High Clinical Exome Sequencing Diagnostic Rates and Novel Phenotypic Expansions for Nonisolated Microphthalmia, Anophthalmia, and Coloboma

Bhavana Kunisetty,[1] Bailey A. Martin-Giacalone,[2,3] Xiaonan Zhao,[1,4] Pamela N. Luna,[1] Brian P. Brooks,[5] Robert B. Hufnagel,[5] Chad A. Shaw,[1] Jill A. Rosenfeld,[1] A. J. Agopian,[6] Philip J. Lupo,[3] and Daryl A. Scott[1,7]

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, United States
[2]Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St. Louis, Missouri, United States
[3]Section of Hematology-Oncology, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, United States
[4]Baylor Genetics, Houston, Texas, United States
[5]Ophthalmic Genetics & Visual Function Branch, National Eye Institute, NIH, Bethesda, Maryland, United States
[6]Department of Epidemiology, Human Genetics & Environmental Sciences, UTHealth School of Public Health, Houston, Texas, United States
[7]Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, Texas, United States

Correspondence: Daryl A. Scott, R813, One Baylor Plaza, BCM225, Houston, TX 77030, USA; dscott@bcm.edu.

Received: November 22, 2023
Accepted: February 26, 2024
Published: March 19, 2024

Citation: Kunisetty B, Martin-Giacalone BA, Zhao X, et al. High clinical exome sequencing diagnostic rates and novel phenotypic expansions for nonisolated microphthalmia, anophthalmia, and coloboma. *Invest Ophthalmol Vis Sci.* 2024;65(3):25. https://doi.org/10.1167/iovs.65.3.25

Purpose. A molecular diagnosis is only made in a subset of individuals with nonisolated microphthalmia, anophthalmia, and coloboma (MAC). This may be due to underutilization of clinical (whole) exome sequencing (cES) and an incomplete understanding of the genes that cause MAC. The purpose of this study is to determine the efficacy of cES in cases of nonisolated MAC and to identify new MAC phenotypic expansions.

Methods. We determined the efficacy of cES in 189 individuals with nonisolated MAC. We then used cES data, a validated machine learning algorithm, and previously published expression data, case reports, and animal models to determine which candidate genes were most likely to contribute to the development of MAC.

Results. We found the efficacy of cES in nonisolated MAC to be between 32.3% (61/189) and 48.1% (91/189). Most genes affected in our cohort were not among genes currently screened in clinically available ophthalmologic gene panels. A subset of the genes implicated in our cohort had not been clearly associated with MAC. Our analyses revealed sufficient evidence to support low-penetrance MAC phenotypic expansions involving nine of these human disease genes.

Conclusions. We conclude that cES is an effective means of identifying a molecular diagnosis in individuals with nonisolated MAC and may identify putatively damaging variants that would be missed if only a clinically available ophthalmologic gene panel was obtained. Our data also suggest that deleterious variants in *BRCA2*, *BRIP1*, *KAT6A*, *KAT6B*, *NSF*, *RAC1*, *SMARCA4*, *SMC1A*, and *TUBA1A* can contribute to the development of MAC.

Keywords: exome sequencing, microphthalmia, anophthalmia, machine learning, candidate genes

Microphthalmia, anophthalmia, and coloboma (MAC) are related structural, congenital eye malformations that display a spectrum of severity and account for approximately 15% to 20% of severe visual impairment and blindness in children worldwide.[1–3] Microphthalmia describes an eye with reduced volume and a total axial length that is less than two standard deviations below the population age-adjusted mean, while anophthalmia is defined as the clinical absence of ocular tissue with no visible sign of a globe.[4] Coloboma describes a defect in which tissue is missing from the eyelid, cornea, iris, lens, ciliary body, zonules, retina, choroid, and/or optic nerve.[5] Colobomas that are situated inferonasally are thought to arise from incomplete fusion of the optic fissure, while colobomas located in other sites are considered atypical and likely to have a different developmental etiology. MAC prevalence estimates per 100,000 births range from 2 to 17 for microphthalmia, 0.6 to 4.2 for anophthalmia, and 2 to 14 for coloboma.[3]

Abnormalities in single genes have been shown to cause both isolated MAC and nonisolated (syndromic)

MAC that occurs in association with a spectrum of extraocular defects. For example, autosomal recessive variants in *STRA6*, which encodes a transmembrane receptor responsible for vitamin A and retinoic acid metabolism, cause both microphthalmia, isolated, with coloboma 8 (MIM #601186) and microphthalmia, syndromic 9 (MIM #601186; Matthew–Wood syndrome), in which MAC occurs in the setting of pulmonary, diaphragm, and/or cardiac defects.[6]

Whole-exome sequencing has proven to be an effective means of identifying putatively damaging single-nucleotide variants in individuals with MAC, but whole-exome sequencing is not always ordered on a clinical basis for individuals with these eye anomalies.[7,8] In some cases, commercially available ophthalmologic gene panels may be used as an alternative mode of genetic testing, mostly due to their lower cost.[9] We also note that, although many genes associated with the development of MAC have been identified, most cases continue to have an unknown genetic etiology.[8] This suggests the need for further evidence to support the use of clinical (whole) exome sequencing (cES) in individuals with MAC and the need to continue to identify genes that contribute to the development of these disorders.

Here, we analyze data from a clinical database to determine the efficacy of cES in individuals with nonisolated MAC, and we evaluate the coverage of commercially available ophthalmologic gene panels in light of the variants identified by cES. We then use cES data, MAC-specific rank annotation scores generated by a machine learning algorithm, and previously published expression data, case reports, and animal models to identify low-penetrance phenotypic expansions involving MAC that can be easily overlooked in traditional gene discovery efforts.

Methods

Human Subjects Research

This work was approved by the institutional review board of Baylor College of Medicine (protocol H-47546) and conducted in accordance with the ethical standards of this institution's committee on human research and international standards.

Clinical Exome Sequencing

All cES data were generated at Baylor Genetics—a Clinical Laboratory Improvement Amendments–certified laboratory. The following quality control metrics for cES were generally achieved: >70% of reads are aligned to the target with >95% of targeted bases covered at >20 reads, >85% of targeted bases are covered at >40 reads, and mean coverage of targeted bases is >100 reads.

Subject Identification

We searched deidentified data from ~17,000 individuals referred to Baylor Genetics for cES (December 2011 to June 2020) for cases in which the indication for testing included one or more MAC-related phenotypes. Cases in which a chromosome anomaly was likely to have contributed to one or more of the phenotypes listed in the indication were excluded. A total of 189 individuals with "microphthalmia," "anophthalmia," or "coloboma" listed in their indication for testing were identified. Within this cohort, 91 individuals (subjects S1–S91) had one or more variants reported back to the referring physicians as having the potential to contribute to all or a subset of the phenotypes listed in their indication for testing (see Supplementary Table S1). All of these individuals had nonisolated MAC except for subject S59.

Variant Reclassification

Variants reported back to the referring physicians of subjects S1 to S91 as having the potential to contribute to all or a subset of the phenotypes listed in their indication for testing were reclassified as pathogenic, likely pathogenic, a variant of unknown significance (VUS), likely benign, or benign, based on the American College of Medical Genetics and Genomics (ACMG) standards for the interpretation of sequence variants by a clinical laboratory geneticist (November 2022; see Supplementary Table S1 for criteria).[10]

Data Access

All previously unpublished variants identified in this study were submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/; SUB13969089).

Determination of Diagnostic Certainty

Molecular and clinical data were reviewed to determine the diagnostic certainty—"definitive," "probable," or "provisional"—associated with the variant(s) reported in each individual based on criteria proposed by Scott et al.[11] These criteria consider the ACMG variant classification, variant inheritance pattern, variant configuration (*cis* versus *trans*), proband sex, and the similarity between the phenotypes listed in the indication and phenotypes known to be associated with disorder(s) caused by deleterious variants in the affected gene(s).

Statistics

Two-tailed Fisher's exact tests were used to compare the cES efficacy between MAC subgroups. *P* values were generated for these comparisons using the 2 × 2 contingency table tool from GraphPad (https://www.graphpad.com/quickcalcs/contingency1/). Box plots were generated using the Alcula.com Statistical Calculator: Box Plot program (http://www.alcula.com/calculators/statistics/box-plot/). Two-tailed unpaired *t*-tests were performed using the T test calculator tool from GraphPad (https://www.graphpad.com/quickcalcs/ttest1/?format=C).

Commercially Available Ophthalmologic Gene Panels

We manually curated a list of genes screened on three commercially available ophthalmologic gene panels whose descriptive labels were "Comprehensive Ocular Disorders Panel," "Microphthalmia, Anophthalmia, and Anterior Segment Dysgenesis Panel," and "MAC and Anterior Segment Dysgenesis Panel" (September 2023).

Investigative Ophthalmology & Visual Science

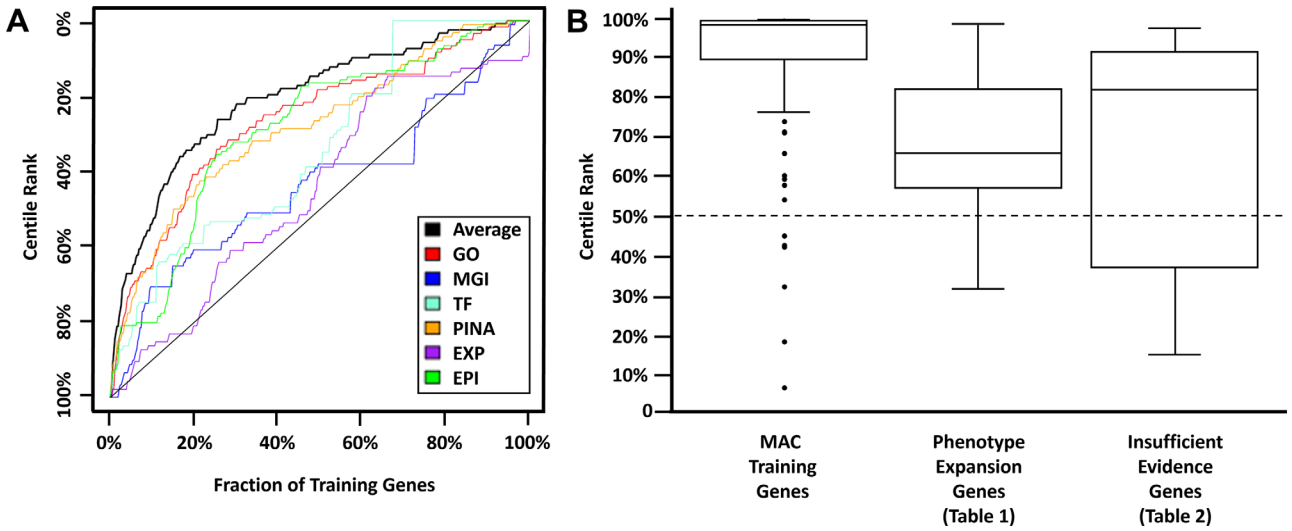

FIGURE. Using machine learning to rank all RefSeq genes based on their similarity to genes known to cause MAC. (**A**) A previously published machine learning algorithm was trained using 122 human genes known to cause MAC and/or the human homologs of genes that cause MAC in mice.[13,14] Receiver operating characteristic style curves were generated based on a leave-one-out validation analysis performed for each knowledge source (*colored lines*) and an omnibus score (*black line*) whose positive deviation indicates that our algorithm can identify genes in the training set more effectively than chance (*diagonal line*). (**B**) Box plots showing the algorithmically generated MAC-specific rank annotation scores for MAC training genes, candidate genes for which there is sufficient evidence to suggest a phenotypic expansion involving MAC (Table 1), and selected candidate genes with some evidence suggesting an association with MAC but insufficient evidence to suggest a phenotypic expansion involving MAC (Table 2). The median rank annotation scores of these groups—99%, 69.1%, and 87.1%, respectively—were greater than what would be expected by chance alone (50%; *dotted line*).

Curation of Gene Sets

We manually curated a training set of 122 human genes known to cause MAC and/or the human homologs of genes that cause MAC in mice (see Supplementary Table S2).

Genes sets used for comparison included 35 genes known to cause anomalous pulmonary venous return curated by Huth et al.,[12] 83 genes known to cause epilepsy curated by Campbell et al.,[13] and 332 human olfactory genes that were manually curated.

Machine Learning

We manually curated a training set of 122 human genes known to cause MAC and/or the human homologs of genes that cause MAC in mice (see Supplementary Table S2). We used these genes to train a previously published machine learning algorithm designed to compare the similarity of all RefSeq genes to a set of training genes based on data from Gene Ontology, the Mouse Genome Database phenotype annotation, the Protein Interaction Network Analysis platform, the GeneAtlas expression distribution, and transcription factor binding and epigenetic histone modifications data from the NIH Roadmap Epigenomics Mapping Consortium.[13,14]

Leave-one-out cross-validation studies were performed by iteratively excluding a single training gene and fitting the machine learning model using the remaining training genes. For each cross-validation instance, an evaluation of all RefSeq genes was performed, including the excluded training gene. The resulting scores were studentized, and the score of the excluded gene was recorded. The procedure was repeated so that each training gene received a cross-validated score. We then compared the studentized cross-validated scores of the training set genes to the scores of all other RefSeq genes derived from applying the machine learning model constructed using all training genes. Receiver operating characteristic style curves were generated from this comparison in which the effectiveness of the procedure corresponds to the area under the curve and above the diagonal line, which represents the result that would be generated by chance alone (Fig. A). An omnibus curve produced using fit data from all knowledge sources was positive, indicating that the algorithm could distinguish between the MAC-associated genes in the training set and all other RefSeq genes better than random chance.

After validating the algorithm, we generated MAC-specific rank annotation scores ranging from 0% to 100% for all RefSeq genes (see Supplementary Table S3). The MAC-specific rank annotation scores of the 122 genes in the training set ranged from 100% to 10.7% with a median score of 99% (Fig. B; see Supplementary Table S4 and Supplementary Fig. S1). In contrast, the median score for all RefSeq genes, by definition, is 50%.

To confirm that these MAC-specific rank annotation scores have specificity for the MAC phenotype, we determined the MAC-specific rank annotation scores of 35 genes known to cause a different structural birth defect, anomalous pulmonary venous return (APVR); 83 genes known to cause a neurologic disorder, epilepsy; and 332 olfactory receptor genes (see Supplementary Table S4).[12,13] As expected, the median MAC-specific rank annotation scores of each group of genes were progressively lower: 83.6% for the APVR genes, 76.5% for the epilepsy genes, and 5.2% for the olfactory receptor genes (see Supplementary Fig. S1). When the MAC-specific rank annotation scores of the MAC training gene set were compared to those of the APRV, epilepsy, and olfactory receptor genes using a two-tailed unpaired *t*-test, they were found to be significantly greater ($P < 0.0001$).

Investigative Ophthalmology & Visual Science

This suggests that the MAC-specific rank annotation scores are more highly associated with genes that cause MAC than genes that cause other phenotypes.

Results

Exome Sequencing Efficacy in Individuals With Nonisolated MAC

In 189 individuals with nonisolated MAC, a definitive diagnosis was made in 46 individuals and a probable diagnosis was made in 15, making the cES efficacy rate 32.3% (61/189). If the 30 individuals who received a provisional diagnosis were included, the cES efficacy rate would be 48.1% (91/189).

The cES efficacy rates for individuals with microphthalmia, anophthalmia, or coloboma were 29.4% (25/85), 7.7% (1/13), and 33.3% (43/129), respectively, based on definitive and probable diagnoses, or 43.5% (37/85), 15.4% (2/13), and 51.2% (66/129) if provisional diagnoses were included. No statistically significant difference was seen in the cES efficacy rates between these MAC subgroups at the $P < 0.05$ level, except between the anophthalmia and coloboma subgroups when provisional diagnoses were included ($P = 0.0182$).

Coverage of Commercially Available Ophthalmologic Gene Panels

To determine if gene panels represent a reasonable alternative to cES in individuals with nonisolated MAC, we determined the percentage of genes carrying variants identified by cES in our cohort that would have been screened in one of three clinically available ophthalmologic gene panels. Of the 83 genes with variants reported in this study, the number screened in the ophthalmologic gene panels ranged from 13 (15.7%) to 25 (30.1%). Among the 48 genes by which a definitive or probable diagnosis was made in at least one individual in our cohort, the number screened in the ophthalmologic gene panels ranged from 11 (22.9%) to 13 (27.1%).

Known MAC Genes and MAC Candidate Genes

Some of the variants reported back to referring physicians occurred in genes whose association with one or more MAC-related phenotypes is reported in OMIM (https://www.omim.org/), GeneReviews (https://www.ncbi.nlm.nih.gov/books/NBK1116/), published review articles, or multiple reported individuals as revealed by manual curation (see Supplementary Table S1). These genes included *ACTB* (n = 1), *ASXL1* (n = 1), *BCOR* (n =2), *C2CD3* (n = 1), *CCDC22* (n = 1), *CHD7* (n = 3), *COL18A1* (n = 1), *CREBBP* (n = 2), *CRPPA* (n = 1), *CTNNB1* (n = 1), *DHCR7* (n = 1), *EFTUD2* (n = 1), *FANCD2* (n = 1), *FGFR1* (n = 1), *FOXL2* (n = 1), *FREM2* (n = 1), *FZD5* (n = 1), *GJA1* (n = 1), *GLI2* (n = 1), *KDM6A* (n = 1), *KMT2A* (n = 1), *KMT2D* (n = 4), *MAB21L2* (n = 1), *NAA10* (n = 1), *OTX2* (n = 1), *PACS1* (n = 3), *PAX2* (n = 1), *PAX6* (n = 1), *PORCN* (n = 2), *PRR12* (n = 2), *PUF60* (n = 1), *RAB3GAP1* (n = 2), *RAB3GAP2* (n = 1), *RARB* (n = 3), *RERE* (n = 1), *TMEM138* (n = 1), and *TMEM67* (n = 1). The remainder of the variants identified by cES were in "MAC candidate genes," whose association with the development of MAC in humans has not been clearly established.

Comparison of MAC-Specific Rank Annotation Scores Between Cohorts

The MAC-specific rank annotation scores of genes associated with a definitive or probable diagnosis in our MAC cohort ranged from 23.3% to 100% with a median score of 80.3% (see Supplementary Table S5 and Supplementary Fig. S2). The MAC-specific rank annotation scores of genes associated with a definitive or probable diagnosis in a previously published cohort of individuals with a different structural birth defect, APVR had a similar range, from 27.9% to 97.8%, but only had a median score of 58.2% (see Supplementary Table S5 and Supplementary Fig. S2).[12] When the MAC-specific rank annotation scores of the MAC cohort were compared to those of the APVR cohort using a two-tailed unpaired *t*-test, they were found to be significantly

Table 1. Genes With Sufficient Evidence to Suggest a Low-Penetrance Phenotypic Expansion

Gene	Subject, MAC Type, Diagnostic Certainty	Other Cases of MAC in the Literature?	Expressed in the Developing Mammalian Visual System?*	Animal Models With MAC?	MAC-Specific Rank Annotation Score
BRCA2	S10, C, Definitive	M in individuals with Fanconi anemia[16]	Yes	No	53%
BRIP1	S11, M, Definitive	M in individuals with Fanconi anemia[16]	Yes	No	33.5%
KAT6A	S38, C, Definitive	M[18]	Yes	No	64%
KAT6B	S39, M, Definitive	No	Yes	Mouse, M[19,20]	73%
NSF	S54, M, Definitive	No	Yes	Zebrafish, M[27]	98.8%
RAC1	S73, C, Probable	No	Yes	Mouse, MA (IMPC)	70.6%
SMARCA4	S82, C, Probable	MA[31]	Yes	Zebrafish, M[33,34] Mouse, A (B2B/CvDC)	63.1%
SMC1A	S83, C, Definitive S84, C, Provisional	MC in individuals with Cornelia de Lange syndrome[35,36]	Yes	No	69.1%
TUBA1A	S89, C, Definitive	M,[39] C[38]	Yes	Zebrafish, M[40]	80.1%

* Based on mouse expression studies catalogued in the Mouse Genome Informatics (MGI) database. A, anophthalmia; B2B/CvDC, NHLBI Cardiovascular Development Consortium, Bench to Bassinet Program; C, coloboma; IMPC, International Mouse Phenotyping Consortium; M, microphthalmia.

Table 2. Genes for Which There Is Currently Insufficient Evidence to Suggest a Low-Penetrance Phenotypic Expansion

Gene	Subject, MAC Type, Diagnostic Certainty	Other Cases of MAC in the Literature?	Expressed in the Developing Mammalian Visual System?*	Animal Models With MAC?	MAC-Specific Rank Annotation Score
AFF4	S4, C, Provisional	No	Yes	Mouse, MA (IMPC)	67%
ANK3	S5, C, Provisional	C[41]	Yes	No	94.1%
BRAT1	S9, MC, Provisional	M[42]	Yes	No	35.9%
KIF2A	S42, C, Probable	M[43]	Yes	No	37.6%
LRP4†	S48, C, Provisional	No	Yes	Mouse, M[44]	96.8%
TCTN3†	S48, C, Provisional	No	Yes	Mouse, A[45]	99.2%
RP1L1	S53, C, Provisional	C[46]	Yes	Zebrafish, M[47]	86.6%
SKI	S79, MA, Provisional	No	Yes	Mouse, MC[48]	87.5%
SLC12A5	S80, M, Provisional	No	Yes	Mouse, M (IMPC)	92.3%
THOC2	S85, C, Provisional	C[49]	Yes	Zebrafish, M[50]	18.2%

* Based on mouse expression studies catalogued in the Mouse Genome Informatics (MGI) database. A, anophthalmia; C, coloboma; IMPC, International Mouse Phenotyping Consortium; M, microphthalmia.
† These variants occurred in the same individual.

greater ($P < 0.0001$). This provides additional evidence of the specificity of the MAC-specific rank annotation scores. The only gene leading to a definitive or probable diagnosis in both cohorts was *EFTUD2*, which is known to cause both MAC and cardiac defects, including APVR, in the setting of mandibulofacial dysostosis with microcephaly, Guion–Almeida type (MIM# 610536).[12]

Identification of Phenotypic Expansions Involving MAC

Variants in some MAC candidate genes may be present in individuals in our cohort by chance and/or may be contributing only to their non-MAC phenotypes. To distinguish these genes from genes that are likely to contribute to the development of MAC, we determined if (1) there were published reports of individuals with MAC who carried a putatively damaging variant in the candidate gene, (2) changes in the candidate gene's homolog(s) cause MAC-related phenotypes in animal models, (3) the candidate gene is known to be expressed in the developing visual system, and/or (4) the candidate gene has a high MAC-specific rank annotation score.

Among candidate genes associated with a definitive or probable diagnosis in our cohort, there is sufficient evidence to suggest a low-penetrance phenotypic expansion involving MAC for *BRCA2*, *BRIP1*, *KAT6A*, *KAT6B*, *NSF*, *RAC1*, *SMARCA4*, *SMC1A*, and *TUBA1A* (Table 1). As a group, these genes have higher MAC-specific rank annotation scores than would be expected by chance with a range of 98.8% to 33.5% and a median score of 69.1% (Fig. B).

Among candidate genes associated with a probable or provisional diagnosis in our cohort, we were able to identify additional evidence to suggest a possible association with MAC for *AFF4*, *ANK3*, *BRAT1*, *KIF2A*, *LRP4*, *TCTN3*, *RP1L1*, *SKI*, *SLC12A5*, and *THOC2* (Table 2). Although these genes currently do not have sufficient evidence to support a low-penetrance phenotypic expansion involving MAC, it is possible that they will ultimately be shown to cause MAC in humans in future studies. As a group, these genes have higher MAC-specific rank annotation scores than would be expected by chance with a range of 99.2% to 18.2% and a median score of 87.1% (Fig. B). This is consistent with this group being enriched for bona fide MAC genes.

Discussion

Obtaining a molecular diagnosis though genetic testing has the potential to provide benefits to affected individuals and their families.[15] cES is not universally ordered on individuals with nonisolated MAC. This may be due to uncertainty regarding the efficacy of cES in this population and the exclusive use of alternative genetic testing modes such as commercially available ophthalmologic gene panels. We also note that a molecular diagnosis is only identified in a minority of individuals with MAC that undergo genetic testing.[8] This suggests that some causative genes have yet to be identified. In this study, we used a machine learning approach to identify low-penetrance phenotypic expansions involving MAC that can be easily overlooked in traditional gene discovery efforts.

Efficacy of cES and Comparison to Ophthalmologic Gene Panels

The efficacy of cES in our cohort of 189 individuals with nonisolated MAC was 32.3% (61/189) when considering only individuals with a definitive or probable diagnosis and 48.1% (91/189) if individuals with a provisional diagnosis are included. This suggests that cES has a high efficacy rate among individuals with nonisolated MAC. The cES data analyzed in this study were generated between December 2011 and June 2020. It is likely that cES studies ordered after this time period would have even higher efficacy rates due to advancements in cES technology and our ever-increasing knowledge of gene/phenotype associations. cES is limited in its ability to detect structural variants, copy number variants, and deep intronic variants. These variants are more easily detected by clinical genome sequencing, which we would expect would have a higher diagnostic rate.

To determine if gene panels would have yielded comparable results to cES in our cohort, we determined how many of the 83 genes affected by variants in this study are screened in three clinically available ophthalmologic gene panels. We found that only a subset of these affected genes—between 13 (15.7%) and 25 (30.1%) genes—were screened on these panels. Similar results were obtained when we considered only the 48 genes by which a definitive or probable diagnosis was made—between 11 (22.9%) and 13 (27.1%) genes. This suggests that cES has the potential to identify putatively

damaging variants in individuals with nonisolated MAC that would be missed if only a clinically available ophthalmologic gene panel was obtained.

BRCA2 and BRIP1

Ophthalmic phenotypes, including microphthalmia, are a common feature of Fanconi anemia.[16] However, pathogenic autosomal recessive variants in *BRCA2* and *BRIP1*, which cause Fanconi anemia complementation group D1 (FANCD1; MIM #605724) and J (FANCJ; MIM #609054), respectively, have not been previously reported in individuals with MAC. In this study, subject S10 had iris coloboma and biparental c.[4965C>G];[7007G>C], p.[(Y1655*)];[(R2336P)] pathogenic variants in *BRCA2*, and subject S11 has microphthalmia and biparental c.[2392C>T];[2392C>T]; p.[(R798*)];[(R798*)] pathogenic variants in *BRIP1*. This suggests that biallelic pathogenic variants in *BRCA2* and *BRIP1* can cause MAC in the setting of Fanconi anemia.

KAT6A and KAT6B

KAT6A and *KAT6B* encode lysine (K) acetyltransferases that have been shown to play a critical role in transcriptional regulation and various developmental processes.[17] Heterozygous variants in *KAT6A* cause Arboleda–Tham syndrome (MIM #616268). Heterozygous variants in *KAT6B* cause genitopatellar syndrome (MIM #606170) and Ohdo syndrome, SBBYS variant (MIM #603736) that have overlapping features. Subject S38 has a left-sided coloboma of the optic nerve with visual impairment and phenotypes consistent with Arboleda–Tham syndrome caused by a de novo pathogenic c.3385C>T, p.(R1129*) [NM_006766.3] variant in *KAT6A*. Microphthalmia has been previously reported in a man with phenotypes consistent with Arboleda–Tham syndrome who carried a c.5924A>G, p.(N1975S) [NM_006766.3] variant in *KAT6A* that was passed to his affected daughter, who did not have MAC.[18]

Subject S39 carries a de novo pathogenic c.5646del, p.(N1883Tfs*2) variant in *KAT6B* and has microphthalmia and phenotypes consistent with Ohdo syndrome, SBBYS variant. Although we could not identify additional individuals with MAC who have *KAT6B*-related disorders, we note that mice who are homozygous for a hypomorphic *Kat6b* allele that results in a 90% reduction in *Kat6b* mRNA have microphthalmia.[19,20]

BRPF1 is a chromatin reader that binds to and activates KAT6A and KAT6B.[21] Heterozygous loss-of-function *BRPF1* variants cause intellectual developmental disorder with dysmorphic facies and ptosis (IDDDFP; MIM #617333). Recently, two individuals with IDDDFP and MAC have been described—one with microphthalmia and coloboma who carried a de novo c.1756_1757insT, p.(Glu586Valfs*12) [NM_001003694.1] *BRPF1* variant and one with coloboma who carried a c.655G>T, p.(Glu219*) [NM_001003694.1] *BRPF1* variant.[21,22] Taken together, these data suggest that variants affecting the BRPF1-KAT6A/KAT6B complex can cause microphthalmia and coloboma.

NSF

N-ethylmaleimide-sensitive factor (*NSF*) is a homohexameric AAA ATPase that plays a role in intracellular vesicle transport, membrane fusion, and synaptic transmissions.[23,24] In 2019, Suzuki et al.[24] reported two Japanese girls who carried missense variants in *NSF* whose phenotypes included microcephaly, hypotonia, early-onset seizures, global developmental delay, and respiratory insufficiency. This disorder was subsequently named developmental and epileptic encephalopathy 96 (DEE96; MIM #619340). Subject S54 is a 5-month-old male who carries the same de novo c.1688C>T, p.(Pro563Leu) [NM_006178.4] seen in patient 2 from Suzuki et al.[24] He represents the third individual reported with DEE96. His phenotypes include microphthalmia, microcornea, seizures, abnormal brain architecture, cerebral hypomyelination, seizures, respiratory failure, micropenis, and dysmorphic features.

Suzuki et al.[24] also reported that expression of wild-type *NSF* in the *Drosophila* eye using the *GAL4-UAS* system did not result in any detectable phenotype, but expression of *NSF* carrying the two pathogenic variants associated with DEE96 resulted in eye defects with the p.(Pro563Leu) variant causing complete obliteration of the eyes associated with high levels of cell death being observed in the developing eye discs.[24] Ectopic expression of the mutated *NSF* genes under the control of the pan-neuronal *nSyb-GAL4* driver resulted in embryonic or first instar larval lethality. Suzuki et al.[24] concluded that the *NSF* variants that cause autosomal dominant DEE96 generate dominant negative alleles, although a different gain-of-function mechanism was not excluded.

The International Mouse Phenotyping Consortium (IMPC) has reported that $Nsf^{+/-}$ mice have abnormal retinal morphology, but MAC was not reported, and complete loss of *Nsf* causes embryonic lethality.[25,26] However, Hanovice et al.[27] have reported that zebrafish with variants in *nsfb*—one of two *nsf* genes in zebrafish—have mild microphthalmia, retinal pigment epithelium hypopigmentation, and lens defects.

These findings, combined with *NSF*'s high MAC-specific rank annotation score of 98.8%, lead us to conclude that MAC-related phenotypes may be caused by deleterious variants in *NSF*.

RAC1

RAC1 encodes a RHO GTPase involved in modulation of the cytoskeleton and plays a role in phagocytosis, mesenchymal-like migration, neuronal polarization, axonal growth, adhesion, differentiation, cellular growth, and cell cycle regulation.[28] Pathogenic variants in *RAC1* cause intellectual development disorder, autosomal dominant 48, whose cardinal features include global developmental delay, intellectual disability, structural brain anomalies, and dysmorphic facial features (MIM #617751). Subject S73 carries a de novo likely pathogenic c.198A>T, p.(R66S) [NM_018890.4] variant in *RAC1*. She has a chorioretinal coloboma, delayed motor milestones, delayed speech, intellectual disability, hypotonia, coarse dysmorphic features, short stature, hyperextensibility, and failure to thrive. The IMPC has reported that $Rac1^{+/-}$ mice have microphthalmia and anophthalmia.[26] This, combined with *RAC1*'s high MAC-specific rank annotation score of 80.5%, leads us to conclude that MAC-related phenotypes can be seen in individuals with deleterious *RAC1* variants.

SMARCA4

SMARCA4 encodes a catalytic subunit of SWI/SNF complexes that play a role in regulating gene expression and has been shown to associate with the Fanconi anemia–related gene *FANCA* in the nucleus.[29,30] Heterozygous vari-

ants in *SMARCA4* have been shown to cause Coffin–Siris syndrome 4 (MIM #614609), a neurodevelopmental syndrome. Subject S82 carries a de novo likely pathogenic c.2738C>T, p.(P913L) [NM_001128849.1] *SMARCA4* variant and has optic nerve coloboma, delayed motor milestones, hypotonia, agenesis of the corpus callosum, and other phenotypes commonly seen in individuals with Coffin–Siris syndrome 4. Errichiello et al.[31] reported an individual with microphthalmia and mild Coffin–Siris syndrome 4 who carried a c.2935C>T, p.(Arg979*) loss-of-function variant in *SMARCA4*.

SMARCA4 is expressed in the developing eye, and the NHLBI Cardiovascular Development Consortium, Bench to Bassinet Program (B2B/CvDC; https://benchtobassinet.com/) has reported that mice that are homozygous for a c.2381C>T, p.(Thr794Ile) [NM_011417] *SmarcA4* variant have anophthalmia.[32] Similarly, smarca4a$^{a50/a50}$ (*yng*) zebrafish also have small eyes.[33,34] Taken together, these findings lead us to conclude that the MAC-related phenotype can be seen in individuals with deleterious *SMARCA4* variants.

SMC1A

Cornelia de Lange syndrome is a multiple congenital anomaly syndrome caused by variants in at least six genes.[35] A variety of ophthalmologic phenotypes have been described in individuals with Cornelia de Lange syndrome, including microphthalmia and coloboma.[35,36] However, MAC has not been specifically described in individuals with Cornelia de Lange syndrome 2, which is caused by variants in *SMC1A*, a gene on the X chromosome. Subject S83 is a male with a right-sided coloboma who carries a maternally inherited c.1114-2A>G, p.(?) [NM_006306.4] pathogenic variant in *SMC1A*. Subject S84 is a male with bilateral iris colobomas who carries a maternally inherited c.1829A>G, p.(Q610R) [NM_006306.4] VUS in *SMC1A*. This suggests that deleterious variants in *SMC1A* can cause MAC.

TUBA1A

TUBA1A encodes for α-tubulin, a major component of microtubules within the eukaryotic cytoskeleton. Mutations in *TUBA1A* have been shown to cause lissencephaly 3 (MIM #611603), a neurodevelopmental syndrome characterized by structural brain malformations, congenital microcephaly, developmental delay, intellectual disability, a lack of language development, and diplegia/tetraplegia.[37] Subject S89 carries a de novo pathogenic c.641G>A, p.(R214H) [NM_006009.3] variant in *TUBA1A*. Her phenotypes include coloboma, morning glory disc anomaly, epileptic encephalopathy, delayed speech and language development, profound motor delay, ventriculomegaly, nystagmus, epileptic encephalopathy, photosensitivity, short stature, 11 ribs, and bilateral sensorineural hearing impairment. Oegema et al.[38] reported a 4-year-old girl with bilateral coloboma who carried a c.641G>A, p.(Arg214His) [NM_006009.3] *TUBA1A* variant, and Myers et al.[39] reported a boy with microphthalmia and cataracts who carried a c.808G>T, p.(Ala270Ser) [NM_006009.3] *TUBA1A* variant.

In zebrafish, *tuba1a* is essential for the development of anterior structures, including the brain and retina.[40] Although morpholino *tuba1a* knockdown could be partially rescued with the addition of *tuba1a* mRNA designed to escape morpholino knockdown, the resulting fish at 52 hours postfertilization still had small eyes compared to controls.[40] *TUBA1A* is expressed in the developing eye. These results, combined with *TUBA1A*'s high MAC-specific rank annotation score of 80.1%, lead us to conclude that MAC-related phenotypes can be seen in individuals with deleterious *TUBA1A* variants.

Other MAC Candidate Genes

There is currently insufficient evidence to conclude that deleterious variants in *AFF4*, *ANK3*, *BRAT1*, *KIF2A*, *LRP4*, *TCTN3*, *RP1L1*, *SKI*, *SLC12A5*, and *THOC2* are associated with the development of MAC (Table 2). However, it is possible that some or all these genes will ultimately be shown to be bona fide MAC genes. In particular, we note that *LRP4*, *TCTN3*, *RPIL1*, *SKI*, and *SLC12A5* are all expressed in the developing mammalian eye, their zebrafish and/or mouse homologs have been associated with the development of MAC, and they have high MAC-specific rank annotation scores (≥80%).

Conclusions

We conclude that cES is an effective means of identifying a molecular diagnosis in individuals with nonisolated MAC and may identify putatively damaging variants that would be missed if only a clinically available ophthalmologic gene panel was obtained. Our data also suggest that deleterious variants in *BRCA2*, *BRIP1*, *KAT6A*, *KAT6B*, *NSF*, *RAC1*, *SMARCA4*, *SMC1A*, and *TUBA1A* can contribute to the development of MAC.

Acknowledgments

Supported, in part, by National Institutes of Health/Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant R01HD098458 (DAS) and Grant R01HD093660 (AJA and PJL); National Institutes of Health/National Eye Institute Grant U01EY032403 (PJL and BPB); the National Cancer Institute Grant T32CA190194, the Foundation for Barnes-Jewish Hospital, and the Siteman Cancer Center (BMG); and the intramural program of the National Eye Institute, National Institutes of Health.

The Department of Molecular & Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing completed at Baylor Genetics.

Disclosure: **B. Kunisetty**, None; **B.A. Martin-Giacalone**, None; **X. Zhao**, None; **P.N. Luna**, None; **B.P. Brooks**, None; **R.B. Hufnagel**, None; **C.A. Shaw**, None; **J.A. Rosenfeld**, None; **A.J. Agopian**, None; **P.J. Lupo**, None; **D.A. Scott**, None

References

1. Hornby SJ, Gilbert CE, Rahi JK, et al. Regional variation in blindness in children due to microphthalmos, anophthalmos and coloboma. *Ophthalmic Epidemiol.* 2000;7:127–138.
2. Gilbert C, Foster A. Childhood blindness in the context of VISION 2020—the right to sight. *Bull World Health Organ.* 2001;79:227–232.
3. Skalicky SE, White AJ, Grigg JR, et al. Microphthalmia, anophthalmia, and coloboma and associated ocular and systemic features: understanding the spectrum. *JAMA Ophthalmol.* 2013;131:1517–1524.
4. Shah SP, Taylor AE, Sowden JC, et al. Anophthalmos, microphthalmos, and coloboma in the United Kingdom:

clinical features, results of investigations, and early management. *Ophthalmology.* 2012;119:362–368.
5. Warburg M. Classification of microphthalmos and coloboma. *J Med Genet.* 1993;30:664–669.
6. Chassaing N, Golzio C, Odent S, et al. Phenotypic spectrum of STRA6 mutations: from Matthew-Wood syndrome to nonlethal anophthalmia. *Hum Mutat.* 2009;30:E673–681.
7. Li J, Yang W, Wang YJ, et al. Exome sequencing identifies genetic variants in anophthalmia and microphthalmia. *Am J Med Genet A.* 2022;188:2376–2388.
8. Haug P, Koller S, Maggi J, et al. Whole exome sequencing in coloboma/microphthalmia: identification of novel and recurrent variants in seven genes. *Genes (Basel).* 2021;12:65.
9. Saudi Mendeliome G. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol.* 2015;16:134.
10. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–424.
11. Scott TM, Campbell IM, Hernandez-Garcia A, et al. Clinical exome sequencing data reveal high diagnostic yields for congenital diaphragmatic hernia plus (CDH+) and new phenotypic expansions involving CDH. *J Med Genet.* 2021;59:270–278.
12. Huth EA, Zhao X, Owen N, et al. Clinical exome sequencing efficacy and phenotypic expansions involving anomalous pulmonary venous return. *Eur J Hum Genet.* 2023;12:1430–1439.
13. Campbell IM, Rao M, Arredondo SD, et al. Fusion of large-scale genomic knowledge and frequency data computationally prioritizes variants in epilepsy. *PLoS Genet.* 2013;9:e1003797.
14. Callaway DA, Campbell IM, Stover SR, et al. Prioritization of candidate genes for congenital diaphragmatic hernia in a critical region on chromosome 4p16 using a machine-learning algorithm. *J Pediatr Genet.* 2018;7:164–173.
15. ACMG Board of Directors. Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. *Genet Med.* 2015;17:505–507.
16. Graf CM, Nichele S, Siviero RB, et al. Ocular manifestations in patients with fanconi anemia: a single-center experience including 106 patients. *J Pediatr.* 2022;242:228–234.e221.
17. Wiesel-Motiuk N, Assaraf YG. The key roles of the lysine acetyltransferases KAT6A and KAT6B in physiology and pathology. *Drug Resist Updat.* 2020;53:100729.
18. Trinh J, Huning I, Yuksel Z, et al. A KAT6A variant in a family with autosomal dominantly inherited microcephaly and developmental delay. *J Hum Genet.* 2018;63:997–1001.
19. Thomas T, Voss AK, Chowdhury K, Gruss P. Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development. *Development.* 2000;127:2537–2548.
20. Kraft M, Cirstea IC, Voss AK, et al. Disruption of the histone acetyltransferase MYST4 leads to a Noonan syndrome-like phenotype and hyperactivated MAPK signaling in humans and mice. *J Clin Invest.* 2011;121:3479–3491.
21. Zu G, Liu Y, Cao J, Zhao B, Zhang H, You L. BRPF1-KAT6A/KAT6B complex: molecular structure, biological function and human disease. *Cancers (Basel).* 2022;14:4068.
22. Demeulenaere S, Beysen D, De Veuster I, Reyniers E, Kooy F, Meuwissen M. Novel BRPF1 mutation in a boy with intellectual disability, coloboma, facial nerve palsy and hypoplasia of the corpus callosum. *Eur J Med Genet.* 2019;62:103691.
23. Glick BS, Rothman JE. Possible role for fatty acyl-coenzyme A in intracellular protein transport. *Nature.* 1987;326:309–312.
24. Suzuki H, Yoshida T, Morisada N, et al. De novo NSF mutations cause early infantile epileptic encephalopathy. *Ann Clin Transl Neurol.* 2019;6:2334–2339.
25. Xie MJ, Iwata K, Ishikawa Y, et al. Autistic-like behavior and impairment of serotonin transporter and AMPA receptor trafficking in N-ethylmaleimide sensitive factor gene-deficient mice. *Front Genet.* 2021;12:748627.
26. Dickinson ME, Flenniken AM, Ji X, et al. High-throughput discovery of novel developmental phenotypes. *Nature.* 2016;537:508–514.
27. Hanovice NJ, Daly CM, Gross JM. N-ethylmaleimide-sensitive factor b (nsfb) is required for normal pigmentation of the zebrafish retinal pigment epithelium. *Invest Ophthalmol Vis Sci.* 2015;56:7535–7544.
28. Reijnders MRF, Ansor NM, Kousi M, et al. RAC1 missense mutations in developmental disorders with diverse phenotypes. *Am J Hum Genet.* 2017;101:466–477.
29. Khavari PA, Peterson CL, Tamkun JW, Mendel DB, Crabtree GR. BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature.* 1993;366:170–174.
30. Otsuki T, Furukawa Y, Ikeda K, et al. Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex. *Hum Mol Genet.* 2001;10:2651–2660.
31. Errichiello E, Mustafa N, Vetro A, et al. SMARCA4 inactivating mutations cause concomitant Coffin-Siris syndrome, microphthalmia and small-cell carcinoma of the ovary hypercalcaemic type. *J Pathol.* 2017;243:9–15.
32. Blake JA, Bult CJ, Kadin JA, Richardson JE, Eppig JT; Mouse Genome Database Group. The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. *Nucleic Acids Res.* 2011;39:D842–848.
33. Leung YF, Ma P, Link BA, Dowling JE. Factorial microarray analysis of zebrafish retinal development. *Proc Natl Acad Scie USA.* 2008;105:12909–12914.
34. Link BA, Fadool JM, Malicki J, Dowling JE. The zebrafish young mutation acts non-cell-autonomously to uncouple differentiation from specification for all retinal cells. *Development.* 2000;127:2177–2188.
35. Deardorff MA, Noon SE, Krantz ID. Cornelia de Lange syndrome. In: Adam MP, Feldman J, Mirzaa GM, et al., eds. *GeneReviews*. Seattle, WA: University of Washington, Seattle; 2005:1993–2024.
36. Nallasamy S, Kherani F, Yaeger D, et al. Ophthalmologic findings in Cornelia de Lange syndrome: a genotype-phenotype correlation study. *Arch Ophthalmol.* 2006;124:552–557.
37. Bahi-Buisson N, Poirier K, Boddaert N, et al. Refinement of cortical dysgeneses spectrum associated with TUBA1A mutations. *J Med Genet.* 2008;45:647–653.
38. Oegema R, Cushion TD, Phelps IG, et al. Recognizable cerebellar dysplasia associated with mutations in multiple tubulin genes. *Hum Mol Genet.* 2015;24:5313–5325.
39. Myers KA, Bello-Espinosa LE, Kherani A, Wei XC, Innes AM. TUBA1A mutation associated with eye abnormalities in addition to brain malformation. *Pediatr Neurol.* 2015;53:442–444.
40. Veldman MB, Bemben MA, Goldman D. Tuba1a gene expression is regulated by KLF6/7 and is necessary for CNS development and regeneration in zebrafish. *Mol Cell Neurosci.* 2010;43:370–383.
41. Owen N, Toms M, Young RM, et al. Identification of 4 novel human ocular coloboma genes ANK3, BMPR1B, PDGFRA, and CDH4 through evolutionary conserved vertebrate gene analysis. *Genet Med.* 2022;24:1073–1084.

Investigative Ophthalmology & Visual Science

42. Vercellino F, Valerio M, Dusio MP, Spano A, D'Alfonso S. BRAT1 mutation retrospective diagnosis: a case report. *Cureus.* 2023;15:e35655.
43. Tian G, Cristancho AG, Dubbs HA, Liu GT, Cowan NJ, Goldberg EM. A patient with lissencephaly, developmental delay, and infantile spasms, due to de novo heterozygous mutation of KIF2A. *Mol Genet Genomic Med.* 2016;4:599–603.
44. Tanahashi H, Suzuki T. Deletion of Lrp4 increases the incidence of microphthalmia. *Biophysl Res Commun.* 2018;506:478–484.
45. Thomas S, Legendre M, Saunier S, et al. TCTN3 mutations cause Mohr-Majewski syndrome. *Am J Hum Genet.* 2012;91:372–378.
46. Davidson AE, Sergouniotis PI, Mackay DS, et al. RP1L1 variants are associated with a spectrum of inherited retinal diseases including retinitis pigmentosa and occult macular dystrophy. *Hum Mutat.* 2013;34:506–514.
47. Liu YP, Bosch DG, Siemiatkowska AM, et al. Putative digenic inheritance of heterozygous RP1L1 and C2orf71 null mutations in syndromic retinal dystrophy. *Ophthalmic Genet.* 2017;38:127–132.
48. Traboulsi E, Gupta P, Colmenares C. Ocular abnormalities in ski-deficient mice. *Invest Ophthalmol Vis Sci.* 2002;43:1998–1998.
49. Kumar R, Palmer E, Gardner AE, et al. Expanding clinical presentations due to variations in THOC2 mRNA nuclear export factor. *Front Mol Neurosci.* 2020;13:12.
50. Sadler KC, Amsterdam A, Soroka C, Boyer J, Hopkins N. A genetic screen in zebrafish identifies the mutants vps18, nf2 and foie gras as models of liver disease. *Development.* 2005;132:3561–3572.

Investigative Ophthalmology & Visual Science