AMERICAN ACADEMY OF OPHTHALMOLOGY®

AlphaMissense Predictions and ClinVar Annotations: A Deep Learning Approach to Uveal Melanoma

David J. Taylor Gonzalez, MD,[1] Mak B. Djulbegovic, MD, MSc,[2] Meghan Sharma, MD, MPH,[1] Michael Antonietti, BS,[1] Colin K. Kim, BS,[1] Vladimir N. Uversky, PhD, DSc,[3] Carol L. Karp, MD,[1] Carol L. Shields, MD,[4] Matthew W. Wilson, MD[5]

Objective: Uveal melanoma (UM) poses significant diagnostic and prognostic challenges due to its variable genetic landscape. We explore the use of a novel deep learning tool to assess the functional impact of genetic mutations in UM.

Design: A cross-sectional bioinformatics exploratory data analysis of genetic mutations from UM cases.

Subjects: Genetic data from patients diagnosed with UM were analyzed, explicitly focusing on missense mutations sourced from the Catalogue of Somatic Mutations in Cancer (COSMIC) database.

Methods: We identified missense mutations frequently observed in UM using the COSMIC database, assessed their potential pathogenicity using AlphaMissense, and visualized mutations using AlphaFold. Clinical significance was cross-validated with entries in the ClinVar database.

Main Outcome Measures: The primary outcomes measured were the agreement rates between AlphaMissense predictions and ClinVar annotations regarding the pathogenicity of mutations in critical genes associated with UM, such as *GNAQ, GNA11, SF3B1, EIF1AX*, and *BAP1*.

Results: Missense substitutions comprised 91.35% (n = 1310) of mutations in UM found on COSMIC. Of the 151 unique missense mutations analyzed in the most frequently mutated genes, only 40.4% (n = 61) had corresponding data in ClinVar. Notably, AlphaMissense provided definitive classifications for 27.2% (n = 41) of the mutations, which were labeled as "unknown significance" in ClinVar, underscoring its potential to offer more clarity in ambiguous cases. When excluding these mutations of uncertain significance, AlphaMissense showed perfect agreement (100%) with ClinVar across all analyzed genes, demonstrating no discrepancies where a mutation predicted as "pathogenic" was classified as "benign" or vice versa.

Conclusions: Integrating deep learning through AlphaMissense offers a promising approach to understanding the mutational landscape of UM. Our methodology holds the potential to improve genomic diagnostics and inform the development of personalized treatment strategies for UM.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2025;5:100673 Published by Elsevier on behalf of the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.ophthalmologyscience.org.

Uveal melanoma (UM) is the most common primary intraocular cancer among adults, with an age-adjusted incidence rate of 5.1 per million individuals.[1,2] Originating from melanocytes within various segments of the uvea, including the choroid, ciliary body, and iris, UM accounts for 85% of all ocular melanomas.[2] The development of this malignancy is influenced by factors such as light skin and iris color, along with a susceptibility to sunburn.[3,4] Diagnosis often occurs incidentally during routine eye examinations, although >70% of patients may experience visual disturbances like flashes or floaters.[4] The diagnostic process integrates a thorough ophthalmic examination with advanced imaging techniques, including fundus photography, ultrasound at 15 to 20 MHz, OCT, and fluorescein angiography.[5–7] In certain scenarios, a fine-needle aspiration biopsy may be essential to conclusively diagnose the condition.[4,8]

The clinical course of UM is significantly driven by genetic mutations. Recent advances in genetic medicine have allowed researchers and clinicians to understand the molecular basis of UM, highlighting the role of specific mutations in dictating the behavior of this malignancy.[9] Gene expression profiling has significantly advanced the prognostic landscape of UM by analyzing the expression levels of specific genes within the tumor, enabling a classification based on metastatic risk. Gene expression profiling classifies tumors into class 1 (low metastatic risk) and class 2 (high metastatic risk).[9] Notably, mutations in

https://doi.org/10.1016/j.xops.2024.100673
ISSN 2666-9145/25

BAP1 have been closely associated with the more aggressive class 2 UMs, with BAP1 mutations specifically linked to heightened metastatic risk.[9–12] Recent findings from the Collaborative Ocular Oncology Group indicate that the Preferentially Expressed Antigen in Melanoma expression status, when combined with 15-gene expression profiling, enhances prognostic accuracy for both class 1 and 2 UMs.[13] Specifically, Preferentially Expressed Antigen in Melanoma–positive tumors in both classes are associated with a higher risk of metastasis and poorer survival outcomes, further stratifying these classes based on metastatic risk.[13]

In contrast, The Cancer Genome Atlas, a worldwide effort developed by the National Cancer Institute and the National Human Genome Research Institute, offers a robust method for classifying UM by identifying chromosomal changes and somatic mutations to classify UM.[14,15] This classification delineates 4 prognostic types based on chromosomal alterations rather than gene expression levels. Type A is characterized by chromosome 3 and 8 disomy and often features a mutation in EIF1AX, which encodes the eukaryotic translation initiation factor 1A X-linked.[16,17] Type B also shows disomy of chromosome 3 but with a gain in chromosome 8q and may harbor mutations in SF3B1, encoding the splicing factor 3B subunit 1A.[16–18] Both types A and B are typically associated with a favorable prognosis, with type B carrying a slightly higher risk than type A. Conversely, types C and D represent more aggressive forms of UM, associated with higher metastatic risk and poorer outcomes.[14,16] Type C is identified by either chromosome 3 monosomy or gain in chromosome 8q, and it is often associated with inactivation of BAP1 (which encodes BRCA1-associated protein 1 located on chromosome 3).[14,16] Type D escalates in severity with chromosome 3 monosomy and multiple gains in chromosome 8q. This most aggressive form of UM is associated with an inflammatory phenotype.[12,14,16] The complexity of UM's genetic landscape highlights the crucial role of genetic analysis in enhancing our understanding of the disease and paving the way for the development of targeted therapeutic strategies based on mutation-driven pathogenicity.

Computational biology and deep learning technologies have emerged as pivotal tools to deepen our understanding of the genetic landscape of UM. AlphaFold, developed by Google's DeepMind, represents a cornerstone in this evolution, accurately predicting protein structures on a proteome-wide scale.[19] Building on this foundation, DeepMind developed AlphaMissense, a tool designed to assess the functional impact of missense mutations on proteins by leveraging the structural insights provided by AlphaFold.[20] This advancement allows researchers to visualize protein structures with unprecedented accuracy and to predict how mutations within these structures might influence their function. AlphaMissense adapts AlphaFold's predictive capabilities, fine-tuned on human and primate variant population frequency databases, to forecast the pathogenicity of missense mutations as it relates to protein function.[20] Integrating AlphaMissense into genomic research on UM offers a leap forward, enhancing our capacity to discern the functional and clinical ramifications of missense mutations.

The primary aim of our research is to evaluate the capabilities of AlphaMissense in conjunction with comprehensive genomic databases such as the Catalogue of Somatic Mutations in Cancer (COSMIC) and ClinVar, a free public archive that contains reports detailing the relationships between human genetic variations and phenotypes, along with supporting evidence.[21,22] Our first objective is to predict the pathogenicity of missense mutations observed in UM with a deep learning analysis via AlphaMissense, which involves predicting the functional impacts of amino acid substitutions on protein behavior, a key factor in the UM's progression and response to treatment. Our second objective is to compare these computational predictions with the clinical annotations provided by the ClinVar database to evaluate the clinical relevance of our findings, ensuring that the mutations identified as pathogenic indeed correlate with observed patient outcomes. Through these objectives, our study aims to deepen the understanding of the genetics of UM, possibly paving the way for more accurate diagnostic and prognostic tools in the future.

Methods

We conducted a comprehensive examination of genetic mutations associated with UM using the COSMIC database (v99, released 11/28/23).[22] The COSMIC database serves as a resource for researchers and clinicians by offering access to a vast repository of data derived from scientific literature and large-scale studies, thereby facilitating a deeper understanding of cancer genomics and aiding translational research efforts.[22] Our approach involved the COSMIC Cancer Browser tool, specifically focusing on the "Eye" tissue category and the "Uveal tract" subtissue, targeting "Malignant melanoma" with the subhistology filters set to "Include All," including "Epithelioid," "Spindle," "Mixed," "Ring," and "NS" or "Not Specified" (accessed on 03/10/24, https://cancer.sanger.ac.uk/cosmic). Through this approach, we retrieved comprehensive datasets of gene mutations related to UM from COSMIC, integrating both targeted sequencing and whole-genome sequencing data. We then assessed the most frequently occurring genetic abnormalities, including a thorough analysis of gene alterations, mutation types, and their specific loci within UM samples.

We utilized AlphaMissense to assess the pathogenicity of missense mutations in genes associated with UM found on COSMIC. The tool is a deep learning model that predicts the functional impact of amino acid substitutions on protein stability across the human proteome. Utilizing its published database of 216 million missense protein variant predictions, we assessed all potential amino acid substitutions within our genes of interest.[20] AlphaMissense calculates a pathogenicity score ranging from 0 to 1, where scores nearer to 1 suggest a higher probability of a variant being pathogenic and values closer to 0 indicate benign mutations.[20] Predictions with scores in the intermediate range (0.2–0.8) are considered less reliable and ambiguous.[20] We adhered to AlphaMissense's predefined cutoffs that were validated to achieve 90% precision (positive predictive value) on ClinVar data according to the original AlphaMissense publication, classifying mutations as pathogenic, ambiguous, or benign.[20] Pathogenic variants compromise protein function and organismal fitness, whereas benign variants generally have minimal impact.

We calculated the average pathogenicity score for each gene to provide a holistic view of each gene's susceptibility to mutations with potential pathogenic effects, aiding in the identification of mutations that may play crucial roles in the pathology of UM. We averaged the scores assigned by AlphaMissense to all possible amino acid substitutions within the gene. This preliminary analysis offered an initial understanding of the potential pathogenicity of mutations across the genome. For visual representation of pathogenic mutations, we processed structures predicted by AlphaFold2, a state-of-the-art deep learning model developed by Google's DeepMind that predicts the 3-dimensional structures of proteins with high accuracy.[19] These structures, obtained from the Universal Protein Resource (UniProt), were visualized using ChimeraX, and each residue was colored according to their average pathogenicity scores as predicted by AlphaMissense.[23–26] The structures were colored on a spectrum from blue (score of 0, indicating nonpathogenic) through white (score of 0.5, ambiguous) to red (score of 1, indicating pathogenic).

To substantiate our findings, we utilized the ClinVar database, a resource maintained by the National Institutes of Health that provides a public archive of human genetic variants along with detailed interpretations of their significance concerning various diseases.[21] ClinVar is widely recognized for its comprehensive collection of variant annotations. These interpretations are based on multiple lines of evidence, including clinical testing results, research studies, population data, and expert guidelines, making ClinVar a robust resource for variant classification.[21] Submissions to ClinVar are curated, with each record assigned a review status reflecting the level of consensus and quality of the evidence provided.[21] This diversity of sources offers valuable insights into variant–disease relationships and classification. Despite its recognized utility, ClinVar does have inherent limitations, such as variability in the quality of annotations and a significant number of variants classified as "uncertain significance" due to limited or conflicting evidence. We used ClinVar to compare our AlphaMissense-derived predictions with clinically recognized annotations of genetic variants, particularly to corroborate the clinical relevance of the mutations identified in UM through the COSMIC database (https://www.ncbi.nlm.nih.gov/clinvar/, accessed on 03/12/24). While alternative resources such as the Human Gene Mutation Database and the Leiden Open Variation Database are also available, these options face similar challenges in providing consistent and comprehensive clinical annotations. ClinVar was chosen for this study because it represents one of the most widely accepted and utilized repositories for variant interpretation, offering clinically relevant and expert-curated data that are freely accessible and align well with our research objectives. This methodology reinforced the accuracy and relevance of our mutation analysis in the broader context of genetic research and patient care.

This study was conducted in accordance with the ethical standards set forth in the Declaration of Helsinki. Given that the data analyzed were publicly available and fully deidentified, formal approval from an institutional review board or ethics committee was not required. Furthermore, as no patient-identifiable information was accessed and there was no direct involvement of human subjects, the need for informed consent was waived.

Results

In our study, we detailed genetic mutations in UM by leveraging the comprehensive resources of the COSMIC database (version 99). We observed that among the 1434 unique UM samples identified, a variety of mutation types were observed (Table 1). Missense substitutions emerged as the most prevalent, accounting for 91.35% (1310 samples) of the mutations. Other mutations included frameshift deletions (8.93%, 128 samples), nonsense substitutions (5.79%, 83 samples), synonymous substitutions (1.60%, 23 samples), inframe deletions (3.42%, 49 samples), frameshift insertions (2.16%, 31 samples), inframe insertions (0.56%, 8 samples), complex mutations (0.35%, 5 samples), and mutations that were labeled as "Other" (7.32%, 105 samples). It is worth noting that mutations identified through both targeted sequences and whole-genome sequencing are included, accounting for the differences in the number of samples tested for each gene. Specifically, 195 samples underwent whole-genome sequencing screening; however, we did not stratify this information further in our analysis, maintaining a broad overview of genetic variations within the dataset.

The genes most frequently mutated are displayed in Table 2. The genes *GNAQ* and *GNA11* stood out due to their high mutation frequencies, with 526 of 1429 samples tested showing mutations for *GNA11* (36.8%) and 588 of 1685 UM samples tested showing mutations in *GNAQ* (34.9%). Other genes included *BAP1*, with mutations found in 311 of 887 samples tested (35.1%); *SF3B1*, with mutations in 158 of 942 samples tested (16.8%); and *EIF1AX*, with mutations in 27 of 307 samples tested (8.80%). Our dataset also identified *BRAF* mutations in 42 of 1197 samples tested. However, given the known association of *BRAF* mutations predominantly with cutaneous and conjunctival melanoma and considering the potential for misidentification of melanoma type, we elected to exclude *BRAF* from further analysis.[27]

Following the categorization of genetic mutations in UM using COSMIC, we conducted a detailed pathogenicity prediction assessment using the AlphaMissense tool. The goal of this portion of our analysis was to evaluate the impact of all possible amino acid substitutions within each gene of interest on protein function and stability. The mutations were categorized as pathogenic, ambiguous, or benign (Table 3).

For *GNAQ*, among 359 amino acid positions analyzed, we identified 5242 pathogenic mutations (76.85%), 666

Table 1. Mutation Types and Distribution in Uveal Melanoma Samples

Mutation Type	Number of Samples (%)
Missense substitution	1310 (91.35%)
Frameshift deletion	128 (8.93%)
Nonsense substitution	83 (5.79%)
Inframe deletion	49 (3.42%)
Frameshift insertion	31 (2.16%)
Synonymous substitution	23 (1.60%)
Inframe insertion	8 (0.56%)
Complex mutation	5 (0.35%)
Other	105 (7.32%)
Total unique samples	1434

Summary of the distribution of mutation types observed in uveal melanoma samples from the Catalogue of Somatic Mutations in Cancer database.

Table 2. Frequency of Mutations in Key Genes Associated with Uveal Melanoma

Gene	Mutated Samples	Samples Tested
GNAQ	588	1685
GNA11	526	1429
BAP1	311	887
SF3B1	158	942
BRAF	42	1197
EIF1AX	27	307

The frequency of mutations observed in key genes associated with uveal melanoma was extracted from the Catalogue of Somatic Mutations in Cancer database.

ambiguous mutations (9.76%), and 913 benign mutations (13.39%), resulting in an average pathogenicity score of 0.779. *GNA11* showed a similar distribution across 359 amino acid positions, with 4956 pathogenic mutations (72.66%), 689 ambiguous mutations (10.1%), and 1176 benign mutations (17.24%), and an average pathogenicity score of 0.744. *BAP1*, analyzed across 729 amino acid positions, had 8694 pathogenic mutations (62.77%), 1094 ambiguous mutations (7.9%), and 4063 benign mutations (29.33%), with an average pathogenicity score of 0.666. In *SF3B1*, with the highest number of amino acid positions analyzed (1304), we found 19 525 pathogenic mutations (78.81%), 1998 ambiguous mutations (8.06%), and 3253 benign mutations (13.13%), yielding an average pathogenicity score of 0.800. Lastly, *EIF1AX* analysis covered 144 amino acid positions, revealing 2179 pathogenic mutations (79.64%), 223 ambiguous mutations (8.15%), and 334 benign mutations (12.21%), with an average pathogenicity score of 0.803.

In the visualization of pathogenicity across protein structures, a distinct pattern emerges where core or highly structured regions exhibit higher pathogenicity scores, indicated by the prevalence of red hues. In comparison, regions of disorder or less structured areas tend to display lower pathogenicity scores, as denoted by blue coloring (Fig 1). The structured segments, often comprising α-helices and β-sheets, are predominantly colored red, suggesting that if these regions carry mutations, they could contribute to a higher pathogenic potential. GNAQ (Fig 1A) and GNA11 (Fig 1B) display a similar distribution of pathogenicity, with the majority of their structured domains highlighted in red, corresponding to an average pathogenicity score of 0.779 and 0.744, respectively. In the case of SF3B1 (Fig 1C), a protein with a more complex tertiary structure, there is a significant presence of red within the core regions, aligning with its higher pathogenicity score of 0.800. BAP1 (Fig 1D), despite its lower average score of 0.666, still shows this trend of red within its structured regions. Similarly, EIF1AX (Fig 1E) has a high pathogenicity score of 0.803, with red segments dispersed throughout its structured core. Contrastingly, areas that appear to be flexible or disordered, which often correspond to loops or termini extending from the core structure, are mostly blue, signifying a lower pathogenicity score.

Following the initial mutation type analysis and assessment of pathogenicity scores by quantitative and qualitative measures, we identified all unique mutations found in UM samples from the COSMIC database. Utilizing the predictive capabilities of AlphaMissense, we assessed the pathogenicity of missense substitutions in the genes most frequently mutated in UM (Table 4). Through the COSMIC database, for *GNAQ,* we found 21 unique missense mutations, with 18 (85.71%) predicted to be pathogenic, 2 (9.52%) ambiguous, and 1 (4.76%) benign. *GNA11* presented a distinct pattern, where all 13 unique missense mutations identified were predicted to be pathogenic, highlighting its significant role in disease progression. The *BAP1* gene showed a broader diversity of mutation effects, with 53 of 83 unique missense mutations (63.86%) predicted as pathogenic, 7 (8.43%) as ambiguous, and 23 (27.71%) as benign. In the case of *SF3B1*, of 18 unique missense mutations, 17 (94.44%) were deemed pathogenic and 1 (5.56%) benign. Similarly, *EIF1AX* demonstrated a uniform pathogenicity prediction, with all 16 unique missense mutations identified as pathogenic.

Gene-Specific Mutation Analysis and Clinical Correlations

In the following sections, we detail our findings for each gene analyzed in our study of UM samples from the COSMIC database. This includes a comprehensive overview of unique mutations identified, their classifications via AlphaMissense predictions, and the distribution of these mutations across our samples. We also compare these findings with ClinVar annotations to provide a clear snapshot of the genetic landscape within our dataset.

All information discussed can be found in Supplementary Table S5 (available at www.ophthalmologyscience.org) "Pathogenicity Analysis of Mutations Commonly Mutated in Uveal Melanoma." This table presents the AlphaMissense-predicted pathogenicity for mutations in the

Table 3. Predicted Pathogenicity of All Possible Mutations in Key Genes Associated with Uveal Melanoma by AlphaMissense

Gene	Amino Acid Count	Pathogenic	Ambiguous	Benign	Average Pathogenicity Score
GNAQ	359	5242 (76.85%)	666 (9.76%)	913 (13.39%)	0.779
GNA11	359	4956 (72.66%)	689 (10.1%)	1176 (17.24%)	0.744
BAP1	729	8694 (62.77%)	1094 (7.9%)	4063 (29.33%)	0.666
SF3B1	1304	19 525 (78.81%)	1998 (8.06%)	3253 (13.13%)	0.8
EIF1AX	144	2179 (79.64%)	223 (8.15%)	334 (12.21%)	0.803

Distribution of mutations categorized as pathogenic, ambiguous, or benign for the genes *GNAQ*, *GNA11*, *BAP1*, *SF3B1*, and *EIF1AX*.

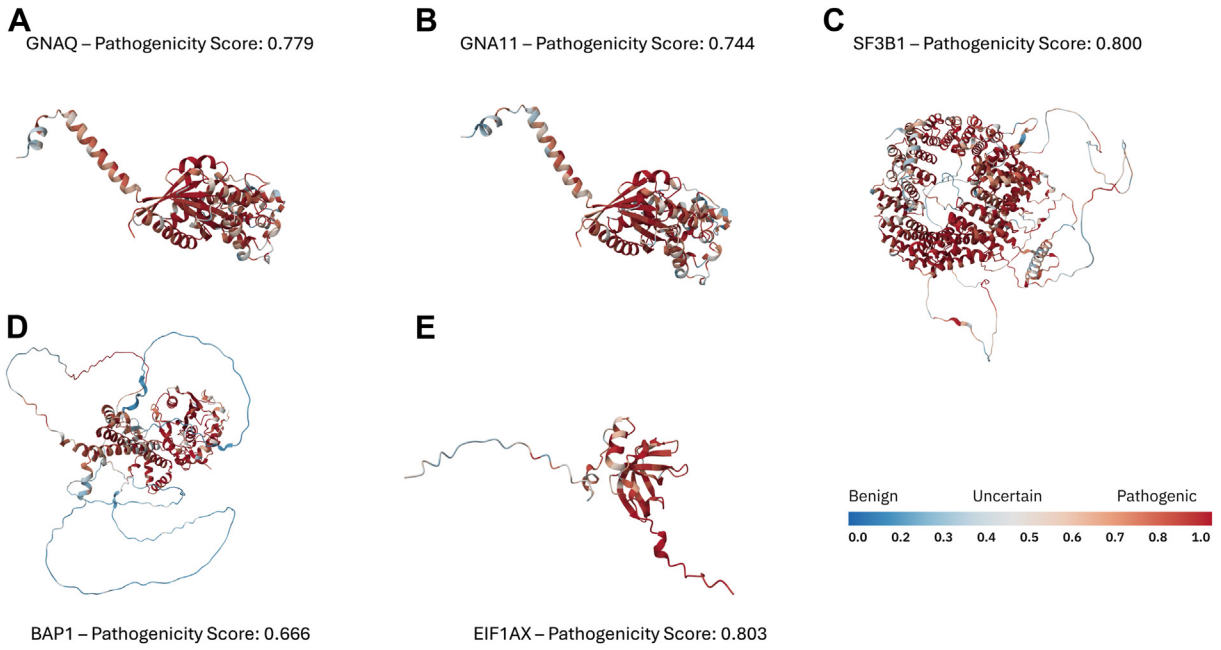

Figure 1. Visualization of pathogenicity through protein structures. This figure displays the AlphaFold2-predicted molecular structures of 5 proteins, each annotated with their average pathogenicity score as predicted by AlphaMissense. (**A**) GNAQ with an average pathogenicity score of 0.779, (**B**) GNA11 with a score of 0.744, (**C**) SF3B1 with a score of 0.800, (**D**) BAP1 with a score of 0.666, and (**E**) EIF1AX with a score of 0.803. Residues are colored according to their individual pathogenicity scores, ranging from blue (benign, score of 0) to red (pathogenic, score of 1), with a gradient through white (ambiguous, score of 0.5) for intermediate scores. The rightmost horizontal bar serves as a color reference, correlating color intensity with the degree of pathogenicity.

GNAQ, GNA11, BAP1, SF3B1, and *EIF1AX* genes, including gene names, amino acid mutations, mutation counts, protein identifiers, AlphaMissense pathogenicity scores, AlphaMissense classifications, and ClinVar classifications. Mutations marked with an asterisk (*) indicate amino acid changes that are unknown or unspecified (e.g., Q209X*).

GNAQ Mutation Profile and Pathogenicity Analysis. In our in-depth analysis of the *GNAQ* gene within UM samples from COSMIC, we identified 21 unique mutations. The most prevalent missense mutations identified were Q209P, found in 328 samples, Q209L in 202 samples, and R183Q in 23 samples, highlighting their significant occurrence within the analyzed dataset. Notably, 1 mutation was labeled as Q209X, indicating a change to an unspecified amino acid, with this variant found in 4 of our samples.

Table 4. Predicted Pathogenicity of Missense Mutations in Key Genes Associated with Uveal Melanoma Samples

Gene Name	Benign	Ambiguous	Pathogenic
GNAQ	1 (4.76%)	2 (9.52%)	18 (85.71%)
GNA11	0 (0.0%)	0 (0.0%)	13 (100.0%)
BAP1	23 (27.71%)	7 (8.43%)	53 (63.86%)
SF3B1	1 (5.56%)	0 (0.0%)	17 (94.44%)
EIF1AX	0 (0.0%)	0 (0.0%)	16 (100.0%)

The predicted pathogenicity of missense mutations identified in uveal melanoma samples from the Catalogue of Somatic Mutations in Cancer database was assessed using AlphaMissense.

Among all unique mutations identified, 18 (85.71%) were predicted to be pathogenic, 2 (9.52%) ambiguous, and 1 (4.76%) benign, as determined by the AlphaMissense predictive analysis.

All 5 missense mutations in our samples with available ClinVar data were consistently identified as pathogenic by both AlphaMissense and ClinVar, affirming the predictive accuracy of AlphaMissense against established clinical data (100%). Three unique mutations, Q209X (4 samples), Q209Y (1 sample), and Q209K (1 sample), were not directly documented in ClinVar. However, mutation sites similar to these, such as Q209H, Q209L, Q209P, Q209R, and Q209L, are recognized by ClinVar as pathogenic sites, suggesting a potential pathogenic nature for these variants.

Interestingly, 1 mutation predicted to be benign, P170S, found in our dataset, lacked corresponding documentation in ClinVar, leaving a gap in our ability to validate its predicted benign nature through external clinical annotations. Similarly, 2 labeled as ambiguous by AlphaMissense, V182I and Q176R, also lacked ClinVar data for further verification, highlighting areas where the predictive model's outcomes could not be directly corroborated with clinical significance annotations.

GNA11 Mutation Profile and Pathogenicity Analysis. Within the *GNA11* gene within UM samples, we identified a total of 13 unique missense mutations. The most prevalent missense mutations identified were Q209L, found in 483 samples; R183C in 24 samples; and Q290P in 6 samples. Among all unique missense mutations, 13 (100%) were predicted to be pathogenic and none were predicted to be ambiguous or benign, as determined by the

AlphaMissense predictive analysis. Similar to *GNAQ*, 1 missense mutation was labeled as Q209X, indicating a change to an unspecified amino acid, with this variant found in 2 of the samples.

All 4 unique missense mutations in our UM COSMIC samples with available ClinVar data were identified as pathogenic by both AlphaMissense and ClinVar (100%). Four unique missense mutations, Q209A, Q209H, Q209V, and Q209X (X = unknown amino acid change), were not directly documented in ClinVar; however, similar to *GNAQ*, these are well-known missense mutational hotspots for this protein, and ClinVar documented other missense mutations such as Q209L and Q209P as pathogenic. Similarly, R1893F had no ClinVar data but was a known pathogenic site according to the annotations, such as R183C. Other mutations, such as R214M, R213Q, E234K, E221D, and E191G, were also not documented in the ClinVar data.

One benign missense mutation, P170S, found in our dataset, lacked corresponding documentation in ClinVar, leaving a gap in our ability to validate its predicted benign nature through external clinical annotations. Similarly, 2 labeled as ambiguous by AlphaMissense, V182I and Q176R, also lacked ClinVar data for further verification, highlighting areas where the predictive model's outcomes could not be directly corroborated with clinical significance annotations.

BAP1 Mutation Profile and Pathogenicity Analysis. In our comprehensive analysis of the *BAP1* gene within our dataset, we observed a total of 83 unique missense mutations. Among these, classifications by AlphaMissense revealed a distinct distribution: 53 (63.86%) were predicted to be pathogenic, 23 (27.71%) benign, and 7 (8.43%) ambiguous. Notably, our analysis further delved into the agreement between AlphaMissense predictions and ClinVar data.

For mutations classified as pathogenic by AlphaMissense, 26 of 27 (96.30%) were also classified as ambiguous in the ClinVar database, and remarkably, only 1 (3.70%) was in agreement as pathogenic. This highlights a significant discrepancy in classification between predictive analysis and established clinical data.

In the benign category, we identified only 1 of 14 (7.14%) mutations predicted as benign by AlphaMissense that was annotated as benign in the ClinVar database, whereas 13 (92.86%) were classified as ambiguous. This indicates a substantial level of uncertainty in the benign classification when compared with clinical annotations. The ambiguous classification by AlphaMissense showed a total of 5 instances, all of which were consistently classified as ambiguous in the ClinVar database (100%), demonstrating a complete alignment in this category.

SF3B1 Mutation Profile and Pathogenicity Analysis. In our investigation of the *SF3B1* gene, we analyzed 18 unique mutations, revealing a predominance of pathogenic classifications according to the AlphaMissense predictive model, with 17 mutations (94.44%) categorized as pathogenic and only 1 mutation (5.56%) classified as benign. Interestingly, our dataset did not include any mutations classified as ambiguous by AlphaMissense. The most prevalent mutations identified were R625H, found in 74 samples; R625C in 56; and R625L, found in 9. We note the presence of ClinVar data for 6 of these mutations, 4 of them being the most common mutations in GNA11.

For the mutations that did have ClinVar data available, we found a high degree of agreement between the AlphaMissense predictions and ClinVar classifications for pathogenic mutations. Specifically, 5 of the 6 mutations with classification data were confirmed as pathogenic in ClinVar, accounting for 83.33% of these mutations. This indicates a robust alignment between predictive analyses and clinical data for the majority of pathogenic classifications. However, 1 mutation was classified as ambiguous in ClinVar, representing 16.67% of mutations with available data, suggesting a discrepancy in classification for a small subset of mutations.

The solitary benign mutation identified by AlphaMissense did not have corresponding classification data in ClinVar, which precludes a direct comparison of classifications for this mutation and underscores a limitation in our ability to validate the benign prediction through external clinical annotations.

EIF1AX Mutation Profile and Pathogenicity Analysis. During our study on the *EIF1AX* gene, a total of 16 unique mutations were identified, with each of these mutations classified as pathogenic according to the AlphaMissense predictive model. Our dataset revealed a limited intersection with ClinVar data, with only 1 of these mutations having corresponding classification information available.

This solitary mutation with ClinVar data was classified as ambiguous, marking a point of divergence between the predictive classification by AlphaMissense and the clinical annotation provided by ClinVar. Most of the mutations, 15 of 16 (93.75%), lacked ClinVar data, underscoring a significant gap in the external validation available for these genetic variations.

The mutations within the *EIF1AX* gene analyzed in this study did not encompass benign or ambiguous classifications according to the AlphaMissense model, highlighting a unanimous pathogenic prediction across the dataset. The presence of ClinVar data for a singular mutation provides a glimpse into the potential for clinical corroboration or dispute of the predictive classifications, though the overall scarcity of ClinVar annotations for the *EIF1AX* mutations delineates an area ripe for further clinical investigation and annotation efforts.

Performance Analysis of AlphaMissense Compared with ClinVar

Our analysis identified a significant number of mutations that were not present in the ClinVar database. Specifically, of the 151 unique mutations analyzed in the most frequently mutated genes, 40.4% (n = 61) had corresponding data in ClinVar, while 59.6% (n = 90) had none.

Furthermore, out of the predictions made by AlphaMissense, 41 mutations (27.2%) that were classified as either "benign" or "pathogenic" by AlphaMissense were classified as "unknown significance" in ClinVar. Additionally, none of the mutations predicted as ambiguous by AlphaMissense were classified as anything other than ambiguous or "uncertain significance" by ClinVar.

Importantly, no discrepancies were observed between AlphaMissense predictions and ClinVar annotations across all genes analyzed. Specifically, no mutations predicted as "pathogenic" by AlphaMissense were classified as "benign" by ClinVar, and no mutations predicted as "benign" by AlphaMissense were classified as "pathogenic" by ClinVar.

To further evaluate the performance of AlphaMissense compared with ClinVar annotations, we generated confusion matrices for each gene analyzed, as detailed in Table 6. True positives (TPs) refer to mutations that AlphaMissense predicted as "pathogenic" and were also classified as "pathogenic" by ClinVar. True negatives (TNs) represent mutations that AlphaMissense predicted as "benign" and were likewise classified as "benign" by ClinVar. False positives are mutations predicted as "pathogenic" by AlphaMissense but classified as "benign" by ClinVar, while false negatives (FNs) are mutations predicted as "benign" by AlphaMissense but classified as "pathogenic" by ClinVar.

For each gene, sensitivity, specificity, and accuracy were calculated based on these values. Sensitivity was determined by the formula TP/(TP + FN), representing the proportion of actual positives correctly identified by AlphaMissense. Specificity was calculated using TN/(TN + false positive), indicating the proportion of actual negatives correctly identified. Accuracy was calculated as (TP + TN)/(TP + TN + false positive + FN), reflecting the overall correctness of AlphaMissense's predictions.

Across the genes *GNAQ, GNA11, BAP1*, and *SF3B1*, AlphaMissense demonstrated 100% sensitivity, meaning all mutations predicted as "pathogenic" by AlphaMissense were confirmed as "pathogenic" by ClinVar, with no FNs recorded. *BAP1* was the only gene to include a mutation classified as "benign" by AlphaMissense that also had corresponding ClinVar data, thus achieving 100% specificity. For the other genes, specificity was marked as N/A in Table 6 because they had no TNs; in other words, no mutations predicted as "benign" by AlphaMissense were also classified as "benign" by ClinVar. Consequently, the accuracy for *GNAQ, GNA11, BAP1,* and *SF3B1* was 100%, underscoring the predictive reliability of AlphaMissense for these genes.

For *EIF1AX,* 16 unique mutations were identified, all classified as pathogenic by AlphaMissense. However, only 1 mutation had corresponding ClinVar data, which annotated it as ambiguous, so performance metrics such as sensitivity, specificity, and accuracy could not be calculated.

Table 6. Summary of Confusion Matrix Results and Performance Metrics for AlphaMissense Predictions Compared with ClinVar Annotations in Uveal Melanoma Genes

Gene	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
GNAQ	5	0	0	0	100%	N/A	1 (100%)
GNA11	4	0	0	0	100%	N/A	1 (100%)
BAP1	1	1	0	0	100%	100%	1 (100%)
SF3B1	5	0	0	0	100%	N/A	1 (100%)
EIF1AX	0	0	0	0	N/A	N/A	0 (N/A)

FN = false negative; FP = false positive; TN = true negative; TP = true positive.
This table presents a summary of TP, TN, FP, and FN results for each gene analyzed, along with calculated sensitivity, specificity, and accuracy metrics. Only mutations with definitive classifications (i.e., "pathogenic" or "benign") by both AlphaMissense and ClinVar were included. Mutations categorized as "ambiguous" by AlphaMissense or of "uncertain significance" by ClinVar were excluded to ensure a clear evaluation of AlphaMissense predictions.
Sensitivity: calculated as TP/(TP + FN), indicating the proportion of actual pathogenic mutations correctly identified.
Specificity: calculated as TN/(TN + FP), indicating the proportion of actual benign mutations correctly identified.
Accuracy: calculated as (TP + TN)/(TP + TN + FP + FN), representing the overall correctness of predictions.
N/A: not applicable due to the absence of TPs or TNs, making the calculation of sensitivity or specificity impossible for that gene.

Discussion

Our analysis identified key missense mutations in genes recurrently mutated in UM. In our investigation of the genetic underpinnings of UM, we leveraged the comprehensive data of the COSMIC database (version 99) to uncover a detailed mutational landscape of this malignancy. Our analysis highlighted the predominance of missense substitutions, which constituted 91.35% (1310 samples) of mutations across 1434 unique UM samples, underlining the critical role of these mutations in the disease's pathogenesis. This distribution highlighted the predominance of missense mutations in the genetic landscape of UM, prompting a focused analysis of this mutation type due to its significant frequency.

The predictive model provided by AlphaMissense demonstrated high agreement with clinical annotations from ClinVar, indicating its efficacy in distinguishing pathogenic mutations from those that are benign or ambiguous. Furthermore, the study highlighted discrepancies between predicted pathogenicity and existing clinical annotations, suggesting areas for future investigation and potential updates to genetic databases.

Our study assessed key genes frequently mutated in UM, notably *GNAQ* and its paralogue *GNA11*, with mutations present in 588 of 1685 (34.9%) and 526 of 1429 samples tested (36.8%), respectively. Studies have shown that mutations in these 2 genes, which encode guanine nucleotide-binding protein Gα subunits of the Gαq family, occur in 80% to 90% of UM cases, generally in a mutually exclusive pattern, emphasizing their critical role in the disease.[16,28,29] Other significant genes included *BAP1* (311 of 887 samples, 35.1%), *SF3B1* (158 of 942 samples, 16.8%), and *EIF1AX* (27 of 307 samples, 8.8%), pointing toward a complex genetic architecture that underlies UM. The decision to exclude *BRAF* mutations from further analysis, due to their association with other melanoma types, allowed for a focused examination of genes directly implicated in the UM.[27]

Utilizing the deep learning tool AlphaMissense, we performed a pathogenicity assessment of all possible mutations within these genes, categorizing them as pathogenic, ambiguous, or benign. This analysis revealed a high percentage of pathogenic mutations for *GNAQ* (76.85%), *GNA11* (72.66%), *SF3B1* (78.81%), and *EIF1AX* (79.64%),

with slightly lower percentages for *BAP1* (62.77%). These findings underscore the potential of deep learning algorithms in identifying mutations with significant functional impacts. It is important to note that this type of analysis does not inform us about the prevalence or likelihood of these mutations occurring; rather, it provides an average pathogenicity score for the gene as a whole. This approach allows for the exploration of protein regions likely to harbor pathogenic mutations, shedding light on potential targets for future investigation regarding functional changes in specific regions of proteins.

Our analysis also presents a way of visualizing the relationship between protein structure and potential pathogenicity, with color coding serving as a qualitative tool for mapping mutational impacts. The gradient from blue to red across the proteins highlights a trend: mutations within the core structural elements—specifically α-helices and β-sheets—are frequently associated with increased pathogenicity. This pattern suggests that these regions are functionally significant, where mutations might disrupt protein stability or interaction with other biomolecules, potentially leading to disease. Interestingly, not all structured regions follow this pattern strictly; some retain blue or intermediate hues, indicating lower pathogenicity scores. These exceptions may represent regions where the protein has a degree of tolerance to mutations, which could be due to inherent flexibility within the fold or redundant functional roles that can compensate for structural changes. Conversely, regions that are less structured, often on the protein surface or in flexible loops, typically exhibit lower pathogenicity scores, colored blue. It is conceivable that these areas are under less selective pressure, allowing a broader spectrum of amino acid substitutions without leading to adverse effects. Alternatively, these regions may be involved in less critical aspects of protein function or may possess an inherent adaptability, allowing them to maintain function despite mutational changes. The presence of high pathogenicity scores in disordered or unstructured regions, while less common, raises interesting questions. Such anomalies could indicate regions where even minor alterations might have a disproportionate effect on protein function, possibly by affecting regulatory features like posttranslational modifications or transient interaction sites.

Overall, this study underscores the utility of integrating structural predictions with pathogenicity data to interpret the functional consequences of protein mutations. It serves as a starting point for more detailed investigations into specific regions of interest and may inform future mutagenesis experiments aimed at elucidating the precise functional ramifications of these mutations. Understanding the spatial distribution of pathogenic potential within proteins can aid in identifying critical regions relevant to disease, which may become targets for therapeutic intervention.

The gaps identified in the clinical annotations, where 59.6% of the mutations analyzed were not present in the ClinVar database, highlight the limitations of current clinical resources in capturing the full mutational spectrum of UM. This demonstrates the potential of tools like AlphaMissense to fill these gaps by identifying and categorizing mutations that may not yet be documented or understood within clinical settings. By addressing these previously uncharted mutations, AlphaMissense can contribute to a more comprehensive understanding of the genetic landscape, potentially leading to more accurate diagnoses and targeted therapies.

The agreement between AlphaMissense predictions and ClinVar annotations across multiple genes, particularly for *GNAQ, GNA11, BAP1*, and *SF3B1*, showcases the predictive model's accuracy. In our analysis, AlphaMissense demonstrated 100% sensitivity and accuracy across these genes, meaning that every mutation annotated as "pathogenic" or "benign" by ClinVar was predicted correctly by AlphaMissense. This consistent performance across key UM genes reinforces AlphaMissense's utility in accurately predicting the pathogenicity of missense mutations within these critical genetic loci. This remarkable agreement underscores the utility of AlphaMissense in accurately predicting the pathogenicity of missense mutations within these pivotal genes. This alignment extends to *SF3B1*, where 5 of the 6 unique missense mutations in ClinVar data matched AlphaMissense's pathogenic predictions, except for one classified as ambiguous in ClinVar. While this high level of agreement is impressive, it is important to note that the absence of TNs for most of these genes, excluding *BAP1*, limits the calculation of specificity and highlights a need for broader clinical data to validate these predictions comprehensively.

For *EIF1AX*, while all mutations were predicted to be pathogenic by AlphaMissense, ClinVar data were only available for one, which was marked as ambiguous. However, it is crucial to note that while AlphaMissense showed a strong alignment with ClinVar for well-characterized mutations, its ability to predict mutations in lesser-known genes like *EIF1AX* was limited by the availability of clinical annotations. This discrepancy highlights a gap in our current clinical annotation capabilities, especially evident in the case of *BAP1*.

For *BAP1* (encoding BRCA1-associated protein 1; located on chromosome 3), among missense mutations predicted as pathogenic by AlphaMissense that had corresponding ClinVar data (23 in total), only one was documented as pathogenic in ClinVar, with the remainder listed as "uncertain significance." This demonstrates a significant gap in our understanding and the potential for deep learning tools to bridge this divide, as "uncertain significance" reflects a lack of conclusive data rather than a definitive classification.

Notably, *BAP1* was the only gene in our study where a mutation classified as "benign" by AlphaMissense also had corresponding ClinVar data, achieving 100% specificity. However, only 1 of 14 mutations labeled as benign by AlphaMissense matched a benign annotation in ClinVar, with the rest classified as uncertain significance. This result suggests that while AlphaMissense can reliably identify TNs, there remains a significant disparity between predictive analyses and existing clinical annotations, particularly for benign mutations. This underscores the challenges in accurately predicting pathogenicity in BAP1 and highlights the critical need for further research and validation.

Building on insights from our analysis of key mutations in UM, the role of *BAP1* inactivation demands attention.

Approximately 50% of primary UMs exhibit biallelic inactivation of *BAP1*, often caused by a loss of heterozygosity, often through monosomy 3, with a deleterious somatic mutation of the second *BAP1* allele, adhering to a 2-hit model.[30] Critically, UM with *BAP1* inactivation are associated with a high risk of metastatic relapse, with inactivating somatic mutations identified in 84% of metastasizing tumors.[12] This observation highlights the need for further studies to explore the impact of missense mutations in *BAP1*. Given that missense mutations are not the primary mechanism of advancement in this gene, understanding their role, especially in the context of the *BAP1* pathway, becomes critical for assessing metastatic risk and identifying potential therapeutic targets.

Furthermore, the observation that 27.2% of the mutations predicted as either "benign" or "pathogenic" by AlphaMissense were classified as "unknown significance" in ClinVar highlights the tool's capability to provide more definitive predictions for mutations where clinical data are currently inconclusive. This potential capacity to refine ambiguous classifications suggests that AlphaMissense could be instrumental in reevaluating and potentially reclassifying mutations within clinical databases, thus enhancing the utility of these resources in clinical decision-making.

Notably, all mutations identified as ambiguous by AlphaMissense that had corresponding ClinVar data were aligned as "uncertain significance," indicating complete agreement in this category. This observation suggests that while AlphaMissense's predictions often match clinical annotations, a substantial portion of mutations are classified as uncertain significance by ClinVar, pointing to a need for further research and data to confirm these predictions. It is particularly telling that there were no instances where mutations were predicted as pathogenic by AlphaMissense and benign by ClinVar or vice versa.

AlphaMissense's structure-aware predictions may offer a pathway to closing these gaps, particularly for less common mutations lacking extensive documentation or well-researched mutations that have yet to be definitively classified. The study highlights the potential of deep learning in enhancing our understanding of mutation pathogenicity, though it also underscores the crucial need for expanding clinical and experimental data to refine and validate these predictive models further. As the research community continues to bridge these gaps, the hope is to more accurately tailor treatments to the genetic nuances of each patient's tumor, potentially enhancing the care and outcomes for individuals battling UM.

Limitations

Our study provides significant insights into the genetic underpinnings of UM but encountered notable limitations. For example, our analysis did not encompass the clinical outcomes for the 1434 UM samples identified. Assessing clinical outcomes is essential for linking genetic alterations to patient prognosis and the effectiveness of treatments. Additionally, our validation of genetic findings was limited to ClinVar data. While ClinVar is a widely utilized and valuable resource for variant annotations, it has limitations, including variability in annotation quality and a significant proportion of variants labeled as "uncertain significance" due to limited or conflicting evidence. This variability can impact the consistency and reliability of comparisons between computational predictions and clinical annotations, highlighting the need for ongoing curation and validation of ClinVar data to ensure its robustness as a reference standard. This reliance may not fully capture the entire spectrum of pathogenicity, and future studies should incorporate additional sources of clinical validation to strengthen the findings.

Moreover, while computational tools like AlphaMissense offer valuable insights, their predictions need to be validated through experimental studies to avoid overreliance on in silico analyses. Experimental validation is crucial for confirming the functional impact of predicted pathogenic mutations and ensuring that computational predictions translate effectively to clinical relevance. Furthermore, our study focused exclusively on the pathogenicity of missense mutations using AlphaMissense, a tool specifically designed for this purpose. AlphaMissense does not currently provide models to assess other types of mutations, such as frameshift, nonsense, or indel mutations, and these were not part of our study. Future research should explore the development and application of predictive models that can evaluate the pathogenicity of a broader range of genetic mutations to gain a more comprehensive understanding of UM. The capacity of our study to validate the mutations identified was constrained by the limited availability of comprehensive clinical data, with future research efforts anticipated to bridge this gap through both clinical and computational methods, such as AlphaMissense. Moreover, our study did not distinguish between primary tumors and metastases; a separate analysis controlling for the stage of the tumor could illuminate differences in mutational profiles relevant to the disease's progression and metastatic potential.

While incorporating both targeted sequencing and whole-genome sequencing data, with 195 samples undergoing whole-genome sequencing, the study did not offer stratification based on these sequencing methods. Such differentiation could potentially highlight disparities in mutation detection capabilities and provide a richer understanding of the genomic complexity of UM. Additionally, the study did not explore the cooccurrence of multiple mutations across different genes, which could yield insights into compound pathogenic effects and their implications for disease behavior and treatment response.

Another important consideration is that the functional impact of missense mutations can be highly context dependent, influenced by factors such as the genetic background of the individual, environmental exposures, and other molecular interactions. This context dependency means that even accurately predicted pathogenic mutations may have varying impacts depending on these additional factors, underscoring the need for comprehensive studies that take these variables into account.

Despite the predictive model's significant contributions to our understanding of mutation pathogenicity, the reliance on AlphaMissense's predefined pathogenicity score cutoffs for classifications (i.e., pathogenic, ambiguous, or benign) without adjustments specific to UM suggests an area for

future methodological refinement. This adaptation could enhance the model's applicability and accuracy in the context of this malignancy, potentially leading to more precise predictive capabilities.

In summary, this research stands at the forefront of a transformative approach to understanding the genetic intricacies of UM. By integrating the analytical prowess of deep learning with the extensive mutation data contained in genomic databases, we are poised to make significant strides in elucidating the mutational landscape of this complex malignancy. The implications of our work extend beyond academic interest, promising to enhance the precision of genomic diagnostics and inform the development of personalized treatment regimens. As we bridge the gap between computational predictions and clinical realities, our findings have the potential to directly impact patient care. By identifying mutations with true pathogenic potential, scientists and clinicians can develop and tailor therapies to the specific genetic profile of each tumor to improve outcomes for individuals battling UM.

The integration of AlphaMissense with COSMIC and ClinVar databases in our study offers a novel lens through which to view the genetic underpinnings of UM. Our findings particularly emphasize the potential of deep learning tools in bridging the gap between vast genomic datasets and clinical applicability, providing an avenue for patients with UM to receive personalized and effective therapeutic interventions. Our findings call for further clinical investigations that could refine our understanding of mutation-driven pathogenicity in UM as it relates to protein structure and function. Ultimately, this study reaffirms the value of computational biology in advancing genomic diagnostics and charts the course for future explorations that could transform patient care in ocular oncology.

Footnotes and Disclosures

Originally received: May 17, 2024.
Final revision: September 9, 2024.
Accepted: December 3, 2024.
Available online: December 6, 2024. Manuscript no. XOPS-D-24-00153.

[1] Bascom Palmer Eye Institute, University of Miami, Miami, Florida.

[2] Wills Eye Hospital, Thomas Jefferson University, Philadelphia, Pennsylvania.

[3] Department of Molecular Medicine and USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida.

[4] Ocular Oncology Service, Wills Eye Hospital, Thomas Jefferson University, Philadelphia, Pennsylvania.

[5] Hamilton Eye Institute, University of Tennessee Science Center, Memphis, Tennessee.

Accepted as an on-demand poster presentation at the American Academy of Ophthalmology Annual Meeting, 2024, October 18-21, Chicago, Illinois.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have no proprietary or commercial interest in any materials discussed in this article.

Financial support was provided by NIH Center Core Grant P30EY014801, The RPB Unrestricted Award and Career Development Awards, Dr Ronald and Alicia Lepke Grant, The Lee and Claire Hager Grant, The Grant and Diana Stanton-Thornbrough, The Robert Baer Family Grant, The Emilyn Page and Mark Feldberg Grant, The Robert Farr Family Grant, The Jose Ferreira de Melo Grant, Mr. and Mrs. Irwin Friedman Grant, The Roberto and Antonia Menendez Family Grant, The Calvin and Flavia Oak Foundation, The Dr Tim and Cammy Ioannides Grant, The Stephen Takach Grant, The Richard and Kathy Lesser Grant, The Ragheb Family Grant, The Honorable A. Jay Cristol Grant, The Michele and Ted Kaplan Grant, The Zvi Levin Grant, The Christian Kathke Grant, The Carol Soffer Grant, and the Richard Azar Family Grant (institutional grants). This study consisted of molecular graphics and analyses performed with University of California, San Francisco (UCSF), ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the UCSF, with support from the National Institutes of Health (R01-GM129325) and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

HUMAN SUBJECTS: No human subjects were included in the study. This study was conducted in accordance with the ethical standards set forth in the Declaration of Helsinki. Given that the data analyzed were publicly available and fully deidentified, formal approval from an institutional review board or ethics committee was not required. Furthermore, as no patient-identifiable information was accessed and there was no direct involvement of human subjects, the need for informed consent was waived.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Gonzalez, Djulbegovic, Sharma, Karp, Wilson

Data collection: Gonzalez, Djulbegovic, Sharma, Antonietti, Kim, Wilson

Analysis and interpretation: Gonzalez, Djulbegovic, Sharma, Antonietti, Kim, Uversky, Karp, Shields, Wilson

Obtained funding: N/A

Overall responsibility: Gonzalez, Djulbegovic, Sharma, Antonietti, Kim, Uversky, Karp, Shields, Wilson

Support for Open Access publication was provided by the Department of Ophthalmology, University of Tennessee Health Science Center.

Abbreviations and Acronyms:
COSMIC = Catalogue of Somatic Mutations in Cancer; **FN** = false negative; **TN** = true negative; **TP** = true positive; **UM** = uveal melanoma.

Keywords:
AlphaFold, AlphaMissense, ClinVar, COSMIC, Missense mutations, Uveal melanoma.

Correspondence:
Matthew W. Wilson, MD, Hamilton Eye Institute, 930 Madison Ave, Suite 100, Memphis, TN 38103. E-mail: mwilson5@uthsc.edu.

References

1. Egan KM, Seddon JM, Glynn RJ, et al. Epidemiologic aspects of uveal melanoma. *Surv Ophthalmol*. 1988;32:239–251.

2. Singh AD, Turell ME, Topham AK. Uveal melanoma: trends in incidence, treatment, and survival. *Ophthalmology*. 2011;118:1881–1885.

3. Singh AD, Rennie IG, Seregard S, et al. Sunlight exposure and pathogenesis of uveal melanoma. *Surv Ophthalmol*. 2004;49:419−428.
4. Rodríguez A, Dueñas-Gonzalez A, Delgado-Pelayo S. Clinical presentation and management of uveal melanoma. *Mol Clin Oncol*. 2016;5:675−677.
5. Pavan-Langston D. *Manual of ocular diagnosis and therapy*. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
6. Say EAT, Shah SU, Ferenczy S, Shields CL. Optical coherence tomography of retinal and choroidal tumors. *J Ophthalmol*. 2012;2011:385058.
7. Kadakia A, Zhang J, Yao X, et al. Ultrasound in ocular oncology: technical advances, clinical applications, and limitations. *Exp Biol Med (Maywood)*. 2023;248:371−379.
8. McCannel TA. Fine-needle aspiration biopsy in the management of choroidal melanoma. *Curr Opin Ophthalmol*. 2013;24:262−266.
9. Onken MD, Worley LA, Char DH, et al. Collaborative Ocular Oncology Group report number 1: prospective validation of a multi-gene prognostic assay in uveal melanoma. *Ophthalmology*. 2012;119:1596−1603.
10. Djulbegovic MB, Taylor DJ, Uversky VN, et al. Intrinsic disorder in BAP1 and its association with uveal melanoma. *Genes*. 2022;13.
11. Djulbegovic M, Taylor Gonzalez DJ, Antonietti M, et al. Intrinsic disorder may drive the interaction of PROS1 and MERTK in uveal melanoma. *Int J Biol Macromol*. 2023;250:126027.
12. Harbour JW, Onken MD, Roberson ED, et al. Frequent mutation of BAP1 in metastasizing uveal melanomas. *Science*. 2010;330:1410−1413.
13. Harbour JW, Correa ZM, Schefler AC, et al. 15-Gene expression profile and PRAME as integrated prognostic test for uveal melanoma: first report of Collaborative Ocular Oncology Group Study No. 2 (COOG2. 1). *J Clin Oncol*. 2024;24:00447. JCO.
14. Bas Z, Grant-Kels JM, Shields CL. The Cancer Genome Atlas for uveal melanoma is predictive of patient outcomes. *Clin Dermatol*. 2024;42:56−61.
15. Shields CL, Dalvin LA, Vichitvejpaisal P, et al. Prognostication of uveal melanoma is simple and highly predictive using the Cancer Genome Atlas (TCGA) classification: a review. *Indian J Ophthalmol*. 2019;67:1959−1963.
16. Jager MJ, Shields CL, Cebulla CM, et al. Uveal melanoma. *Nat Rev Dis Prim*. 2020;6:24.
17. Martin M, Maßhöfer L, Temming P, et al. Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. *Nat Genet*. 2013;45:933−936.
18. Furney SJ, Pedersen M, Gentien D, et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov*. 2013;3:1122−1129.
19. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583−589.
20. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381:eadg7492.
21. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46:D1062−D1067.
22. Sondka Z, Dhir NB, Carvalho-Silva D, et al. COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res*. 2024;52:D1210−D1217.
23. Consortium TU. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res*. 2022;51:D523−D531.
24. Goddard TD, Huang CC, Meng EC, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*. 2018;27:14−25.
25. Pettersen EF, Goddard TD, Huang CC, et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci*. 2021;30:70−82.
26. Meng EC, Goddard TD, Pettersen EF, et al. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci*. 2023;32:e4792.
27. Cruz III F, Rubin BP, Wilson D, et al. Absence of BRAF and NRAS mutations in uveal melanoma. *Cancer Res*. 2003;63:5761−5766.
28. Onken MD, Worley LA, Long MD, et al. Oncogenic mutations in GNAQ occur early in uveal melanoma. *Invest Ophthalmol Vis Sci*. 2008;49:5230−5234.
29. Besaratinia A, Pfeifer GP. *Uveal melanoma and GNA11 mutations: a new piece added to the puzzle*. Hoboken, NJ: Wiley Online Library; 2011.
30. Onken MD, Worley LA, Ehlers JP, Harbour JW. Gene expression profiling in uveal melanoma reveals two molecular classes and predicts metastatic death. *Cancer Res*. 2004;64:7205−7209.