.0 beadchip (Illumina, Inc., San Diego, CA, USA). Genotypes were determined using the Genotyping Module in the Illumina GenomeStudio v2011.1 software and used for haplotype sharing analysis as previously described (Cipriani, Kalhoro, et al., 2017; Cipriani, Silva, et al., 2017).

2.3 *In silico* analysis

Intergenic candidate variants were investigated for potential effect on gene regulation as a possible mechanism of disease. UCSC browser (Kent et al., 2002) (http://genome.ucsc.edu/, Santa Cruz University, California) was

queried for relevant tracks at loci of interest, namely presence of regulatory elements, repetitive elements and cross species conservation.

The Encyclopaedia of DNA Elements (ENCODE) (Dunham et al., 2012; ENCODE Project Consortium, 2012; Sloan et al., 2016) was interrogated for fetal retina datasets of interest. Bed files from DNA-seq datasets (ENCFF249FGP, ENCFF937NUZ, ENCFF401BCF, ENCFF591NRB, ENCFF265ZNN, John Stamatoyannopoulos lab) were downloaded and exported to R Studio (R version 3.5.0 (2018-04-23)). Coordinates of interest were extracted for analysis (chr6:100,040,000-100,050,000).

Furthermore, the online tool "Search for TF Binding Sites" from UniPROBE (Hume, Barrera, Gisselbrecht, & Bulyk, 2015), as previously described in (Farley, Olson, Zhang, Rokhsar, & Levine, 2016), was used for prediction of transcription factor binding motifs. (online tool available at http://the_brain.bwh.harvard.edu/uniprobe/index.php, last updated on 2018-03-05). The default settings were used (Score Threshold set at 0.45 and all species were selected). The DNA sequence queried was of 20 bp on each flaking side of the identified variant, so total of 41 nt queried per variant. The virtual Hi-C browser (http://promoter.bx.psu.edu/hi-c/chic.php) (Wang et al., 2017) was interrogated for DNA-DNA contacts, namely promoter contacts with identified non-coding SNV coordinates. Datasets used include Capture Hi-C and 4-C from human species (hg19). The target sequence for both datasets was chr6:100,046,783-100,046,804. The extended region was used as default, set at +/-500kb from the queried site.

3 Results

3.1 Clinical findings

The clinical phenotypes of fifteen affected members of family 1 (GC4059, Figure 1a) have been previously described (Godley et al., 1996). Three affected individuals were re-examined and also underwent widefield fundus imaging (Optos) FAF and SD-OCT. Representative imaging is shown in Figure 1. Five affected members of family 2 (GC21086, Figure 1a) (3 generations, aged 2 months to 56 years), from northern France, were ascertained. The longest follow-up of these individuals was over four decades (II:2, proband). Affected members presented with the same progressive stages of disease as described in Family 1 (Godley et al., 1996). Briefly, from birth there was significant visual impairment, photophobia and nystagmus. The main lesion of the fundus is located in the macular region, limited to within the retinal vascular arcades of the posterior pole (stage 1, Figure 1j). Further smaller atrophic lesions develop progressively around this central lesion. Subsequently in stage 2 there is slow progression of the chorioretinal atrophy with structurally intact retina remaining at the peripapillary region and at the extreme periphery (Figure 1g and i). A second smaller focus of chorioretinal atrophy presents later in the second decade of life and overtime there is further progression of the atrophic lesions during stage 3 of disease.

There were two affected individuals, a mother and her child, in family 3 (GC20008, Figure 2). The mother (II:2), a 38-year-old mixed East Asian/Caucasian female born to normally sighted parents, had the typical features of NCMD but the child (III:1), had more severe macular atrophy with

associated nystagmus, altogether findings more typical of PBCRA. At age 3 the child developed a retinal detachment requiring surgical repair. There has been no observed progression of the macular lesion.

3.2 WGS and variant prioritization identifies putative PBCRA-associated variant

55 rare variants (MAF <0.01) were shared across the linkage region (D6S249-D6S301: chr6:98,117,898-103,695,199) in family 1, 72 variants were shared in family 2 over the same genomic interval (Supp.Table S1 and Supp.Table S2). Further filtering focused on identifying ultra-rare heterozygous variants on the disease haplotype in each family. Seven and 11 ultra-rare SNVs (absent from gnomAD) were identified in family 1 and 2 respectively. A single candidate variant was shared by the two families. This heterozygous non-coding SNV (chr6:100,046,804T>C) is located 7.8 kb upstream from the transcriptional start site (TSS) of *PRDM13* (Figure 3).

Given the phenotypic similarity observed in family 3 to both NCMD and PBCRA, it was hypothesised that the disease associated genetic variant may be located within the associated loci. Using the WGS SNV and SV call data and manual interrogation of the paired-end reads using the Integrative Genomics Viewer (IGV (Thorvaldsdóttir, Robinson, & Mesirov, 2013),) all known NCMD associated variants were excluded at both the chr5 and chr6 loci (Bowne et al., 2016; Cipriani, Silva, et al., 2017; Manes et al., 2017; Small et al., 2016). At the PBCRA/MCDR1 locus, 75 SNV variants with MAF<0.01 (and allele count <3 from NIHR dataset), were shared, of which 18 were absent from gnomAD (Supp.Table S3). One ultra-rare SNV was 21

bp away from the variant identified in families 1 and 2 (chr6:100,046,783A>C).

The identified putatively causative SNVs were not seen in any unaffected members of each family available for genotyping (Figure 1a and Supp.Table S4, additional family screening performed by direct Sanger sequencing. Primers available upon request). 3.3 Haplotype sharing analysis suggests the PBCRA variant arose independently in family 1 and 2

Given the rarity of the phenotype and the discovery of the same variant in two not knowingly related families, we sought to determine whether a common ancestral haplotype could account for the disease. Haplotype sharing analysis (Cipriani, Kalhoro, et al., 2017; Cipriani, Silva, et al., 2017) of SNP-array data showed that the families had a different chromosomal background at the 6q linkage region (Supp.Figure S1). Furthermore, WGS enables the characterisation of every SNV (common or rare) in the linkage region in both families. As described above, interrogation of the WGS variant calls showed complete absence of shared low frequency variants other than chr6:100,046,804T>C identified between family 1 and family 2 (Supp.Table S1 and 2). The closest flanking rare (MAF <0.01) SNVs to the single shared variant being discordant between the families were 391,196bp away (centromeric, chr6:99,655,608) and 65,565bp away (telomeric, chr6:100,112,368) in family 1 and 150,890bp (centromeric, chr6:99,895,915) and 80,676bp (telomeric, chr6:100,127,479) in family 2 (Supp.Table S1 and S2). Taken together, these data suggest the variant arose independently in each family.

3.4 Haplotype analysis suggests the variant was *de novo* in family 3

Neither parent of, II:2 was clinically affected. Unaffected family members available for genotyping (I:2, II:3, II:4) did not harbour the variant, therefore we hypothesised that the variant arose in the proband (II:2). To test this, rare (MAF<0.01) SNV calls from the two affected patients in family 3 (II:2 and III:1) were used to identify the haplotype likely to harbour the chr6:100,046,783A>C variant: ten rare heterozygous SNV calls were selected spanning 0.5Mb (chr6:99,702,081-100,201,085) flanking the variant (Supp.Table S3), these SNVs are enriched in East Asian alleles in gnomAD in keeping with the maternal East Asian background. Genotyping of these SNV by direct Sanger sequencing in the available unaffected individuals (I:2, II:3, II:4) confirmed the haplotype was maternally inherited and the variant was *de novo* in the proband (Supp.Table S4). In light of this finding and the discordant phenotypes in the family, possible mosaicism in the proband (II:2) was investigated: the paired-end reads showed abnormal zygosity in II:2 based a skewed read ratio of reference/alternate (31/9 reads, 78/22%) nucleotide at this position in comparison to the affected offspring (III:1, 21/15 reads, 58/42%) (Supp.Figure S2) and the surrounding heterozygous variants (data not shown). This finding was confirmed by orthogonal testing (direct Sanger sequencing) confirming the skewed zygosity (Supp.Figure S2).

3.5 In silico prediction of the effect of the identified variants in retinal development

Because of the intergenic location and close proximity of the putative causative variants to each other, it was hypothesised that the likely

mechanism of disease was their effect on a gene regulatory region. The region encompassing the variants was investigated for transcription factor occupancy, active chromatin signatures, chromosome conformation, and the gene expression patterns of implicated genes.

The affected nucleotide in family 1 and 2 (chr6:100,046,804T) is evolutionarily conserved within primates and some mammals (Figure 3). Transcription factor binding site prediction using the UniPROBE tool suggested the T>C transition causes a gain of CBF1, MLX, Tcfec, Smad3, Max, Bhlhb2, HLH-25, MXL-1_MDL-1 and HLH-1 binding sites (Supp.Figure S3). There were no predicted losses of binding sites. The affected nucleotide identified in family 3 (chr6:100,046,783A) is conserved in all available primates and mammals where this sequence is present (Figure 3). Transcription factor binding site prediction using the same tool suggested a gain of HOXD13 and a loss of several TFs binding sites, including members of Lhx family. In addition, several binding sites are not lost, but the core binding motif is shorter due to the alteration, which would also affect the strength of binding. One such example is the photoreceptor specific CRX site (Supp.Figure S4).

DNase-seq datasets provide information on open chromatin regions, usually associated with active genes and regulatory regions. Datasets from human fetal retina tissue for 5 available developmental stages (Day 74, 79, 85, 103 and 125) were available on ENCODE repository (Dunham et al., 2012; ENCODE Project Consortium, 2012; Sloan et al., 2016). Five different sites were identified to be open/active at two different developmental stages within

the 10 kb interval analysed, specifically at D103 and D125 (Figure 4 and Supp.Table S5). One active site at D103 in particular, encompassed chr6:100,046,783 (identified in Family 3) and is situated 14bp away from variant chr6:100,046,804 (identified in Family 1 and 2). The earlier developmental stages (D74, D79 and D85) had no open sites in the same region (Figure 4 and Supp.Table S5).

Chromosome conformation capture data maps the frequency of contact between distal DNA elements, which is important in the 3D structure of genome organization and critical for normal biological function (Franke et al., 2016). Available data from human embryonic stem cells (hESCs) and neural progenitor cells identified a number of possible interaction regions (Wang et al., 2017) with the candidate variants. The predicted interactions converged on *PRDM13*, *TSTD3*, *USP45*, *FAXC*, *POU3F2*, *SIM1*, *CCNC* and *PNISR*. Of this set only *PRDM13*, *FAXC* and *POU3F2* have enriched expression in retinal tissue, whereas *TSTD3* and *USP45* have a moderately enriched expression in fetal retinal pigment epithelium (RPE) (Supp.Figure S5, EyeIntegration, (Hufnagel, Brooks, & McGaughey, 2017)).

4 Discussion

We report findings of genomic investigations in three unrelated families with PBCRA or a related developmental macular phenotype and identify putative disease associated variants located 7.8 Kb from the *PRDM13* transcription start site and 5.7 Kb from the previously reported V1-V3 variant cluster associated with NCMD (Small et al., 2016)(Figure 3).

PRDM13 is part of the highly conserved positive regulatory (PR) domain containing protein family, which drives context specific transitions fundamentally important in developmental cues (Hohenauer & Moore, 2012). PRDM members act either by directing histone methyltranferase activity (PR domain) or by recruiting histone-modifying complexes, including co-repressors, to target promoters by sequence specific DNA binding-zinc fingers (Hanotel et al., 2014; Hohenauer & Moore, 2012). PRDM13 has been shown to be expressed in hindbrain, spinal cord, retina and olfactory placode, in developing zebrafish (Sun et al., 2008). In the mouse, expression is reported in dorsal spinal cord progenitor (Kinameri et al., 2008) and retinal amacrine cells, with a role in determining excitatory and inhibitory neurons (Mona et al., 2017; Watanabe et al., 2015). *PRDM13* is also found highly transcribed in microarray data from fetal human retina (Kozulin & Provis, 2009).

PRDM13 has been previously reported as the likely gene affected in NCMD (Small et al., 2016). The spectrum of variants associated with disease at this locus suggests that rather than loss of function, *PRDM13* expression may be spacio-temporally dysregulated. Overexpression models suggest PRDM13 has a toxic effect on photoreceptors in pre-natal mice with no perturbation of amacrine cell lineages (Goodson et al., 2017); it was also shown to impair eye development in *Drosophila* and *Xenopus* (Bessodes, Parain, Bronchain, Bellefroid, & Perron, 2017; Manes et al., 2017). Conversely, murine knockout of *Prdm13* results only in alteration of amacrine cell development without

functional vision loss (Watanabe et al., 2015). Nevertheless, the effect of the disease associated variants on the expression of PRDM13 remains to be characterised.

In light of this, we investigated the potential of the ultra-rare variants to alter *PRDM13* transcription in human developing retina, *in silico*. DNase-seq datasets from *wild type* human fetal retina (Dunham et al., 2012; ENCODE Project Consortium, 2012; Sloan et al., 2016, Figure 4) showed that in contrast to other cell types and tissues, the NCMD V1-V3 variants region was not marked as active. In fact, the closest DHS to the V1-V3 region is at chr6:100,040,740–100,040,890 (16 bp from V1 and 150 bp from V3) and is active at D103. Similarly, one DHS site that harbours the *de novo* variant from family 3 and is located 14 bp from the PBCRA variant (at chr6:100,046,640-100,046,790) is active at the same developmental stage (D103). This stage is compatible with when retinal progenitor cells of the central retina exit mitosis (J M Provis, van Driel, Billson, & Russell, 1985) and commit to lineage. The proximity of the variants to DHS sites open at the same restricted developmental stage may provide a clue to the disease mechanism and warrants further investigation.

The disease-associated SNVs could potentially alter the target sequence for DNA-binding transcriptional regulators and thus dysregulate transcription of *PRDM13*. The gain of a Bhlhb2 site with the PBCRA variant (Supp.Figure S4) may be of relevance given the finding that PRDM13 and other PRDM members co-bind DNA with Ebox proteins/BHLHb members to exert their repression function (Mona et al., 2017; Ross et al., 2012). Different members

of PRDM family have self and auto-regulation strategies (Mona et al., 2017). Furthermore, LHX1/5 have been shown as part of the downstream indirect targets of PRDM13 in neural tube formation (Chang et al., 2013; Mona et al., 2017) and in Family 3, the putative SNV causes a loss of a binding site for members of this family of neural regulators (Supp.Figure S5). Without a unifying mechanism however, it is difficult to conclude which transcription factor binding alteration (if any) may be relevant. This uncertainty is further compounded by the predictive nature of this method and by the lack of human data available for developing retinal tissue. Further experimental approaches are therefore required in order to characterise the precise TF occupancy in the context of human macular development.

It is noteworthy that the tandem duplications encompassing *PRDM13* lead to a similar NCMD phenotype as V1-V3 and not PBCRA. This may suggest that the milder non-progressive NCMD may be a result of over-expression of *PRDM13* during development whereas the more severe and progressive PBCRA may result from a more dramatic dysregulation of *PRDM13*. In addition, the variability of the disorder in family 3 suggests that NCMD/PBCRA are part of a disease spectrum, and evidence of mosaicism in the proband with a milder disease may provide additional testimony for a toxic effect of over expression of PRDM13 with a stronger effect of the PBCRA variants. There is however phenotypic heterogeneity within NCMD families with many affected individuals being asymptomatic, this may indicate that the *trans* allele or other modifiers may play a role in disease severity at least for NCMD associated variants.

Macular development is a challenging pathway to analyse given its primate-specificity and extended developmental period from the early foetus to final maturation during the first years of life (Jan M Provis, Dubis, Maddess, & Carroll, 2013). It remains to be elucidated at which point(s) during development expression of *PRDM13* is pertinent for macular development. Recent evidence suggests Carnegie Stage (CS) 22 is the start of human macular specification influenced by retinoic acid signalling (da Silva & Cepko, 2017) and hinting at the initial timepoints to investigate the start of these disorders, and potential solutions to use human-derived retinal organoids to model high acuity regions (da Silva & Cepko, 2017).

In conclusion, we provide evidence for two disease-causing variants in three unrelated families with a phenotypically related congenital macular dystrophies. Taken together with previous findings in NCMD families, this suggests a key role for non-coding variants upstream of *PRDM13* in developmental macular dystrophy. Exploring the function and precise target of these putative regulatory regions will be essential for mechanistic understanding of developmental macular dystrophy and for investigating the potential role of *PRDM13* in normal retinal and macular development.

Acknowledgments

Conflicts of interest

The authors declare no conflict of interest.

Figures

Figure 1 – Family pedigrees and retinal imaging of PBCRA families 1 and 2. **A:** Simplified pedigrees of two families (family 1: GC4059 (Godley et al., 1996), family 2: GC21086) demonstrating autosomal dominant inheritance of PBCRA, black solid and open symbols indicate affected and unaffected individuals respectively, arrows indicate the probands, * indicate subjects of WGS analysis. **B:** Optos widefield imaging, **C:** Optos widefield FAF and **D:** horizontal and vertical SD-OCT scans of GC4059 III:8 (proband) showing bifocal chorioretinal atrophic lesions and intact retina in the unaffected peripapillary region. **E:** Optos widefield imaging and **F:** Optos widefield FAF in the affected daughter (IV:5) showing similar distribution of chorioretinal atrophy. **G:** Color fundus montage, **H:** Heidelberg FAF and **I:** SD-OCT scan of GC21086 III:3 at age 18 years illustrating stage 2 of the disease with bifocal chorioretinal atrophy and intact retinal structure at the peripapillary region. **J:** Color fundus montage images of stage 1 chorioretinal atrophic lesions in IV:1 and IV:2 at age 7 months.
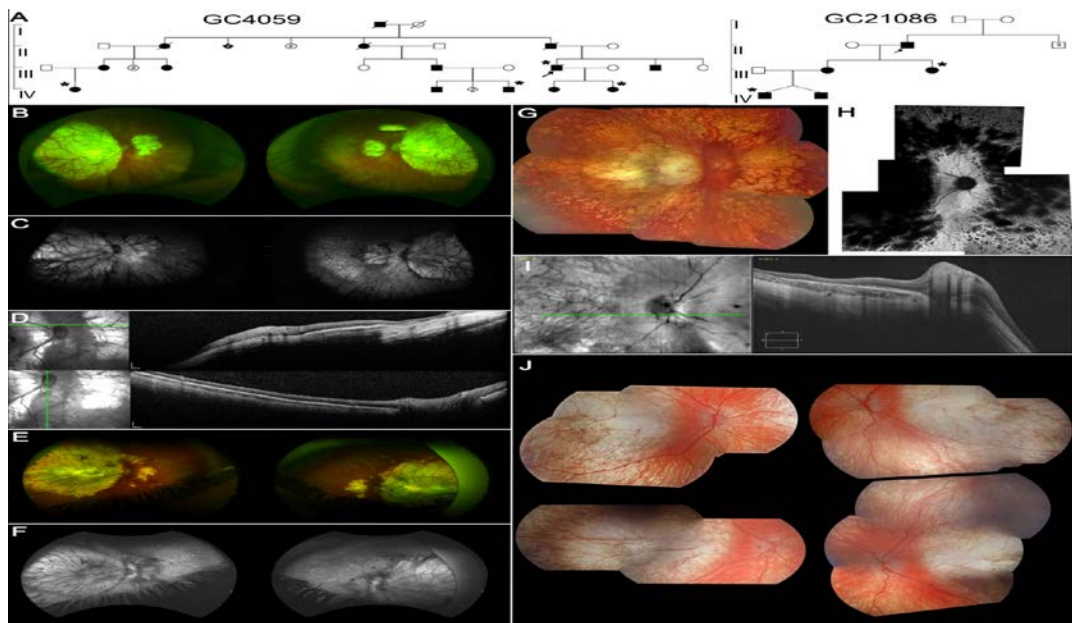
Figure 2. **Retinal imaging and pedigree of NCMD/PBCRA family 3. A:** Optos widefield imaging, **B:** Optos widefield FAF and **C:** Heidelberg SD-OCT imaging of II:2 (proband) demonstrating extensive macular NCMD-like lesion. **D:** Topcon fundus imaging of III:1 at age 8 demonstrating PBCRA-like lesion extending peripheral to the retinal arcades. **E:** pedigree of family 3 (GC20008) * indicates subjects that underwent WGS analysis.
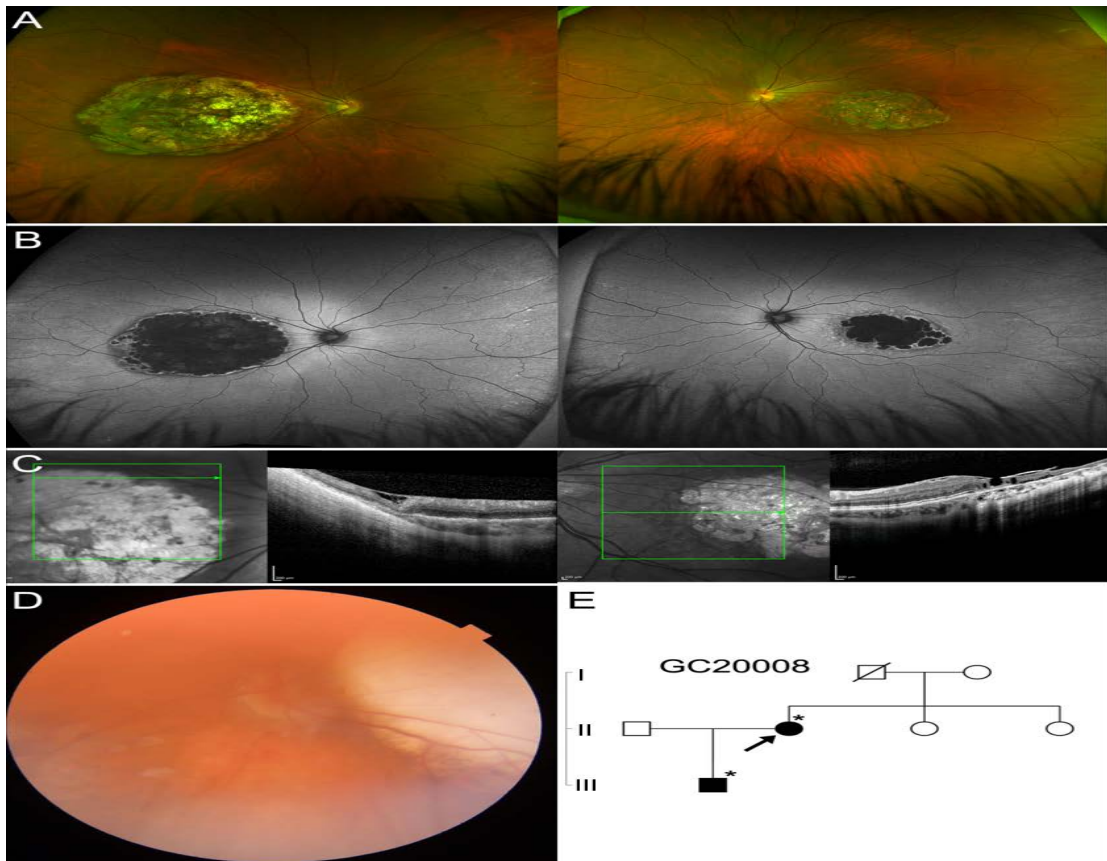
Figure 3 – Schematic representation of variants found in NCMD and PBCRA at the 6q16 locus within the context of genome sequence and conservation. Two vertical bars highlighted in grey indicate the variants identified in Famiies 1-3 (GRCh37/hg19 chr6:100,046,804T in Family 1 and 2; chr6:100,046,783A in Family 3) 7.8 Kb from the *PRDM13* transcription start site and 5.7 Kb downstream of the three non-coding single-nucleotide variants (V1-V3,Small et al., 2016) (chr6:100,040,906, chr6:100,040,987, chr6:100,041,040, respectively). V4, V6 and V7 denote previously identified tandem duplications that overlap V1-V3 SNVs and the genes *CCNC* and *PRDM13* (Bowne et al., 2016; Manes et al., 2017; Small et al., 2016).V5 is not depicted on this schematic since it is located at the MCDR3 locus (chr5q, Small et al., 2016) . V6 and V7 refer to recently identified SV associated with NCMD (references). A zoomed region is presented in the bottom panel comparing sequence similarity, including across primate species with macula.
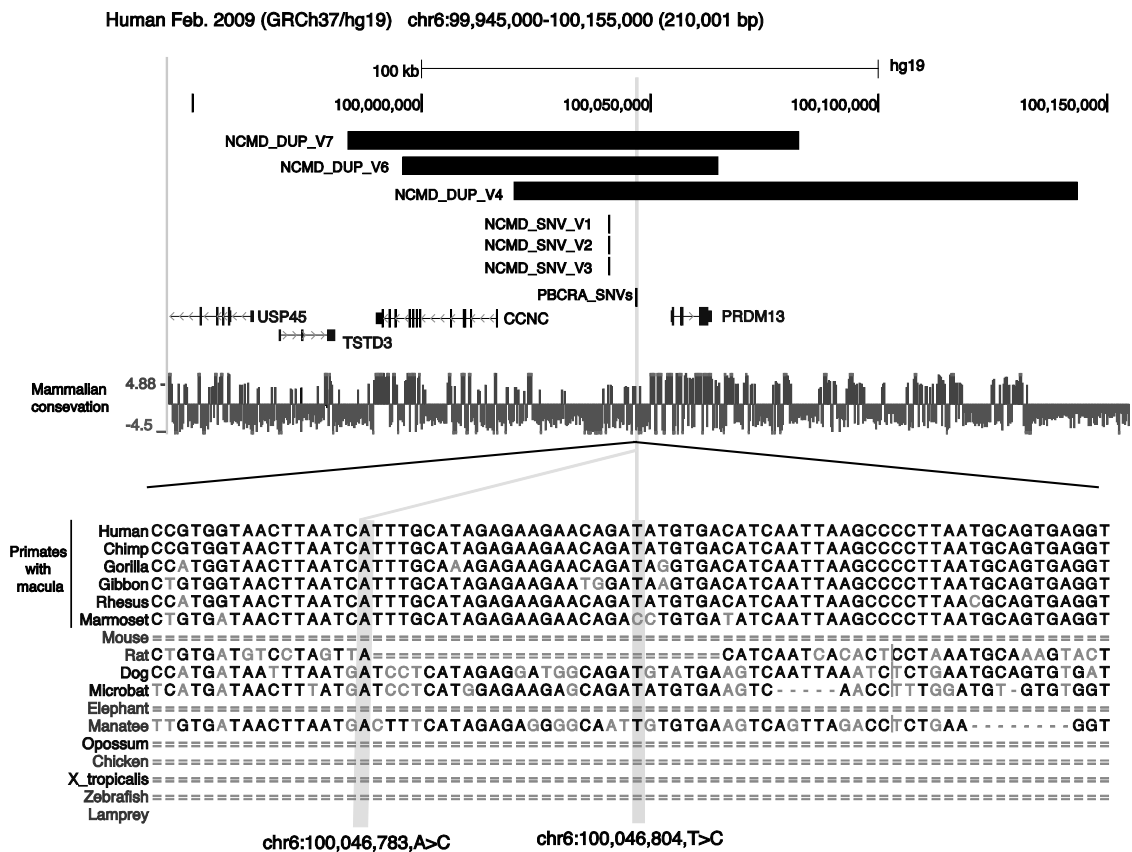
Figure 4 –Open chromatin sites in human fetal retina in chr6:100,040,000-100,050,000 differ to open sites in 125 human cell-lines. The putative variants identified are highlighted in a vertical grey bar. From the 5 available ENCODE DNAse-seq datasets (Dunham et al., 2012; ENCODE Project Consortium, 2012; Sloan et al., 2016), DNase Hypersensitive Sites (DHS) were only identified as active at the upstream non-coding region of *PRDM13* for fetal stage on day 103 (D103) and day 125 (D125). One site active at D103 is located 14 bp from the PBCRA variant. UCSC browser was used to visualise and create the display and viewing window (chr6:100,039,950-100,050,050). The track for repetitive elements from Repeat Masker and DHS sites inferred from 125 cell-lines is also displayed (DNase I HS). Interestingly, NCMD V1-V3, located in a DHS site in these cell lines, were not active in any of the 5 developmental stages investigated in human fetal retina. NCMD-V1 is located 16 bp from the nearest DHS active in fetal retina, also at D103. Genome build GRCh37/hg19.

Human Feb. 2009 (GRCh37/hg19) chr6:100,039,950-100,050,050 (10,101 bp)
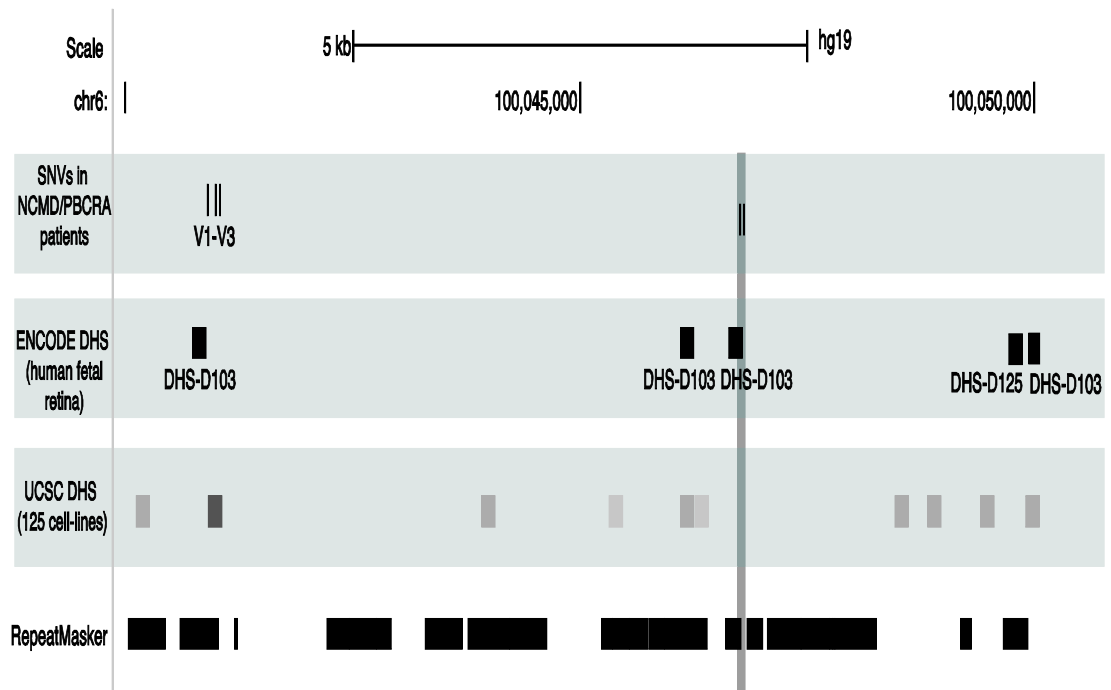
References

Bessodes, N., Parain, K., Bronchain, O., Bellefroid, E. J., & Perron, M. (2017). Prdm13 forms a feedback loop with Ptf1a and is required for glycinergic amacrine cell genesis in the Xenopus Retina. *Neural Development*, *12*(1), 16. https://doi.org/10.1186/s13064-017-0093-2

Bowne, S. J., Sullivan, L. S., Wheaton, D. K., Locke, K. G., Jones, K. D., Koboldt, D. C., … Daiger, S. P. (2016). North Carolina macular dystrophy (MCDR1) caused by a novel tandem duplication of the PRDM13 gene. *Molecular Vision*, *22*, 1239–1247. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/27777503

Carss, K. J., Arno, G., Erwood, M., Stephens, J., Sanchis-Juan, A., Hull, S., … Raymond, F. L. (2017). Comprehensive Rare Variant Analysis via Whole-Genome Sequencing to Determine the Molecular Pathology of Inherited Retinal Disease. *The American Journal of Human Genetics*, *100*(1), 75–90. https://doi.org/10.1016/j.ajhg.2016.12.003

Chang, J. C., Meredith, D. M., Mayer, P. R., Borromeo, M. D., Lai, H. C., Ou, Y.-H., & Johnson, J. E. (2013). Prdm13 Mediates the Balance of Inhibitory and Excitatory Neurons in Somatosensory Circuits. *Developmental Cell*, *25*(2), 182–195. https://doi.org/10.1016/j.devcel.2013.02.015

Cipriani, V., Kalhoro, A., Arno, G., Silva, R. S., Pontikos, N., Puech, V., … Puech, B. (2017). Genome-wide linkage and haplotype sharing analysis implicates the MCDR3 locus as a candidate region for a developmental macular disorder in association with digit abnormalities. *Ophthalmic Genetics*, 1–9. https://doi.org/10.1080/13816810.2017.1289544

Cipriani, V., Silva, R. S., Arno, G., Pontikos, N., Kalhoro, A., Valeina, S., … Moore, A. T. (2017). Duplication events downstream of IRX1 cause North Carolina macular dystrophy at the MCDR3 locus. *Scientific Reports*, *7*(1), 7512. https://doi.org/10.1038/s41598-017-06387-6

da Silva, S., & Cepko, C. L. (2017). Fgf8 Expression and Degradation of Retinoic Acid Are Required for Patterning a High-Acuity Area in the Retina. *Developmental Cell*, *42*(1), 68–81.e6. https://doi.org/10.1016/j.devcel.2017.05.024

Douglas, A. A., Waheed, I., & Wyse, C. T. (1968). Progressive bifocal chorio-retinal atrophy. A rare familial disease of the eyes. *The British Journal of Ophthalmology*, *52*(10), 742–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5686965

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., … Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

ENCODE Project Consortium, T. E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Farley, E. K., Olson, K. M., Zhang, W., Rokhsar, D. S., & Levine, M. S. (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *Proceedings of the National Academy of Sciences*

of the United States of America, *113*(23), 6508–13. https://doi.org/10.1073/pnas.1605085113

Franke, M., Ibrahim, D. M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., … Mundlos, S. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, *538*(7624), 265–269. https://doi.org/10.1038/nature19800

Glusman, G., Caballero, J., Mauldin, D. E., Hood, L., & Roach, J. C. (2011). Kaviar: an accessible system for testing SNV novelty. *Bioinformatics*, *27*(22), 3216–3217. https://doi.org/10.1093/bioinformatics/btr540

Godley, B. F., Tiffin, P. A., Evans, K., Kelsell, R. E., Hunt, D. M., & Bird, A. C. (1996). Clinical features of progressive bifocal chorioretinal atrophy: a retinal dystrophy linked to chromosome 6q. *Ophthalmology*, *103*(6), 893–898. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8643244

Goodson, N. B., Nahreini, J., Randazzo, G., Uruena, A., Johnson, J. E., & Brzezinski, J. A. (2017). Prdm13 is required for Ebf3+ amacrine cell formation in the retina. *Developmental Biology*, *434*(1), 149–163. https://doi.org/10.1016/j.ydbio.2017.12.003

Hanotel, J., Bessodes, N., Thélie, A., Hedderich, M., Parain, K., Driessche, B. Van, … Bellefroid, E. J. (2014). The Prdm13 histone methyltransferase encoding gene is a Ptf1a-Rbpj downstream target that suppresses glutamatergic and promotes GABAergic neuronal fate in the dorsal neural tube. *Developmental Biology*, *386*(2), 340–357. https://doi.org/10.1016/j.ydbio.2013.12.024

Hohenauer, T., & Moore, A. W. (2012). The Prdm family: expanding roles in stem cells and development. *Development (Cambridge, England)*, *139*(13), 2267–82. https://doi.org/10.1242/dev.070110

Hufnagel, J. M., Brooks, R. B., & McGaughey, B. P. (2017). Identifying core biological processes distinguishing human eye tissues with systems-level gene expression analyses and weighted correlation networks. https://doi.org/10.1101/136960

Hume, M. A., Barrera, L. A., Gisselbrecht, S. S., & Bulyk, M. L. (2015). UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein–DNA interactions. *Nucleic Acids Research*, *43*(D1), D117–D122. https://doi.org/10.1093/nar/gku1045

Kelsell, R. E., Godley, B. F., Evans, K., Tiffin, P. A. C., Gregory, C. Y., Plant, C., … Hunt, D. M. (1995). Localization of the gene for progressive bifocal chorioretinal atrophy (PBCRA) to chromosome 6q. *Human Molecular Genetics*, *4*(9), 1653–1656. https://doi.org/10.1093/hmg/4.9.1653

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006. https://doi.org/10.1101/gr.229102. Article published online before print in May 2002

Kinameri, E., Inoue, T., Aruga, J., Imayoshi, I., Kageyama, R., Shimogori, T., & Moore, A. W. (2008). Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse

neurogenesis. *PLoS ONE*, *3*(12). https://doi.org/10.1371/journal.pone.0003859

Kozulin, P., & Provis, J. M. (2009). Differential gene expression in the developing human macula: Microarray analysis using rare tissue samples. *Journal of Ocular Biology, Diseases, and Informatics*, *2*(4), 176–189. https://doi.org/10.1007/s12177-009-9039-1

Manes, G., Joly, W., Guignard, T., Smirnov, V., Berthemy, S., Bocquet, B., … Meunier, I. (2017). A novel duplication of PRMD13 causes North Carolina macular dystrophy: overexpression of PRDM13 orthologue in drosophila eye reproduces the human phenotype. *Human Molecular Genetics*, *26*(22), 4367–4374. https://doi.org/10.1093/hmg/ddx322

Mona, B., Uruena, A., Kollipara, R. K., Ma, Z., Borromeo, M. D., Chang, J. C., & Johnson, J. E. (2017). Repression by PRDM13 is critical for generating precision in neuronal identity. *ELife*, *6*. https://doi.org/10.7554/eLife.25787

Provis, J. M., Dubis, A. M., Maddess, T., & Carroll, J. (2013). Adaptation of the central retina for high acuity vision□: Cones , the fovea and the avascular zone. *Progress in Retinal and Eye Research*, *35*, 63–81. https://doi.org/10.1016/j.preteyeres.2013.01.005

Provis, J. M., van Driel, D., Billson, F. A., & Russell, P. (1985). Development of the human retina: patterns of cell distribution and redistribution in the ganglion cell layer. *The Journal of Comparative Neurology*, *233*(4), 429–51. https://doi.org/10.1002/cne.902330403

Ross, S. E., McCord, A. E., Jung, C., Atan, D., Mok, S. I., Hemberg, M., … Greenberg, M. E. (2012). Bhlhb5 and Prdm8 Form a Repressor Complex Involved in Neuronal Circuit Assembly. *Neuron*, *73*(2), 292–303. https://doi.org/10.1016/j.neuron.2011.09.035

Sloan, C. A., Chan, E. T., Davidson, J. M., Malladi, V. S., Strattan, J. S., Hitz, B. C., … Cherry, J. M. (2016). ENCODE data at the ENCODE portal. *Nucleic Acids Research*, *44*(D1), D726-32. https://doi.org/10.1093/nar/gkv1160

Small, K. W., DeLuca, A. P., Whitmore, S. S., Rosenberg, T., Silva-Garcia, R., Udar, N., … Stone, E. M. (2016). North Carolina Macular Dystrophy Is Caused by Dysregulation of the Retinal Transcription Factor PRDM13. *Ophthalmology*, *123*(1), 9–18. https://doi.org/10.1016/j.ophtha.2015.10.006

Sun, X. J., Xu, P. F., Zhou, T., Hu, M., Fu, C. T., Zhang, Y., … Chen, Z. (2008). Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. *PLoS ONE*, *3*(1). https://doi.org/10.1371/journal.pone.0001499

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Wang, Y., Zhang, B., Zhang, L., An, L., Xu, J., Li, D., … Yue, F. (2017). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. *BioRxiv*, 112268. https://doi.org/10.1101/112268

Watanabe, S., Sanuki, R., Sugita, Y., Imai, W., Yamazaki, R., Kozuka, T., …
Furukawa, T. (2015). Prdm13 regulates subtype specification of retinal
amacrine interneurons and modulates visual sensitivity. *The Journal of
Neuroscience*, *35*(20), 8004–20. https://doi.org/10.1523/JNEUROSCI.0089-
15.2015