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

Valentina Cipriani,[1,2,3] Ambreen Kalhoro,[1,2] Gavin Arno,[1,2] Raquel S Silva,[1,2] Nikolas Pontikos,[1,3] Virginie Puech,[4] Michelle E McClements,[5] David M Hunt,[6] Veronica van Heyningen,[1] Michel Michaelides,[1,2] Andrew R Webster,[1,2] Anthony T Moore,[1,2,7] Bernard Puech[4]

[1]UCL Institute of Ophthalmology, London, UK, [2]Moorfields Eye Hospital, London, UK, [3]UCL Genetics Institute, London, UK, [4]Service d'Exploration de la Vision CHU, Lille, France, [5]Nuffield Department of Clinical Neurosciences (Ophthalmology), University of Oxford, Oxford, UK, [6]Lions Eye Institute and School of Animal Biology, University of Western Australia, Perth, Australia, [7]Ophthalmology Department, UCSF School of Medicine, San Francisco, CA, USA

Corresponding authors and address for reprint requests: Valentina Cipriani, PhD, UCL Institute of Ophthalmology, Department of Ocular Biology and Therapeutics, University College London, Wolfson Building (3rd floor, Room 3/5), 11-43 Bath Street, London, EC1V 9EL, United Kingdom; v.cipriani@ucl.ac.uk; Phone: +44 (0)207 608 4042; Fax: +44 (0)207 608 6830;

Bernard Puech, MD, Hôpital Roger Salengro, SEVNO, Rue Emile Laine, 59037, Lille, Cedex, France; bernard.puech@yahoo.fr; Phone: + 33 320446286; Fax: +33 320446288.

Running title: MCDR3 and macular disorder with digit anomalies

ABSTRACT

Background: Developmental macular disorders are a heterogeneous group of rare retinal conditions that can cause significant visual impairment from childhood. Among these disorders, autosomal dominant North Carolina macular dystrophy (NCMD) has been mapped to 6q16 (MCDR1) with recent support for a non-coding disease mechanism of *PRDM13*. A second locus on 5p15-5p13 (MCDR3) has been implicated in a similar phenotype, but the disease-causing mechanism still remains unknown. **Methods**: Two families affected by a dominant developmental macular disorder that closely resembles NCMD in association with digit abnormalities were included in the study. Family members with available DNA were genotyped using the Affymetrix GeneChip Human Mapping 250K Sty array. A parametric multipoint linkage analysis assuming a fully penetrant dominant model was performed using MERLIN. Haplotype sharing analysis was carried out using the non-parametric Homozygosity Haplotype method. Whole-exome sequencing was conducted on selected affected individuals. **Results**: Linkage analysis excluded MCDR1 from the candidate regions (LOD<-2). There was suggestive linkage (LOD=2.7) at two loci, including 9p24.1, and 5p15.32 that overlapped with MCDR3. The haplotype sharing analysis in one of the families revealed a 5 cM shared IBD segment at 5p15.32 (p-value=0.004). Whole-exome sequencing did not provide conclusive evidence for disease-causing alleles. **Conclusions**: These findings do not exclude that this phenotype may be allelic with NCMD MCDR3 at 5p15 and leave the possibility of a non-coding disease mechanism, in keeping with recent findings on 6q16. Further studies, including whole-genome sequencing, may help elucidate the underlying genetic cause of this phenotype and shed light on macular development and function.

KEYWORDS

1. Congenital limb deformities

2. Genetic linkage

3. Macular dystrophy, retinal 1, human

4. North Carolina macular dystrophy

5. Retinal dysplasia

INTRODUCTION

Developmental disorders of the macula are a rare cause of visual impairment and identifying the causative genes will help to understand the biological pathways involved in human macular development and function. One such disorder is North Carolina macular dystrophy (NCMD), an autosomal dominant macular disorder of congenital onset that was first described in the early 1970's in a large kindred from North Carolina[1,2] and revisited by Small, many years later, when it was defined as a non-progressive disorder with highly variable expressivity.[3] The bilaterally symmetrical clinical manifestations of NCMD can be classified into three non-progressive groups: Grade 1 is characterised by small drusen-like lesions in the central macula, Grade 2 presents with larger confluent drusen, and Grade 3 exhibits well demarcated macular chorioretinal atrophy. The severity of visual impairment is dependent upon the grade of the retinal phenotype, with central vision being poorest in association with Grade 3 lesions. Dark adaptation and colour vision, as well as the electro-oculogram (EOG) and electroretinogram (ERG), are normal, suggesting that this is an isolated macular phenotype and making NCMD of particular interest for improving our understanding of macular biology.[4]

Sorsby in 1935 reported a British family with autosomal dominant bilateral chorioretinal macular dysplasia in association with apical dystrophy of the hands and feet.[5] The retinal phenotype observed in these patients closely resembled the phenotype subsequently described in affected members of NCMD families. Further generations of the British family were reported in 1988 by Thompson and Baraitser[6] who referred to this condition as 'Sorsby syndrome' (family GC16334 in this report). Some affected individuals also presented with unilateral renal agenesis, double uterus and vagina, severe sensorineural or mixed hearing loss and accessory ribs. A family of French origin with an autosomal dominant developmental macular disorder and digit abnormalities similar to the phenotype previously observed by Sorsby was described in 1991[7] (family GC16500 in this report). To our knowledge, no genetic investigations have been reported on the Sorsby syndrome family or other families with a similar phenotype, and the underlying disease variants remain to be identified.

In contrast, the molecular genetics of NCMD has been extensively investigated with the disorder mapped to chromosome 6q16 (MCDR1, MIM:136550) in the early 1990's.[8] The identification of many additional NCMD families of different ethnic origins allowed the linkage region to be substantially narrowed,[9–15] but early sequencing studies of the coding regions of genes within the MCDR1 interval failed to identify exonic disease-causing variants.[16] Recently, however, using whole-genome sequencing, support for a genetic disease mechanism that may involve non–coding modification of the expression of the retinal transcription factor *PRDM13* in macular development has been provided with the identification of three rare non-coding variants that lie in a DNase 1 hypersensitivity site (DHS) upstream of *PRDM13*, *TSD3*, and *CCNC* and two duplications of *PRDM13* and the upstream region that included the same DHS, in 12 NCMD MCDR1 families.[17,18]

Interestingly, Small et al.[18] also identified a 900-kb tandem duplication that included the entire coding sequence of *IRX1* in all affected members from a Danish NCMD family that had been mapped to chromosome 5p15-5p13 (MCDR3, MIM: 608850),[19] first identified by Michaelides et al.[20] However, no additional variants in other MCDR3 families have been identified and evidence for a causative role of the duplicated *IRX1* coding sequence is weakened by the fact that, unlike *PRDM13, IRX1* revealed no variation in expression in the first 100 days of development in normal iPSC-derived human retinal cells.[18] The disease-causing mechanism at the NCMD MCDR3 locus still remains unknown.

In this report we present genetic linkage, haplotype and exome sequencing analyses in the previously reported family of French origin[7] and members of the originally reported Sorsby syndrome family[5,6] affected by an autosomal dominant developmental macular disorder in association with digit abnormalities. Our aim was to assess whether the disorder maps to any of the previously identified NCMD loci or to a new locus, and identify the underlying genetic cause.

MATERIALS AND METHODS

Families

Family GC16500 was ascertained at the Centre Hospitalier Régional Universitaire de Lille, France.[7] Members from younger generations of the originally reported British family[6] (GC16334) were seen at Moorfields Eye Hospital, London, United Kingdom (Figure 1). Patients and family members underwent a full physical and ophthalmic examination, including visual acuity testing and dilated fundus examination. When possible, retinal imaging was undertaken with colour fundus photography and/or fundus autofluorescence imaging, and X-rays of the hands and feet were obtained. Selected individuals underwent electrophysiological assessment, including ERG. Blood samples were taken for DNA extraction and genotyping and sequencing analyses. The study protocol was approved by the local ethics committee and conformed to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants, or their parents, before inclusion in the study.

Genotyping

Genomic DNA was extracted from whole blood and genotyped using the Affymetrix GeneChip Human Mapping 250K Sty array (Affymetrix, Inc., Santa Clara, CA, USA). Genotypes were determined using the Bayesian Robust Linear Model with Mahalanobis (BRLMM) genotyping algorithm implemented in the manufacturer's GTYPE 4.1 software (Affymetrix, Inc., Santa Clara, CA, USA).

Genome-wide linkage analysis

LINKDATAGEN[21] was used to remove markers incompatible with Mendelian inheritance and to select a subset of informative single nucleotide polymorphisms (SNPs) with high heterozygosity spaced approximately 0.3 cM apart across the genome. MERLIN[22] was used to identify and remove genotyping errors based on inferred excessive and unlikely recombination events between tightly linked markers. A parametric genome-wide multipoint linkage analysis assuming a fully penetrant

autosomal dominant disease model with a disease allele frequency of 0.0001 was performed using MERLIN. Marker allele frequencies from the HapMap phase III Caucasian European in Utah (CEU) population were used (LINKDATAGEN annotation file). The most likely haplotypes inferred by MERLIN were imported into Haplopainter[23] for graphical visualisation.

Haplotype sharing analysis

As a subsidiary analysis to the parametric genome-wide linkage search, we used the non-parametric Homozygosity Haplotype (HH) method proposed by Miyazawa et al.[24] The HH is a type of haplotype described by the homozygous SNPs only (all heterozygous SNPs are removed) and, therefore, it can be uniquely determined on each chromosome. Since affected family members who inherited the same mutation from a common ancestor share a chromosomal segment identical by descent (IBD) around the disease gene, they should not have discordant homozygous calls in the IBD region and thus they should share the same Homozygosity Haplotype. The HH approach predicts IBD regions through the identification of Regions with a Conserved HH (RCHH) defined as those regions with a shared HH among patients and a genetic length longer than a certain cut-off value (recommended cut-off value for Affymetrix GeneChip Human Mapping 250K Sty array is 5.0 cM). The HH method was further applied by using affected and unaffected family members as cases and controls respectively, as suggested by Jiang et al.[25] An autosomal interval is firstly divided into minute regions. The RCHH shared by the largest number of patients in the patient pool is then selected as the representative RCHH for each small region. The numbers of subjects who share the representative RCHH are then counted for both the patient pool and the control pool and the significance of each representative RCHH is calculated with a z-test.

Whole-exome sequencing and variant prioritisation

Four patients underwent whole-exome sequencing (WES, AROS Applied Biotechnology, Aarhus, Denmark) using the Illumina TruSeq exome capture kit and the Illumina HiSeq 2000 sequencer (Illumina Inc.). After hybridization and indexing, samples were pooled and 100bp paired end

sequencing was performed (Illumina HiSeq 2000 sequencer). Reads were aligned to the hg19 human reference sequence (build GRCh37) using Novoalign version 2.08 (Novocraft, Selangor, Malaysia). Duplicate reads were marked using Picard tools MarkDuplicates (http://broadinstitute.github.io/picard). Calling was performed using GATK,[26] creating gVCF formatted files for each sample. The individual gVCF files for the exomes discussed in this study, in combination with 4,539 in-house clinical exomes (UCL-ex database), were combined into merged VCF files for each chromosome containing on average 100 samples each. The final variant calling was performed using the GATK GenotypeGVCFs module jointly for all samples (cases and controls). Variant quality scores were then re-calibrated according to the GATK best practices separately for indels and SNPs. Resulting variants were processed with Exomiser,[27] a variant prioritisation software that annotates, filters and prioritises likely causative variants starting from a VCF file and a set of phenotypes encoded using the Human Phenotype Ontology (HPO).[28] The functional annotation of variants was handled by Jannovar[29] that is embedded within Exomiser and uses UCSC KnownGene transcript definitions and hg19 genomic coordinates. Variants listed in the VCF files were first filtered according to user-defined criteria on rarity (minor allele frequency (MAF) $\leq$ 0.001 in either the publicly available 1000 Genomes Project Phase 1 component of dbSNP (1000G), the NHLBI GO Exome Sequencing Project and the Exome Aggregation Consortium datasets (ExAC)), autosomal dominant mode of inheritance (heterozygote) and quality (Phred Quality score Q > 30). The filtered variants were then ranked on the basis of their deleteriousness (predicted pathogenicity data as extracted from the dbNSFP resource,[30] and phenotypic relevance: that is how closely the given HPO-encoded phenotype matches the known phenotype of disease genes from human,[31,32] mouse[33,34] and zebrafish[35] model data (cross-species phenotype comparisons performed by PhenoDigm tool.[36] Finally, the list of ranked variants obtained from Exomiser was filtered further, based on the variant frequency observed in the in-house exome database (fewer than 5 heterozygotes and no homozygote seen in the UCL-ex database).

Sanger sequencing validation of candidate variants

Segregation analysis of the candidate variants identified by WES was performed in affected individuals and available family members. Oligonucleotide primers were specifically designed to amplify the DNA fragment containing the candidate variants using standard polymerase chain reaction (PCR) amplification and confirmed by electrophoresis (primer sequences are described in Supplementary Table 1). Positive reactions were followed by bidirectional DNA Sanger sequencing with dye termination chemistry. Results were visualised using SnapGene Viewer.

RESULTS

Clinical phenotype description

Fourteen members (8 female, 6 male) from family GC16500 and four members (2 female, 2 male) from family GC16334 (Figure 1) underwent detailed physical and ophthalmic examination. There was no reported consanguinity in either family. Autosomal dominant inheritance was supported by the presence of affected members in each generation, males and females equally severely affected (in both families), equal numbers of affected males and females and male to male transmission (in family GC1650). *Family GC16500*

After clinical examination, eight members of family GC16500 were assigned affected status (Figure 1). Visual acuity in affected individuals ranged from 6/12 to 6/60. The macular defect was variable in severity from mild retinal pigment epithelium (RPE) changes with pigment deposition to excavated chorioretinal atrophic lesions. Individuals II:2, II:5 and III:1 underwent full field ERG which was normal. Four patients displayed additional systemic abnormalities affecting their digits. In particular, individual III:1 and IV:4 presented with ectrodactyly (split hand/split foot malformation that occurs due to aplasia of the central rays of the hand/foot associated with abnormal phalanges and irregular skin clefts) and individual III:3 and IV:1 showed fifth finger clinodactyly - (Table 1, Figure 2). There were no family members with isolated digit abnormalities (and a normal macula).

Family GC16334

Clinical examination confirmed affected status of three individuals (Figure 1). Individual II:2 corresponds to individual IV:3 in the original report by Thompson and Baraitser.[6] Visual acuity in affected individuals ranged from 6/60 to 3/60. All three affected individuals demonstrated bilateral symmetrical excavated chorioretinal atrophic macular lesions and presented with bilateral hand and foot brachydactyly (shortening and deformity of the fingers and toes due to aplasia and hypoplasia of middle and terminal phalanges) and skin syndactyly (two digits are fused together) (Table 1, Figure 2). There were no family members with isolated digit abnormalities (and a normal macula).

Table 1 Clinical phenotype characteristics of affected individuals of two families with a developmental macular disorder and digit abnormalities (Figure 1).

Family	Patient identifier	Sex	Age (y)[a]	Visual acuity[b]	Congenital macular lesions[c]	Digit abnormalities
GC16500	II:2	F	76	OD 6/24; OS 6/18	Chorioretinal atrophy	Normal hands and feet
GC16500	II:3	F	73	OD 6/12; OS 6/12	Granular changes with RPE atrophy	Normal hands and feet
GC16500	II:5	F	61	OD 6/24; OS 6/24	Chorioretinal atrophy	Normal hands and feet
GC16500	III:1	M	51	OD 6/12; OS 6/24	Chorioretinal atrophy	Bilateral split hand malformation
GC16500	III:3	M	36	OD 6/60; OS 6/60	Chorioretinal atrophy	Bilateral fifth finger clinodactyly
GC16500	IV:1	M	29	OD 6/12; OS 6/24	Chorioretinal atrophy	Left fifth finger clinodactyly
GC16500	IV:4	F	27	OD 6/60; OS 6/24	Chorioretinal atrophy	Bilateral split hand and split foot malformation
GC16500	V:1	F	4	Not available	Granular pigmentary changes	Normal hands and feet
GC16334	II:2	F	66	OD 3/60; OS 3/60	Chorioretinal atrophy	Bilateral hand and foot brachydactyly and skin syndactyly
GC16334	III:2	M	30	OD 6/60; OS 6/60	Chorioretinal atrophy	Bilateral hand and foot brachydactyly and skin syndactyly
GC16334	III:3	M	29	OD 4/60; OS 4/60	Chorioretinal atrophy	Bilateral hand and foot brachydactyly and skin syndactyly

[a]At last examination

[b]OD = right eye; OS = left eye

[c]All congenital macular lesions are bilateral and relatively symmetrical

Genome-wide linkage analysis

Eleven individuals (8 affected, 1 unaffected and 2 unaffected married-in) from family GC16500 and 4 individuals (3 affected, 1 unaffected) from family GC16334 were genotyped using the Affymetrix GeneChip Human Mapping 250K Sty array (Figure 1) and parametric linkage analysis under a fully penetrant dominant model was performed in the two families both separately (allowing for locus heterogeneity), and simultaneously (assuming locus homogeneity).

Linkage at the NCMD MCDR1 locus (6q16)[8,15] was excluded in all linkage analyses with LOD scores less than -2. The linkage analysis of family GC16500 used a subset of 11,368 informative markers and identified 8 regions with variable support for linkage (LOD score range = 1.2 – 2.1), including chromosomes 3q23-24, 5p15.33-32 and 9p24.1, that achieved a maximum LOD score of 2.1 (Table 2). The analysis of family GC16334 (11,189 informative markers) showed linkage peaks at 60 regions on all autosomal chromosomes with a maximum LOD score of 0.6, reflecting the small number of available genotyped individuals (Supplementary Table 2).

When we analysed the two families together (10,955 informative markers) the linkage signals were refined to two peaks at chromosomes 5p15.32 and 9p24.1 with a maximum LOD score of 2.7 (Table 2). Notably, the 5p15.32 region partially overlaps with the previously identified NCMD MCDR3 locus (5p15-5p13).[19,20] Haplotype analysis supported the linkage results showing segregation of haplotypes with the disease at both the 5p15.32 (Supplementary Figure 1) and 9p24.1 (Supplementary Figure 2) loci in each family. Recombination events in unaffected individual III:1 of family GC16334 on the telomeric side and in affected individual II:3 of family GC16500 on the centromeric side narrowed down the critical 5p15.32 region to a 3 cM region between SNP marker rs861512 and rs4702303 (GRCh37/hg19 chr5:g.5173776-5986111) (Supplementary Figure 1). Recombination events in affected individual II:3 and IV:1 in family GC16500 suggested the critical 9p24.1 region is located between 15.81 and 17.89 cM (SNP marker rs10976175 and rs9650711; GRCh37/hg19 chr9:g.7275482-8098790) (Supplementary Figure 2). We observed different haplotypes at both the 5p15.32 and 9p24.1 loci across the two families.

Table 2 Linkage analysis in family GC16500, and including family GC16334 (assuming locus homogeneity).

Chr	**Genetic position (cM)**		**Physical position (bp)**		**SNP**		**Linkage analysis**
	Start	**End**	**Start**	**End**	**Start**	**End**	**max HLOD**
Family GC16500							
3q23-24	14.85	15.15	141954700	143919831	rs6769004	rs9861939	2.1
5p15.33-32	5p-telomere	15.92	5p-telomere	5986111	5p-telomere	rs4702303	2.1
5p13.3-2	55.05	55.87	33595996	33808956	rs10461924	rs1037104	1.2
5q12.3-13.1	78.05	79.98	64723022	67023975	rs4700082	rs7448274	2.0
8q24.22	14.19	14.20	132774750	132942751	rs11995633	rs4644223	1.3
9p24.1	14.74	17.89	6976267	8098790	rs912207	rs9650711	2.1
14q32.33	12.49	12.78	105669360	107179847	rs7157285	rs2007467	1.6
16p12.1	51.48	54.82	26502518	29159435	rs8051570	rs252496	1.8
Family GC16500 and family GC16334							
5p15.32	12.24	15.92	4976629	5986111	rs489095	rs4702303	2.7
9p24.1	14.74	17.89	6976267	8098790	rs912207	rs9650711	2.7

Genomic coordinates refer to GRCh37/hg19 assembly.

Haplotype sharing analysis

As a subsidiary approach to the parametric linkage analysis, a search for a shared IBD chromosomal segment among affected individuals within each family was performed using the HH method. The analysis of the four affected members of family GC16334 showed evidence for haplotype sharing at a number of regions on all autosomal chromosomes reflecting the availability of only a small number of closely related family members (Supplementary Figure 3). The HH analysis of the eight affected members of family GC16500 revealed four RCHH (Figure 3A), of which only a large one of approximately size 5 cM (GRCh37/hg19 chr5:g.4970365-6304617) overlapped with a linkage signal, i.e., at 5p15.32 (Table 2). The linkage signal at 9p24.1 did not receive support from the HH analysis: the region with a conserved HH was of approximate size 0.5 cM (GRCh37/hg19 chr9:g.7279636-7442546) and did not pass the recommended significance cut-off. We then performed the HH analysis using both the patient pool (8 affected members) and the control pool (3 unaffected members) and the representative RCHH at the 5p15 locus obtained the highest significance (p-value=0.004) (Figure 3B).

Whole-exome sequencing analysis

To identify the underlying genetic cause, the DNA of four individuals including III:3 and V:1 in family GC16500 and III:2 and III:3 in family GC16334 were analysed using WES. After variant filtering, no variants were shared by all four affected individuals and, significantly, no genes harboured rare variants across the two families.

In family GC16500, individual III:3 and V:1 shared two exonic variants in a heterozygous state: a novel missense change (c.32A>C:p.Glu11Ala) in *CLCN2* (MIM:600570), and a rare frameshift deletion (c.517del:p.His173ThrfsTer79) in *ZNF774* that is seen once out of 121,390 alleles in ExAC database and once in a heterozygous state in a patient affected by rare corneal dystrophy in our in-house UCL-ex cohort. Neither of these two variants is located within a predicted linkage interval (Table 2). However, the HH analysis showed that the *ZNF774* gene is embedded in a region with a conserved HH that is shared by all 8 affected members of family GC16500 (Figure 3). Segregation

analysis by direct Sanger sequencing confirmed that the *ZNF774* variant does not segregate with the affected phenotype (present in all 8 affected family members in a heterozygous state, absent from three married-in individuals and unaffected member V:2, but present in the unaffected individual II:4) (Supplementary Figure 4).

WES in family GC16334 revealed that affected individuals III:2 and III:3 shared 34 heterozygous variants that passed filtering (Supplementary Table 3). Neither of these variants is located within a predicted linkage interval (Table 2). We then assessed the shared variants against the linkage results from the analysis of the family GC16334 only, allowing for locus heterogeneity. Twenty variants reside in a region of excluded linkage (LOD score < -2). Segregation analysis of the remaining 14 variants in all 4 family members for whom DNA was available (affected mother, two affected siblings and one unaffected daughter) (Figure 1) determined that only eight variants segregate with the affected phenotype (Table 3). Two missense variants in *IKBKB* (MIM:603258) (p.Gly209Ser) and *C22orf15* (p.Thr96Ser) were absent from all databases used for variant filtering, while the remaining six variants were seen in sporadic heterozygote carriers in either the publicly available databases or our in-house UCL-ex database. Overall, the variant p.Gly209Ser in *IKBKB* received the highest pathogenicity scores and Exomiser score (i.e., 0.75, using HPO terms reported in Supplementary Table 4) as well as the maximum value (i.e., 1) for the ExAC index pLI (i.e., the probability that a gene is intolerant to a loss of function mutation.[37]

Table 3 Shared exonic variants by affected individuals III:2 and III:3 in family GC16334, after variant filtering and segregation analysis in two additional family members (III:1 and II:2)

Variant (chr:position:ref>alt)	Variant type	Gene	Linkage	Kaviar frequency (database source)[a]	In-house UCL-ex allele count	Pathogenicity score (PolyPhen; SIFT)[b]	ExAC pLI[c]	Exomiser score[d]
c.625G>A:p.Gly209Ser (chr8:42166476:G>A)	Missense	*IKBKB*	LOD=0.6	Not seen	Not seen	0.999 (D); 0 (D)	1	0.75
c.1436A>G:p.Gln479Arg (chr4:170042051:A>G)	Missense	*SH3RF1*	LOD=0.6	0.000006 (ExAC)	Not seen	0.992 (D); 0.05 (D)	0.05	0.46
c.1534C>A:p.Pro512Thr (chr8:27401706:C>A)	Missense	*EPHX2*	LOD=0.6	0.000019 (dbsnp, ExAC, 1000G)	Not seen	0.997 (D); 0.005 (D)	0	0.46
c.1877A>G:p.Glu626Gly (chr16:2505557:A>G)	Missense	*CCNF*	LOD=0.6	0.000013 (dbsnp, ExAC)	Not seen	0.133 (B); 0.001 (D)	0.02	0.46
c.454C>A:p.Leu152Met (chr19:3770858:C>A)	Missense	*RAX2*	LOD=0.6	0.000006 (GEUVADIS)	1[e]	0.964 (D); 0 (D)	0.51	0.19
c.442G>A:p.Val148Ile (chr5:10618520:G>A)	Missense	*ANKRD33B*	LOD=0.6	Not seen	1[f]	0.922 (P); 0.20 (T)	N/A	0.02
c.167C>T:p.Ser56Leu(chr6:27100363:G>A)	Missense	*HIST1H2BJ*	LOD=0.6	0.000013 (dbsnp, ExAC)	Not seen	0.620 (P); 0.016 (D)	0.48	$4.7x10^{-3}$
c.287C>G:p.Thr96Ser (chr22:24106847:C>G)	Missense	*C22orf15*	LOD=0.6	Not seen	Not seen	0.395 (B); 0.502 (T)	0.08	$1.6x10^{-4}$

Genomic co-ordinates refer to GRCh37/hg19 assembly

[a]Kaviar database is available at http://db.systemsbiology.net/kaviar

[b]B=Benign, D=Deleterious, P=Probably/Possibly damaging, T=Tolerated

[c]ExAC pLI=probability that a gene is intolerant to a loss of function mutation[37]

[d]Exomiser score[27] based on HPO terms in Supplementary Table 4. Full results from the Exomiser analysis are displayed in Supplementary Table 5

[e]Seen in a heterozygote state in a patient affected by rare non-syndromic retinitis pigmentosa

[f]Seen in a heterozygote state in a patient affected by rare non-syndromic unilateral cleft lip and palate

DISCUSSION

We report the results of a molecular genetic investigation of two families affected by an autosomal dominant developmental macular disorder that closely resembles NCMD in association with digit abnormalities (Figure 2). Through a combination of linkage and haplotype analyses in the two families we were able to exclude the NCMD MCDR1 locus[15,18] at chromosome 6q16 from the candidate disease regions (LOD score < -2).

Given the similarity of the phenotype across the two families and the exceptional rarity of the macular lesions observed in combination with digit anomalies, these two families are likely to share the same genetic cause. The different haplotypes observed at both linkage loci (Table 2) across the two families (Supplementary Figures 1 and 2) suggest the presence of allelic heterogeneity. Notably, in the larger family GC16500 we were able to exclude the 9p24 locus from the candidate regions with the use of the HH approach, an effective method for the identification of disease susceptibility loci in inherited monogenic disorders.[24,25] The 5p15 locus was the only region to overlap with a linkage signal and showed the most significant evidence (p-value=0.004) for a shared IBD chromosomal segment of a considerable size (approximately 5 cM) among affected members. Overall, these findings do not exclude the possibility that this phenotype may be allelic with the developmental NCMD MCDR3 at the 5p15 locus,[19,20] or alternatively caused by variants in two different adjacent developmental genes. Significantly, WES did not reveal any shared variants or genes that harboured rare variants across the two families, and neither of the shared variants within each family was located in any of the two linkage regions (Table 2). Although we cannot exclude the presence of undetected exonic variants, the failure of WES to produce plausible disease-causing alleles may indicate that the underlying disease mechanism is non-coding, in keeping with the identification of rare non-coding variants in a number of 6q16 NCMD families.[17,18]

Locus heterogeneity and, ultimately, the presence of two different disorders across the two families remain another possibility to consider: there are differences in the non-ocular phenotype between the two families. In the French family the developmental abnormalities are confined to the eye and digits,

whereas in the British family some affected individuals have a more complex phenotype (e.g., individuals I:1, II:2 and III:2 present with unilateral kidney).[5,6] While the only two shared exonic variants after strict filtering in two distantly related members of family GC16500 were ruled out on the basis of the linkage, haplotype and segregation analyses, WES in two affected siblings from family GC16334 led to the identification of two novel missense variants, p.Thr96Ser in *C22orf15* and p.Gly209Ser in *IKBKB* (Table 3), that resided in a region of linkage (when analysing the British family only) (Supplementary Table 3) and segregated with the phenotype. However, we acknowledge that the power of the genetic analyses in the British family was limited by the availability of only closely related family members. *C22orf15* is an uncharacterized gene of unknown function, mildly expressed in the retina. The missense variant p.Thr96Ser is conserved from human to chicken and predicted to be 'benign' and 'tolerated' by in silico prediction tools PolyPhen and SIFT, respectively.

If the phenotypes observed in the two families are not allelic, we argue that the missense variant in *IKBKB* is the most likely genetic cause for the disorder seen in the British family. The variant p.Gly209Ser in *IKBKB* was scored at the top of our variant prioritisation analysis (highest Exomiser score, predicted to be deleterious and pLI=1) and alters a highly conserved amino acid across species. *IKBKB* is highly expressed in the retina and has been shown to play a role in skeletal development during embryogenesis.[38,39] Conditional knockout in murine osteoblasts and chondrocytes led to abnormalities in the skeletal and limb development in both heterozygous and homozygous mice.[39] Nevertheless, deletion of *IKBKB* in the mouse retina caused no obvious defect in retinal development or function.[40] It could be speculated that since the retinal phenotype seen in this family is specific to the macula, mouse models may not be informative. Further experiments on human cell models would be required to investigate this possibility. Moreover, the identification of additional families with similar phenotype and *IKBKB* variants would further strengthen the evidence for pathogenicity.

If the phenotypes observed in the two families are allelic, given the close similarity of their retinal phenotype with NCMD and the evidence for the presence of a 5p15 shared IBD chromosomal segment of size 5 cM among the affected members in the larger family, we conclude that the most

likely genetic cause is to be searched among undetected variants in the 5p15 region that partially overlaps with the NCMD MCDR3 locus. Taking advantage of the advances in whole-genome sequencing technologies is the ideal step forward for furthering the research on these two families. A comprehensive screening of both coding and non-coding sequences across the genome will likely help elucidate the underlying genetic cause of their phenotype that may consist in variation similar to the 900-kb tandem duplication that has been recently found in a single NCMD MCDR3 family of Danish origin.[18] Finally, the identification of the genetic cause of the NCMD phenotype seen in these two families will shed light on macular development and function with potential wide-reaching implications given the fundamental role of the macula in human vision.

FIGURE LEGENDS

Figure 1 Pedigree structure of two families affected by an autosomal dominant developmental macular disorder in association with digit abnormalities. Solid and open symbols indicate affected and unaffected individuals respectively. Presence of digit abnormalities in addition to macular lesion (Digit) and availability of DNA from individuals that were genotyped for linkage analysis (DNA) are indicated.

Figure 2 Colour fundus photographs, fundus fluorescein angiography images and X-rays and photographs of hands and feet performed on affected family members (Figure1). Family GC16500: a) individual IV:1, b) individual III:1, d) individual II:5, e), f) individual II:2 present with chorioretinal atrophy; c) individual V:1 presents with granular pigmentary changes; g) individual II:3 presents with granular changes with RPE atrophy; h), i), l), m) IV:4 presents with bilateral split hand and split foot malformation; n), o) individual III:1 presents with bilateral split hand malformation; p) individual IV:1 presents with minimal left hand clinodactyly. Family GC16334: q) individual III:2, r) individual III:3, s) individual II:1 present with chorioretinal atrophy; t), u) individual II:1 presents with bilateral hand and foot brachydactyly and skin syndactyly.

Figure 3 Identification of the candidate regions for family GC16500 using the Homozygosity Haplotype (HH) approach.[24] (A) Eight affected family members were included in the analysis. A densitogram of the genomic Regions with a Conserved Homozygosity Haplotype (RCHHs) is depicted. The darker the colour, the more individuals share a HH in the region. Black regions indicate RCHH that are shared by all 8 affected individuals. (B) A densitogram of -log(P) values for the representative RCHH shared by the 8 affected family members with the 3 unaffected family members as controls. The darker the colour, the more significant the RCHH is.

Acknowledgements

We thank Dr. Thomas Scerri and Prof. Melanie Bahlo for their help with the use of LINKDATAGEN, and Dr. Vincent Plagnol for his help with the whole-exome sequencing data analysis.

Funding

This work was supported by grants from the National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital National Health Service Foundation Trust and UCL Institute of Ophthalmology (UK), Fight for Sight including the Mercer Fund (UK), the Macular Society (UK), Moorfields Eye Hospital Special Trustees (UK), Moorfields Eye Charity (UK), the Foundation Fighting Blindness (FFB) (USA), Retinitis Pigmentosa Fighting Blindness (UK). Ambreen Kalhoro was the recipient of an Iris Fund Clinical Research Fellowship and Michel Michaelides is the recipient of an FFB Career Development Award.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the writing and content of this article.

References

1. Frank HR, I MBL, Williams RJ, Sidbury JB. A New Dominant Progressive Foveal Dystrophy. *Am J Ophthalmol*. 1974;78(6):903-916.

2. Lefler WH, Wadsworth JAC, Sidbury JB. Hereditary Macular Degeneration and Amino-Aciduria. *Am J Ophthalmol*. 1971;71(1):224-230.

3. Small KW. North Carolina Macular Dystrophy, Revisited. *Ophthalmology*. 1989;96(12):1747-1754.

4. Michaelides M, Jeffery G, Moore A. Developmental macular disorders: phenotypes and underlying molecular genetic basis. *Br J Ophthalmol*. 2012;96:917-924.

5. Sorsby A. Congenital coloboma of the macula. *Br J Ophthalmol*. 1935;19:65-90.

6. Thompson EM, Baraitser M. Sorsby syndrome: a report on further generations of the original family. *J Med Genet*. 1988;25(5):313-321.

7. Puech B. Colobomes maculaires héréditaires. In: *Les Dystrophies Héréditaires de La Macula. Bulletin d'ophtalmologie des Sociétés de France. Rapport annuel*, 1991:237-244.

8. Small KW, Weber JL, Roses A, Lennon F, Vance JM, Pericak-Vance M a. North Carolina macular dystrophy is assigned to chromosome 6. *Genomics*. 1992;13(3):681-685.

9. Pauleikhoff D, Sauer CG, Müller CR, Radermacher M, Merz A, Weber BH. Clinical and genetic evidence for autosomal dominant North Carolina macular dystrophy in a German family. *Am J Ophthalmol*. 1997;124(3):412-415.

10. Rabb MF, Mullen L, Yelchits S, Udar N, Small KW. A North Carolina macular dystrophy phenotype in a Belizean family maps to the MCDR1 locus. *Am J Ophthalmol*. 1998;125(4):502-508.

11. Reichel MB, Kelsell RE, Fan J, et al. Phenotype of a British North Carolina macular dystrophy family linked to chromosome 6q. *Br J Ophthalmol*. 1998;82(10):1162-1168.

12. Rohrschneider K, Blankenagel A, Kruse FE, Fendrich T, Völcker HE. Macular function testing in a German pedigree with North Carolina macular dystrophy. *Retina*. 1998;18(5):453-459.

13. Small KW, Puech B, Mullen L, Yelchits S. North Carolina macular dystrophy phenotype in France maps to the MCDR1 locus. *Mol Vis*. 1997;3(310):1.

14. Small K, Garcia C, Gallardo G, Udar N, Yelchits S. North Carolina macular dystrophy (MCDR1) in Texas. *Retina*. 1998;18(5):448-452.

15. Small KW, Udar N, Yelchits S, et al. North Carolina macular dystrophy (MCDR1) locus: a fine resolution genetic map and haplotype analysis. *Mol Vis*. 1999;5:38.

16. Z. Y, Z. T, L.J. C, et al. Clinical characterization and genetic mapping of North Carolina macular dystrophy. *Vision Res*. 2008;48(3):470-477.

17. Bowne SJ, Sullivan LS, Wheaton DK, et al. North Carolina macular dystrophy (MCDR1) caused by a novel tandem duplication of the PRDM13 gene. 2016;22:1239-1247.

18. Small KW, DeLuca AP, Whitmore SS, et al. North Carolina Macular Dystrophy Is Caused by Dysregulation of the Retinal Transcription Factor PRDM13. *Ophthalmology*. 2016;123:9-18.

19. Rosenberg T, Roos B, Johnsen T, et al. Clinical and genetic characterization of a Danish family with North Carolina macular dystrophy. *Mol Vis*. 2010;16:2659-2668.

20. Michaelides M, Johnson S, Tekriwal AK, et al. An early-onset autosomal dominant macular dystrophy (MCDR3) resembling North Carolina macular dystrophy maps to chromosome 5. *Investig Ophthalmol Vis Sci*. 2003;44(5):2178-2183.

21. Bahlo M, Bromhead CJ. Generating linkage mapping files from Affymetrix SNP chip data. *Bioinformatics*. 2009;25(15):1961-1962.

22. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet*. 2002;30:97-101.

23. Thiele H, Nürnberg P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics*. 2005;21(8):1730-1732.

24. Miyazawa H, Kato M, Awata T, et al. Homozygosity haplotype allows a genomewide search for the autosomal segments shared among patients. *Am J Hum Genet*. 2007;80(6):1090-1102.

25. Jiang H, Orr A, Guernsey DL, et al. Application of homozygosity haplotype analysis to genetic mapping with high-density SNP genotype data. *PLoS One*. 2009;4(4).

26. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.

27. Robinson PN, Köhler S, Oellrich A, et al. Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res*. 2014;24(2):340-348.

28. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: A Tool for Annotating and Analyzing Human Hereditary Disease. *Am J Hum Genet*. 2008;83(5):610-615.

29. Jäger M, Wang K, Bauer S, Smedley D, Krawitz P, Robinson PN. Jannovar: a java library for exome annotation. *Hum Mutat*. 2014;35(5):548-555.

30. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat*. 2011;32(8):894-899.

31. Amberger J, Bocchini C, Hamosh A. A new face and new challenges for Online Mendelian Inheritance in Man (OMIM®). *Hum Mutat*. 2011;32(5):564-567.

32. Maiella S, Rath A, Angin C, Mousson F, Kremp O. Orphanet et son réseau : où trouver une information validée sur les maladies rares. *Rev Neurol (Paris)*. 2013;169:S3-S8.

33. Bult CJ, Eppig JT, Blake JA, Kadin JA, Richardson JE, Mouse Genome Database Group. The mouse genome database: genotypes, phenotypes, and models of human disease. *Nucleic Acids Res*. 2013;41(Database issue):D885-91.

34. Smith CL, Goldsmith C-AW, Eppig JT. The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome Biol*. 2005;6(1):R7.

35. Van Slyke CE, Bradford YM, Westerfield M, Haendel MA. The zebrafish anatomy and stage ontologies: representing the anatomy and development of Danio rerio. *J Biomed Semantics*. 2014;5(1):12.

36. Smedley D, Oellrich A, Köhler S, et al. PhenoDigm: analyzing curated annotations to associate animal models with human diseases. *Database*. 2013;2013:bat025.

37. Samocha KE, Robinson EB, Sanders SJ, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet*. 2014;46(9):944-950.

38. Bakkar N, Wang J, Ladner KJ, et al. IKK/NF-kappaB regulates skeletal myogenesis via a signaling switch to inhibit differentiation and promote mitochondrial biogenesis. *J Cell Biol*. 2008;180(4):787-802.

39. Swarnkar G, Zhang K, Mbalaviele G, et al. Constitutive Activation of IKK2/NF-κB Impairs Osteogenesis and Skeletal Development. Awad HA, ed. *PLoS One*. 2014;9(3):e91421.

40. Lu H, Lu Q, Gaddipati S, et al. IKK2 Inhibition Attenuates Laser-Induced Choroidal Neovascularization. Chen J, ed. *PLoS One*. 2014;9(1):e87530.