9-15-2017

Mutations of conserved non-coding elements of PITX2 in patients with ocular dysgenesis and developmental glaucoma.

Meredith E. Protas
Departments of Ophthalmology and Anatomy and Institute for Human Genetics, University of California, San Francisco, meredith.protas@dominican.edu

Eric Weh
Department of Pediatrics, Medical College of Wisconsin

Tim Footz
Department of Medical Genetics, University of Alberta

Jay Kasberger
Celgene Quanticel Research

Scott C. Baraban
Department of Neurological Surgery, University of California, San Francisco

See next page for additional authors

https://doi.org/10.1093/hmg/ddx251

Human Molecular Genetics, 2017, Vol. 26, No. 18 3630–3638

doi: 10.1093/hmg/ddx251
Advance Access Publication Date: 3 July 2017
Original Article

ORIGINAL ARTICLE

Mutations of conserved non-coding elements of *PITX2* in patients with ocular dysgenesis and developmental glaucoma

Meredith E. Protas[1,†], Eric Weh[2], Tim Footz[3], Jay Kasberger[4], Scott C. Baraban[5], Alex V. Levin[6], L. Jay Katz[7], Robert Ritch[8], Michael A. Walter[3,*], Elena V. Semina[2,*] and Douglas B. Gould[1,*]

[1]Departments of Ophthalmology and Anatomy and Institute for Human Genetics, University of California, San Francisco, CA 94143, USA, [2]Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI 53226, USA, [3]Department of Medical Genetics, University of Alberta, Edmonton, AB T6G 2H7, Canada, [4]Celgene Quanticel Research, San Francisco, CA 94158, USA, [5]Department of Neurological Surgery, University of California, San Francisco, CA 94143, USA, [6]Pediatric Ophthalmology and Ocular Genetics, Wills Eye Hospital, Thomas Jefferson University, Philadelphia, PA 19107, USA, [7]Glaucoma Service, Wills Eye Hospital, Thomas Jefferson University, Philadelphia, PA 19107, USA and [8]Einhorn Clinical Research Center, The New York Eye and Ear Infirmary of Mount Sinai, New York, NY 10003, USA

*To whom correspondence should be addressed. Tel: 780 4924172; Fax: 780 4921998; Email: mwalter@ualberta.ca (M.A.W.); Tel: 414 9554996; Fax: 414 9556329; Email: esemina@mcw.edu (E.V.S.); Tel: 415 4763592; Fax: 415 4760336; Email: gouldd@vision.ucsf.edu (D.B.G.)

Abstract

Mutations in *FOXC1* and *PITX2* constitute the most common causes of ocular anterior segment dysgenesis (ASD), and confer a high risk for secondary glaucoma. The genetic causes underlying ASD in approximately half of patients remain unknown, despite many of them being screened by whole exome sequencing. Here, we performed whole genome sequencing on DNA from two affected individuals from a family with dominantly inherited ASD and glaucoma to identify a 748-kb deletion in a gene desert that contains conserved putative *PITX2* regulatory elements. We used CRISPR/Cas9 to delete the orthologous region in zebrafish in order to test the pathogenicity of this structural variant. Deletion in zebrafish reduced *pitx2* expression during development and resulted in shallow anterior chambers. We screened additional patients for copy number variation of the putative regulatory elements and found an overlapping deletion in a second family and in a potentially-ancestrally-related index patient with ASD and glaucoma. These data suggest that mutations affecting conserved non-coding elements of *PITX2* may constitute an important class of mutations in patients with ASD for whom the molecular cause of their disease have not yet been identified. Improved functional annotation of the human genome and transition to sequencing of patient genomes instead of exomes will be required before the magnitude of this class of mutations is fully understood.

[†]Present address: Department of Natural Sciences and Mathematics, Dominican University of California, San Rafael, CA 94901, USA.
Received: April 20, 2017. **Revised:** June 9, 2017. **Accepted:** June 27, 2017

Introduction

The ocular anterior segment comprises the cornea, iris, lens, ciliary body and iridocorneal angle tissues. Aqueous humor is secreted by the epithelium of the ciliary body and serves to nourish the avascular tissues of the lens and cornea (1). Aqueous humor must egress through the trabecular meshwork to Schlemm's canal in order to maintain normal intraocular pressure. Anterior segment dysgenesis (ASD) describes a genetically heterogeneous group of developmental disorders that affects the structure and function of these tissues (2,3). ASD can include corneal opacity, posterior embryotoxon (prominent and anteriorly displaced Schwalbe's line demarking the outer limit of the corneal Descemet membrane), iris hypoplasia, corectopia, polycoria, iridolenticular adhesions and cataract. Alterations of the iridocorneal angle drainage structures can impair aqueous humor outflow and lead to glaucoma. Some patients with ASD have particularly severe forms of very early onset glaucoma that can be refractory to many therapeutic approaches.

The most commonly identified genetic causes of ASD are mutations in the genes encoding the winged helix/forkhead transcription factor *Forkhead Box C1 (FOXC1)* and the bicoid-like homeodomain protein *Pituitary Homeobox 2 (PITX2)* (4,5). Still, the underlying mutations accounting for a large proportion of patients with ASD have yet to be identified. Previous work has also shown that proximal deletions upstream of the *PITX2* gene, but not affecting the coding region, can be pathogenic (4,6,7). In this study, we performed whole genome sequencing (WGS) of two patients from a four generation family diagnosed with ASD and glaucoma. We found a deletion on chromosome 4q25 in a gene desert that contains conserved putative regulatory elements of *PITX2*. We excised the orthologous DNA in zebrafish using the CRISPR/Cas9 system and found reduced *pitx2* expression during development and a mild ocular phenotype. We screened additional patients with ASD and glaucoma for similar structural variants and identified a smaller, but overlapping, deletion in a second family and a potentially-related index case. We propose that deletions of conserved non-coding elements upstream of *PITX2* are likely causative of the disease in these patients and suggest that other patients with ASD of unknown etiology will harbor mutations affecting regulatory elements of *PITX2* or other known ASD-causing genes.

Results

We studied a four-generation pedigree comprising five individuals with dominantly inherited ASD and glaucoma (Fig. 1A). Affected individuals exhibited classical anterior segment findings, including iris hypoplasia, correctopia and posterior embryotoxon. Elevated intraocular pressure developed by the second decade of life (range: 1 month to 18 years), and in one individual required multiple surgeries, and ultimately an enucleation unilaterally. This histology report (from 1986) documents ectropion uveae with iris stromal atrophy, although the pathological sample was unavailable for review. In addition, three affected individuals (II-3, III-3 and III-7) exhibited dental anomalies (microdontia), with incomplete primary dentition (e.g. 'milk teeth' just present on the mandible) that in some cases necessitated surgical removal. We performed WGS on two affected individuals (Fig. 1A) and prioritized coding variants that were shared between them and rare or absent from the 1000 Genomes Project data set. We identified 27 possible causative variants, however, literature searches and expression data available at EMAGE did not implicate obvious functional candidate genes. Moreover, when we extended genotyping of the coding variants to additional members of the pedigree we found that none segregated with the disease phenotype.

To assist in prioritizing the many non-coding variants, we performed an SNP array on the ten individuals in the pedigree for whom DNA was available (Fig. 1A) and identified the fifteen largest regions of copy number variation that were shared by the affected individuals (Table 1). Notably, we identified a 750kb region (chr4:84977588-110877045) that includes *PITX2*. Within this interval there are 72 SNPs for which the unaffected individuals had heterozygosity of 20% or greater. However, each affected individual lacked heterozygosity for the SNPs across this region. Upon re-evaluating the individual copy number variant files in the WGS data for both affected patients we discovered a ~750kb region (that may extend as far as chr4:110871123-111627123) of hemizygosity that was ~250kb upstream of *PITX2*. Together the SNP array data and the copy number files both suggest that the affected individuals carry a deletion upstream of *PITX2*.

The deletion breakpoints were not originally detected in the WGS data because the predicted junction falls within highly repetitive intergenic sequences. Based on the ploidy data, two 6kb intervals that flanked the predicted breakpoints were examined (chr4:110871123-110877123 and chr4:111622123-111628123; GRCh38). Five sequencing reads were found in one individual and seven in the other individual, that mapped to both intervals (Supplementary Material, Table S1). These reads narrowed the breakpoints to approximately 300bp on either side of the junction. Primers flanking this narrowed junction were used to amplify and sequence a single band present in affected individuals but absent in unaffected individuals that allowed us to resolve a 748,461bp deletion defined as chr4:110875898-111624359 (GRCh38).

A series of elements with conservation of sequence and position between humans and zebrafish have been identified in the ~1.5Mb gene desert upstream of *PITX2* (7). Some of these elements are capable of driving reporter gene expression in zebrafish in patterns consistent with *pitx2*, suggesting that these are conserved *PITX2/pitx2* regulatory elements, and large deletions upstream of *PITX2* (without affecting *PITX2* coding regions) have been reported in patients diagnosed with Axenfeld-Rieger syndrome (4,6,7). A new BLAST2 comparison of the human and zebrafish genomes was performed to refine the CEs and identify new conserved elements (CE14-16) (Fig. 2; Supplementary Material, Table S2). The deletion identified in Family 1 includes conserved elements (CEs) 5 through 11 while previously published studies reported larger deletions (CE4 or CE5 to a point beyond CE16, or CE4 through at least CE9) (Fig. 2).

To determine if this deletion influenced *pitx2* expression and ocular development, we used CRISPR/Cas9 genome editing to remove the orthologous DNA sequences in zebrafish. Single guide RNAs (sgRNAs) that bind within CE5 and CE11 were co-injected with mRNA encoding Cas9 into one-cell zebrafish embryos. A mosaic founder with germ line transmission was identified which carried an allele containing a deletion between CE5 and CE11 (chr14:35674630-35806484; GRCz10). This structural rearrangement creates a hybrid element consisting of the first 164bp of CE5 and the last 231bp of CE11. Embryos from intercrossed fish carrying the CE5-11 deletion were grossly normal out to 96 hours post fertilization (hpf). However, a subset of the progeny appeared to have shallow ocular anterior chambers (36/132; 27%) with no other visible ocular defects (Fig. 3). Genotyping revealed that all 36 embryos were homozygous for the CE5-11 deletion. To quantify our observation, we

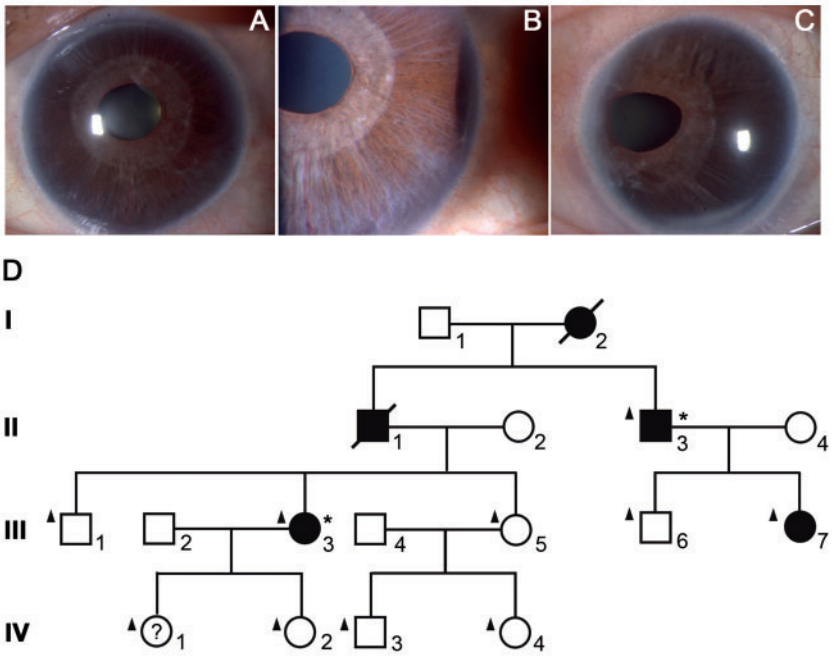

Figure 1. Pedigree and ocular anterior segment photographs for Family 1. (**A**) Left eye, close-up of left eye, and right eye of Individual III-3 showing partial hypoplasia of the anterior iris stroma, leading to a more prominent iris sphincter, and mild corectopia. (**B**) Pedigree of a family with autosomal dominant inheritance of ASD and glaucoma. Filled symbols represent affected individuals with ASD and glaucoma. Individuals selected for WGS are indicated with asterisks. Individuals that were genotyped with the SNP array are indicated with black triangles. The "?" indicates that the phenotype is unknown and the slash indicates a deceased individual.

Table 1. A list of the 15 largest regions of non-heterozygosity in affected individuals of Family 1

Chr	Start location	End location	Genes of interest (ASD)
1	9362376	67567086	
1	119766818	145846884	
2	80375399	106803953	
3	129025892	175262925	
3	73563798	111535372	
3	28716903	54569489	
4	84977588	110877045	*PITX2*
8	63779855	104457745	
9	38769006	68577592	
11	15128238	87971202	
13	59676709	107169802	
16	31843036	48027227	
18	10465316	45682394	
20	8681575	31939262	
21	15658558	37919640	

intercrossed heterozygous fish to generate homozygous, heterozygous and wild-type progeny and an investigator masked to the genotypes measure anterior chambers in cohorts at 72-, 96- and 120-hpf. At all three ages, fish homozygous for the CE5-11 deletion had significantly smaller anterior chambers while heterozygous fish had an intermediate phenotype (Fig. 3G and Supplementary Material, Fig. S1). Because abnormal craniofacial development is often associated with *PITX2* mutations in humans, we stained 120-hpf embryos with alcian blue to identify possible defects in craniofacial cartilage development. No differences were identified between heterozygous and homozygous embryos (data not shown). Examination of adult homozygous mutant *pitx2* fish identified normal survival rates and no obvious morphological defects (Fig. 3H–M).

Zebrafish *pitx2* encodes two alternatively spliced isoforms (*pitx2a* and *pitx2c*) that share the same homeodomain but differ in the N-terminal sequence. Both isoforms are expressed during ocular development (7). We used qRT-PCR of cDNA to measure levels of *pitx2a* and *pitx2c* in whole embryos at 24-hpf and 72-hpf. While no differences were detected at 24-hpf, there were marked decreases of both *pitx2* isoforms at 72-hpf whereby heterozygous and homozygous embryos had approximately 50% and 25% of normal expression levels, respectively (Fig. 4A). To analyze the expression with greater anatomic resolution we dissected whole eyes, head (with no ocular tissue) and trunk from 24-, 48-, 72- and 96-hpf embryos. At 24-hpf, no significant changes in expression of *pitx2a* or *pitx2c* were detected in any tissue, however, at 48-hpf, both *pitx2* isoforms were significantly reduced in eyes of homozygous embryos and by 72-hpf all tissues examined had decreased *pitx2a* and *pitx2c* levels; at 96-hpf expression in the developing eye remains to be low for both isoforms (but with a statistically significant difference for *pitx2c* only) while *pitx2* levels in other tissues appear to recover and reach normal amounts (Fig. 4B).

To investigate the frequency of this class of regulatory element deletions we screened additional patients that had been diagnosed with ASD and who were negative for coding mutations in *FOXC1* and *PITX2*. Family 2, of Puerto Rican heritage, demonstrated autosomal dominant inheritance of typical Axenfeld-Rieger spectrum disorder (Fig. 5). Ocular abnormalities at initial presentation included posterior embryotoxon, iridocorneal adhesions, iris atrophy, corectopia, polycoria and bilateral glaucoma. Systemic findings included dental anomalies and an umbilical defect. DNA samples from three family members (two affected and one unaffected; I:1, II:1, and II:6 in Fig. 5) were previously sent to the Beijing Genomics Institute for whole-exome sequencing (WES). The average coverage for the samples over two merged runs was 82x-90x, although coverage for certain

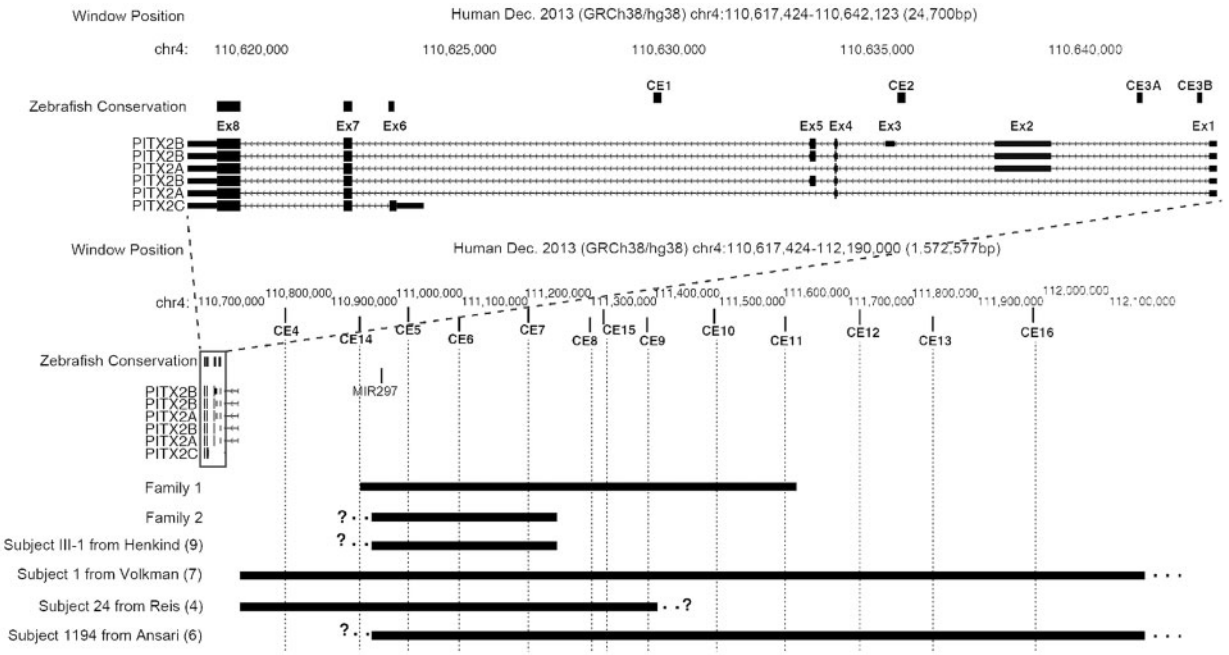

Figure 2. Annotated map of the *PITX2* locus on 4q25. The transcript isoforms of *PITX2* are the RefSeq-annotated versions in UCSC Genome Browser. Exons 2 and 3, which are subject to alternative splicing, are located in the 5' untranslated region of the indicated transcripts (which encode protein isoforms PITX2A and PITX2B; PITX2B contains the alternatively-spliced coding exon 5). Exon 6 is the alternative first exon of the transcript that encodes protein isoform PITX2C. The deletion of Subject 1 described in Volkman et al. (7) removes elements CE4-16 extending at least 7.7Mb upstream of PITX2. The deletion of Subject 24 in Reis et al. (4) removes elements CE4-9 and extends an unknown distance further upstream. The deletion of Subject 1194 in Ansari et al. (6) removes elements CE5-16 extending at least 3.7Mb upstream. The deletions identified in Families 1 and 2 (and the potentially related index patient) of the current study encompass conserved elements CE5-11 and CE5-7, respectively. CE1–13 = conserved elements described in Volkmann et al. CE14–16 = new conserved elements identified in this study.

genes associated with ASD (*COL4A1*, *FOXC2*, *FOXE3*, *JAG1*, *PAX6* and *PITX3*) was incomplete and we used Sanger sequencing to exclude pathogenic variants. A total of 8518 variants were detected in the WES data that segregated with the phenotype in the three sequenced family members. These variants were filtered and prioritized based on their novelty, subjective local sequence alignment quality and predicted detriment to protein function. Of six variants of interest that were confirmed by Sanger sequencing, none co-segregated with the phenotype in the remaining family members.

Upon re-evaluation of the WES data, four synonymous or non-coding SNPs were identified on 4q25 at the *PITX2* locus for which the affected individuals appeared to share a haplotype. The genetic marker D4S2945 was tested for all available family members to determine linkage of this locus. The data suggested that D4S2945 is linked to the phenotype in this family (Fig. 5). Further, SNPs rs2278782 and rs2278783 in intron 7 of *PITX2* were amplified and sequenced; all affected members were heterozygous for the reference C allele and the rare T allele whereas the unaffected members were homozygous for the reference C allele (Fig. 5). These results suggest co-segregation of the *PITX2* locus with the disease phenotype. We sequenced the untranslated regions of *PITX2* transcripts (A/B and C isoforms) and no variants were found. We developed primers to amplify CEs (while avoiding binding sites containing known SNPs) and used qPCR to investigate copy number changes in the transcribed and upstream regions of the *PITX2* locus (8). We screened an affected individual with 14 amplicons (in triplicate) and discovered ~50% dosage for CE5-CE7 (Fig. 2 and Supplementary Material, Table S3) which represents a deletion of approximately 189-360kb. Elements CE14 and CE5-CE8 were tested in the remaining available family members and confirmed the segregation of the small deletion with the phenotype.

An additional 20 cases diagnosed with ASD (and without *FOXC1* or *PITX2* coding mutations) were subjected to qPCR of elements CE5-7 in the *PITX2* upstream region. One individual displayed ~50% dosage of the CE5-7 region but full dosage of CE14 and CE8. This case is individual III:1 from a published family and, like Family 2, is of Puerto Rican heritage (9). The affected individuals from both families share the same disease-associated haplotype (for markers D4S2945, rs2278782 and rs2278782) consistent with the possibility they have common ancestry and inherited the same CE5-7 deletion.

Discussion

The goal of this study was to discover novel genetic causes of ASD. We found two overlapping deletions that include conserved regulatory elements upstream of *PITX2* that are likely causative mutations. Family 1 had mild ASD, juvenile glaucoma, dental anomalies and a 748kb deletion that deletes CE5-11. Family 2 and an additional patient who is possibly related have a deletion of no more than 360kb that encompasses CE5-7. This represents the smallest known deletion, with likely pathogenicity, upstream of *PITX2*. We modeled the larger deletion in zebrafish using CRISPR/Cas9 and found dose-dependent decreases of *pitx2a* and *pitx2c* and a mild defect in the formation of the anterior chamber which was detectable by 72-hpf and persisted through at least 120-hpf. These data support the conclusion that the loss of CE5-11 upstream of *PITX2* is pathogenic in humans. Deletion of CE5-11 in zebrafish did not cause complete loss of *pitx2* expression (homozygous embryos still had approximately 25% of normal expression) and it is likely that additional elements coordinate development of the anterior segment and craniofacial regions. Notably, recent work to completely disrupt

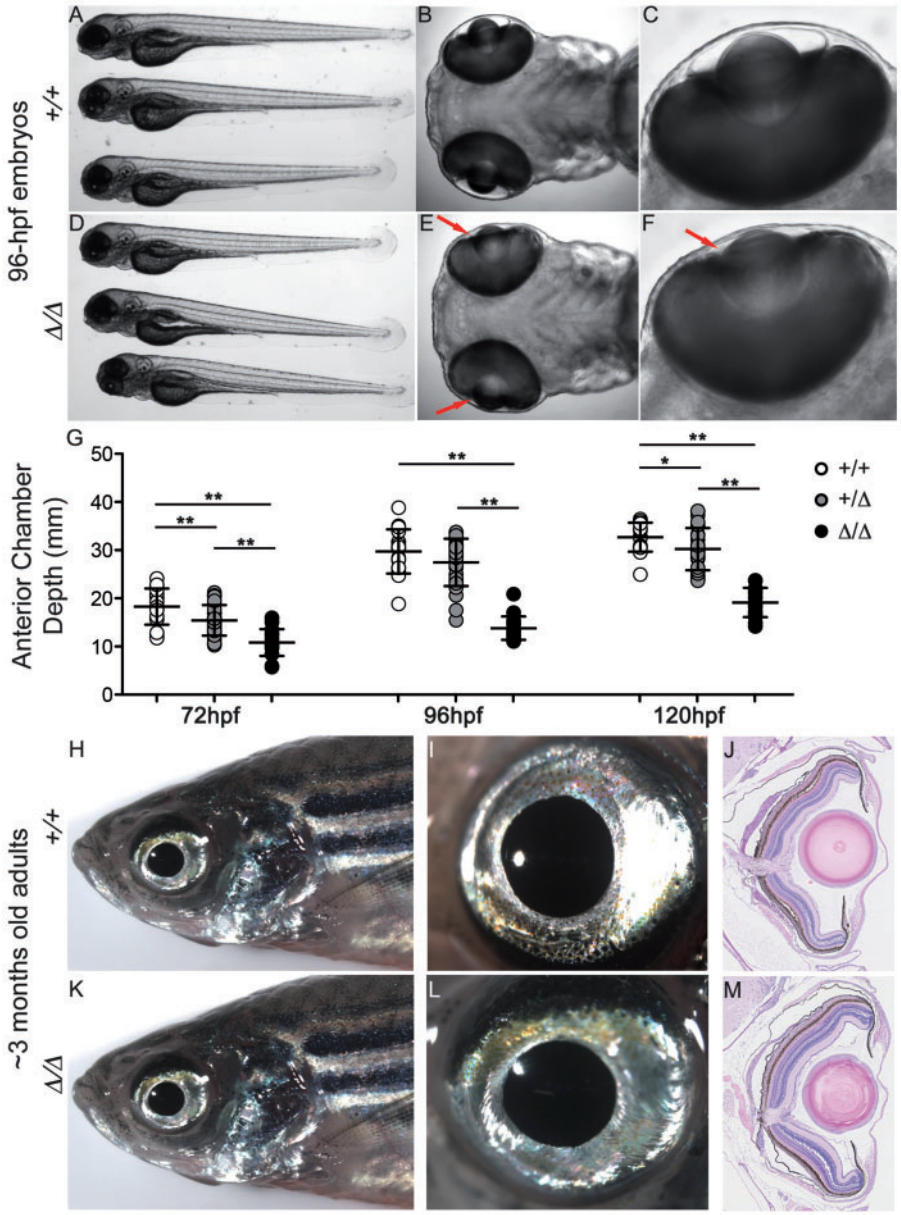

Figure 3. Characterization of embryos and adult zebrafish homozygous for the CE5–11 mutation. Wild-type (**A–C**) and mutant (**D–F**) embryos at 96hpf are morphologically similar with the notable exception that anterior chambers of the mutant fish appear shallow and underdeveloped (red arrows in E and F, $n = 36$). (**G**) Quantification of anterior chamber depths revealed significantly smaller anterior chambers in homozygous mutant fish and an intermediate phenotype in heterozygous mutant fish (mean +/– standard deviation; *$P<0.05$; ***$P<0.001$, using a two-tailed, unpaired student's *t*-test). Eyes of wild-type (**H–J**) and mutant (**K–M**) zebrafish at 3 months appeared to be indistinguishable ($n = 3$ fish per genotype).

pitx2 function in zebrafish reveals an embryonic phenotype consisting of anterior segment dysgenesis and craniofacial malformations (10).

Understanding the mechanisms by which *pitx2* expression is controlled by individual and combinatorial contributions of conserved regulatory elements will be important for understanding the regulation of *pitx2* during development of different tissues. Moreover, the number and functions of deleted elements could influence the phenotypic spectrum in patients. To this end, reporter constructs under the control of different elements expressed GFP in patterns that were consistent with *pitx2* expression in various developing structures including brain, heart, oral cavity and eyes (7). Within the minimal deleted region (CE5-7) CE5 and CE7 directed GFP expression in the periocular mesenchyme of embryonic zebrafish, whereas CE6 controlled expression in developing brain (7). One might expect individuals with greater numbers of conserved elements deleted to have more severe phenotypes or additional systemic defects, however, our data did not show such a trend and both Family 1 and Family 2 (with disruption of conserved elements 5–11 and 5–7, respectively) had ASD and extraocular features including dental abnormalities.

The genetic cause remains unknown for many patients with ASD. One possibility is that mutations in a large number of novel

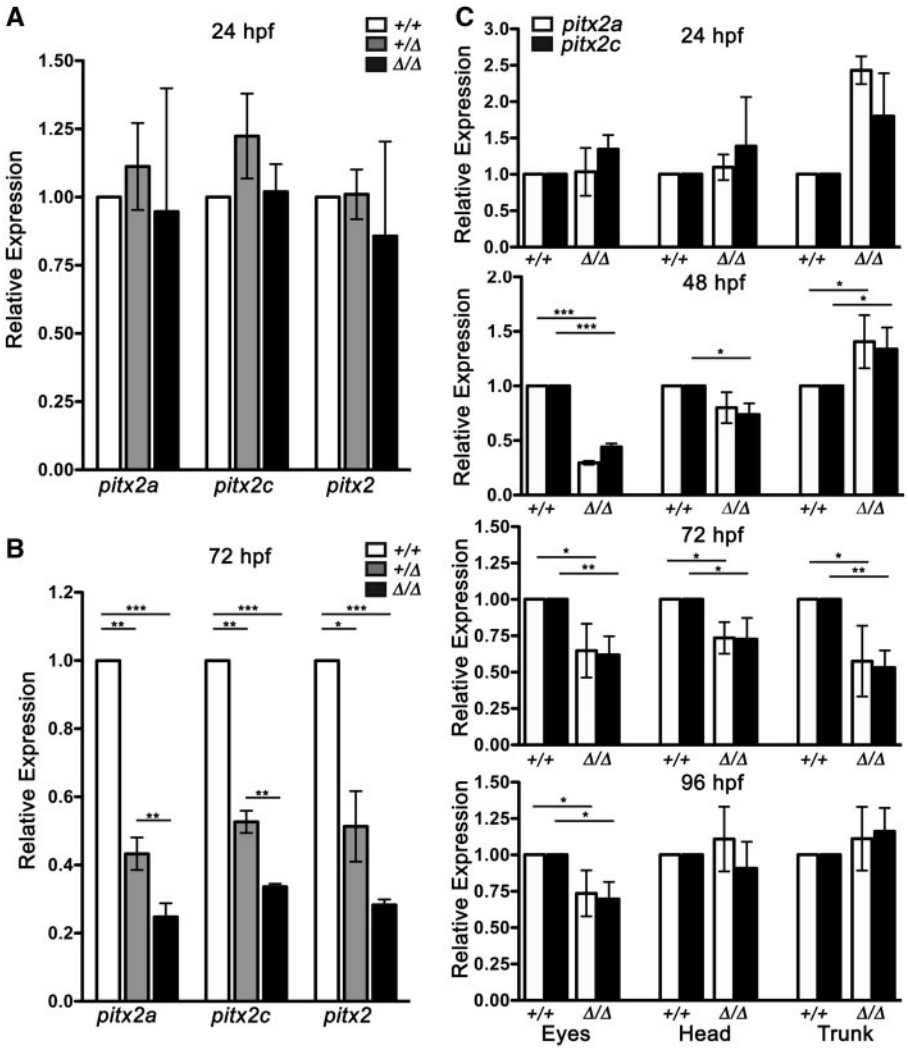

Figure 4. Zebrafish embryos with the CE5–11 mutation have reduced pitx2 expression. (A) Expression levels of *pitx2* isoforms in 24hpf embryos that were heterozygous or homozygous for the CE5–11 mutation were not significantly different from those in wild-type embryos. (B) In contrast, in 72hpf embryos expression of *pitx2* isoforms were significantly reduced in mutant embryos. (C) Separation of wild-type or homozygous CE5–11 mutant embryos into different anatomic regions indicated that the strongest effect of the mutation on *pitx2* expression was in eyes at 48hpf (mean +/– standard deviation; *$P<0.05$; **$P<0.01$; ***$P<0.001$, using a two-tailed, unpaired student's *t*-test).

genes will each account for a small proportion of these patients and exome sequencing will reveal the underlying mutations. Although it is almost certainly true that novel genes remain to be identified, it is also likely that mutations within non-coding regulatory elements of genes already identified will underlie a proportion of these cases. These results suggest that *PITX2* may play an even greater role in ASD than is currently appreciated and that mutations in regulatory elements of already known genes, rather than coding mutations in novel genes, might explain a substantial proportion of cases awaiting molecular diagnosis. For example, *FOXC1* regulatory elements are not well characterized but with the advances in analyzing non-coding regions, this might also soon be possible. Further investigation into the possibility of other anterior segment developmental genes having regulatory sequence changes may reveal other similar pathogenic changes.

Most of the variation between humans is found within the ~98% of the human genome comprised of non-coding DNA (11). While the interpretation of non-coding variants remains difficult, recent advances such as the ENCODE project (ENCyclopedia of DNA Elements) have greatly improved the potential for identification and annotation of functional elements (12). Increasing appreciation for the roles of promoters, enhancers, regulatory RNAs (microRNA, circular RNA, long non-coding RNAs), cryptic splice sites and other non-coding elements will reveal additional categories of mutations contributing to disease (13). However, pathogenic mutations in non-coding elements will remain under-represented without a shift to WGS and advanced functional annotation of the genome. Even then, certain types of mutations may still go undetected. For example, if the deletion in Family 1 had been smaller it may have gone overlooked. Mutations in gene deserts populated by repeat elements, complex rearrangements or very small perturbations of conserved elements (especially single base pair changes) will remain a challenge even for WGS. Our study demonstrates the power of using genome editing in zebrafish for functionally testing potentially pathogenic non-coding variants and further supports the function of the previously identified conserved elements as ocular enhancers of *pitx2*. Our work demonstrates that non-coding mutations may explain some cases of ASD for which candidate gene sequencing and WES have not identified the genetic etiology of the phenotype.

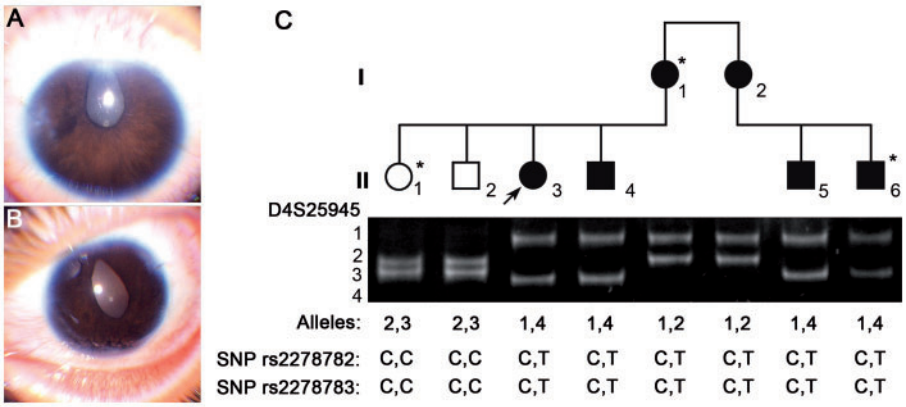

Figure 5. Family 2 with ASD that co-segregates with *PITX2*. (**A,B**) Corectopia and polycoria observed in the proband (II-3) of Family 2. (**C**) The pedigree for this family shows autosomal dominant inheritance of ASD. Filled symbols represent affected individuals with a diagnosis of ASD and glaucoma. Individuals I:1, II:1 and II:6 were selected for WES. The phenotype in this family segregates with markers linked to *PITX2* (Allele 1 for marker D4S2945; the T allele for marker rs2278782; the T allele for marker rs2278783).

Materials and Methods

Patients

Blood was collected from ten members of Family 1 consisting of three affected and seven unaffected individuals. For Family 2, blood was collected from six affected and two unaffected individuals. These, and other blood samples from sporadic patients, were collected with informed consent in accordance with the University of Alberta Ethics Board. DNA was prepared using standard detergent lysis and salting out protocols.

Analysis of genome sequencing

We performed whole-genome sequencing of affected individuals II-3 and III-3 (Fig. 1) and received high quality data with 96% and 97% of the exome and genome covered, respectively (Complete Genomics, Mountain View, CA). We used the Complete Genomics standard pipeline (version 1.12) and quality control methodology (14) to detect and analyze variants (15). Variants of interest were checked for abnormally low or high coverage and cross-referenced with read quality scores for each locus in order to eliminate variants one standard deviation from the mean quality score and for each genome. No variants were disqualified for these reasons. Using an LD-thinned ($r < 0.2$) subset of the genomes, we identified regions of identity-by-descent using the fastIBD algorithm in the BEAGLE package (16). We used the gene-annotation mode of ANNOVAR (19) to locate potentially pathogenic exonic variants within the regions of identity-by-descent, and SIFT to predict deleterious amino acid substitutions at a confidence level of 0.05 (17,18). Minor allele frequency estimates were obtained from the CEU population within the 1000 Genomes Project data set and the data were analyzed sequentially using frequency thresholds of 0.10, 0.05, and 0.01. Genes with prioritized variants were investigated by literature search and EMAGE (www.emouseatlas.org) to identify functional candidates involved in ocular or craniofacial development.

Identifying an upstream deletion of *PITX2*

We performed SNP array genotyping using an Illumina HumanCytoSNP-12 v2.1 DNA Analysis Kit with 500ng of DNA from the 11 members of Family 1 and used Genome studio to convert the files to Excel. Data were sorted to identify blocks of regions with at least one shared allele in all affected individuals. Any regions where at least one allele was not in common between all of the affected individuals were excluded. Using the genome sequence data, we estimated the location of the deletion breakpoints and used a script provided by Complete Genomics to isolate the raw reads that mapped to this interval. Then, we used a second script provided by Complete Genomics to find the reads in common to both predicted sides of the junction (Supplementary Material, Table S1). PCR primers were designed to this narrowed region and used to amplify from DNA, using a standard PCR protocol, from each member of the pedigree for whom we had a DNA sample.

Generation of deletions in zebrafish using CRISPR genome editing

All experiments were conducted under protocols approved by the Institutional Animal Care and Use Committees at UCSF and Medical College of Wisconsin. Zebrafish, *Danio rerio*, were maintained in a 14/10-h light/dark cycle. The embryos were raised at 28.5 °C in E2 medium containing phenylthiourea to inhibit pigment formation and extend the period of transparency of the fish during development. The developmental stage of each embryo was assessed by previously described morphological criteria (19). We used CRISPR/Cas9 genome editing (20,21) to create an orthologous deletion of the upstream *pitx2* regulatory regions to test the functional consequences of the deletion identified in Family 1. We utilized MLM3613 (Addgene plasmid # 42251, Addgene, Cambridge, MA) and DR274 (Addgene plasmid # 42250) plasmids (gifts from Keith Joung) to generate Cas9 mRNA and single guide RNAs (sgRNAs) respectively (22). We designed sgRNAs oligonucleotides using ZiFiT Targeter software (http://zifit.partners.org/ZiFiT/ChoiceMenu.aspx) (22,23). The "GG" constraints on the two 5' nucleotides were relaxed to identify an sgRNA target site in the middle of the conserved regulatory elements (CEs) 5 and 11 (7). We injected 300 ng/μl of Cas9 RNA and 50ng/μl of sgRNAs into one-cell zebrafish embryos. sgRNA efficacy was evaluated by amplifying targeted CEs from day 1 to 3 injected embryos ($n = 20$) using the primers listed (Supplementary Material, Table S4). The products were then sequenced to verify that we were able to induce targeted disruptions. We chose 1–2 successful sgRNAs per element (Supplementary Material, Table S5) and performed injections to delete the genomic region between CE5 and CE11. To test for successful deletions we screened injected day 1–3 embryos

($n = 20$) using PCR primers specific to deletion events occurring between the targeted elements. Injected clutch-mates of embryos that were positive for genomic rearrangements were kept and grown to adulthood. Once mosaic founders reached sexual maturity, they were mated in pairs and DNA was extracted from 20 embryos of each mating. Embryos were screened for editing events at CE5 and 11 as well as for a large genomic deletion between the two elements as described above. Founders that carried germ line deletions were crossed and the offspring were raised to maturity. F1 adult fish that were heterozygous for the deletion between CE5 and 11 were identified and crossed to produce embryos homozygous for the deletion.

Molecular and morphologic analysis of CE5-11 deletion fish

Embryos generated from intercrosses of adults that were heterozygous for the CE5-11 deletion were collected and examined beginning at 24 hpf until 6 days post fertilization for gross morphological defects. Embryos were anesthetized using tricane (Sigma-Aldrich, St. Louis, MS, USA) and carefully positioned to ensure both eyes were in the same focal plane. Images were taken using a Zeiss SteREO Discovery V12 microscope (Carl Zeiss AG, Oberkochen, Germany) with either a 1.0X stereo objective or a 10X compound objective for anterior chamber imaging. The area of the anterior chamber was measured using ImageJ using the polygon selection tool to outline the interior surface of the anterior chamber. To normalize for differences in embryo size the widest part of the eye from rostral to caudal was used. Three separate measurements were averaged for area and eye width. The average of the anterior chamber area was divided by the average of the width of the eye to generate a ratio in units of length. To mitigate bias, embryos were not genotyped until after the measurements were obtained. Alcian blue staining was performed to examine craniofacial development as previously described (10,24). Whole embryos or dissected tissues (whole eyes, head tissue, trunk) from 11 to 30 embryos at 24-, 48-, 72- or 96-hpf embryos were solubilized in TRI Reagent (Zymo Research, Irvine, CA). Total RNA was extracted using the Direct-zol kit (Zymo Research, Irvine, CA), examined for degradation by non-denaturing agarose gel electrophoresis (1%) and quantified using a Nanodrop 1000 (Nanodrop Products, Wilmington, DE). Total RNA (150 ng) was used as template for cDNA synthesis using the Super Script III kit (Thermo Fisher Scientific). RT-qPCR was used to determine relative expression levels of *pitx2a* and *pitx2c* as previously described (25) with the following primers: *pitx2a* qPCR F and *pitx2a* qPCR R (product = 84 bp) and *pitx2c* qPCR F and *pitx2c* qPCR R (product= 145 bp) (Supplementary Material, Table S3). The primers used for RT-qPCR were verified to produce a single product and their amplification efficiencies were determined. The RT-qPCR reaction was performed using the Power Up SYBR Green Master Mix (Thermo Fisher Scientific) following the manufacturer's recommendations and a Bio-Rad CFX Connect 96 or 384 well thermocycler (Bio-Rad, Hercules, CA) using the following conditions: 95 °C for 2 min followed by 40 cycles of 94 °C for 20 s, 60 °C for 20 s and 72 °C for 20 s. The final concentration of each primer used was 0.4 μM. *β-actin* was used to normalize cycle values (10). Wildtype expression levels were used as the reference in comparisons and data were compared using a two-tailed, unpaired student's t-test.

Genotyping of Family 2

A standard touchdown PCR protocol (95 °C for 3 min, 5 cycles of 95 °C for 30 s, 64–56 °C (2 °C touchdown per cycle) for 30 s, 72 C for 30 s, 30 cycles of 95 °C for 30 s, 54 °C for 30 s, 72 °C for 30 s) was used for all amplifications from genomic DNA with the following reagents: 100 ng genomic DNA, 0.3 pmol of each of forward and reverse primers, 1x FailSafe™ PCR PreMix J buffer (Epicentre, Madison, WI) and 1Unit of *Taq* DNA Polymerase. Amplified fragments were purified with QIAquick columns (QIAGEN Inc., Valencia, CA) and directly sequenced at The Applied Genomics Core (University of Alberta, Edmonton, AB). Amplification and sequencing primers are listed in Supplementary Material, Table S3. For exome sequencing of three individuals from Family 2 (I-1, II-1, and II-6 from Fig. 5) libraries were prepared at The Centre for Applied Genomics at The Hospital for Sick Children (Toronto, ON) using the SureSelect Human All Exons 50Mb v4 kit (Agilent Technologies, Mississauga, ON), and sequenced on a 5500xl SOLiD™ System instrument (Life Technologies, Burlington, ON), following the manufacturer's recommendations.

Screening sporadic patients for deletions of conserved elements

We investigated copy number changes in *PITX2* and upstream conserved elements using SYBR Green qPCR as described (8). We analyzed previously published functional conserved elements (7) as well as any additional conserved sequenced determined from BLAST comparison of the ~1.5Mb genomic sequence upstream of human *PITX2* (version GRCh38 chr4:110617424-112145396) before the nearest annotated gene (*C4ORF32*) with the corresponding zebrafish *pitx2* locus (version GRCz10, chr14:35571919-35889512). We designed qPCR primers to amplify sequences encompassing twelve upstream elements plus exons 2 and 8 (see Fig. 2; Supplementary Material, Table S4) according to the recommendations of PrimerBank (26). For empirical validation, amplification efficiency (E) was determined for each primer set via generation of standard curves using serial dilution of genomic DNA from an unrelated unaffected individual and incorporated into the dosage calculation for Normalized Relative Quantity $\left[\mathrm{NRQ} = E^{\Delta ct}_{(GJA5\_exon2\ reference)} / E^{\Delta ct}_{(PITX2\ amplicon)}\right]$, where Ct is the cycle of amplification at which fluorescence reached the threshold value of 2.0 in the linear phase on the ABI PRISM® 7900HT Sequence Detection System (Life Technologies), and $\Delta Ct = Ct_{(control\ DNA)} - Ct_{(patient\ DNA)}$. The assay efficiencies ranged from 1.8 to 2.0 and amplicon specificity was confirmed by melting curve analysis. PCR reagents were as follows: 20 ng of genomic DNA, 0.3 μM of each primer, and 5 μl of 2x QuantiTect SYBR Green PCR Master Mix (QIAGEN Inc.) in 10 μl total volume. Reactions were performed in triplicate with the following conditions: 15 min at 95 °C; 40 cycles of 15 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C. Replicate sample values were excluded from analysis if the Ct differed by >0.5 cycles from the median of the triplicate, and entire samples were repeated if two values of a triplicate were excluded. Twelve normal control samples were used to determine the normal variation for each qPCR assay (Supplementary Material, Table S5) and patient samples were deemed to harbor no copy number variation if the NRQ was within 2 standard deviations of the normal average (two copies). Values below the minimum threshold were tested in triplicate and then manually scrutinized to assess if they ranked within the distribution of theoretical values for 50% dosage (one copy).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We would like to thank Dr. Ordan Lehmann for clinical samples. We would like to thank Dr. Cassandra Labelle-Dumais, Kendall Hoff, Dr. Gabriela Hortopan, Matthew Dinday, Dr. Brian Grone, Dr. Nadav Ahituv, and Karl Murphy for advice and assistance. We would also like to thank Dr. Birgit Crain from Complete Genomics for providing scripts and advice to identify breakpoints of the deletion.

Conflict of Interest statement. None declared.

Funding

National Institutes of Health (R01EY019887 to D.B.G., R01EY015518 to E.V.S.); Bright Focus Foundation (D.B.G.); Research to Prevent Blindness (D.B.G.); That Man May See (D.B.G.) and The Foerderer Fund and Robison D. Harley, MD Endowed Chair in Pediatric Ophthalmology and Ocular Genetics (AVL), Denise B. Evans Endowed Chair in Ophthalmology (DBG). Additional support was provided by a core grant from National Institute of Health (EY02162) and Research to Prevent Blindness unrestricted grants to UCSF Department of Ophthalmology.

References

1. To, C.H., Kong, C.W., Chan, C.Y., Shahidullah, M., and Do, C.W. (2002) The mechanism of aqueous humour formation. *Clin. Exp. Optom.*, **85**, 335–334.
2. Gould, D.B., Smith, R.S. and John, S.W. (2004) Anterior segment development relevant to glaucoma. *Int. J. Dev. Biol.*, **48**, 1015–1029
3. Idrees, F., Vaideanu, D., Fraser, S.G., Sowden, J.C. and Khaw, P.T. (2006) A review of anterior segment dysgeneses. *Surv. Ophthalmol.*, **51**, 213–231.
4. Reis, L.M., Tyler, R.C., Volkmann Kloss, B.A., Schilter, K.F., Levin, A.V., Lowry, R.B., Zwijnenburg, P.J., Stroh, E., Broeckel, U., Murray, J.C. *et al.* (2012) PITX2 and FOXC1 spectrum of mutations in ocular syndromes. *Eur. J. Hum. Genet.*, **20**, 1224–1233.
5. Chang, T.C., Summers, C.G., Schimmenti, L.A. and Grajewski, A.L. (2012) Axenfeld-rieger syndrome: New perspectives. *Br. J. Ophthalmol.*, **96**, 318–322.
6. Ansari, M., Rainger, J., Hanson, I.M., Williamson, K.A., Sharkey, F., Harewood, L., Sandilands, A., Clayton-Smith, J., Dollfus, H., Bitoun, P. *et al.* (2016) Genetic analysis of 'PAX6-negative' individuals with aniridia or gillespie syndrome. *PLoS One*, **11**, e0153757.
7. Volkmann, B.A., Zinkevich, N.S., Mustonen, A., Schilter, K.F., Bosenko, D.V., Reis, L.M., Broeckel, U., Link, B.A. and Semina, E.V. (2011) Potential novel mechanism for axenfeld-rieger syndrome: Deletion of a distant region containing regulatory elements of PITX2. *Invest. Ophthalmol. Vis. Sci.*, **52**, 1450–1459
8. Seifi, M., Footz, T., Taylor, S.A., Elhady, G.M., Abdalla, E.M. and Walter, M.A. (2016) Novel PITX2 gene mutations in patients with axenfeld-rieger syndrome. *Acta Ophthalmol.*, **94**, e579.
9. Henkind, P., Sigel, I.M. and Carr, R.E. (1965) Mesodermal dysgenesis of the anterior segment: Rieger's anomaly. *Arch. Ophthalmol.*, **73**, 810–817
10. Liu, Y. and Semina, E.V. (2012) Pitx2 deficiency results in abnormal ocular and craniofacial development in zebrafish. *PLoS One*, **7**, e30896.
11. Spielmann, M. and Mundlos, S. (2013) Structural variations, the regulatory landscape of the genome and their alteration in human disease. *Bioessays*, **35**, 533–543.
12. Qu, H. and Fang, X. (2013) A brief review on the human encyclopedia of DNA elements (ENCODE) project. *Genomics Proteomics Bioinformatics*, **11**, 135–141.
13. Ward, L.D. and Kellis, M. (2012) Interpreting noncoding genetic variation in complex traits and human disease. *Nat. Biotechnol.*, **30**, 1095–1106.
14. Carnevali, P., Baccash, J., Halpern, A.L., Nazarenko, I., Nilsen, G.B., Pant, K.P., Ebert, J.C., Brownley, A., Morenzoni, M., Karpinchyk, V. *et al.* (2012) Computational techniques for human genome resequencing using mated gapped reads. *J. Comput. Biol.*, **19**, 279–292.
15. Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G. *et al.* (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, **327**, 78–81.
16. Browning, B.L. and Browning, S.R. (2011) A fast, powerful method for detecting identity by descent. *Am. J. Hum. Genet.*, **88**, 173–182.
17. Ng, P.C. and Henikoff, S. (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.*, **31**, 3812–3814.
18. Kumar, P., Henikoff, S. and Ng, P.C. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.*, **4**, 1073–1081
19. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. and Schilling, T.F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.*, **203**, 253–310.
20. Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X. *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/cas in zebrafish. *Nucleic Acids Res.*, **41**, e141.
21. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., Sander, J.D., Joung, J.K., Peterson, R.T. and Yeh, J.R. (2013) Heritable and precise zebrafish genome editing using a CRISPR-cas system. *PLoS One*, **8**, e68708.
22. Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K. and Dobbs, D. (2010) ZiFiT (zinc finger targeter): An updated zinc finger engineering tool. *Nucleic Acids Res.*, **38**, 462.
23. Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F. and Dobbs, D. (2007) Zinc finger targeter (ZiFiT): An engineered zinc finger/target site design tool. *Nucleic Acids Res.*, **35**, 599.
24. Barrallo-Gimeno, A., Holzschuh, J., Driever, W. and Knapik, E.W. (2004) Neural crest survival and differentiation in zebrafish depends on mont blanc/tfap2a gene function. *Development*, **131**, 1463–1477.
25. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C(T)) method. *Methods*, **25**, 402–408.
26. Spandidos, A., Wang, X., Wang, H. and Seed, B. (2010) PrimerBank: A resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Res.*, **38**, 792.