North Carolina macular dystrophy (MCDR1) caused by a novel tandem duplication of the PRDM13 gene

Sara J. Bowne  
*University of Texas Health Science Center at Houston*

Lori S. Sullivan  
*University of Texas Health Science Center at Houston*

Dianna K. Wheaton  
*Retina Foundation of the Southwest*

Kirsten G. Locke  
*Retina Foundation of the Southwest*

Kaylie D. Jones  
*Retina Foundation of the Southwest*

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.wustl.edu/open_access_pubs](https://digitalcommons.wustl.edu/open_access_pubs)  
Please let us know how this document benefits you.

**Recommended Citation**  
https://digitalcommons.wustl.edu/open_access_pubs/5379

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
North Carolina macular dystrophy (MCDR1) caused by a novel tandem duplication of the PRDM13 gene

Sara J. Bowne,1 Lori S. Sullivan,1 Dianna K. Wheaton,2 Kirsten G. Locke,2 Kaylie D. Jones,2 Daniel C. Koboldt,3 Robert S. Fulton,3 Richard K. Wilson,3 Susan H. Blanton,4 David G. Birch,2,5 Stephen P. Daiger1,6

(The first two authors contributed equally to this work.)

1Human Genetics Center, School of Public Health, The University of Texas Health Science Center (UTHealth), Houston, TX; 2Retina Foundation of the Southwest, Dallas, TX; 3McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO; 4Dr. John T. Macdonald Foundation Department of Human Genetics, Hussman Institute for Human Genomics, University of Miami, Miami, FL; 5Dept. of Ophthalmology, University of Texas Southwestern Medical Center, Dallas, TX; 6Ruiz Dept. of Ophthalmology, University of Texas Health Science Center Houston (UTHealth), Houston, TX

Purpose: To identify the underlying cause of disease in a large family with North Carolina macular dystrophy (NCMD). Methods: A large four-generation family (RFS355) with an autosomal dominant form of NCMD was ascertained. Family members underwent comprehensive visual function evaluations. Blood or saliva from six affected family members and three unaffected spouses was collected and DNA tested for linkage to the MCDR1 locus on chromosome 6q12. Three affected family members and two unaffected spouses underwent whole exome sequencing (WES) and subsequently, custom capture of the linkage region followed by next-generation sequencing (NGS). Standard PCR and dideoxy sequencing were used to further characterize the mutation. Results: Of the 12 eyes examined in six affected individuals, all but two had Gass grade 3 macular degeneration features. Large central excavation of the retinal and choroid layers, referred to as a macular caldera, was seen in an age-independent manner in the grade 3 eyes. The calderas are unique to affected individuals with MCDR1. Genome-wide linkage mapping and haplotype analysis of markers from the chromosome 6q region were consistent with linkage to the MCDR1 locus. Whole exome sequencing and custom-capture NGS failed to reveal any rare coding variants segregating with the phenotype. Analysis of the custom-capture NGS sequencing data for copy number variants uncovered a tandem duplication of approximately 60 kb on chromosome 6q. This region contains two genes, CCNC and PRDM13. The duplication creates a partial copy of CCNC and a complete copy of PRDM13. The duplication was found in all affected members of the family and is not present in any unaffected members. The duplication was not seen in 200 ethnically matched normal chromosomes. Conclusions: The cause of disease in the original family with MCDR1 and several others has been recently reported to be dysregulation of the PRDM13 gene, caused by either single base substitutions in a DNase 1 hypersensitive site upstream of the CCNC and PRDM13 genes or a tandem duplication of the PRDM13 gene. The duplication found in the RFS355 family is distinct from the previously reported duplication and provides additional support that dysregulation of PRDM13, not CCNC, is the cause of NCMD mapped to the MCDR1 locus.

Families with North Carolina macular dystrophy were first identified more than 50 years ago [1]. Families initially were described with various terms for macular dystrophy and were linked only when genealogy showed a connection between them. Because the putative founder effect originated in North Carolina, the disease was named North Carolina macular dystrophy (NCMD) [2-4]. The mapping of NCMD in 1992 to chromosome 6q13-q16, called the MCDR1 locus, led the way for many more families with NCMD to be identified via linkage analysis [5-8]. Affected members of many of these families shared a common haplotype on chromosome 6q supporting the concept of a founder mutation.

NCMD is an autosomal dominant form of non-progressive macular impairment, originating during prenatal development with varying degrees of affectation or grades in adulthood [4,9]. Despite extensive studies over numerous years and by many scientists, no cause of disease has been found in the genes located in the chromosome 6 linkage region [10].

Recently, the cause of NCMD was identified using whole genome sequencing of individuals from two families mapped to the MCDR1 locus [9]. Sequence analysis identified a nucleotide substitution at chr6:100,040,906 G>T (hg19) that was subsequently found in nine families. In addition, a C>T...
substitution at chr6:100,041,040 was found in an additional family with NCMD. These two nucleotide substitutions are located in a DNase I hypersensitive site 5′ of two genes, PR/SET domain-containing zinc finger protein 13 (PRDM13; Gene ID 59336, OMIM 616741) and cyclin-C protein (CCNC; Gene ID 892, OMIM 123838). In addition to these two nucleotide substitutions, a third family was found to have a 123 kb duplication of a region that includes the PRDM13 gene and the DNase I hypersensitive site. Experiments using induced pluripotent stem cells (iPSCs) suggested that the disease was likely due to the dysregulation of PRDM13 [9].

This study describes the clinical characterization and mutation identification of another NCMD family, RFS355, which maps to the MCDR1 locus. The tandem duplication identified in this family is different from that reported by Small et al., further strengthening the premise that dysregulation of the PRDM13 gene, not the CCNC gene, is the cause of NCMD (MCDR1) [9].

METHODS

This study adhered to the Declaration of Helsinki and was approved by the Institutional Review Board at each participating institution. Written informed consent was obtained from each participant before the examination and genetic studies were performed.

Clinical characterization: Affected family members (n = 6) spanning three generations underwent ophthalmic exams that included visual acuity, retinal imaging, and fundus grading (Figure 1 and Figure 2).

Visual acuity: Electronic visual acuity (EVA) was measured with a computerized version of the electronic Early Treatment Diabetic Retinopathy Study (E-ETDRS), which has been described previously [11]. Briefly, single high-contrast black letters are randomly displayed on a computer screen surrounded by crowding bars spaced a letter width around the letter. At the 3 m test distance, the letters are displayed from 20/800 (1.6 LogMAR) to 20/12 (∼0.2 LogMAR).

Retinal imaging: Frequency domain optical coherence tomography (fDOCT) retinal imaging, including line and volume scans, was performed (Spectralis HRA+OCT, Heidelberg Engineering, Heidelberg, Germany). The fDOCT measurements included the height, width, and depth of the macular calderas and the width of the papillomacular area, from the temporal edge of the disk to the nasal border of the caldera.

Color fundus photography used a 60 degree digital fundus camera (CF-60UD, Canon, Melville, NY, with Sonomed Escalon image capture software) to capture macula and optic nerve images of both eyes from most subjects.

Three subjects also underwent autofluorescence imaging of the macula.

Fundus grading: The retinal lesion of each patient was characterized as Gass grade 1, 2, or 3 based on the fundus examination and the color fundus photographs (Table 1) [12,13]. Varying degrees of drusen-like deposits were also noted. The clinical features of each Gass grade are as follows:

Grade 1: Limited to small (<50 μm) yellow drusen in the macula with possible mild RPE disturbances. Patients have visual acuity of 20/30 or better.

Grade 2: Larger elevated confluent yellow drusen with RPE atrophy and/or disciform scars with pigment clumping. Patients typically have visual acuity in the range from 20/25 to 20/60, although old disciform scars from choroidal neovascularization may lead to acuity in the 20/100 to 20/400 range.

Grade 3: Large central atrophic excavation of the retina and choroid referred to as a macular caldera. Patients have visual acuity in the range from 20/40 to 20/200.

Linkage analysis: DNA was extracted from peripheral blood and saliva using standard protocols [14]. Peripheral blood was collected by venipuncture into vacutainer tubes containing EDTA. Whole blood was stored at room temperature or buffy coats were stored at -80 °C. DNA was extracted using the Gentra Puregene Blood Kit (Qiagen, Valencia, CA) and the manufacture’s protocol with slight modifications as described previously [14]. Saliva was collected in Oragene DISCOVER (OR-G500) tubes from DNA Genotek (Ottawa, Canada) and stored at room temperature until DNA extraction could be performed. The collection tubes containing saliva were incubated at 50°C and DNA extracted using the manufacture’s protocol with slight modifications as described previously [14]. DNAs from six affected individuals and three unaffected spouses were genotyped at the Hussman Institute for Human Genomics Center for Genome Technology using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, CA). Processing and analysis were performed as previously described [15]. Briefly, DNA was digested with NspI and StyI restriction enzymes (New England Biolabs, Ipswich, MA), ligated to adapters and amplified using adapter-specific primers. PCR products were fragmented, labeled and loaded on the SNP 6.0 arrays. Hybridization was performed in a GeneChip Hybridization Oven (Affymetrix, Santa Clara, CA). Arrays were washed and stained with streptavidin phycoerythrin and scanned on a GeneChip Scanner 3000 7G (Affymetrix).

Single nucleotide polymorphism (SNP) genotype calls and copy number variation (CNV) analysis was performed on the raw Affymetrix Genome-Wide Human SNP 6.0 array.
data using Genotyping Console™ (Affymetrix). PLINK was used for quality control and assessment [16]. Heterozygosity and inter-SNP distance were used to select a subset of SNPs for linkage analysis. All chosen SNPs were at least 0.2 cM apart and had average heterozygosity of 0.5. Centre d’Etude du Polymorphisme Humain (CEPH) Caucasian (CEU) data set allele frequencies were used for all calculations [17]. MERLIN was used to perform multipoint linkage analysis using a dominant model with 90% penetrance in heterozygotes and disease allele frequency of 0.0001 [18,19].

Additional haplotype analysis was performed using short tandem repeat (STR) markers D6S1610, D6S300, D6S1671, and D6S434. Fluorescently labeled primers for amplifying the STRs were obtained from Applied Biosystems (Carlsbad, CA). Genomic DNA from all available family members was amplified and separated on a 3500xL DNA Sequencer (Applied Biosystems). STR alleles were determined using GeneMapper V3.7 (Applied Biosystems).

Genomic paired-end library construction: Genomic DNA (1 μg) was used to make Illumina paired-end libraries according to the manufacturer’s protocol (Illumina Inc., San Diego, CA), with slight modifications as described previously [15].

Exome capture next-generation sequencing: Exome capture was performed using a Nimblegen SeqCap EZ Human

Figure 1. RFS355 pedigree and duplication typing. A: Pedigree of NCMD family RFS355. Circles indicate female individuals while squares indicate male individuals. Diagonal lines indicate a deceased individual. Filled symbols indicate affected individuals while unfilled symbols indicate unaffected individuals. Individual ID numbers are located below the symbols. B: PCR product analysis of the presence or absence of the 69 kb duplication. PCR products of the three amplimers were combined and separated on an agarose gel. Results from each family member are in the lane directly underneath their symbol in the pedigree. The presence of the bottom band (shown by the arrow) indicates these individuals have the duplication on chromosome 6q.
Exome Library v2.0 (Roche, Madison, WI) according to the manufacturer’s protocol. Illumina paired-end sequencing (2×100 bp), alignment, and variant calling were performed as described previously [20].

**Targeted capture next-generation sequencing:**

**Custom capture and next-generation sequencing**—A custom library was designed to capture genomic DNA found in the published MCDR1 locus. This custom library targeted all non-repetitive coding and non-coding DNA on chromosome 6 from 81,907,766–101,730,611 bp (hg19) [4,7,10,21,22].

Genomic libraries were hybridized with the MCDR1 capture library according to the manufacturer’s protocol [14]. The KAPA SYBR FAST qPCR Kit (KAPA Biosystems, Woburn, MA) was used for library quantification to determine the amount required to generate 180,000 clusters on a single lane of the Illumina GAIIx platform. Three lanes of 2×100 bp paired-end sequence were generated for each patient library using the SBS Sequencing Kit Ver. Three (Illumina).

**Alignment and variant identification**—Illumina reads were mapped to the human reference sequence (GRCh7-lite) using BWA v0.5.9 with parameters -t 4 -q 5; duplicates were

---

**Table 1. Summarization of physical exams performed.**

<table>
<thead>
<tr>
<th>Pt ID</th>
<th>Age</th>
<th>Sex</th>
<th>Eye</th>
<th>Visual Acuity</th>
<th>OCT Focus Diopter</th>
<th>Caldera (Grade 3) Height Width Depth</th>
<th>Papillo-Caldera Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>8603</td>
<td>5</td>
<td>M</td>
<td>OD</td>
<td>20/80</td>
<td>1.61</td>
<td>3549 4170 671</td>
<td>2784</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/50</td>
<td>1.71</td>
<td>2930 3617 398</td>
<td>3137</td>
</tr>
<tr>
<td>8687</td>
<td>7</td>
<td>M</td>
<td>OD</td>
<td>20/80</td>
<td>N/A</td>
<td>5095 6617 *1191+</td>
<td>3087**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/63</td>
<td>0.31</td>
<td>4889 6302 *1195+</td>
<td>3585</td>
</tr>
<tr>
<td>8602</td>
<td>39</td>
<td>M</td>
<td>OD</td>
<td>20/63</td>
<td>0.58</td>
<td>3990 5028 1098</td>
<td>2170</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/63</td>
<td>1.07</td>
<td>4309 4987 1222</td>
<td>2041</td>
</tr>
<tr>
<td>8686</td>
<td>39</td>
<td>M</td>
<td>OD</td>
<td>20/80</td>
<td>0.72</td>
<td>3802 4289 611</td>
<td>2292</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/63</td>
<td>0.24</td>
<td>3304 3498 379</td>
<td>2612</td>
</tr>
<tr>
<td>9395</td>
<td>44</td>
<td>F</td>
<td>OD</td>
<td>20/100</td>
<td>−0.91</td>
<td>3468 3083 669</td>
<td>1968</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/16</td>
<td>−0.15</td>
<td>Grade 1</td>
<td></td>
</tr>
<tr>
<td>8688</td>
<td>68</td>
<td>F</td>
<td>OD</td>
<td>20/63</td>
<td>0.91</td>
<td>2666 3219 570</td>
<td>2490</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OS</td>
<td>20/125</td>
<td>1.14</td>
<td>Grade 2</td>
<td></td>
</tr>
</tbody>
</table>

See Figure 1 for patient ID’s. N/A-data not available. *The +indicates an approximate measurement as both edges and bottom of caldera could not be visualized at same time. **Distance measure on fundus color image as no OCT scan was available of the papillo-macular bundle areas.
marked using Picard v1.46. Reads with a mapping quality of zero, or that were marked as duplicates by Picard, were excluded from further analysis.

Putative SNPs and indels were called in the exome data using VarScan 2 and the following thresholds: coverage ≥8X, phred base quality ≥15, minimum variant allele frequency ≥10%, and Fisher Exact Test p value <0.05 [23]. False positives were removed from paralogous alignments, local misalignments, sequencing errors, and other factors by filtering the SNVs to remove any with strand bias, read position bias, or multiple high-quality mismatches in supporting reads. Predicted indels were filtered to remove small events around homopolymers, which likely are false positives.

**Copy number analysis**—On average, a 181X depth was achieved for affected individuals and a 185X depth for unaffected individuals across the approximate 19.8 Mbp region of interest on chromosome 6. To search for copy number changes, the region of interest was divided into non-overlapping 100-bp segments and average read depth for affected (n = 3) and unaffected (n = 2) individuals were computed according to SAMtools (r982) mpileup. The magnitude and direction of the copy number change were computed as log2 (affected_depth/unaffected_depth). The resulting log2 values were segmented using the “DNAcopy” package in R with smoothing (data.type="logratio," undo.SD=4).

**PCR confirmation and segregation analysis:** PCR primers were designed to span the predicted CNV break points and the resulting junctions of the normal and duplicated genomic regions (Figure 3). PCR amplifications were performed using AmpliTaqGold™ 360 Master Mix (Fisher, Waltham, MA) and the following amplification conditions: 95 °C 5 min; 95 °C 1 min, 56 °C 1 min, 72 °C 1 min x35; 72 °C 5 min; 8 °C hold. Upon amplification, each PCR product was sequenced using standard dideoxy fluorescence Sanger sequencing and a 3500XL Genetic Analyzer (Applied Biosystems). Products were combined and run on agarose gels to confirm duplication segregation with disease (Figure 1). Primer sequences are available upon request.

**RESULTS**

Six affected members of the RFS355 family underwent comprehensive clinical exams. Clinical data are presented in Table 1. The six affected individuals were tested at several different ages spanning the first to sixth decades of life. Each individual had good to moderate visual acuity ranging from 20/16 to 20/125. Technical details related to the acquisition of OCT in these patients have been reported previously [24].

A single eye of individual RFS355–9395 had a grade 1 Gass score while a single eye of RFS355–8688 was grade 2 (Figure 2). All the other eyes had grade 3 Gass phenotypes with varying size calderas, where the smallest were only half the size of the largest. Calderas were centered in the macula at the location of the missing fovea. In spite of the wide range of sizes, the distances between the disk and the nasal edge of the caldera were relatively similar, suggesting that the expansion during caldera formation was primarily in the temporal direction. The preserved nasal retina typically contained the preferred locus of fixation, which probably explains the good vision in these patients. The size and depth of the calderas were not age dependent, as the youngest individual tested (RFS355–8687) had the largest lesion when compared to older family members.

Genome-wide linkage analysis was performed on all available affected and relevant unaffected family members (n = 9) using Affymetrix Genome-Wide Human SNP Array 6.0 and MERLIN analyses. The majority of the genome had logarithm (base 10) of odds (LOD) scores of less than −2. A few genomic regions were also found to have LOD scores between −2 and 0. Only one region was found to have a LOD score above zero. This region on chromosome 6 had a maximum LOD score of 1.8. Chromosome 6q STR markers confirmed that a consistent haplotype existed in all affected family individuals and was not present in any unaffected individuals. The common haplotype was located on chromosome 6q14-16 in a region consistent with the reported North Carolina macular dystrophy locus, MCDR1 [25].

On the initial examination of the exome data, no potentially deleterious point mutations or small indels within coding sequences of the genes in the linkage region were observed to segregate with disease in the family. To increase coverage and examine any possible copy number variations, a custom capture library was designed to capture as much of the MCDR1 locus as possible, including the coding sequences and introns and other non-coding, non-repetitive DNA sequences.

Analysis of the variants identified in the custom capture region yielded no new candidates. However, analysis of the data for the CNVs showed an increased copy number in a region of chromosome 6 indicating a duplication. In the 68.8 Kbp region of duplication, the read depths were 1.5-fold higher in the affected individuals (283X for affected, 195X for unaffected, log2 ratio = 0.563; Figure 4).

When the duplication was detected, there was no convincing way to determine if or how this duplication was related to disease. Subsequently, the recent paper by Small et al. identified a duplication in another family with NCMD.
Figure 3. Pictorial representation of the wild-type and pathogenic duplicated regions on chromosome 6. Green bars indicate the wild-type region on chromosome 6 while the yellow bars indicate the duplicated region. Red asterisks indicate the DNase I hypersensitivity site. Small arrows below the chromosome bar pictorials are the three pairs of PCR primers that produce different size products. The PCR product generated from these primers was used to determine the presence or absence of the deletion and the exact sequence of the duplication in family members (Figure 1B). Red and blue primers should amplify from the wild-type and duplicated chromosomes. The orange primers amplify only on the duplicated chromosome. Sequencing of the PCR product resulting from the orange depicted primers determined the exact duplication break point and the inclusion of 5 bp of exogenous DNA, TCCTG, between the wild-type and duplicated regions. Large horizontal black arrows with gene names above illustrate the strand and transcriptional direction of each gene. Vertical lines and blocks represent exons. The duplication found of the CCNC gene is only from exons 1 through 9 making it unlikely that any protein is generated from the duplicated region. The presence of the entire PRDM13 gene downstream from the DNase hypersensitivity site likely causes a change in PRDM13 protein levels compared to wild-type levels.
at approximately the same region as the RFS355 family [9]. Given evidence that CNVs are a cause of NCMD, our data suggested the duplication identified in RFS355 could be a likely cause of disease in this family.

Based on the targeted capture CNV results, the genomic DNA duplication in the NCMD family is located from approximately 99,996,220–100,065,140 bp (hg19) on chromosome 6q16.2 (Figure 3). To confirm the duplication and determine its exact size, two pairs of PCR primers were designed to flank the wild-type sites; the primers amplify the complete region in affected and unaffected individuals, regardless of whether the duplication is present or not. A third set of PCR primers was designed to flank the 6q telomeric region of the normal chromosome and the centromeric region of the duplicated 6q regions (Figure 3). PCR products from one affected and one unaffected individual were sequenced to determine the exact duplication locations and their relationship to canonical genomic DNA. These amplifications showed that the duplication segregated with disease in the RFS355 family. Sequencing of the PCR products determined the exact break points and the 5 bp of inserted, exogenous DNA sequence (Figure 1 and Figure 3)

**DISCUSSION**

We investigated a family with a retinal phenotype similar to that of other families with NCMD. Linkage analysis in this family mapped the disease locus to 6q14-16, which contains the MCDR1 locus. Targeted capture NGS and follow-up PCR-based fluorescent sequencing determined that the cause of disease in this family lies within the MCDR1 locus region. The retinal disease in this family is caused by a 68,912 bp duplication which contains an entire intact PRDM13 gene copy, a partial CCNC gene copy, and a copy of the DNase I hypersensitive site located between the 5′ ends of both genes. The distinct mutation in this family with NCMD further demonstrates that multiple mutations are responsible for the NCMD phenotype.
As there is only a partial copy of the CCNC gene, it is probable that any mRNA generated undergoes nonsense-mediated decay and no additional CCNC protein is produced, regardless of the presence of the extra DNase I hypersensitive site. The only gene in the duplication that is likely to make intact mRNA and protein is PRDM13. Further, it is likely that the duplication of the DNase I hypersensitive region affects the duplicated PRDM13 gene and may even affect the wild-type PRDM13 gene.

PRDM proteins are tissue-specific transcription factors [26]. Research has shown that in induced pluripotent cells, PRDM13 mRNA levels decrease as stem cells become mature neural cells [9]. Recently, it was shown that Prdm13 is expressed in developing and mature mouse retinal amacrine cells. Prdm13−/− mouse retinas show a decreased number of amacrine cells in the inner nuclear layer and are thought to regulate amacrine subtype specification [27].

The duplicated copy of the PRDM13 gene and hypersensitive DNase site found in RFS355 might lead to over-expression of PRDM1, which may be the cause of NCMD symptoms, possibly through amacrine cell regulation in the neural retina. Results presented here further strengthen the argument presented by Small et al. that dysregulation of the neural retina. Results presented here further strengthen the argument presented by Small et al. that dysregulation of the

ACKNOWLEDGMENTS

We thank the RFS355 family members for their participation in this study. We also thank Cheryl Avery and Elizabeth Cadena for technical assistance. Funding for this work was provided by The Foundation Fighting Blindness (SPD and DBG), R01EY007142 (SPD), R01EY09076 (DGB), and by The William Stamps Farish Fund (SPD) and Hermann Eye Fund (SPD).

REFERENCES


