

ARTICLE



Evaluation of somatic and/or germline mosaicism in congenital malformation of the eye

Bertrand Chesneau^{1,2}, Véronique Ivashchenko¹, Christophe Habib¹, Véronique Gaston¹, Frédéric Escudé³, Godelieve Morel⁴, Yline Capri⁵, Catherine Vincent-Delorme⁶, Patrick Calvas^{1,2}, Nicolas Chassaing^{1,2} and Julie Plaisancie^{1,2,7}✉

© The Author(s), under exclusive licence to European Society of Human Genetics 2022, corrected publication 2022

Microphthalmia, Anophthalmia and Coloboma (MAC) form a spectrum of congenital eye malformations responsible for severe visual impairment. Despite the exploration of hundreds of genes by High-Throughput Sequencing (HTS), most of the patients remain without genetic diagnosis. One explanation could be the not yet demonstrated involvement of somatic mosaicism (undetected by conventional analysis pipelines) in those patients. Furthermore, the proportion of parental germline mosaicism in presumed de novo variations is still unknown in ocular malformations. Thus, using dedicated bioinformatics pipeline designed to detect mosaic variants, we reanalysed the sequencing data obtained from a 119 ocular development genes panel performed on blood samples of 78 probands with sporadic MAC without genetic diagnosis. Using the same HTS strategy, we sequenced 80 asymptomatic parents of 41 probands carrying a disease-causing variant in an ocular development gene considered de novo after Sanger sequencing of both parents. Reanalysis of the previously sequencing data did not find any mosaic variant in probands without genetic diagnosis. However, HTS of parents revealed undetected *SOX2* and *PAX6* mosaic variants in two parents. Finally, this work, performed on two large cohorts of patients with MAC spectrum, provides for the first time an overview of the interest of looking for mosaicism in ocular development disorders. Somatic mosaicism does not appear to be frequent in MAC spectrum and might explain only few diagnoses. Thus, other approaches such as whole genome sequencing should be considered in those patients. Parental mosaicism is however not that rare (around 5%) and challenging for genetic counselling.

European Journal of Human Genetics (2023) 31:526–530; <https://doi.org/10.1038/s41431-022-01234-3>

INTRODUCTION

Microphthalmia, Anophthalmia and Coloboma (MAC) form a spectrum of related congenital eye malformations [1]. Microphthalmia refers to a reduced axial length of the eye of various severity (below two standard deviations of the age-adjusted population mean). Microphthalmia can be simple, when the eye is of reduced size, but anatomically intact, or complex, when associated with anterior segment dysgenesis (ASD) or defect in the posterior segment of the eye. Anophthalmia corresponds to the total absence of any tissue of the eye. Ocular coloboma refers to a segmental defect affecting all or parts of the iris, choroid, retina and optic nerve. These ocular defects can be unilateral or bilateral and are frequently responsible for severe visual impairment. Moreover, extraocular features, mainly neurodevelopmental disorders, are associated with the ocular anomaly in 33–95% of patients [2]. MAC are typically caused by absent or insufficient growth of the eye or by a failure of the optic fissure to close during early eye development [1]. The exact pathophysiology remains however poorly understood. Their aetiology can include environmental factors, but genetic alterations represent the major cause [1]. Despite the use of high-throughput DNA sequencing (HTS) approaches such as

exome sequencing, 20–80% of patients remain without a genetic diagnosis after analysis [1, 3, 4]. Of note, the diagnosis rate is generally better when the ocular defect is severe, bilateral and syndromic.

Mosaicism is the consequence of a post-zygotic event resulting in the presence of a variant in a proportion of cells of the whole body [5]. This may result in clinical signs and/or risk of recurrence if the variant is present in the soma, the germline or both [6].

Detecting a mosaic variant with high throughput sequencing (HTS) requires sequencing at a greater depth of coverage (>100×) than required for the detection of constitutive variants (30–50×) and specific bioinformatics pipelines should be used. We can therefore miss a disease-causing variant when sequencing by using standard procedures designed for constitutive variant detection. As mosaic variants are known to have an impact on human genetic diseases [6–9], we wanted to investigate if an undetected mosaic pathogenic variant could explain part of unsolved cases of patients with a diagnosis of MAC. We therefore reanalysed previously sequenced data obtained with a panel of 119 ocular genes performed in 78 probands with a diagnosis of MAC using a dedicated pipeline conceived for mosaic variant detection.

¹Service de Génétique Médicale, Hôpital Purpan, CHU de Toulouse, Toulouse, France. ²Centre de Référence pour les Affections Rares en Génétique Opthalmologique (CARGO), CHU de Toulouse, Toulouse, France. ³Département d'anatomopathologie, IUCT Oncopole, Toulouse, France. ⁴Service de Génétique Médicale, CHU de Rennes, Rennes, France. ⁵Service de Génétique Médicale, Hôpital Robert Debré, APHP, Paris, France. ⁶Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU de Lille, Lille, France. ⁷INSERM U1214, ToNIC, Université Toulouse III, Toulouse, France. ✉email: plaisancie.j@chu-toulouse.fr

Received: 24 August 2022 Revised: 17 October 2022 Accepted: 1 November 2022
Published online: 21 November 2022

On the other hand, parental germline (or gonosomal) mosaicism can lead to the recurrence of an apparently de novo disease-causing variant. This has already been occasionally described in ocular defects [10–13] as in many other genetic conditions and should be taken into account in genetic counselling. Parental samples test are usually performed on leucocytes' DNA with Sanger sequencing, which cannot usually detect variants below 10%. We selected 41 probands with a diagnosis of ocular defect carrying an apparently de novo disease-causing variant after parental samples test by Sanger sequencing. We then sequenced parental samples of these 41 probands by using the same HTS strategy to check the pipeline's ability to detect low mosaicism of the offspring's variant missed by Sanger sequencing.

METHODS

Probands with MAC and no genetic diagnosis

We selected 78 individuals (38 females and 40 males) with a diagnosis of unilateral (25 individuals) or bilateral (53 individuals) ocular defects belonging to the MAC spectrum. Among the 41 patients with a diagnosis of microphthalmia, six had simple microphthalmia, 16 had colobomatous microphthalmia and 19 had complex microphthalmia. Thirty-four patients had a diagnosis of simple coloboma and only one patient had a diagnosis of anophthalmia. Of note, extra-ocular features, mostly developmental delay, learning difficulties and growth retardation, were reported in one third of patients (25/78). None of them had a family history of ocular defects.

They were previously explored by the mean of a HTS panel of 119 genes of ocular development (Supplementary Data 1) without any disease-causing variant identified. DNA was extracted from blood using the MagnaPure system (Roche Applied Science, Germany), except for three individuals whose DNA was extracted from other tissues (two from amniotic liquid, one from muscle foetal tissue). Capture probes were designed with SureDesign (Agilent, USA). A library of all coding exons and intron-exon boundaries was prepared using the SureSelect XT HS and XT Low Input Enzymatic Fragmentation Kit and the SureSelect XT HS Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library (Agilent, USA) following the manufacturer's instructions. Further sequencing was performed on a NextSeq500 platform (Illumina Inc., CA). Sequence alignment was performed with BWA 0.7.10, picard-tools-2.18.23, elprep4 (Indel realignment, base recalibration). Then, the variant calling was made with GATK-3.3 (HaplotypeCaller) and VarScan2.3.7. Annotations were made with SnpEff-4.3 with additional information from gnomAD, ClinVar and dbSNP151.

Concerning the analysis of Copy Number Variations (CNV), screening was performed using a custom in-house pipeline previously described [14].

Prior to reanalysis, eight individuals were found to have a heterozygous pathogenic or likely pathogenic variant in a gene with an autosomal recessive inheritance of the ocular disease (Supplementary Data 2). Thus, they were not sufficient alone to explain patient's ocular phenotype.

Parental samples test of apparently de novo pathogenic variants

We selected 41 probands with disease-causing variants in ocular development genes previously identified in our laboratory (*PAX6*, *SOX2*, *RARB*, *OTX2*, *PAX2*, *PRSS56*, *CHD7*, *TFAP2A*, *RAX*, *GJA8*, *BMP4*, *CRYAB*, *PITX2*, and *ITPR1*) (genotypes are detailed in Supplementary Data 3). They displayed various ocular defects belonging to the MAC (25/41), aniridia (8/41), or ASD spectrum (8/41) (Supplementary Data 3). Parents were all asymptomatic after ophthalmic examination. A recurrence in a proband's sibling was noted in only two families (SG040829, SG110088).

For all probands, parental samples test was performed on both parents using Sanger sequencing and did not find offspring's disease-causing variants. Sanger sequencing was performed after targeted PCR amplification using specific primers on the proband's and parents' genomic lymphocytes' DNA. PCR products were then sequenced using an ABI3130XL or an ABI3500XL (Applied Biosystems, Foster City, California, U.S.A.). DNA sequence variants were identified using SeqScape™ Software v3.0 (ThermoFisher scientific, Waltham, Massachusetts, U.S.A.) and Sequencing Analysis Software V7.0 (ThermoFisher scientific).

In order to improve the detection of a potential parental gonosomal mosaicism that could have been missed by Sanger sequencing, parents were sequenced by the mean of a targeted HTS panel (described above).

In two families, only one parent was resequenced: (i) In family LA090161, the proband had a heterozygous deletion of the autosomal recessive disease *RAX* gene inherited from his mother and a heterozygous nonsense variant [c.665C>A p.(Ser222*)] involving *RAX* found in none of the parental samples. Only the father's DNA was resequenced as the mother transmitted the heterozygous whole gene deletion. (ii) In family SG040829, the *PAX6* pathogenic variant [c.991C>T p.(Arg331*)] was found in a half-sister and half-brother with the same asymptomatic father. Thus, only the father was resequenced.

In total, 80 parents were sequenced using the 119 genes panel.

Mosaic detection using a targeted HTS panel: development of a new dedicated pipeline

Our lab already explored 78 probands using our 119 genes HTS panel with no result allowing a diagnosis. We therefore directly analysed the raw data (FASTQ files) previously obtained in our laboratory. We performed the parental samples test of the 80 parents using the same sample used for Sanger sequencing (DNA extracted from blood). We applied the same procedure for the high throughput sequencing (HTS) step (described above) for parental samples test to generate the raw sequencing data.

We reanalysed raw HTS data (FASTQ files) from all patients using a pipeline dedicated to mosaic variant detection. The overall pipeline performs trimming, sequence alignment (GRCh37) and duplicates suppression with CutAdapt [15], BWA-mem [16] and MarkDuplicate (<http://broadinstitute.github.io/picard/>) respectively. Variant Calling step used three different variant callers in parallel: Mutect2 [17], VarDict [18] and FreeBayes [19]. The pipeline was able to detect a variant as low as 4% of allelic frequency at 100X depth in a previous test performed in our laboratory with known mosaic variants (data not shown). The standardisation of VCF files follows the VCF 4.3 specifications, which allows secondary fusion of co-occurring variants and haplotype normalisation and the production of a unique variant calling file (VCF) by sample. Annotation was made using VEP tool (Ensembl). We analysed annotated VCFs without filters on allelic frequency. Of note, we analysed the parents' annotated VCF with a focus on their offspring's pathogenic variant to avoid unwanted secondary findings. To make sure that our pipeline did not miss any parental mosaic, their binary alignment map (BAM) files were visualised at offspring's variant position using Integrative Genomics Viewer (IGV).

RESULTS

Probands with MAC and no genetic diagnosis

After reanalysis of previously sequenced data of a targeted 119 genes of ocular development HTS panel, we did not find any mosaic variant in the 78 individuals with MAC. All previously detected heterozygous variants of interest were found (Supplementary Data 2). The quality values of the run were there to allow mosaic variant screening with a mean depth at 716x, a 30x coverage at 99.65% and 100x coverage at 98.30%.

Parental test of apparently de novo pathogenic variants

Resequencing of the 80 parents revealed two parental gonosomal mosaics (Fig. 1). The *PAX6* c.52G > C pathogenic variant associated with bilateral complex microphthalmia in the proband (SG110088) was found around 10% of cells in blood's DNA of her asymptomatic mother (allelic frequency: 13/252 = 5%) (Fig. 1A et B). This mosaic variant has led to a recurrence identified on prenatal diagnosis. In another family, the *SOX2* c.70_89del pathogenic variant associated with right anophthalmia and left Peters' Anomaly in the proband (ADN200107) was found around 4% cells in blood's DNA of her asymptomatic mother (allelic frequency: 9/441 = 2%) (Fig. 1C). Sanger sequencing of *PAX6* or *SOX2* performed previously using the same DNA samples failed to detect these two variants (Fig. 1BE). However, reanalysis of previously obtained sequences (.ab1 files) using *Minor Variant Finder* Software (Applied Biosystems) successfully found the 10% mosaic *PAX6* variant but failed to detect the 5% mosaic *SOX2* variant. Parental samples test of other parents did not reveal any other mosaic variant, despite high depth at each variant base in each family (mean depth: 1213X, minimal depth: 278X). Of note, in one family, germline mosaicism was highly suspected because of

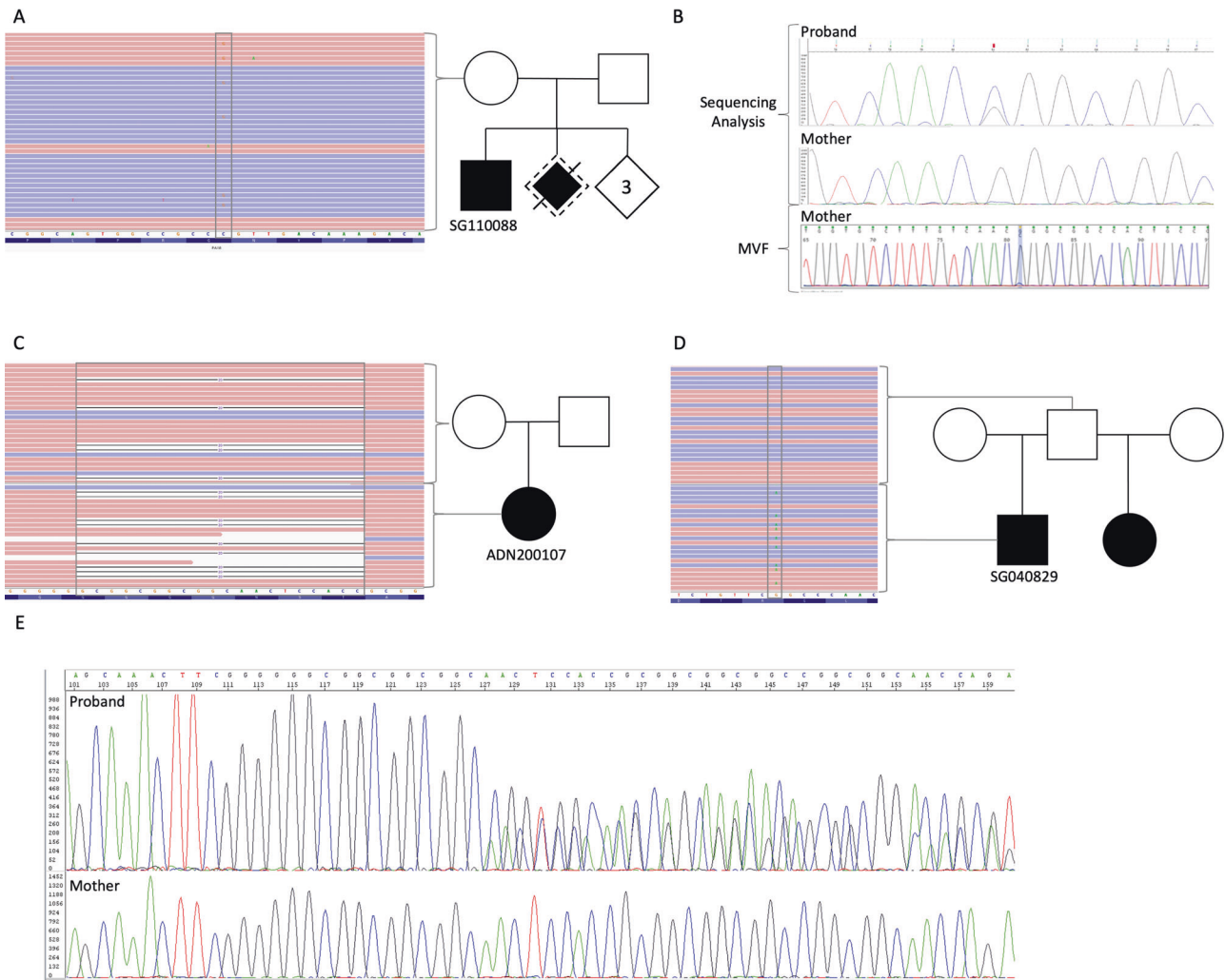


Fig. 1 Pedigrees and molecular results of the three families with suspected or confirmed parental mosaicism. **A** Pedigree of SG110088 with a pathogenic *PAX6* c.52G > C variant showing recurrence observed on prenatal diagnosis. Bam file visualisation on Integrative Genome Viewer (IGV) of the mosaic variant in asymptomatic mother's blood. **B** Sequence visualisation of the *PAX6* variant using Sequencing Analysis (up) and Minor Variant Finder software (MVF; down). Note that the mother's mosaic variant (that corresponds to the background noise level with Sequencing Analysis) is detected by MVF. **C** Pedigree of ADN200107 with a pathogenic *SOX2* c.70_89del variant with Bam files visualisation on IGV of the mosaic variant in asymptomatic mother's blood (up) and her affected daughter (down). **D** Pedigree of SG040829 with a pathogenic *PAX6* c.991C > T variant found in both children with aniridia but not found in their affected father as shown in Bam files visualisation on IGV (up). **E** Sequence visualisation of the *SOX2* variant of ADN200107 using Sequencing Analysis showing absence of distinguishable variant in mother's sample.

the recurrence in a girl (SG040829) and her paternal half-brother of the nonsense c.991C > T *PAX6* variant associated with bilateral aniridia in both (Fig. 1D). HTS resequencing of the father did not detect the variant found in his offspring despite a good coverage (depth of 1965X at variant base) suggesting mosaicism confined to germline lineage. In this family, samples' identities were confirmed by PowerPlex® 16 HS System (Promega, Madison, Wisconsin, U.S.A.).

DISCUSSION

In a unique cohort of 78 patients with a MAC diagnosis, we did not find any additional genetic diagnosis after reanalysis with a pipeline conceived for mosaic variant detection. The use of targeted HTS allowed to have a sufficient depth for mosaic variants screening (98.30% at 100X, mean depth 716X). Despite this high coverage, we cannot exclude missing a variant with a very low allelic frequency (VAF < 0.1). The implication of such variant in patient's phenotype is however difficult to ascertain [20].

Furthermore, tissue specific mosaicism with a variant absent from patient's blood but present in other tissues (such as ocular tissue) could also have been missed as we used only leucocytes' DNA. This phenomenon has previously been described in structural variations [21] and some syndromes such as Cornelia de Lange syndrome [22, 23]. Furthermore, Daich Varela et al. [24] described a suspected gonosomal *SOX2* variant mosaicism absent from blood in a mother with asymptomatic uveal coloboma. Her son had a *SOX2* variant associated with microphthalmia and developmental delay. Parental samples test analysis using exome sequencing did not find offspring's *SOX2* variant in mother's blood nor in her saliva. In a large cohort of probands with severe developmental disorders (DDD Study), there was however no significant variation of the levels of mosaicism between saliva and blood [20]. Although tissue specific mutation is conceivable in MAC disorders, we did not have access to other tissues, in particular the ocular ones.

In a recent review, Ohuchi et al. [25] underlined the potential role of mosaicism in congenital eye anomalies, partly explaining

phenotypic variability. Several cases of parental mosaicism have in fact been reported with phenotypes of various severity in parents ranging from asymptomatic to severe, without any clear correlation with the mosaic rate [11, 12, 26]. In example, Ragge et al. [27], described the segregation of an *OTX2* disease-causing variant in two children with bilateral complex microphthalmia and a less severe phenotype observed in their mother who had the variant at mosaic state (around 20% in blood) and display pigmentary retinopathy without ocular malformation. Thus, somatic mosaicism could result in less typical or milder clinical presentation and therefore no exploration is done before the occurrence of a more severe phenotype in offspring. In the same way, somatic mosaicism might partly explain unilateral defects that are frequent in congenital eye defects but with no pathophysiological explanation to date. In literature, all mosaic Single Nucleotide Variations (SNVs) reported to date in MAC phenotypes are parental mosaics and were found after the identification of the heterozygous variant in offspring. However, mosaic CNV in *PAX6* have previously been described in a cohort of patients with aniridia [28]. Of note, in this last study, patients with a mosaic variant (30–62% of cells) did not show significant clinical differences with patients harbouring a constitutive variant.

Thus, somatic mosaicism does not appear to be a frequent mechanism of variant apparition in MAC and systematic screening for mosaic variants, using the same sample, after negative molecular screening does not seem relevant and other techniques or samples should be preferred. In order to increase the diagnosis rate in those patients, Whole genome sequencing would allow analysis of previously not explored genomic regions (other genes, introns, intergenic regions). Furthermore, long-read sequencing could also help finding more diagnoses with a better genome coverage than short-read sequencing and detection of structural variations [29].

On the other hand, parental gonosomal mosaicism seems to be not that rare in ocular defects [25]. In our cohort of 80 parents of 41 probands with a disease-causing variant, presumed de novo after parental Sanger sequencing, we found two variants at low mosaic state in asymptomatic parents. This means that at least 5% (2/41) of the asymptomatic parents could be mosaic. This percentage is probably underestimated as illustrated by the probable mosaic associated with a recurrence that was not detected in the paternal sample (SG040829). These three parental mosaics (two confirmed, one highly suspected) represent 7% (3/41) of our cohort. Demonstration of parental mosaicism is important for genetic counselling as it induces a higher recurrence risk as shown in family SG110088. Prenatal diagnosis should therefore be discussed with the couples, even in case of negative parental testing. The use of HTS confirms the limits of Sanger sequencing to detect low mosaic variants. Performing exome and whole genome sequencing usually requires trio sequencing of the proband and both parents, which will theoretically rule out the limitations of Sanger sequencing. HTS however still have some limitations to be taken in consideration such as non-detection of parental low mosaicism by standard pipelines or insufficient depth [20]. Furthermore, when analysing only peripheral blood DNA, we could have underestimated the risk of recurrence, missing a germline mosaicism absent from blood. This is the case in one family (SG040829) with a paternal highly suspected germline mosaicism of a *PAX6* disease-causing variant not detected in blood but confirmed by the recurrence in two of his children (with different mother). Although it is the most used sample in routine, blood might not be the best-suited sample to detect parental mosaicism in ocular development diseases. In a family with an Alport syndrome-affected child, Dai et al. [30] showed paternal germline mosaicism not found in blood sample, but detected at 2.65% cells in the father's sperm sample. Thus, it would be interesting to look for parental mosaicism in other tissues (saliva, urine, sperm samples) to see if another easily accessible tissue is

more appropriate for mosaics' detection. However, although screening of multiple tissues would help to refine the risk of recurrence, it would never allow the exclusion of germline mosaicism and recurrence risk in clinical practice.

In conclusion, mosaic variants detection represents a challenge for both diagnosis and genetic counselling, especially in MAC disorders. Firstly, when considering the fact that the clinically relevant tissue is absent or inaccessible in those disorders, this study represents a unique work showing that mosaic variants in blood are rare in patients with MAC. Secondly, we show here that parental mosaicism is higher (at least 5% and probably underestimated) than the theoretical risk of recurrence usually gave in consultation to the parents of around 1%.

DATA AVAILABILITY

All data generated or analysed during this study are included in this published article and its supplementary information files. Newly described variants have been added to ClinVar database (SCV002583597 - SCV002583623).

REFERENCES

1. Plaisancié J, Ceroni F, Holt R, Zazo Seco C, Calvas P, Chassaing N, et al. Genetics of anophthalmia and microphthalmia. Part 1: non-syndromic anophthalmia/microphthalmia. *Hum Genet.* 2019;138:799–830.
2. Slavotinek A. Genetics of anophthalmia and microphthalmia. Part 2: syndromes associated with anophthalmia–microphthalmia. *Hum Genet.* 2019;138:831–46.
3. Patel N, Khan AO, Alsahli S, Abdel-Salam G, Nowilaty SR, Mansour AM, et al. Genetic investigation of 93 families with microphthalmia or posterior microphthalmos. *Clin Genet.* 2018;93:1210–22.
4. Haug P, Koller S, Maggi J, Lang E, Feil S, Włodarczyk A, et al. Whole exome sequencing in coloboma/microphthalmia: identification of novel and recurrent variants in seven genes. *Genes.* 2021;12:65.
5. Veltman JA, Brunner HG. De novo mutations in human genetic disease. *Nat Rev Genet.* 2012;13:565–75.
6. Spinner NB, Conlin LK. Mosaicism and clinical genetics. *Am J Med Genet C Semin Med Genet.* 2014;166:397–405.
7. Stosser MB, Lindy AS, Butler E, Retterer K, Piccirillo-Stosser CM, Richard G, et al. High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet Med.* 2018;20:403–10.
8. Erickson RP. Somatic gene mutation and human disease other than cancer: an update. *Mutat Res.* 2010;705:96–106.
9. Acuna-Hidalgo R, Bo T, Kwint MP, van de Vorst M, Pinelli M, Veltman JA, et al. Post-zygotic point mutations are an underrecognized source of de novo genomic variation. *Am J Hum Genet.* 2015;97:67–74.
10. Tarilonte M, Morin M, Ramos P, Galdós M, Blanco-Kelly F, Villaverde C, et al. Parental mosaicism in *PAX6* causes intra-familial variability: implications for genetic counseling of congenital aniridia and microphthalmia. *Front Genet.* 2018;9:479.
11. Chassaing N, Gilbert-Dussardier B, Nicot F, Fermeaux V, Encha-Razavi F, Fiorenza M, et al. Germinal mosaicism and familial recurrence of a *SOX2* mutation with highly variable phenotypic expression extending from AEG syndrome to absence of ocular involvement. *Am J Med Genet A.* 2007;143A:289–91.
12. Faivre L, Williamson KA, Faber V, Laurent N, Grimaldi M, Thauvin-Robinet C, et al. Recurrence of *SOX2* anophthalmia syndrome with gonosomal mosaicism in a phenotypically normal mother. *Am J Med Genet A.* 2006;140:636–9.
13. Suzuki J, Azuma N, Dateki S, Soneda S, Muroya K, Yamamoto Y, et al. Mutation spectrum and phenotypic variation in nine patients with *SOX2* abnormalities. *J Hum Genet.* 2014;59:353–6.
14. Plaisancié J, Tarilonte M, Ramos P, Jeanton-Scaramouche C, Gaston V, Dollfus H, et al. Implication of non-coding *PAX6* mutations in aniridia. *Hum Genet.* 2018;137:831–46.
15. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17:10.
16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–60.
17. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling Somatic SNVs and Indels with Mutect2. 2019:861054.
18. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 2016;44:e108–e108.
19. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012.arXiv:1207.3907v2 [q-bio.GN].

20. Wright CF, Prigmore E, Rajan D, Handsaker J, McRae J, Kaplanis J, et al. Clinically-relevant postzygotic mosaicism in parents and children with developmental disorders in trio exome sequencing data. *Nat Commun.* 2019;10:2985.
21. King DA, Sifrim A, Fitzgerald TW, Rahbari R, Hobson E, Homfray T, et al. Detection of structural mosaicism from targeted and whole-genome sequencing data. *Genome Res.* 2017;27:1704–14.
22. Nizon M, Henry M, Michot C, Baumann C, Bazin A, Bessières B, et al. A series of 38 novel germline and somatic mutations of NIPBL in Cornelia de Lange syndrome. *Clin Genet.* 2016;89:584–9.
23. Kline AD, Moss JF, Selicorni A, Bisgaard A-M, Deardorff MA, Gillett PM, et al. Diagnosis and management of Cornelia de Lange syndrome: first international consensus statement. *Nat Rev Genet.* 2018;19:649–66.
24. Daich Varela M, Hufnagel RB, Guan B, Blain D, Sapp JC, Gropman AL, et al. Clinical diagnosis of presumed SOX2 gonadosomatic mosaicism. *Ophthalmic Genet.* 2021;42:320–5.
25. Ohuchi H, Sato K, Habuta M, Fujita H, Bando T. Congenital eye anomalies: more mosaic than thought? *Congenit Anom.* 2019;59:56–73.
26. Shima H, Ishii A, Wada Y, Kizawa J, Yokoi T, Azuma N, et al. SOX2 nonsense mutation in a patient clinically diagnosed with non-syndromic hypogonadotropic hypogonadism. *Endocr J.* 2017;64:813–7.
27. Ragge NK, Brown AG, Poloschek CM, Lorenz B, Henderson RA, Clarke MP, et al. Heterozygous mutations of OTX2 cause severe ocular malformations. *Am J Hum Genet.* 2005;76:1008–22.
28. Robinson DO, Howarth RJ, Williamson KA, Heyningen V, van, Beal SJ, Crolla JA. Genetic analysis of chromosome 11p13 and the PAX6 gene in a series of 125 cases referred with aniridia. *Am J Med Genet A.* 2008;146A:558–69.
29. Mitsuhashi S, Matsumoto N. Long-read sequencing for rare human genetic diseases. *J Hum Genet.* 2020;65:11–19.
30. Dai C, Cheng D, Li W, Zeng S, Lu G, Zhang Q. Identification of paternal germline mosaicism by MicroSeq and targeted next-generation sequencing. *Mol Genet Genomic Med.* 2020;8. <https://doi.org/10.1002/mgg3.1394>.

ACKNOWLEDGEMENTS

We appreciated the generous support received from patients and their families and are grateful for the contributions of physicians who evaluated or followed up patients with ocular malformation.

AUTHOR CONTRIBUTIONS

All authors have actively participated in data collection or interpretation and have revised and approved the manuscript. BC and JP designed experiments and write the

first version of manuscript. GM, YC, CVD, PC collected data from patients. VG performed sequencing. VI, CH, and FE performed bioinformatic analysis. BC, NC, and JP performed analysis and interpretation.

FUNDING

Patients were recruited through the Rare Diseases Cohorts (RaDiCo) programme which is funded by the French National Research Agency under the specific programme “Investments for the Future”, Cohort grant agreement ANR-10-COHO-0003.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

Patients gave informed consent to genetic analyses and research. This study is a retrospective study based on the exploitation of usual care data. A submission at an ethic committee is not required according to the French ethic and regulatory law (public health code).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41431-022-01234-3>.

Correspondence and requests for materials should be addressed to Julie Plaisancié.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.