Identification and functional characterisation of genetic variants in *OLFM2* in children with developmental eye disorders

Holt R[1], Ugur Iseri SA[2], Wyatt AW[3], Bax DA[1], Gold Diaz D[4], Santos C[4], Broadgate S[1, 5], Dunn R[6], Bruty J[7], Wallis Y[7], McMullan D[7], Ogilvie C[8], Gerrelli D[4], Zhang Y[1], Ragge NK[1, 9]

1 Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK

2 Department of Genetics, Aziz Sancar Institute of Experimental Medicine, Istanbul University, Istanbul, Turkey

3 Department of Urologic Sciences, University of British Columbia, Vancouver, Canada

4 Neural Development Unit, Institute of Child Health, University College London, London, UK

5 Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

6 Department of Genetics, Viapath, Guy's Hospital, London, UK

7 West Midlands Regional Genetics Laboratory, Birmingham Women's Hospital, Birmingham, UK

8 Department of Cytogenetics, Guy's and St Thomas' NHS Foundation Trust, London, UK

9 Clinical Genetics Unit, West Midlands Regional Genetics Service, Birmingham Women's Hospital, Birmingham, UK

*Corresponding author: Professor Nicola Ragge, MD DM FRCP FRCOphth FRCPCH

School of Life Sciences, Oxford Brookes University, Gipsy Lane, Oxford, OX3 0BP, UK; Clinical Genetics Unit, West Midlands Regional Genetics Service, Birmingham Women's Hospital, Mindelsohn Way, Birmingham, B15 2TG

Tel: +44 1865 484413

Fax: +44 1865 742177

E-mail: nragge@brookes.ac.uk; nicola.ragge@bwnft.nhs.uk

Acknowledgements

We would like to thank the patients and their families for their participation in our study. We are grateful to the following: Youming Zhang, Imperial College London, for contributing vectors; Stephanie Halford, University of Oxford, for providing cDNA and cell lines; Andrew Gallagher (paediatrician), Worcestershire Royal Hospital for initial patient referral; Electrophysiology Department, Great Ormond St Hospital, for electrophysiology studies; M Parulekar (Birmingham Children's Hospital), K Nischal, Y Abou-Rayyah for clinical care; Thalia Antoniadi (West Midlands Regional Genetics Laboratory) for assistance with the gene panel sequencing. This work was supported by grants from Baillie Gifford, Visually Impaired Children Taking Action (VICTA) (http://www.victa.org.uk/) and Microphthalmia, Anophthalmia and Coloboma Support (MACS) (www.macs.org.uk). The human embryonic and fetal material was provided by the Joint Medical Research Council (MRC)/Wellcome Trust (grant #099175/Z/12/Z) Human Developmental Biology Resource (www.hdbr.org).

Abstract

Anophthalmia, microphthalmia and coloboma are a genetically heterogeneous spectrum of developmental eye disorders and affect around 30 per 100,000 live births. *OLFM2* encodes a secreted glycoprotein belonging to the noelin family of olfactomedin domain-containing proteins that modulate the timing of neuronal differentiation during development. *OLFM2* SNPs have been associated with open angle glaucoma in a case-control study, and knockdown of *Olfm2* in zebrafish results in reduced eye size. From a cohort of 258 individuals with developmental eye anomalies, we identified two with heterozygous variants in *OLFM2*: an individual with bilateral microphthalmia carrying a *de novo* 19p13.2 microdeletion involving *OLFM2*, and a second individual with unilateral microphthalmia and contralateral coloboma who had a novel single base change in the 5' untranslated region. Dual luciferase assays demonstrated that the latter variant causes a significant decrease in expression of *OLFM2*. Furthermore, *in situ* hybridisation experiments using human developmental tissue revealed expression in relevant structures, including the lens vesicle and optic cup. Our study indicates that *OLFM2* is likely to be important in mammalian eye development and disease and should be considered as a gene for human ocular anomalies.

Key words: OLFM2, anophthalmia, microphthalmia, coloboma, eye abnormalities, 5' untranslated regions

Introduction

Developmental eye anomalies, including anophthalmia, microphthalmia and coloboma (AMC) are a genetically heterogeneous group of disorders affecting around 30 per 100,000 live births (Campbell et al., 2002). A rapidly growing number of genes have been identified for these anomalies, including *SOX2* Fantes et al., 2003), *OTX2* (Ragge et al., 2005) and *BMP4* (Bakrania et al., 2008). Following the identification of a *de novo* 19p13.2 microdeletion involving *OLFM2* in an individual with bilateral microphthalmia, we investigated the role of this gene in developmental eye disorders.

Single nucleotide polymorphisms (SNPs) in *OLFM2* have been previously associated with open angle glaucoma (Funayama et al., 2006). *OLFM2* has a 31% homology to *MYOC* (*Myocilin*), mutations of which account for 3-4% of primary open angle glaucoma and a significant proportion of juvenile open angle glaucoma (Mukhopadhyay et al., 2004).

OLFM2 encodes a secreted glycoprotein belonging to the noelin family of olfactomedin domain-containing proteins that modulate the timing of neuronal differentiation during development (Gautier et al., 2009; Tomarev and Nakaya, 2009). OLFM2 is present in the nucleus and is predicted to contain a DNA binding domain and may play a role as a transcriptional coactivator (Shi et al., 2014). It is expressed in human (Mukhopadhyay et al., 2004), mouse (Sultana et al., 2011) and zebrafish eyes, zebrafish knockdowns exhibiting an average 24% decrease in eye area (Lee et al., 2008). This reduction is possibly mediated by decreased expression of *Pax6*, which is observed in zebrafish *Olfm2* knockdowns (Lee et al., 2008). Heterozygous mutations in *PAX6* cause aniridia, Peters' anomaly, foveal hypoplasia, keratitis, cataract, coloboma and microphthalmia (OMIM #607108). OLFM2 may play a role in TGFβ signaling, which includes BMP4, during early eye development since OLFM2 expression has been shown to increase after TGFβ induction in human embryonic stem cell–derived mesenchymal cells (Shi et al., 2014).

In addition to eye size, *Olfm2* zebrafish knockdowns show altered nervous system development, including optic tectum, defects of the midbrain–hindbrain boundary, and altered axonal trajectories and extension (Lee et al., 2008). Interestingly, a distinct cutoff in *OTX2* expression at the midbrain-hindbrain boundary is observed in the human embryo at CS14 (Ragge et al., 2005). Expression of *Olfm2* is initially present in the entire zebrafish eye, becoming increasingly localised to the retinal ganglion cell layer as development progresses.

Therefore, *OLFM2* represents an excellent candidate gene for AMC. We describe an individual with bilateral microphthalmia and a *de novo* deletion on chromosome 19p13.2 affecting *OLFM2* and *COL5A3*. We screened *OLFM2* for variants in 258 patients with developmental eye anomalies, identifying a second individual with a 5' untranslated region (UTR) variant that decreases *OLFM2* expression. We demonstrate expression of *OLFM2* in early human embryonic tissue including in the optic cup and lens vesicle.

Materials and Methods

Cohort Description

A UK cohort with ocular anomalies, principally AMC, was recruited as part of a national 'Genetics of Eye and Brain anomalies' study (REC study number 04/Q0104/129). Patients were screened by indication in known AMC genes. From this cohort 258 individuals without a genetic diagnosis were screened for variants in *OLFM2*. Of these, 32 individuals had bilateral anophthalmia or unilateral anophthalmia with contralateral microphthalmia, 43 had bilateral microphthalmia (+/- coloboma), 146 had unilateral anophthalmia/microphthalmia (+/- contralateral defects), and 37 had other phenotypes, including anterior segment dysgenesis.

Array-CGH

Array CGH was performed using the Agilent (USA) 44K oligonucleotide platform (design 017457), as previously described (Ahn et al., 2010).

Mutation Analysis

All six coding exons of *OLFM2* (NM_058164.3), the alternate first exon of NM_001304348.1, and flanking sequences were screened for mutations by high resolution melting curve analysis on a LightScanner (Idaho Technology) followed by Sanger sequencing of samples with aberrant melting curves (see Supplementary Methods and Supplementary Table 1 for detailed description) and sequence data analysed using CodonCode Aligner (CodonCode Corporation, Dedham, MA) or Chromas (Technelysium).

Gene Panel Sequencing

Next generation sequencing was performed for Case 1 using a custom designed Nextera based chemistry (Illumina) panel of 546 genes divided between 351 known eye disorder genes and 195 candidate genes, selected from animal models, pathophysiological pathways, and preliminary studies. Sequencing was performed using a custom designed Nextera based chemistry (Illumina) containing in excess of 10,225 probes targeting 1,380,306bp. Sequence capture, sample barcoding and pre-enrichment sample pooling were performed according to manufacturer's instructions (Illumina) and paired-end sequencing performed on the MiSeq sequencing platform (Illumina). Sequences were aligned to the reference human genome (UCSC hg19) using BWA-MEM (http://bio-bwa.sourceforge.net/) and Novoalign v3.01.00 (Novocraft). The mean coverage was 230x. Identified variants were filtered based on sequence quality, minor allele frequency and predicted effect. Filters were applied to remove variants with low quality ('scThreshold': 0.94999999999999996, 'abThreshold': 0.001, 'minMapQual': 20, 'minBaseQual': 20, 'badReadsThreshold': 20, 'hapScoreThreshold': 4). In addition, all synonymous variants and those with a frequency of at least 1% in the 1000 Genomes

Project database, or those which occurred in multiple patients run on this custom panel were removed.

Transcription Factor Binding Bioinformatical Analysis

Potential effects of the c.-57C>G *OLFM2* 5' UTR variant were examined using the online program TFbind (http://tfbind.hgc.jp/) (Tsunoda and Takagi, 1999). Sequences containing the wildtype and variant alleles plus 15 bases of flanking sequencing were used as input data, and the program run using default settings.

***In Situ* Hybridisation**

Nonradioactive RNA *in situ* hybridization was performed on human embryo sections at Carnegie Stages (CS) 15 and 17 as described elsewhere (Bakrania et al., 2008). Probes were designed to span the end of the coding region and part of the 3' UTR of the *OLFM2* to bind all known splice variants and generated using primers GCTCGGCGGCCGCTTCACGTCATCAGCACCTCT and GCTCGGTCGACTCCATGCACTCAACTCCTGG. Human embryos were obtained from the MRC/Wellcome Trust Human Developmental Biology Resource, UCL, with full ethical approval.

Cell Culture and cDNA Generation

Y79 (human retinoblastoma) cells cultured in RPMI1640 media supplemented with 10mmol HEPES, 1mmol sodium pyruvate, 4.5g/L glucose, 2mmol L-glutamine, 20U/ml penicillin and 0.1mg/ml streptomycin and 20% fetal bovine serum (Sigma, Gillingham, UK). Cells were cultured at 37°C and 5% CO_2. RNA was extracted using the RNeasy Mini Kit (Qiagen, UK). cDNA was generated using the QuantiTect Reverse Transcription kit (Qiagen).

cDNA PCR

Primers were designed to span the unannotated 5' UTR variant (Supplementary Table 2). PCRs were performed using 1:10 diluted cDNA, 1X HotShot^{TM} Mastermix (Cadama Medical), 10% v/v DMSO (Sigma) and 0.25µM of each primer in a final volume of 10µl. PCRs were 35 cycles and used a 55^{o}C annealing temperature. *OLFM2* expression was confirmed using a second PCR with primers for exon 6 that were also used for the LightScanner analysis (Supplementary Table 2).

Dual Luciferase Assays

Constructs were generated to incorporate the unannotated *OLFM2* 5' UTR (NM_058164.3) variant from case 2 using primers (GCTCGGGTACCGCAACAAAGACTCGGAGCG and GCTCGGAGCTCCGTGAGCGGCCACATGAC) and cloned in the pGL3-Promoter vector (Promega). Vectors containing wild type and variant inserts were identified by Sanger sequencing, cultured overnight and DNA extracted.

Transfections were performed using plasmids containing wild type, mutant or no insert. Normalisation was performed using pRL-TK plasmids (Promega). Per reaction, 2.5µg plasmid was diluted in 500µl Opti-MEM® I Reduced Serum Medium (Gibco) and 5µl Plus^{TM} Reagent (Invitrogen), incubated at room temperature 15 minutes, 5µl Lipofectamine® LTX Reagent (Invitrogen) added and incubated for 15 minutes at room temperature. 500µl plasmid/lipofectamine mix was added to each well containing 3 x 10^5 Y79 cells in 2ml RPMI1640 media (Sigma) and incubated at 37^{o}C overnight. Cell lysates were obtained using Passive Lysis Buffer (PLB) (Promega) (manufacturer's instructions), resuspended in 60µl PLB and stored at -80^{o}C. Assays were performed using the Dual-Luciferase Reporter Assay System (Promega) (manufacturer's instructions) and measured on a MicroBeta TriLux using Micro Beta Windows Workstation software (PerkinElmer).

Results

Case 1

A male infant was born to non-consanguineous Caucasian parents by normal delivery at 41 weeks' gestation weighing 3.046kg following an uneventful pregnancy. He had apparently isolated bilateral mild microphthalmia with cloudy vascularised sclerocornea. There was a reasonable view to a normal right disc and retina and no fundal view on the left. Cranial ultrasound examination was normal. His initial growth was slow, along the 2nd centile, length 25th centile, and head circumference 0.4th centile at 8 months of age. At 2 months of age, ocular ultrasound examination demonstrated the eyes had axial lengths of 15.1mm and 15.3mm (average for age 16.5mm). Flash visual evoked potentials and electroretinograms were elicited from both left and right eyes. He developed a mild bilateral ptosis, early hearing problems and had narrow ear canals. His speech was delayed initially. At six years-of-age, his height was 107cm (2nd centile), weight 16kg (2nd centile) and head circumference 49 cm (0.4th centile). He had a relatively flat face with prominent chin and paucity of facial movements, small teeth, and unusual low set ears with slightly tilted earlobes and a bifid uvula. His cognitive abilities appear to be normal, considering his visual impairment.

Array CGH analysis revealed a *de novo* 19p13.2 microdeletion of between 46kb (chr19:9,913,222–9,959,613) and 81kb (chr19:9,896,892–9,978,107) (GRCh38/hg38) partially involving *OLFM2* and *COL5A3*. The minimal deleted region is predicted to remove the first exon of two splice forms of *OLFM2* (NM_058164.3 and NM_001304347.1) and part of the 3' UTR of the final exon of *COL5A3*. In contrast, while the maximal deleted region is predicted to affect the same two exons of *OLFM2*, it also removes the final 26 exons of *COL5A3* (Figure 1a). Long range PCR across the region failed to generate a product (data not shown). G-banded chromosome analysis detected a *de novo* balanced reciprocal translocation between chromosomes 3 and 13 (46,XY,t(3;13)(p23;q32)dn), detailed array CGH did not reveal any copy number variation in the regions involved in the translocation. Parental karyotyping was normal and targeted aCGH of the deleted region showed no copy number variation. As the

translocation was balanced, the deletion was considered more likely to be relevant to ocular phenotype. In addition, after panel sequencing of 546 known and candidate eye development genes using a new purpose designed gene panel, Case 1 was found to be heterozygous for a variant (rs143247685, NM_004380.2: c.2728A>G, p.Thr910Ala) in *CREBBP*, alterations in which can cause Rubenstein-Taybi syndrome 1. However, as this is a known polymorphism with a population frequency of 0.226%, is not located within a functional domain of the gene and affects a non-conserved amino acid, it was deemed unlikely to affect the phenotype. No other plausible pathogenic variants were identified.

As *OLFM2* is expressed in the human eye, associated with glaucoma, and zebrafish knockdowns demonstrate microphthalmia, with no comparable reports for *COL5A3* suggesting a role in developmental eye anomalies, we prioritised *OLFM2* and screened it for variants in our patient cohort.

Case 2

Case 2 is a male with unilateral microphthalmia with a unilateral cyst affecting the same eye and a contralateral coloboma, born to non-consanguineous parents at full term by normal delivery. His birth weight was 2.89kg. He is registered blind, has normal growth and cognitive development, and no other medical issues. No causal variants were identified in *SOX2* (Bakrania et al. 2007), *OTX2* (Ragge et a., 2005), *SHH* (Bakrania et al. 2010), *FOXE3* (Iseri et al., 2009), *BMP4* (Bakrania et al., 2008) or *BMP7* (Wyatt et al., 2010) via a combination of Sanger sequencing, fluorescence in situ hybridisation (FISH) and Multiplex Ligation-dependent Probe Amplification (MLPA) analysis (data not shown), and his chromosome analysis was normal. We identified a novel single base substitution of c.-57C>G in the 5' UTR of *OLFM2* (see below).

Mutation analysis

OLFM2 is located on chr19:9,853,718-9,936,552 (GRCh38/hg38), spans 82,835bp and comprises six exons. It has three splice forms with alternate first exons of 2019bp

(NM_058164.3), 1692bp (NM_001304348.1) and 1993bp (NM_001304347.1) encoding proteins of 454, 376 and 478 amino acids, respectively. A cohort of 258 patients were screened for variations in the exons and flanking intronic sequences, except NM_001304347.1 exon 1 (no working assay). We identified ten annotated variants, none reported pathogenic (ClinVar database); rs143959085, rs201189277, rs79341807, rs200752028, rs116666369, rs544945664, rs2303100, rs11556087, rs11556088, rs34961482 (Supplementary Table 3). Of these, one was intronic (rs116666369), four were synonymous (rs79341807, rs11556088, rs200752028, rs3496148). One variant was in the 5' UTR (rs544945664), but due to a minor allele frequency nearing 0.01, it was considered unlikely to have a pathogenic role. The remaining four variants were missense changes. However, rs2303100 and rs11556087 have minor allele frequencies greater than 0.01 (dbSNP147). In addition, while the minor alleles for rs201189277 and rs143959085 are rare, they are not predicted damaging by SIFT (http://sift.jcvi.org/) (Ng and Henikoff 2003) or PolyPhen (,http://genetics.bwh.harvard.edu/pph/) (Adzhubei et al. 2010) and are in regions of limited conservation. rs201189277 was present in one individual with unilateral anophthalmia, now deceased, for whom parental DNA was unavailable. rs143959085 was identified in one individual with anophthalmia with contralateral normal eye, bronchiectasis, congenital heart disease, skeletal problems, undescended testes, umbilical hernia, inguinal hernia, and polydactyly of both hands, but was inherited from the asymptomatic father. Therefore, these variants were considered unlikely to contribute to the patient phenotypes.

A previously unreported 5' UTR c.-57C>G variant (chr19:9,936,451 [GRCh38/hg38]) was identified in Case 2, validated by Sanger sequencing, and inherited from his asymptomatic mother. Analysis using TFbind (http://tfbind.hgc.jp/) (Tsunoda and Takagi, 1999) identified potential effects on transcription factor binding, including General Transcription Factor IIIA, Sp1 Transcription Factor and Early Growth Response 3.

Luciferase Assays

OLFM2 expression in Y79 cells was confirmed by PCR from cDNA. Two sets of triplicate transfection reactions were performed using plasmids containing wild type, mutant or no inserts, and normalised to *Renilla* luciferase expression from a co-transfected pRL-TK plasmid containing no insert. A 21% decrease in expression from the mutant insert relative to wild type was identified (2 tailed unpaired t-test $P = 0.0209$) (Figure 1b).

In Situ Hybridisations

In situ hybridisations were performed on transverse sections of human CS15 and CS17 brain and eye tissue. At CS15 *OLFM2* expression was observed in structures including the rhombencephalon, vagus, tissue surrounding the adenohypophyseal pouch, hypothalamus, otic vesicle, lens vesicle and optic cup. At CS17 we observed expression in the lens vesicle, cerebral vesicle, spinal ganglion, ventricular zone, intermediate zone, marginal zone and the tissue surrounding the diencephalic superventricle (Figure 2a-m).

Discussion

The identification of a *de novo* deletion affecting *OLFM2* and *COL5A3* in a boy with bilateral microphthalmia and sclerocornea led to a positional candidate gene approach to investigate *OLFM2* in human eye development. We identified a further case with unilateral microphthalmia plus cyst and contralateral coloboma with an unannotated mutation in the 5' UTR of *OLMF2*, which significantly decreases expression of *OLFM2 in vitro. OLFM2* expression was demonstrated in multiple regions of the developing human embryo, including the lens vesicle and optic cup, but also primitive hypothalamus-pituitary axis, spinal ganglion and otic vesicle.

Case 1 had a *de novo* deletion affecting *OLFM2* and *COL5A3*. Although microdeletions spanning this region in other cases have been reported (Wong et al., 2007; Shaikh et al., 2009), it is unknown whether affected individuals were screened for eye phenotypes. Given prior reports of the role of *OLFM2* in developmental eye disease (Mukhopadhyay

et al., 2004; Funayama et al., 2006; Lee et al., 2008), we screened *OLFM2* in our cohort of 258 patients with developmental eye anomalies.

The unannotated heterozygous substitution in the 5' UTR of *OLFM2* in Case 2 was predicted to alter transcription factor binding, and demonstrated a 21% decrease in *in vitro* expression. The asymptomatic status of the carrier mother suggests this variant exhibits variable penetrance, often seen in eye development genes. The role of UTR and promoter variants in human disease is increasingly recognised. For example, a 5' UTR substitution in *PDE6H* in a patient with cone dystrophy increases *PDE6H* expression leading to 4-5 times more protein, possibly resulting in cone degeneration by elevating cGMP *in vivo* (Piri et al., 2005). Zhou *et al* identified a novel heterozygous *SOX2* 3' UTR variant in an individual with bilateral anophthalmia, suggesting this may affect expression (Zhou et al., 2008). Investigation of variants located within control regions of developmental eye genes is likely to provide an important source of causal changes and aid delineation of eye gene networks.

The potential contribution of Case 1's deletion and Case 2's unannotated variant to their phenotype is supported by the expression of *OLFM2* in the developing human optic cup and lens vesicle. *OLFM2* was also expressed in the developing ear, hypothalamic-pituitary axis, and other parts of brain and neural tube, which may partly explain narrow ear canals, and reduced growth observed in Case 1. This indicates that more highly penetrant *OLFM2* variants may cause wide-ranging phenotypes and could be important for developmental disorders affecting other structures. Findings, including the bifid uvula, may relate to effects of *COL5A3* since this is seen in other collagen genes (Mark et al., 2011). The relatively mild phenotype observed in Case 2 is consistent with the moderate effect of the variant on expression, and the modest reduction in eye size seen in zebrafish models (Lee et al., 2008). *OLFM2* may function by regulating transcription as an enhancer via the TGFβ cascade in eye development (Shi et al., 2014) and by interaction with PAX6 (Lee et al., 2008). The TGFβ cascade has been implicated in eye

development (Lovicu et al., 2011), and *PAX6* is a well-known control gene involved in this process (Hever et al., 2006). Thus alterations in expression caused by mutations may have cascading effects through these pathways. However, further variants within additional patient cohorts will be required to more accurately ascertain the range of phenotypes associated with changes in *OLFM2*.

Our investigations have not excluded the role of the partial *COL5A3* deletion contributing to the phenotype of Case 1, which includes a number of non-ophthalmologic features. *COL5A3* is closely related to *COL5A1*, which encodes a protein that forms a heterotrimer with COL5A1 and COL5A2 to generate one form of type V collagen. In the mouse embryo *Col5a3* is primarily expressed in developing muscle and ligaments. In humans, *COL5A3* is expressed in multiple tissues in both the adult and fetus, including multiple regions of the brain, while in 15.5dpc mice it is primarily expressed in muscle, ligaments and connective tissues (Imamura et al., 2000). During zebrafish development, *Col5a3* is expressed in the notochord, as well as the retinal pigment epithelium, and mesenchyme cells of the eye and lens (Fang et al., 2010). Therefore, it is possible that the loss of the 3' end of *COL5A3* may contribute to the eye phenotype of Case 1. In addition, both *COL5A1* and *COL5A2* (but not *COL5A3*) have been implicated in classical Ehlers-Danlos syndrome, a complex disorder with multiple phenotypes involving the connective tissue, including skin, eyes, ears and teeth (Imamura et al., 2000; Hoffman et al. 2008) (OMIM #130000). This is of particular interest, as the Case 1 phenotype includes small ears and low set ears with slightly tilted earlobes. Therefore, further investigation of the role of this gene is warranted.

In conclusion, we identify genetic variants of *OLFM2* in two individuals with developmental eye disorders and reveal expression in the developing human eye, indicating that it is an important candidate for screening in other cohorts with developmental eye anomalies, and highlighting the importance of screening control

regions both to identify causal variants and to aid understanding of the networks of genes involved in eye development.

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010). A method and server for predicting damaging missense mutations. Nat Methods 7:248-249.

Ahn JW, Mann K, Walsh S, Shehab M, Hoang S, Docherty Z, Mohammed S, Mackie Ogilvie C. 2010. Validation and implementation of array comparative genomic hybridisation as a first line test in place of postnatal karyotyping for genome imbalance. Mol Cytogenet 3:9.

Bakrania P, Robinson DO, Bunyan DJ, Salt A, Martin A, Crolla JA, Wyatt A, Fielder A, Ainsworth J, Moore A, Read S, Uddin J, Laws D, Pascuel-Salcedo D, Ayuso C, Allen L, Collin JR, Ragge NK. 2007. SOX2 anophthalmia syndrome: 12 new cases demonstrating broader phenotype and high frequency of large gene deletions. Br J Ophthalmol. 91(11):1471-6.

Bakrania P, Efthymiou M, Klein JC, Salt A, Bunyan DJ, Wyatt A, Ponting CP, Martin A, Williams S, Lindley V, Gilmore J, Restori M, et al. 2008. Mutations in BMP4 cause eye, brain, and digit developmental anomalies: overlap between the BMP4 and hedgehog signaling pathways. Am J Hum Genet 82:304-319.

Bakrania P, Ugur Iseri SA, Wyatt AW, Bunyan DJ, Lam WW, Salt A, Ramsay J, Robinson DO, Ragge NK. 2010. Sonic hedgehog mutations are an uncommon cause of developmental eye anomalies. Am J Med Genet A. 152A(5):1310-3.

Campbell H, Holmes E, MacDonald S, Morrison D, Jones I. 2002. A capture-recapture model to estimate prevalence of children born in Scotland with developmental eye defects. J Cancer Epidemiol Prev 7:21-28.

Fang M, Adams JS, Mcmahan BL, Brown RJ, Oxford JT. 2010. The expression patterns of minor fibrillar collagens during development in zebrafish. Gene Expr Patterns 10(7-8):315-322.

Fantes J, Ragge NK, Lynch SA, McGill NI, Collin JR, Howard-Peebles PN, Hayward C, Vivian AJ, Williamson K, van Heyningen V, FitzPatrick DR. 2003. Mutations in SOX2 cause anophthalmia. Nat Genet 33:461-463.

Funayama T, Mashima Y, Ohtake Y, Ishikawa K, Fuse N, Yasuda N, Fukuchi T, Murakami A, Hotta Y, Shimada N; Glaucoma Gene Research Group. 2006. SNPs and Interaction Analyses of Noelin 2, Myocilin, and Optineurin Genes in Japanese Patients with Open-Angle Glaucoma. Invest Ophthalmol Vis Sci 47(12):5368-5375.

Gautier M, Flori L, Riebler A, Jaffrézic F, Laloé D, Gut I, Moazami-Goudarzi K, Foulley JL. 2009. A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. BMC Genomics 21:10:550.

Hever AM, Williamson KA, van Heyningen V. 2006. Developmental malformations of the eye: the role of *PAX6*, *SOX2* and *OTX2*. Clin Genet 69(6):459-470.

Hoffman GG, Dodson GE, Cole WG, Greenspan DS. 2008. Absence of apparent disease causing mutations in COL5A3 in 13 patients with hypermobility Ehlers-Danlos syndrome. Am J Med Genet A. 146A(24):3240-1.

Imamura Y, Scott IC, Greenspan DS. 2000. The Pro-α3(V) Collagen Chain: Complete primary structure, expression domains in adult and developing tissues, and comparison to the structures and expression domains of the other types V and XI procollagen chains. J Bio Chem 275(12):8749-8759.

Iseri SU, Osborne RJ, Farrall M, Wyatt AW, Mirza G, Nürnberg G, Kluck C, Herbert H, Martin A, Hussain MS, Collin JR, Lathrop M, Nürnberg P, Ragoussis J, Ragge NK. 2009. Seeing clearly: the dominant and recessive nature of FOXE3 in eye developmental anomalies. Hum Mutat. 30(10):1378-86.

Lee JA, Anholt, RRH, Cole GJ. 2008. Olfactomedin-2 mediates development of the anterior central nervous system and head structures in zebrafish. Mech Dev 125:167-181.

Lovicu FJ, McAvoy JW, de Iongh RU. 2011. Understanding the role of growth factors in embryonic development: insights from the lens. Philos Trans R Soc Lond B Biol Sci 366(1568):1204-1218.

Mark PR, Torres-Martinez W, Lachman RS, Weaver DD. 2011. Association of a p.Pro786Leu variant in COL2A1 with mild spondyloepiphyseal dysplasia congenita in a three-generation family. Am J Med Genet A 155A(1):174-179.

Mukhopadhyay A, Talukdar S, Bhattacharjee A, Ray K. 2004. Bioinformatic approaches for identification and characterization of olfactomedin related genes with a potential role in pathogenesis of ocular disorders. Mol Vis 10:304-314.

Ng PC, Henikoff S (2003). SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812-3814.

Piri N, Gao YQ, Danciger M, Mendoza E, Fishman GA, Farber DB. 2005. A substitution of G to C in the cone cGMP-phosphodiesterase gamma subunit gene found in a distinctive form of cone dystrophy. Ophthalmology 112(1):159-166.

Ragge NK, Brown AG, Poloschek CM, Lorenz B, Henderson RA, Clarke MP, Russell-Eggitt I, Fielder A, Gerrelli D, Martinez-Barbera JP, Ruddle P, Hurst J, et al. 2005. Heterozygous mutations of OTX2 cause severe ocular malformations. Am J Hum Genet 76:1008-1022.

Shaikh TH, Gai X, Perin JC, Glessner JT, Xie H, Murphy K, O'Hara R, Casalunovo T, Conlin LK, D'Arcy M, Frackelton EC, Geiger EA, et al. 2009. High-resolution mapping and

analysis of copy number variations in the human genome: a data resource for clinical and research applications. Genome Res 19(9):1682-1690.

Shi N, Guo X, Chen SY. 2014. Olfactomedin 2, a novel regulator for transforming growth factor-β–induced smooth muscle differentiation of human embryonic stem cell–derived mesenchymal cells. Mol Biol Cell 25(25):4106-4114.

Sultana A, Nakaya N, Senatorov VV, Tomarev SI. 2011. Olfactomedin 2: Expression in the eye and interaction with other olfactomedin domain-containing proteins. Invest Ophthalmol Vis Sci 52(5):2584-2592.

Tomarev SI, Nakaya N. 2009. Olfactomedin Domain-Containing Proteins: Possible Mechanisms of Action and Functions in Normal Development and Pathology. Mol Neurobiol 40(2):122–138.

Tsunoda T, Takagi T. 1999. Estimating Transcription Factor Bindability on DNA. Bioinformatics 15(7/8):622-630.

Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, Lam WL. 2007. A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet 80(1):91-104.

Wyatt AW, Osborne RJ, Stewart H, Ragge NK. 2010. Bone morphogenetic protein 7 (BMP7) mutations are associated with variable ocular, brain, ear, palate, and skeletal anomalies. Hum Mutat. 31(7):781-7.

Zhou J, Kherani F, Bardakjian TM, Katowitz J, Hughes N, Schimmenti LA, Schneider A, Young TL. 2008. Identification of novel mutations and sequence variants in the *SOX2* and *CHX10* genes in patients with anophthalmia/microphthalmia. Molecular Vision 14:583-592.

Figure Legends

Figure 1: Effect of unannotated variants identified within *OLFM2*. a. Copy number variant in Case 1. **b.** Dual-luciferase reporter assay showing reduced expression driven from the novel *OLFM2* 5' UTR variant c.-57C>G (chr19:9,936,451 [GRCh38/hg38]) relative to the wild type sequence. The *Renilla* luciferase expressing plasmid pRL-TK was co-transfected as a control.

Figure 2: OLFM2 *in situ* hybridisation studies in the developing human. a, b: *In situ* hybridisations of coronal sections of a CS15 human fetus using *OLFM2* antisense (a) and sense (negative control) (b) probes reveal expression in structures including the diencephalic superventricle (future third ventricle), optic cup, lens vesicle, Rathke's pouch (future pituitary), trigeminal ganglion, otic vesicle, vagus nerve and rhombencephalon (future fourth ventricle). Highest expression was observed in the dorsal portion of the rhombencephalon. **c, d:** *In situ* hybridisations of coronal sagittal sections of a CS17 human fetus using *OLFM2* antisense (c) and sense (negative control) (d) show expression of *OLFM2* in diencephalic superventricle, optic cup, lens vesicle, neural tube, and in the dorsal root ganglia. **e-j:** High magnification images of *OLFM2* expression in the CS15 human fetus in multiple structures: (e) lens vesicle and retinal ganglion cell layer, (f) trigeminal ganglion, (g) vagus nerve, (h) vestibular cochlear ganglion and otic vesicle, (i) diencephalic superventricle and Rathke's pouch, and (j) rhombencephalon. **k-m:** High magnification images of *OLFM2* expression in the CS17 human fetus in multiple structures: (k) lens vesicle and retinal ganglion cell layer, (l) neural tube and differentiating neurons, and (m) the dorsal root ganglia. Abbreviations: DG = dorsal root ganglion; DN = differentiating neurons of the neural tube; DS = diencephalic superventricle; LV = lens vesicle; NT = neural tube; OC = optic cup; OV = otic vesicle; R = rhombencephalon; RG = retinal ganglion cell layer; RP = Rathke's pouch; TG = trigeminal ganglion; V = vagus nerve; VCG = vestibular cochlear ganglion.

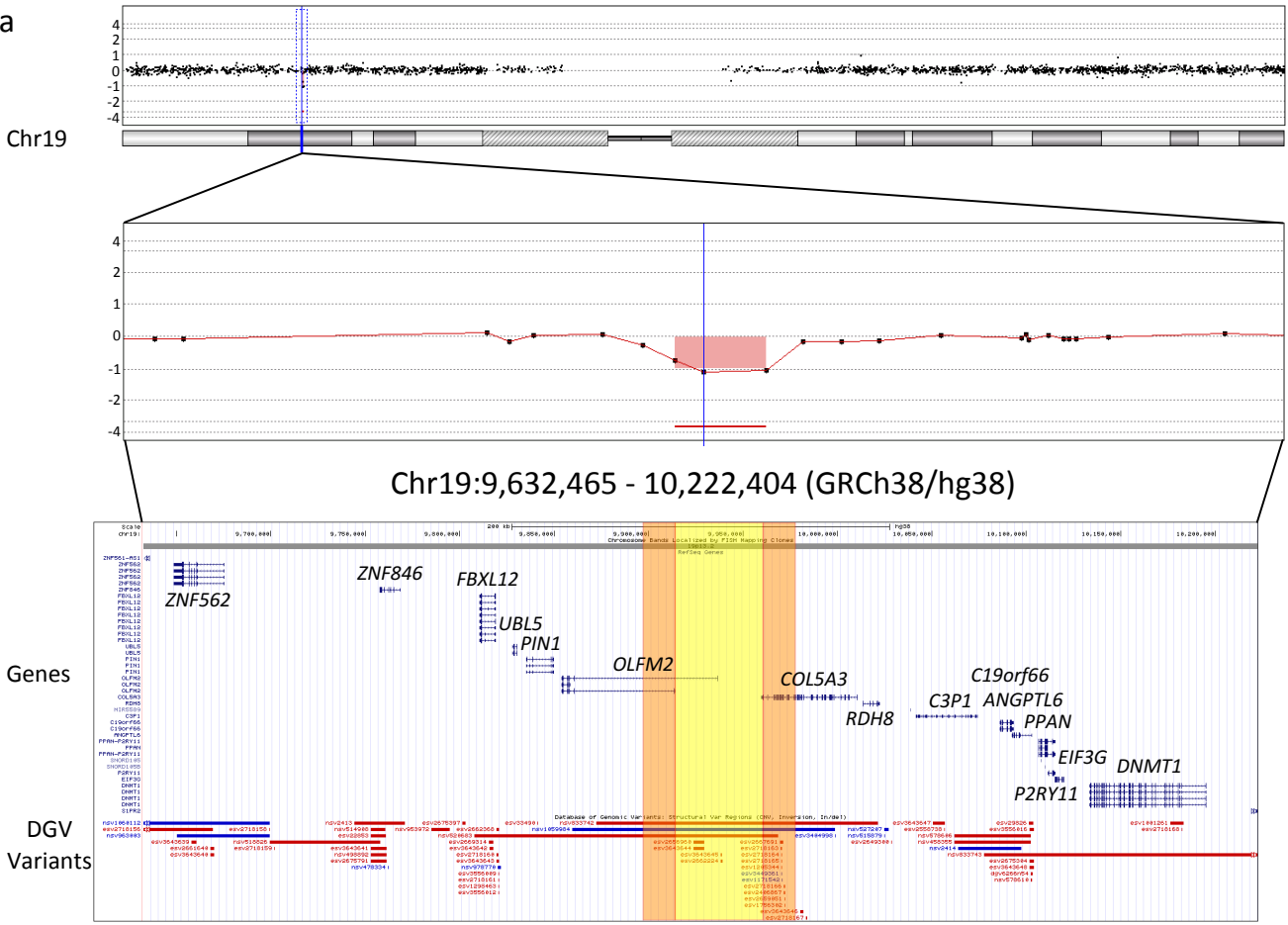
Figure Legends

Figure 1: Effect of unannotated variants identified within *OLFM2*. a. Copy number variant in Case 1. **b.** Dual-luciferase reporter assay showing reduced expression driven from the novel *OLFM2* 5' UTR variant c.-57C>G (chr19:9,936,451 [GRCh38/hg18]) relative to the wild type sequence. The *Renilla* luciferase expressing plasmid pRL-TK was co-transfected as a control.

Figure 2: OLFM2 *in situ* hybridisation studies in the developing human. a, b: *In situ* hybridisations of coronal sections of a CS15 human fetus using *OLFM2* antisense (a) and sense (negative control) (b) probes reveal expression in structures including the diencephalic superventricle (future third ventricle), optic cup, lens vesicle, Rathke's pouch (future pituitary), trigeminal ganglion, otic vesicle, vagus nerve and rhombencephalon (future fourth ventricle). Highest expression was observed in the dorsal portion of the rhombencephalon. **c, d:** *In situ* hybridisations of coronal sagittal sections of a CS17 human fetus using *OLFM2* antisense (c) and sense (negative control) (d) show expression of *OLFM2* in diencephalic superventricle, optic cup, lens vesicle, neural tube, and in the dorsal root ganglia. **e-j:** High magnification images of *OLFM2* expression in the CS15 human fetus in multiple structures: (e) lens vesicle and retinal ganglion cell layer, (f) trigeminal ganglion, (g) vagus nerve, (h) vestibular cochlear ganglion and otic vesicle, (i) diencephalic superventricle and Rathke's pouch, and (j) rhombencephalon. **k-m:** High magnification images of *OLFM2* expression in the CS17 human fetus in multiple structures: (k) lens vesicle and retinal ganglion cell layer, (l) neural tube and differentiating neurons, and (m) the dorsal root ganglia. Abbreviations: DG = dorsal root ganglion; DN = differentiating neurons of the neural tube; DS = diencephalic superventricle; LV = lens vesicle; NT = neural tube; OC = optic cup; OV = otic vesicle; R = rhombencephalon; RG = retinal ganglion cell layer; RP = Rathke's pouch; TG = trigeminal ganglion; V = vagus nerve; VCG = vestibular cochlear ganglion.

Figure 1

a


Chr19
Chr19:9,632,465 - 10,222,404 (GRCh38/hg38)
Genes
DGV Variants
ZNF562
ZNF846
FBXL12
UBL5
PIN1
OLFM2
COL5A3
RDH8
C3P1
C19orf66
ANGPTL6
PPAN
EIF3G
DNMT1
P2RY11

b

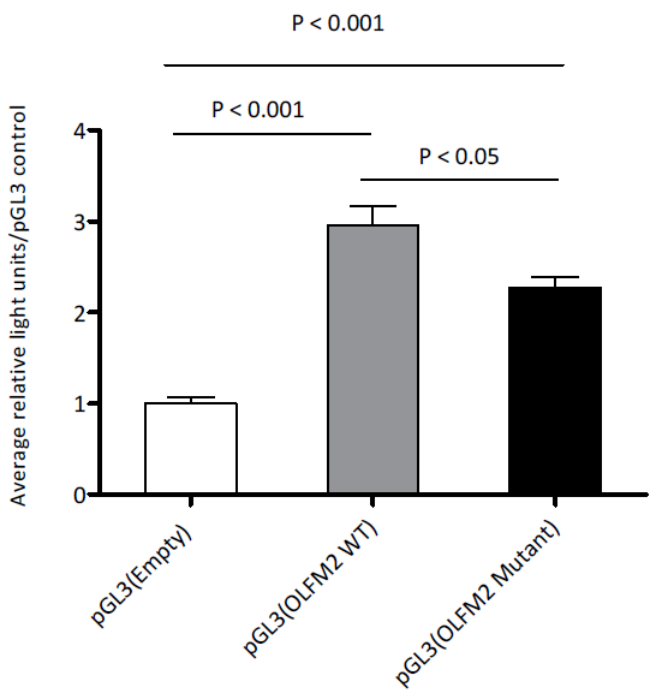
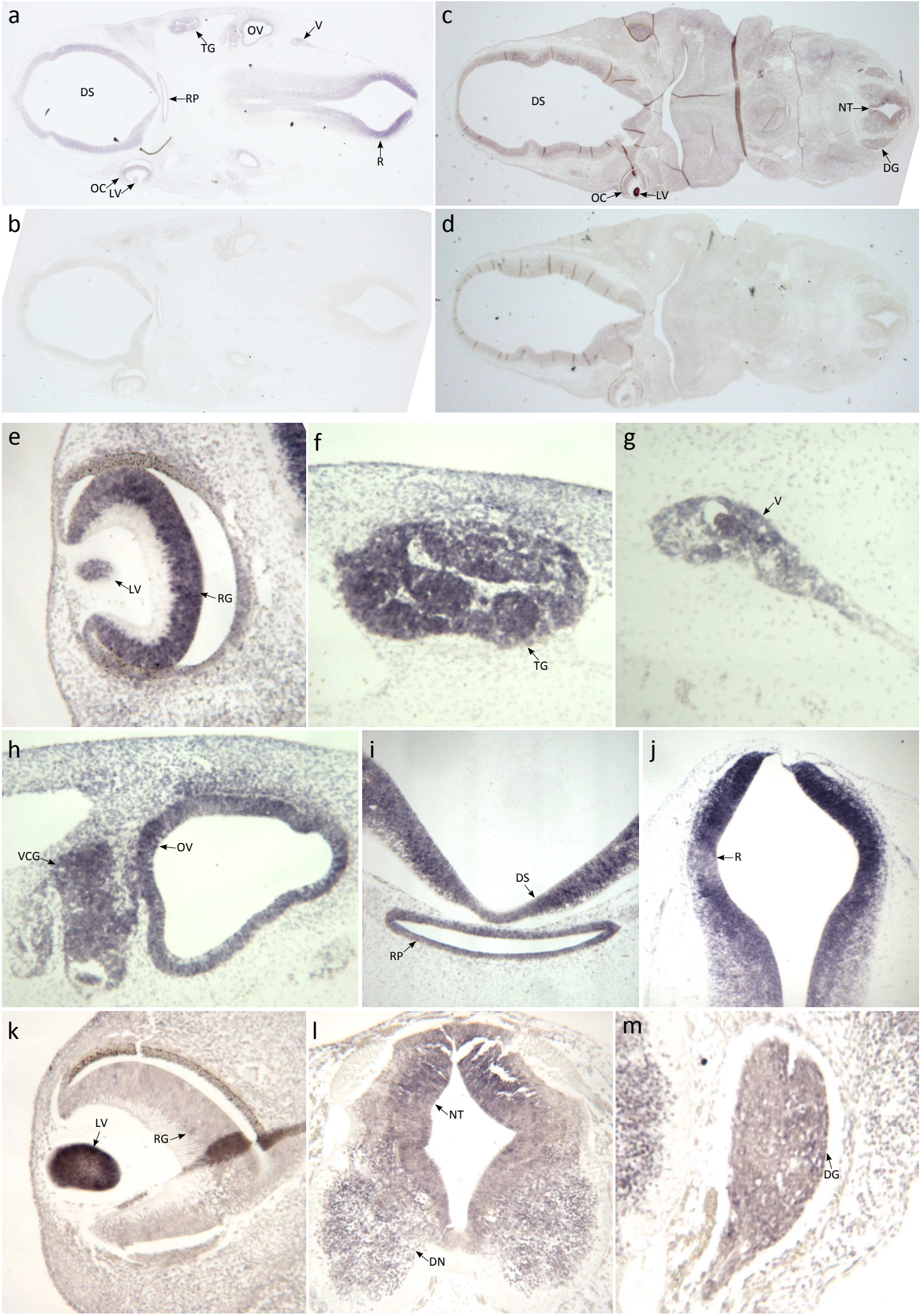

P < 0.001
P < 0.001
P < 0.05
Average relative light units/pGL3 control
pGL3(Empty)
pGL3(OLFM2 WT)
pGL3(OLFM2 Mutant)

Figure 2

Figure 2

Click here to download Figure Fig2.eps



Supplementary Methods

LightScanner Melt Curve Analysis

PCRs for LightScanner melt curve analysis consisted of 10ng of template DNA, and final concentrations of 0.25µM of each primer, 1x HotShot Diamond Master Mix (ClentLife) and 1x LCGreen Plus Dye (BioFire Diagnostics Inc) in a total volume of 10µl. PCRs were typically one cycle of 5 minutes 95^{o}C, 45 cycles of 95^{o}C 30s, annealing temperature 30s and 72^{o}C 45s, followed by a final denaturation step of 95^{o}C for 30s. PCRs were performed in Framestar® 96 (4titude) 96-well plates with a 20µl mineral oil overlay. Melt curves were generated on a LightScanner® (Idaho Technology Inc) using autoexposure, a starting temperature of 75^{o}C and a stop temperature of 98^{o}C. Data was analysed using the LightScanner® Instrument & Analysis Software (Idaho Technology Inc). Melt curves were normalised prior and post the major melt transition and aberrant curves detected using the autogroup function and manual inspection. Samples with an initial fluorescence of less than 600 were excluded from analysis, while those with a starting fluorescence less than 800 were not included during normalisation.

Supplementary Table 1: Primers used for Screening *HMX1*

Exon	Primer	Sequence	PCR	Condition
1	OLFM2_ Ex1F	CAAGCCAGAGAGTGCACGTC	419	60°C, 10% DMSO, 45x
	OLFM2_ Ex1R	gcaacaaagactcggagcga		
1st alternative exon 1	OLFM2_alt1Ex1F	GAAGCACAGGGGTAGAGGG	461	TD, 10% DMSO, 45x
	OLFM2_ alt1Ex1R	TTATAAGAGGAGCCCGCCAG		
2	OLFM2_ Ex2F	CTGGAGAGGAGCTGGATTATCA	388	55°C, 45x
	OLFM2_ Ex2R	CATCTGTGGTTTCCTTGGGC		
3	OLFM2_ Ex3F	TGTGGCTCATATTGGACCCT	336	50°C, 45x
	OLFM2_ Ex3R	AGCTCTTGTCTGTGGCATCT		
4	OLFM2_ Ex4F	GCTTCTAGGCACAAACAGGT	377	60°C, 45x
	OLFM2_ Ex4R	TGAGTCAGAGGTTGGAGTCA		
5	OLFM2_ Ex5F	TCCAGGACACTTTGGGCTAC	295	55°C, 35x
	OLFM2_ Ex5R	CTTCATCTTTGCCTGGCCTC		
6	OLFM2_ Ex6F1	ACAGGCAGAATGAAAAGGGC	300	62°C, 35x
	OLFM2_ Ex6R1	TGCTCTACGTGACCAACTCC		
	OLFM2_ Ex6F2	ACGTCCGTGTACTCGTAACT	372	62°C, 45x
	OLFM2_ Ex6R2	GAGCAACGTGGTGGTCAAAT		
	OLFM2_ Ex6F3	GGAGTAGGGGAAGGTGTTGT	386	50°C, 45x
	OLFM2_ Ex6R3	GTCAACAGAGTTCCCATGACT		

Note: we were unable to amplify the second alternative exon 1

TD = Touchdown PCR from 70°C to 55°C, decreasing by 1°C per cycle

Supplementary Table 2: Primers used to confirm *OLFM2* expression in cDNA

Primer	Sequence	PCR	Condition
Forward	CACATGACGCGCCCCTAG	190	55°C, 10% DMSO, 35x
Reverse	gcaacaaagactcggagcga		

Note: The Ex6F2/R2 primer pair (Supplementary Table 1) were also used to confirm expression.

Supplementary Table 3: Annotated *OLFM2* variants identified via LightScanner mutation analysis within the UK cohort. Changes are detailed according to the NM_058164.3 transcript (NP_477512).

Variant	Position (GRCH38/hg38)	Location	Classification	DNA Change	Amino Acid Change	Minor Allele percentage (dbSNP147)
rs544945664	chr19:9936454	Exon 1	5’ UTR	c.-91_-88delCCGG	N/A	0.879
rs201189277	chr19:9857731	Exon 3	Missense	c.344C>T	p.Ser115Leu	0.008
rs2303100	chr19:9857758	Exon 3	Missense	c.317G>A	p.Arg106Gln	45.336
rs143959085	chr19:9857386	Exon 4	Missense	c.457G>A	p.Val153Met	0.021
rs11556087	chr19:9857463	Exon 4	Missense	c.380C>T	p.Thr127Met	14.807
rs116666369	chr19:9856933	Intron 4	Intronic	c.581-20C>A	N/A	0.372
rs79341807	chr19:9856816	Exon 5	Synonymous	c.678G>A	p.Ala226=	0.511
rs11556088	chr19:9854270	Exon 6	Synonymous	c.1281C>T	p.Arg427=	10.959
rs200752028	chr19:9854588	Exon 6	Synonymous	c.963C>T	p.Ser321=	0.019
rs34961482	chr19:9854666	Exon 6	Synonymous	c.885G>A	p.Ser295=	0.479