AMERICAN ACADEMY OF OPHTHALMOLOGY®

Machine Learning on Prediagnostic Metabolite Data Identifies Etiologic Endotypes of Exfoliation Glaucoma in United States Health Professionals

Akiko Hanyuda, MD, MPH,[1,2,*] *Oana A. Zeleznik, PhD,*[3,*] *Yoshihiko Raita, MD, MPH,*[4] *Kazuno Negishi, MD, PhD,*[1] *Louis R. Pasquale, MD,*[5] *Jessica Lasky-Su, ScD,*[3] *Janey L. Wiggs, MD, PhD,*[6,*] *Jae H. Kang, ScD*[3,*]

Purpose: Exfoliation glaucoma (XFG) etiology is poorly understood. Metabolomics-based etiologic endotypes of XFG may provide novel etiologic insights. We aimed to use unsupervised machine learning on prediagnostic plasma metabolites to characterize etiologic XFG endotypes.

Design: Prospective case-only analysis.

Participants: Among Nurses' Health Study and Health Professionals Follow-up Study participants, 205 (174 female and 31 male) incident XFG cases diagnosed with an average of 11.8 years following blood collection (1989—1995) were included.

Methods: We identified and confirmed incident cases of XFG or XFG suspect (collectively called "XFG" henceforth) through 2016 with medical record review. Liquid chromatography-mass spectrometry was used to profile 341 plasma metabolites. After preprocessing prediagnostic metabolites with adjustment for season, time of blood draw, and fasting status, we computed a distance matrix using Pearson distance and computed gap statistics to identify distinct endotypes.

Main Outcome Measures: Metabolomics-based XFG etiologic endotypes. Metabolomic profiles were compared across endotypes; false discovery rate (FDR) was used to account for multiple comparisons in Metabolite Set Enrichment Analyses. Exfoliation glaucoma environmental risk factors (e.g., lifetime ultraviolet (UV) exposure, folate consumption), a genetic risk score incorporating 8 major single nucleotide polymorphisms for exfoliation syndrome, and clinical presentations were compared across endotypes.

Results: We identified 3 distinct XFG metabolomic endotypes. Compared with the most common endotype 2 (reference group [n = 90; 43.9%]), endotype 1 (n = 56; 27.3%) tended to include more male southern US residents with greater UV exposure and were the least likely to have cardiovascular disease; among women, a higher percentage were postmenopausal. Endotype 3 (n = 59; 28.8%) was associated with being a male northern US resident; a higher prevalence of cardiovascular disease and risk factors such as higher body mass index, diabetes, hypertension, and dyslipidemia; and the lowest genetic susceptibility score. There were no differences in ophthalmic characteristics (e.g., maximum intraocular pressure, bilaterality, age at diagnosis) across endotypes ($P \geq 0.6$). In metabolite class analyses, compared with endotype 2, organic acids and carnitines were positively associated with endotype 1, whereas diacylglycerols and triacylglycerols were positively associated with endotype 3 (FDR <0.05).

Conclusions: Integrated metabolomic profiling can identify distinct XFG etiologic endotypes, suggesting different pathobiological mechanisms.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2025;5:100678 © 2024 by 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.

Exfoliation glaucoma (XFG) is the most prevalent recognizable secondary form of open-angle glaucoma worldwide. It is an ocular manifestation of exfoliation syndrome (XFS) (which is also known as pseudoexfoliation syndrome), an age-related extracellular matrix disorder, with a global prevalence of approximately 60 to 70 million.[1,2] The pathogenesis of XFG is characterized by the accumulation of extracellular matrix (known as exfoliation material) in various ocular structures, including trabecular meshwork, leading to elevated intraocular pressure (IOP) and optic neuropathy.[3]

https://doi.org/10.1016/j.xops.2024.100678
ISSN 2666-9145/25

While large-scale genome-wide association studies (GWASs) support a predominant role of genetic variants, including those in the *LOXL1* (lysyl oxidase-like 1) gene,[4] the same variants are also found in 80% of unaffected populations.[5] Meanwhile, epidemiologic evidence supports the role of environmental risk factors such as lifetime eye UV exposure, residing at greater distances from the equator, and greater time spent outdoors in youth.[6] Moreover, some gene−environment interactions also influence the risk of XFS; for example, the risk associated with the *POMP* (proteasome maturation protein) locus on chromosome 13 becomes stronger with increasing distance from the equator.[6,7] Thus, evidence supports heterogeneity in genetic and environmental etiology in XFG, which implicates different endotypes at the molecular level.

Metabolomic profiling, which involves measuring a set of metabolites that represent the end products of environmental and gene expression factors, has emerged as a valuable tool to uncover the pathobiology of complex diseases, including XFG.[8−10] Exploring metabolomes can provide insights into the etiology of XFG as the exfoliation material is hypothesized to result from various systemic biological reactions such as inflammation, oxidative stress, impaired mitochondrial function, and hypoxia,[11] which in combination may contribute to altered metabolites. Several metabolomic studies have been conducted for XFG with 1 study using aqueous humor samples[9] and 2 measuring plasma metabolites,[8,10] suggesting that specific metabolomic signatures are associated with the development of XFG. However, these studies solely compared metabolites between XFG and non-XFG populations, and to our knowledge, no study has determined distinct etiologic endotypes of XFG among cases only using prediagnostic samples.

Unsupervised machine learning, which can characterize the intrinsic patterns in an unlabeled dataset, is increasingly being used to identify new endotypes or distinct disease subtypes, based on similarities of molecular derivatives in an unbiased approach.[12,13] This approach has been commonly used in clinical settings using postdiagnosis omics data from clinical cases to identify endotypes that may have differing responses to treatment or different trajectories of disease progression.[12,13] We applied similar unsupervised machine learning techniques on prediagnostic plasma metabolites from 205 incident XFG cases to identify molecular etiologic endotypes of XFG; previously,[10] we investigated the prediagnostic metabolomic profile of any XFG overall compared to controls, and, here, we focused on the data from cases only to determine metabolomics-based etiologic endotypes for XFG.

Methods

Study Design and Population

This study included patients with XFG with prediagnostic plasma metabolomic data from 2 US nationwide prospective cohorts (the Nurses' Health Study [NHS] and Health Professionals Follow-up Study [HPFS]). The NHS enrolled 121 700 US female nurses aged 30 to 55 years in 1976.[14] The HPFS began in 1986 with 51 529 US male health professionals aged 40 to 75 years.[15] Participants were tracked biennially with high follow-up rates (>85%) using standardized questionnaires on diet, lifestyle, and medical history, including glaucoma.

Blood specimens were collected from 32 826 NHS participants during 1989−1990 and 18 159 HPFS participants during 1993−1995. All participants provided blood samples that were shipped by overnight courier.[16] After the samples were processed, aliquots of white blood cells, red blood cells, and plasma were stored under liquid nitrogen (≤ −130°C) that have been monitored weekly.[17] Cases of XFG were diagnosed following blood draw until June 1, 2016 (NHS), or January 1, 2016 (HPFS).

For confirmation of XFG, we identified participants with self-reported physician-diagnosed glaucoma and requested permission to receive medical information from their eye care providers. A glaucoma specialist (L.R.P.) reviewed all available visual field (VF) reports and glaucoma-related questionnaires filled out by diagnosing physicians (or medical records), including the presence of exfoliation material or other secondary causes for elevated IOP, maximum untreated IOP, optic nerve features, status of filtration apparatus, ophthalmic surgical history, and earliest VF loss date, to validate the diagnosis of XFG. Cases of XFG or XFG suspect (collectively called "XFG" henceforth) had documented exfoliation material on slit lamp examination in conjunction with (1) a maximum untreated IOP of ≥22 mmHg, (2) a cup-to-disc ratio of ≥0.6 or an intereye difference in a cup-to-disc ratio of ≥0.2, or (3) ≥1 reliable tests for VF showing glaucomatous loss on reliable tests. The exclusion criteria were as follows: exfoliation material but without any glaucomatous signs (such as VF loss, IOP elevation, or abnormal cup-to-disc ratio) to minimize detection bias and to focus on cases with clinically impactful XFS; cancer before diagnosis; and incomplete metabolite data. For the analysis, we included a total of 205 (174 female and 31 male) patients with XFG.

The study protocol was approved by the institutional review boards of Mass Eye and Ear, the Brigham and Women's Hospital, and Harvard T.H. Chan School of Public Health. This research adhered to the tenets of the Declaration of Helsinki. These ongoing cohorts included in our study were initiated in 1986 to 1989, and participants could opt out of the study at any time. The institutional review boards allowed participants' completion of questionnaires to be considered as implied consent.

Metabolite Profiling

Liquid chromatography-tandem mass spectrometry was used to profile plasma metabolites, as described previously (Broad Institute of MIT and Harvard University [Cambridge, MA]).[18] Detailed metabolite profiling methods and selection of metabolites used are provided in Supplementary Materials (available at www.ophthalmologyscience.org). Among a total of 398 known metabolites, 19 were excluded because of poor quality (10 metabolites that failed the delayed blood processing pilot test and 9 that showed a coefficient of variation of >25% in split pilots) and 38 because of missing in >20% of the participant samples, leaving 341 metabolites for final analyses.[18] The included metabolites had good within-person reproducibility over 1 to 2 years.[18] For metabolites with missing values in ≤20% of participants (XFS cases), 50% of the minimum value for that metabolite was used to impute the missing values, as has been done previously.[19,20]

Genotyping, Imputation, and Genetic Risk Scores for XFG

We derived a genetic risk score (GRS) specific for XFS to compare the relative strengths of associations between genetic predisposition

to XFG and any discovered endotypes. Details of the GRS calculation is described in Supplementary Materials. Briefly, the largest GWAS of XFS[7] involved 24 countries and included 9035 XFS cases and 17 008 controls for discovery GWAS, followed by replication in 18 countries with 4803 XFS cases and 93 267 controls. This study identified 8 gene variants related to XFS in 7 genes (*LOXL1*, *CACNA1A*, *POMP*, *TMEM136*, *AGPAT1*, *SEMA6A*, and *RBMS3*) at the genome-wide significant level ($P < 5 \times 10^{-8}$).[7] Based on these 8 variants available in the NHS/HPFS, we constructed the GRS for XFG. Genotypes were derived from 5 genotyping platforms used within the NHS/HPFS (Affymetrix, HumanCoreExome, IlluminaHumanHap, OmniExpress, and OncoArray),[21,22] and we imputed the datasets using the 1000 Genomes phase III release.[23,24] Quality control for genotyping and imputation was standardized within and across the platforms, as described previously.[21,22]

For the GRS for XFG, the 8 individual gene variants were coded for the risk allele dosage associated with XFG based on their imputation score values ranging from zero (for no alleles) up to 2 (signifying 2 alleles). The GRS for XFG was defined as the weighted sum of the products of the effect estimate of each gene variant from the abovementioned GWAS and risk allele dosage across the 8 loci.[25]

Statistical Analyses

All analyses were performed with SAS 9.4 (SAS Institute) and R 3.4.1 (R Foundation for Statistical Computing). Metabolite values were transformed to probit scores to scale to the same range and minimize the influence of skewed distributions. After adjusting for the season, time of blood draw, and fasting status, we computed a distance matrix using Pearson distance.[26] To identify distinct XFG endotypes, we then computed silhouette, elbow, and gap statistics, a commonly used method to compare the change in the within-cluster dispersion with that expected under null distribution.[27] Exfoliation glaucoma risk factors, clinical ophthalmic presentations, and genomic and metabolomic profiles were compared among the identified endotypes by using analysis of variance and chi-square tests as appropriate. To compare the individual metabolite levels across endotypes, odds ratios (ORs) and 95% confidence intervals (CIs) per 1 standard deviation increase in the metabolite level were estimated. To account for multiple comparisons given the high correlation structure of the metabolomics data, the number of effective tests (NEF)[28] was used, with a value of <0.05 considered statistically significant and that of <0.2 as candidates worthy of exploratory analysis.

To identify biologically meaningful pathways, we conducted metabolite class analyses, in which the 341 metabolites were assigned to 17 classes with at least 3 metabolites per class based on chemical taxonomy: steroids and steroid derivatives; carnitines; diacylglycerols; triacylglycerols; cholesteryl esters; lysophosphatidylethanolamines; phosphatidylethanolamines; lysophosphatidylcholines; phosphatidylcholines; phosphatidylethanolamine plasmalogens; phosphatidylcholine plasmalogens; organoheterocyclic compounds; ceramides; carboxylic acids and derivatives; organic acids and derivatives; nucleosides, nucleotides, and analogues; and sphingomyelins. To account for multiple comparisons in metabolite class enrichment analyses, the false discovery rate (FDR) was used, with a value of <0.05 considered statistically significant.[29]

Results

Various methods were used to determine the best number of clusters, which pointed to anywhere from 2 to 7 clusters (Fig 1 and Fig S1, available at www.ophthalmologyscience.org). To optimize the sample size to detect modest differences in characteristics by endotype with >80% power, we chose 3 clusters as they provided at least 50 participants in each group.[30]

Table 1 and Figure S2 (available at www.ophthalmologyscience.org) summarize characteristics overall and across the 3 endotypes as of blood draw as well as patients' systemic and genetic characteristics at diagnosis. Patients with endotype 2 had the highest prevalence of women (91.1%; also, they were most likely to be premenopausal or on postmenopausal hormones currently). They also had the lowest exposure to sunlight in youth, lowest prevalence of nonmelanoma skin cancer and diabetes, but the highest folate intake and highest occurrence of the XFG risk alleles rs11827818 [G] in *TMEM136* and rs4926244 [C] in *CACNA1A*. We considered individuals with endotype 2 as the reference groups as it contained the largest number of individuals (n = 90 of 205 cases; 43.9%).

Compared with individuals in the reference group (endotype 2), those with endotype 1 (n = 56; 27.3%) tended to be males from southern/western US with greater UV exposure in youth and were least likely to have cardiovascular risk factors. Women, in the endotype 1 group, tended to be postmenopausal at blood draw. In contrast, individuals with endotype 3 (n = 59; 28.8%) tended to be men from the northern/eastern US, with a higher prevalence of cardiovascular risk factors such as higher body mass index (BMI), diabetes, hypertension, and dyslipidemia; these individuals had the lowest GRS (the mean ± standard deviation for the GRS was 0.7 ± 0.6 for patients with endotype 3 vs. 0.9 ± 0.4 for those with endotype 2).

As a sensitivity analysis, we considered 4 endotypes and summarized their characteristics (Table S1, available at www.ophthalmologyscience.org). We also cross-classified the 3 main endotypes against the alternative 4 endotypes (Table S2, available at www.ophthalmologyscience.org). In general, compared with the main 3-cluster results, the 4 clusters included the original endotype 1; endotype 4 (n = 38), which was a restricted subset of the original endotype 3; and the second and third clusters of the 4 that came from a mix of those from the original endotypes 2 and 3 in the main analysis. With the reduced sample size and lower statistical power in the 4 endotypes, the characteristics between endotypes 2 and 3 in particular were less distinctive.

The analysis results of the 341 individual metabolites across the 3 endotypes using endotype 2 as the reference group are provided in Table S3 (available at www.ophthalmologyscience.org). Of these, the top 10 most significant across the endotypes are shown in Figure 2 (and Fig S3, available at www.ophthalmologyscience.org where the direct comparison of endotypes 1 and 3 are also provided). Compared with endotype 2, endotype 1 showed significant differences (NEF <0.05) in 128 of 341 metabolites. Among significant metabolites, triacylglycerols, and diacylglycerols were most strongly inversely associated with endotype 1 compared with endotype 2; the OR for XFG was 0.06 (95% CI, 0.02−0.13; NEF = 9.4E−8) for

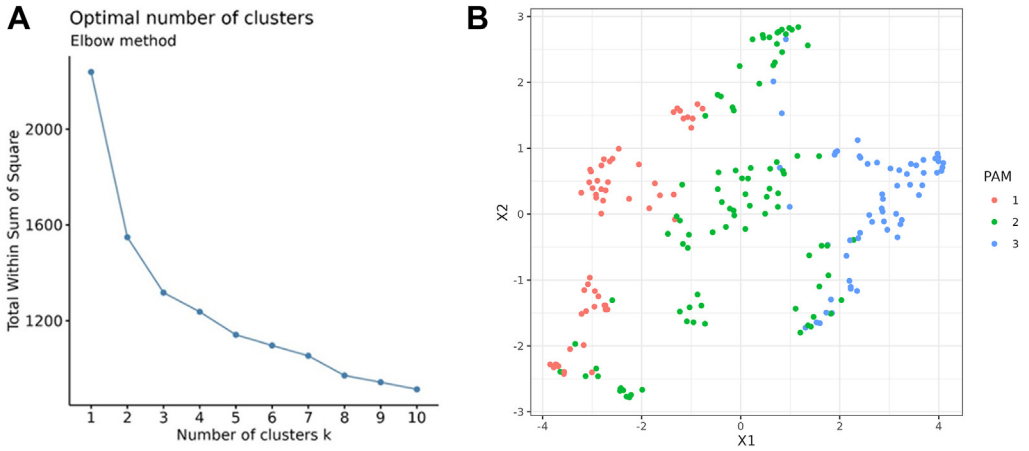

Figure 1. The optimal number of clusters using the elbow methods and endotyping analysis of plasma metabolites. **A,** Average elbow score according to the number of endotypes. Across the different numbers of endotypes (k of 1−10), the optimal numbers of clusters were k = 3−4. **B,** Endotyping analysis of plasma metabolites of patients with XFG based on the partitioning around medoids to measure connectiveness. PAM = partitioning around medoids; XFG = exfoliation glaucoma.

C34:0 diacylglycerols and 0.05 (95% CI, 0.02−0.12; NEF = 2.24E−9) for C52:1 triacylglycerols. When comparing the endotype 3 to 2, there were 109 of 341 metabolites significantly different at an NEF of <0.05. In stark contrast to results observed with endotype 1 vs. 2 comparisons, for endotype 3 vs. 2 comparisons, diacylglycerols and triacylglycerols were strongly associated with endotype 3 compared with endotype 2, where the highest OR was 46.8 (95% CI, 15.8−177; NEF = 1.5E−8) for C32:0 diacylglycerols.

To further explore the potential pathways that may be important for different endotypes, we conducted metabolite class analyses (Fig 3 and Fig S4, available at www.ophthalmologyscience.org). Compared with endotype 2, organic acids and carnitines were positively associated with endotype 1, although diacylglycerols and triacylglycerols were inversely associated (FDR <0.05). In contrast, individuals with endotype 3 compared with those with endotype 2 showed significantly higher levels of diacylglycerols and triacylglycerols but had lower levels of cholesteryl esters and phosphatidylcholine plasmalogens (FDR <0.05).

Finally, we examined whether these metabolite-driven endotypes differed in ophthalmic characteristics (Table 2). However, there were no significant differences in the clinical characteristics of maximum untreated IOP, bilaterality, or age at diagnosis across the 3 endotypes ($P \geq 0.6$).

Discussion

By applying an unsupervised machine learning approach for prediagnostic plasma metabolites among 205 US patients with XFG, we identified 3 distinct XFG endotypes, with different genetic, demographic, medical history, and lifestyle characteristics. Additional pathway analysis revealed some differences in metabolite classes across these endotypes. In particular, compared with the most common endotype (endotype 2), the association of diacylglycerols and triacylglycerols with XFG suggested opposite directions in the other 2 endotypes (significantly higher levels in endotype 3 but lower levels in endotype 1). Although ophthalmic characteristics were generally similar across these endotypes, this study provides evidence for heterogeneous interactions of genetics with modifiable and systemic risk factors in various etiologic pathways associated with XFG.

Integrating molecular endotyping techniques for high-dimensional omics profiles has provided new insights into various disease pathologies, including heart failure,[31] hypertrophic cardiomyopathy,[32] diabetes,[33] and bronchiolitis.[34] These studies demonstrated the potential of molecular endotyping to uncover novel disease pathways and to identify biomarkers for early screening and therapeutic targets. However, to the best of our knowledge, no study to date has been conducted to endotype XFG at the prediagnostic stage. Thus, to address this gap, we applied an unsupervised machine learning method to analyze plasma metabolomics data from 11.8 years before diagnosis in incident XFG cases to investigate etiologic endotypes.

Using this method, we found 3 metabolomic-driven etiologic endotypes. Compared with individuals with endotype 2, those with endotype 1 were most likely to be physically active and had the highest frequency of time spent outdoors in youth. Patients with endotype 1 also had the highest occurrence of the XFG risk allele rs3825942 [G] in *LOXL1* and rs7329408 [T] in *POMP*. In vitro studies have reported that treatment of fibroblasts with UV light can induce a significant increase in expression levels of *LOXL1*,[35] supporting the interaction of genetic susceptibility and UV exposure. Interestingly, the allele rs7329408 [T] in *POMP* has been shown to have the strongest associations with XFG by latitude (OR = 1.35 for 60°−70° latitude N vs. OR = 1.14 for 30°−40° latitude N), which may be related to differences in ocular UV exposure,[7,36] and further support the role of UV exposure in XFG etiology

Table 1. Characteristics as of Blood Draw by the 3 Endotypes of XFG (n = 205)*

Characteristics	All (n = 205)	Endotype 1 (n = 56; 27.3%)	Endotype 2 (n = 90; 43.9%)	Endotype 3 (n = 59; 28.8%)	*P* Value†
Mean age (SD), years	59.3 (6.5)	59.4 (6.6)	58.2 (6.3)	60.8 (6.5)	<0.1
Male individuals, n (%)	31 (15.1)	16 (28.6)	8 (8.9)	7 (11.9)	<0.1
Mean age at diagnosis of XFG (SD), years	71.1 (7.4)	70.9 (7.3)	70.9 (7.1)	71.6 (8.1)	0.8
Scandinavian descent, n (%)	13 (6.3)	4 (7.1)	6 (6.7)	3 (5.1)	0.3
Fasting >8 hrs, n (%)	62 (30.4)	18 (32.7)	25 (27.8)	19 (32.2)	0.8
Mean latitude of residence, °N	39.6 (4.3)	38.8 (4.7)	39.8 (4.4)	40.1 (3.8)	0.2
Mean longitude of residence, °W	81.6 (13.0)	83.5 (14.6)	81.9 (13.5)	79.2 (9.9)	0.2
Family history of glaucoma, n (%)	45 (22.4)	14 (25.0)	19 (21.3)	12 (21.4)	0.9
6+ hrs/wk outdoor sunlight exposure during summers in youth, n (%)	98 (53.3)	34 (66.7)	35 (43.2)	29 (55.8)	<0.1
Nonmelanoma skin cancer, n (%)	28 (13.7)	9 (16.1)	7 (7.8)	12 (20.3)	0.1
Mean caffeine intake (SD), mg/day	299.0 (246.4)	272.7 (260.5)	310.8 (243.8)	306.1 (239.0)	0.6
Mean folate intake (SD), mg/day	436.2 (229.3)	443.2 (210.3)	462.0 (256.6)	390.1 (196.6)	0.2
Mean alcohol intake (SD), g/day	6.9 (10.1)	4.1 (5.7)	7.9 (10.1)	8.2 (12.7)	<0.1
Mean body mass index (SD), kg/m^2	24.9 (4.3)	23.8 (3.8)	24.3 (3.7)	26.9 (4.9)	<0.1
Smoking status, n (%)					0.6
Never smoker	97 (47.3)	29 (51.8)	44 (48.9)	24 (40.7)	
Past smoker	90 (43.9)	21 (37.5)	40 (44.4)	29 (49.2)	
Current smoker	18 (8.8)	6 (10.7)	6 (6.7)	6 (10.2)	
Mean physical activity (SD), MET-h/wk	21.3 (25.2)	25.5 (25.3)	20.7 (26.5)	18.3 (22.8)	0.3
Hypertension, n (%)	47 (22.9)	9 (16.1)	16 (17.8)	22 (37.3)	<0.1
Hyperlipidemia, n (%)	65 (31.7)	13 (23.2)	27 (30.0)	25 (42.4)	0.1
Diabetes, n (%)	4 (2.0)	1 (1.8)	0 (0.0)	3 (5.1)	0.1
Among female individuals					<0.1
Premenopausal, n (%)‡	17 (10.7)	4 (11.4)	13 (16.9)	0 (0.0)	
Postmenopausal and current PMH use, n (%)	59 (37.1)	7 (20.0)	28 (36.4)	24 (51.1)	
Postmenopausal and past PMH use, n (%)	29 (18.2)	5 (14.3)	13 (16.9)	11 (23.4)	
Postmenopausal and no PMH use, n (%)	54 (34.0)	19 (54.3)	23 (29.9)	12 (25.5)	
Genetic risk profiles					
Mean GRS (SD)	0.8 (0.5)	0.9 (0.5)	0.9 (0.4)	0.7 (0.6)	0.2
Gene	**SNP (risk/reference allele)**				
LOXL1	rs1048661 (G/T), n (% GT+GG)	89 (97.8)	22 (95.7)	44 (100.0)	23 (95.8)	0.6
LOXL1	rs3825942 (G/A), n (% GA+GG)	89 (97.8)	23 (100.0)	43 (97.7)	23 (95.8)	0.6
RBMS3	rs12490863 (A/G), n (% AG+AA)	18 (19.8)	5 (21.7)	10 (22.8)	3 (12.5)	0.6
SEMA6A	rs10072088 (T/C), n (% TC+TT)	87 (95.6)	22 (95.7)	41 (93.2)	24 (100.0)	0.7
CACNA1A	rs4926244 (C/T), n (% CT+CC)	24 (26.4)	7 (30.4)	15 (34.1)	2 (8.3)	0.1
FLT-POMP	rs7329408 (T/C), n (% TC+TT)	25 (27.5)	8 (34.8)	13 (29.5)	4 (16.7)	0.3
TMEM136	rs11827818 (G/A), n (% GA+GG)	34 (37.4)	7 (30.4)	18 (40.9)	9 (37.5)	0.1
AGPAT1	rs3130283 (A/C), n (% AC+AA)	29 (31.9)	7 (30.4)	13 (29.5)	9 (37.5)	0.1

GRS = genetic risk score; MET = metabolic equivalent of task; PMH = postmenopausal hormone; SD = standard deviation; SNP = single nucleotide polymorphism; XFG = exfoliation glaucoma.
Values are presented as the mean (SD) for continuous variables and n (%) for categorical variables.
*Blood samples were provided from 32 826 female individuals (Nurses' Health Study) in 1989 to 1990 and from 18 159 male individuals (Health Professionals Follow-up Study) in 1993 to 1995.
†Two-sided raw *P* values. Between-endotype differences were analyzed using analysis of variance, Kruskal–Wallis test, chi-square test, and Fisher exact test, as appropriate.
‡Among women with complete data on menopausal status and menopausal hormone therapy use status.

for this endotype. Additionally, individuals with endotype 1 had significantly higher levels of organic acids and their derivatives and carnitines and significantly lower levels of triacylglycerols with ≤3 double bonds and diacylglycerols, which may be related to a lower prevalence of metabolic syndrome[37,38] and greater UV exposure.[39] In general, the association between triacylglycerols and XFG in the literature to date has been inconsistent. Although several cross-sectional studies observed positive associations between dyslipidemia and prevalent XFG,[40–44] our matched case-control study using prediagnostic metabolomes found an inverse association between triacylglycerols and XFG compared with non-XFG,[10] as lower levels of triacylglycerols can be a proxy for higher physical activity[37,38] and higher ocular UV exposure.[39] Although these studies are not directly comparable due to the substantial heterogeneity in the study design (metabolomics in prediagnostic biospecimens

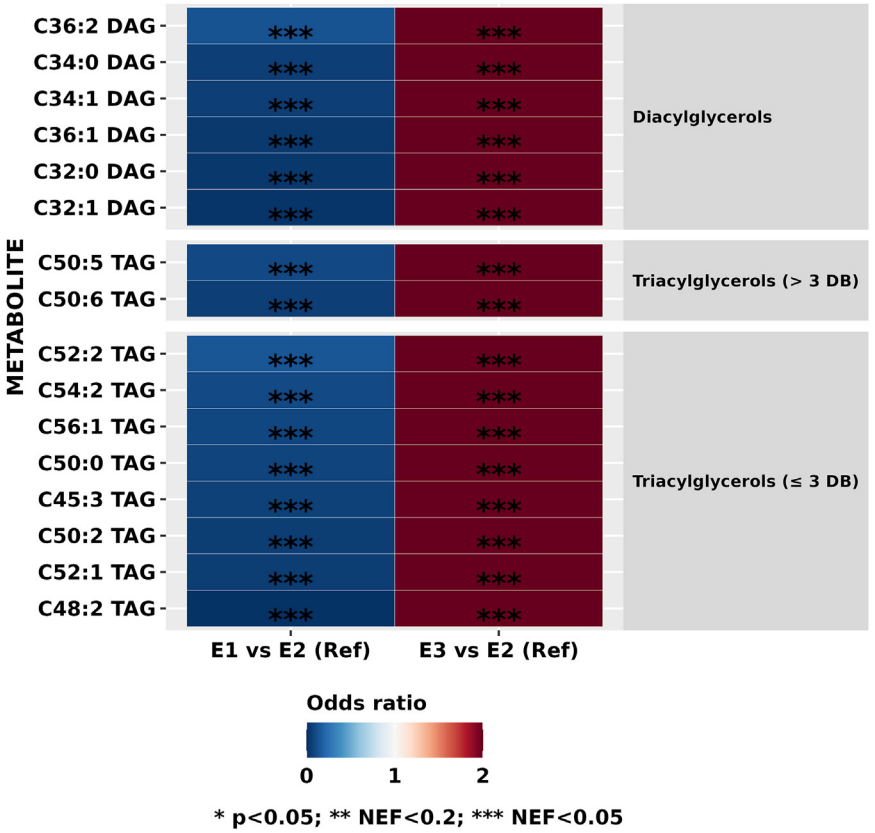

Figure 2. Top 10 most significant metabolites among the 341 individual plasma metabolites evaluated by logistic regression models of either the comparison of E1 vs. E2 or E3 vs. E2 (n = 205 cases). The model was adjusted for season and time of blood draw and fasting status. Nomenclature based on the number of carbon atoms and double bonds (e.g., TAG C53:6) denotes that the fatty acids contained in this metabolite (TAG) include 53 carbon atoms and 6 double bonds. However, the exact fatty acid composition and the exact positions of the double bonds are not available. *$P < 0.05$, ** NEF <0.2, *** NEF <0.05. DAG = diacylglycerol; DB = double bond; E1 = endotype 1; E2 = endotype 2; E3 = endotype 3; NEF = number of effective tests; TAG = triacylglycerol.

vs. metabolomics in those with prevalent disease) and covariates that were adjusted for, we speculate that there is a subtype of participants, particularly male individuals, who are at genetically higher risk of XFG and more susceptible because of greater lifetime UV exposure likely through more lifetime physical activity.

Endotype 3 was distinctive in having the highest prevalence of various cardiovascular risk factors such as low folate intake, low physical activity, higher alcohol intake, higher BMI, greater smoking history, hypertension, hyperlipidemia, and diabetes while showing the lowest genetic susceptibility among the 3 endotypes, indicating a major role for environmental risk factors. Additionally, compared with those in the reference group (endotype 2), individuals with endotype 3 had significantly lower levels of cholesteryl esters and plasmalogens and had significantly higher levels of triacylglycerols with ≤3 double bonds and diacylglycerols. Higher triacylglycerol levels in individuals with this endotype are likely reflective of lower physical activity, higher BMI, and systemic vascular/metabolomic diseases, all of which were observed in those with endotype 3. Lower plasmalogens may also capture a lower level of physical activity[45] and UV exposure.[46]

Individuals with endotype 2, who comprised the largest reference group, were distinguished by female predominance, with female individuals accounting for >90% of this group compared with 85% in the overall cohort. In addition, female individuals with endotype 2 were more likely to be either premenopausal or on postmenopausal hormone therapy, implicating the potential role of estrogens for this endotype. Although the role of estrogen in XFG etiology is unclear, estrogen, when photosensitized by UV, forms adducts that lead to cellular stress, potentially increasing the incidence of Fuchs corneal endothelial dystrophy among women.[47] Considering that cellular stress and the release of elastin stress fibers are pivotal in the development of XFG,[48,49] we speculate that similar interactions between estrogen and UV exposure may be a risk factor for XFG in this endotype. In addition, endotype 2 showed significantly lower levels of organic acids and derivatives (including amino acids) and carnitines (particularly long-chain acylcarnitine) compared with endotype 1. Carnitines have been shown to have neuroprotective effects in murine models of glaucoma,[50] and altered carnitine metabolism was identified in patients with glaucoma, including those with XFG.[8,51] Epidemiologic studies also support a sex-specific association between amino acids and neurological disorders, including

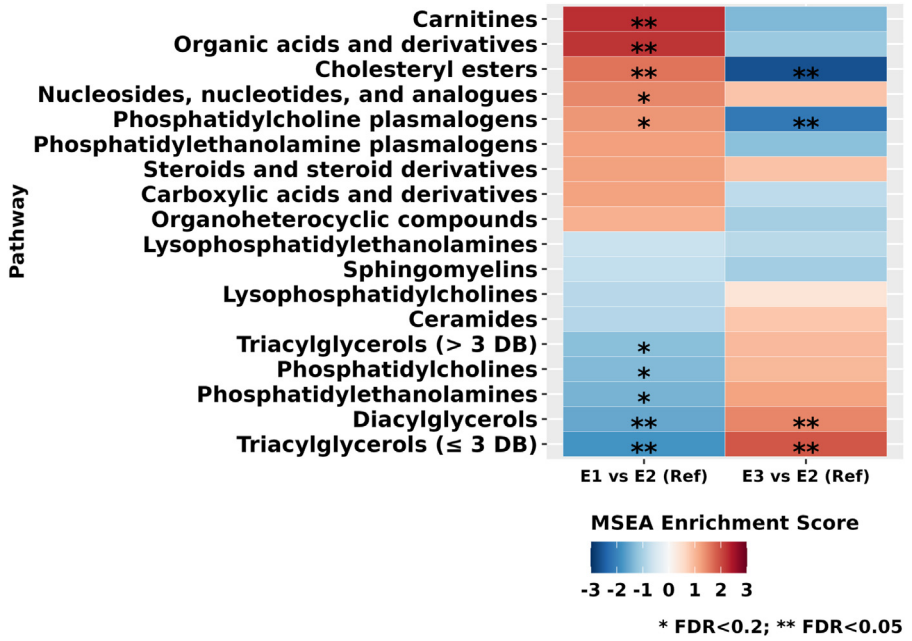

Figure 3. Plasma metabolite classes (n = 17) evaluated in logistic regression analysis of XFG (205 cases) across the endotypes (E1 vs. E2 or E3 vs. E2). The model was adjusted for season and time of blood draw and fasting status. * FDR <0.2, ** FDR <0.05. DB = double bond; E1 = endotype 1; E2 = endotype 2; E3 = endotype 3; FDR = false discovery rate; XFG = exfoliation glaucoma.

glaucoma.[52–54] In general, a stronger effect of decreased amino acids on neurodegeneration has been suggested in women due to lower basal levels in muscles and differences in oxidation kinetics during exercise.[55] The highest occurrence of the risk allele for the *CACNA1A* locus was also notable; estrogen is known to regulate voltage-gated calcium channels,[56] suggesting the possibility of estrogen interacting with calcium signaling in XFG etiology.

The strength of this study was that it addressed the knowledge gap in determining distinct metabolite-driven endotypes of XFG. We also obtained rich prospectively collected information on lifestyle, clinical characteristics, and XFG-genetic risk factors. Additionally, prediagnostic metabolomic information was collected >10 years before the diagnosis of XFG.

Nonetheless, several limitations should be noted. First, the study sample did not include controls as the primary aim of this study was to build from the knowledge gained from the previous analysis[10] of the general metabolomic signature of XFG cases compared to controls by investigating the variations in that XFG signature within XFG endotypes to gain insights on possible distinct etiologic patterns. Second, our study participants were health professionals, mostly females of European descent, which limits the generalizability of our findings. Future studies in other populations with different age, sex, race, and ethnicity distributions are needed. Third, the clustering methods used have the limitation of possibly overfitting the study sample data, and we lacked a replication sample with prediagnostic metabolite data; thus, our results should be

Table 2. Clinical Features of XFG Cases at Diagnosis by Endotype*

Clinical Characteristics	All (n = 205)	Endotype 1 (n = 56; 27.3%)	Endotype 2 (n = 90; 43.9%)	Endotype 3 (n = 59; 28.8%)	*P* Value*
Age at diagnosis (SD), yrs	71.1 (7.4)	70.9 (7.3)	70.9 (7.1)	71.6 (8.1)	0.8
Bilaterality, *n* (%)	95 (46.3)	26 (46.4)	41 (45.6)	28 (47.5)	1.0
High-tension glaucoma, *n* (%)	172 (83.9)	49 (87.5)	75 (83.3)	48 (81.4)	0.7
Severity of XFG					0.9
High IOP only (no optic disc cupping or VF loss), *n* (%)	56 (27.3)	17 (30.4)	24 (26.7)	15 (25.4)	
Optic disc cupping (no VF loss), *n* (%)	56 (27.3)	13 (23.2)	27 (30.0)	16 (27.1)	
VF loss, *n* (%)	93 (45.4)	26 (46.4)	39 (43.3)	28 (47.5)	
Mean maximum IOP (SD), mmHg	27.9 (6.8)	28.3 (7.2)	28.0 (6.6)	27.2 (6.7)	0.6
Mean recent vertical cup-to-disc ratio (SD)	0.57 (0.20)	0.60 (0.19)	0.56 (0.20)	0.54 (0.21)	0.2

IOP = intraocular pressure; SD = standard deviation; VF = visual field; XFG = exfoliation glaucoma.
*XFG cases were diagnosed after blood draw until June 1, 2016 (Nurses' Health Study), or January 1, 2016 (Health Professional's Follow-up Study).

interpreted with caution and as exploratory findings that need confirmation. Fourth, our blood samples were stored for a long time, and we cannot eliminate the possibility of degradation or interactions of these metabolites. Nevertheless, we have previously used metabolomes from NHS and HPFS samples similarly stored for >10 years for studying various health outcomes and observed significant associations,[57,58] which have also been supported in metabolomic studies from different cohorts.[59–61] Therefore, the long storage period might have minimally affected our findings. Fifth, we assessed metabolites once at baseline, without repeated measurements. However, we limited our investigation to those metabolites with high stability over at least 1 year.[18] Finally, the sample size was small, and this could have prevented a clearer resolution of endotypes; in particular, we were unable to find differences in the clinical presentation of XFG across the endotypes.

In conclusion, using metabolomic data, we identified 3 endotypes for XFG among mostly female US health professionals. Although ophthalmic clinical features were generally similar across these endotypes, the observed heterogeneity in prediagnostic plasma metabolites may provide some clues on the differential susceptibility to genetic and environmental modifiable risk factors in the etiology of XFG.

Footnotes and Disclosures

Originally received: August 15, 2024.
Final revision: December 9, 2024.
Accepted: December 10, 2024.
Available online: December 16, 2024. Manuscript no. XOPS-D-24-00300.

[1] Department of Ophthalmology, Keio University School of Medicine, Tokyo, Japan.

[2] Epidemiology and Prevention Group, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan.

[3] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

[4] Department of Nephrology, Okinawa Prefectural Chubu Hospital, Naha, Japan.

[5] Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, New York, New York.

[6] Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts.

*A.H., O.A.Z, J.L.W and J.H.K. contributed equally.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

L.R.P.: Consultant − Twenty Twenty.

J.L.W.: Consultant − Allergan, Editas, and CRISPRTx (Unrelated to this work).

This work was supported by the National Institutes of Health [UM1 CA186107, U01 CA167552, U01 CA176726, R01 EY09611, and R01 EY020928 (J.L.W. and J.H.K.) and R01 EY015473 (L.R.P.)]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. L.R.P. is also supported by the Glaucoma Foundation (NYC), the Barry Family Center for Ophthalmic Artificial Intelligence and Human Health (BFCOAIHH), and Twenty and an unrestricted Challenge Grant from Research to Prevent Blindness (NYC). The funders had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Support for Open Access publication was provided by Keio University School of Medicine.

HUMAN SUBJECTS: Human subjects were included in this study. The study protocol was approved by the institutional review boards of Mass Eye and Ear, the Brigham and Women's Hospital, and Harvard T.H. Chan School of Public Health. All research adhered to the tenets of the Declaration of Helsinki. These ongoing cohorts included in our study were initiated in 1986 to 1989, and participants could opt out of the study at any time. The institutional review boards allowed participants' completion of questionnaires to be considered as implied consent.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Hanyuda, Zeleznik, Raita, Negishi, Pasquale, Lasky-Su, Wiggs, Kang

Data collection: Pasquale, Lasky-Su, Kang

Analysis and interpretation: Hanyuda, Zeleznik, Raita, Pasquale, Wiggs, Kang

Obtained funding: Pasquale, Lasky-Su, Wiggs, Kang

Overall responsibility: Hanyuda, Zeleznik, Raita, Negishi, Pasquale, Lasky-Su, Wiggs, Kang

Abbreviations and Acronyms:
BMI = body mass index; **CI** = confidence interval; **FDR** = false discovery rate; **GRS** = genetic risk score; **GWAS** = genome-wide association studies; **HPFS** = Health Professionals Follow-up Study; **IOP** = intraocular pressure; **LOXL1** = lysyl oxidase-like 1; **NEF** = number of effective tests; **NHS** = Nurses' Health Study; **OR** = odds ratio; **POMP** = proteasome maturation protein; **VF** = visual field; **XFG** = exfoliation glaucoma; **XFS** = exfoliation syndrome.

Keywords:
Endotyping analysis, Exfoliation glaucoma, Metabolome, Omics, Unsupervised machine learning.

Correspondence:
Akiko Hanyuda, MD, MPH, Department of Ophthalmology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. E-mail: akihanyu@keio.jp.

References

1. Schlötzer-Schrehardt U, Naumann GOH. Ocular and systemic pseudoexfoliation syndrome. *Am J Ophthalmol*. 2006;141:921–937.

2. Stein JD, Pasquale LR, Talwar N, et al. Geographic and climatic factors associated with exfoliation syndrome. *Arch Ophthalmol*. 2011;129:1053–1060.

3. Schlötzer-Schrehardt U, Naumann GO. Trabecular meshwork in pseudoexfoliation syndrome with and without open-angle glaucoma. A morphometric, ultrastructural study. *Invest Ophthalmol Vis Sci*. 1995;36:1750−1764.
4. Thorleifsson G, Walters GB, Hewitt AW, et al. Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. *Nat Genet*. 2010;42:906−909.
5. Lemmelä S, Forsman E, Onkamo P, et al. Association of LOXL1 gene with Finnish exfoliation syndrome patients. *J Hum Genet*. 2009;54:289−297.
6. Jiwani AZ, Pasquale LR. Exfoliation syndrome and solar exposure: new epidemiological insights into the pathophysiology of the disease. *Int Ophthalmol Clin*. 2015;55:13−22.
7. Aung T, Ozaki M, Lee MC, et al. Genetic association study of exfoliation syndrome identifies a protective rare variant at LOXL1 and five new susceptibility loci. *Nat Genet*. 2017;49:993−1004.
8. Leruez S, Bresson T, Chao de la Barca JM, et al. A plasma metabolomic signature of the exfoliation syndrome involves amino acids, acylcarnitines, and polyamines. *Invest Ophthalmol Vis Sci*. 2018;59:1025−1032.
9. Myer C, Abdelrahman L, Banerjee S, et al. Aqueous humor metabolite profile of pseudoexfoliation glaucoma is distinctive. *Mol Omi*. 2020;16:425−435.
10. Kang JH, Zeleznik O, Frueh L, et al. Prediagnostic plasma metabolomics and the risk of exfoliation glaucoma. *Invest Ophthalmol Vis Sci*. 2022;63:15.
11. Aboobakar IF, Johnson WM, Stamer WD, et al. Major review: exfoliation syndrome; advances in disease genetics, molecular biology, and epidemiology. *Exp Eye Res*. 2017;154:88−103.
12. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. *J Am Coll Cardiol*. 2018;71:2668−2679.
13. Liang LW, Shimada YJ. Endotyping in heart failure - identifying mechanistically meaningful subtypes of disease. *Circ J*. 2021;85:1407−1415.
14. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health*. 1997;6:49−62.
15. Rimm EB, Giovannucci EL, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet*. 1991;338:464−468.
16. Hankinson SE, Willett WC, Manson JE, et al. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst*. 1998;90:1292−1299.
17. Tworoger SS, Hankinson SE. Use of biomarkers in epidemiologic studies: minimizing the influence of measurement error in the study design and analysis. *Cancer Causes Control*. 2006;17:889−899.
18. Townsend MK, Clish CB, Kraft P, et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. *Clin Chem*. 2013;59:1657−1667.
19. Wei R, Wang J, Su M, et al. Missing value imputation approach for mass spectrometry-based metabolomics Data. *Sci Rep*. 2018;8:663.
20. Zeleznik OA, Eliassen AH, Kraft P, et al. A prospective analysis of circulating plasma metabolites associated with ovarian cancer risk. *Cancer Res*. 2020;80:1357−1367.
21. Lindström S, Loomis S, Turman C, et al. A comprehensive survey of genetic variation in 20,691 subjects from four large cohorts. *PLoS One*. 2017;12:e0173997.
22. Duffy DL, Zhu G, Li X, et al. Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat Commun*. 2018;9:4774.
23. O'Connell J, Gurdasani D, Delaneau O, et al. A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet*. 2014;10:e1004234.
24. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48:1284−1287.
25. Pasquale LR, Aschard H, Kang JH, et al. Age at natural menopause genetic risk score in relation to age at natural menopause and primary open-angle glaucoma in a US-based sample. *Menopause*. 2017;24:150−156.
26. Alloghani M, Al-Jumeily D, Mustafina J, et al. *A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science*. Cham, Switzerland: Springer International Publishing; 2019.
27. Tibshirani R, Walther G, Hastie T. Estimating the number of data clusters via the gap statistic. *J R Stat Soc Ser B*. 2001;63:411−423.
28. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol*. 2008;32:361−369.
29. Benjamini Y, Drai D, Elmer G, et al. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001;125:279−284.
30. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New York, NY: Lawrence Erlbaum Associates; 1988.
31. Woolley RJ, Ceelen D, Ouwerkerk W, et al. Machine learning based on biomarker profiles identifies distinct subgroups of heart failure with preserved ejection fraction. *Eur J Heart Fail*. 2021;23:983−991.
32. Liang LW, Raita Y, Hasegawa K, et al. Proteomics profiling reveals a distinct high-risk molecular subtype of hypertrophic cardiomyopathy. *Heart*. 2022;108:1807−1814.
33. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol*. 2018;6:361−369.
34. Raita Y, Pérez-Losada M, Freishtat RJ, et al. Integrated omics endotyping of infants with respiratory syncytial virus bronchiolitis and risk of childhood asthma. *Nat Commun*. 2021;12:3601.
35. Zenkel M, Krysta A, Pasutto F, et al. Regulation of lysyl oxidase-like 1 (LOXL1) and elastin-related genes by pathogenic factors associated with pseudoexfoliation syndrome. *Invest Ophthalmol Vis Sci*. 2011;52:8488−8495.
36. Kandeeban S, Ishwarya S, Nareshkumar RN, et al. A study on the candidate gene association and interaction with measures of UV exposure in pseudoexfoliation patients from India. *Curr Eye Res*. 2023;48:1144−1152.
37. Healy GN, Matthews CE, Dunstan DW, et al. Sedentary time and cardio-metabolic biomarkers in US adults: nhanes 2003-06. *Eur Heart J*. 2011;32:590−597.
38. Crichton GE, Alkerwi A. Physical activity, sedentary behavior time and lipid levels in the Observation of Cardiovascular Risk Factors in Luxembourg study. *Lipids Health Dis*. 2015;14:87.
39. Solis-Urra P, Cristi-Montero C, Romero-Parra J, et al. Passive commuting and higher sedentary time is associated with vitamin D deficiency in adult and older women: results from Chilean National Health Survey 2016-2017. *Nutrients*. 2019;11:300.
40. Türkyılmaz K, Öner V, Kırbas A, et al. Serum YKL-40 levels as a novel marker of inflammation and endothelial dysfunction in patients with pseudoexfoliation syndrome. *Eye*. 2013;27:854−859.

41. Yilmaz N, Coban DT, Bayindir A, et al. Higher serum lipids and oxidative stress in patients with normal tension glaucoma, but not pseudoexfoliative glaucoma. *Bosn J Basic Med Sci*. 2016;16:21−27.
42. Janićijević K, Kocić S, Pajović S, et al. The importance of developing atherosclerosis in pseudoexfoliation glaucoma. *Vojnosanit Pregl*. 2017;74:8−12.
43. Lesiewska H, Łukaszewska-Smyk A, Odrowąż-Sypniewska G, et al. Chosen vascular risk markers in pseudoexfoliation syndrome: an age-related disorder. *J Ophthalmol*. 2017;2017:5231095.
44. Mirza E. Atherogenic indices in pseudoexfoliation syndrome. *Eye*. 2019;33:1911−1915.
45. Ding M, Zeleznik OA, Guasch-Ferre M, et al. Metabolome-wide association study of the relationship between habitual physical activity and plasma metabolite levels. *Am J Epidemiol*. 2019;188:1932−1943.
46. Jardine A, Bright M, Knight L, et al. Does physical activity increase the risk of unsafe sun exposure? *Heal Promot J Aust*. 2012;23:52−57.
47. Liu C, Miyajima T, Melangath G, et al. Ultraviolet A light induces DNA damage and estrogen-DNA adducts in Fuchs endothelial corneal dystrophy causing females to be more affected. *Proc Natl Acad Sci USA*. 2020;117:573−583.
48. Lee RK. The molecular pathophysiology of pseudoexfoliation glaucoma. *Curr Opin Ophthalmol*. 2008;19:95−101.
49. Schlötzer-Schrehardt U. Molecular pathology of pseudoexfoliation syndrome/glaucoma–new insights from LOXL1 gene associations. *Exp Eye Res*. 2009;88:776−785.
50. Calandrella N, De Seta C, Scarsella G, Risuleo G. Carnitine reduces the lipoperoxidative damage of the membrane and apoptosis after induction of cell stress in experimental glaucoma. *Cell Death Dis*. 2010;1:e62.
51. Zeleznik OA, Kang JH, Lasky-Su J, et al. Plasma metabolite profile for primary open-angle glaucoma in three US cohorts and the UK Biobank. *Nat Commun*. 2023;14:2860.
52. Kimberly WT, Wang Y, Pham L, et al. Metabolite profiling identifies a branched chain amino acid signature in acute cardioembolic stroke. *Stroke*. 2013;44:1389−1395.
53. Arnold M, Nho K, Kueider-Paisley A, et al. Sex and APOE epsilon4 genotype modify the Alzheimer's disease serum metabolome. *Nat Commun*. 2020;11:1148.
54. Hanyuda A, Rosner BA, Wiggs JL, et al. Prospective study of dietary intake of branched-chain amino acids and the risk of primary open-angle glaucoma. *Acta Ophthalmol*. 2022;100:e760−e769.
55. Krumsiek J, Mittelstrass K, Do KT, et al. Gender-specific pathway differences in the human serum metabolome. *Metabolomics*. 2015;11:1815−1833.
56. Subbamanda YD, Bhargava A. Intercommunication between voltage-gated calcium channels and estrogen receptor/estrogen signaling: insights into physiological and pathological conditions. *Cells*. 2022;11:3850.
57. Mayers JR, Wu C, Clish CB, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat Med*. 2014;20:1193−1198.
58. Zeleznik OA, Balasubramanian R, Zhao Y, et al. Circulating amino acids and amino acid-related metabolites and risk of breast cancer among predominantly premenopausal women. *NPJ Breast Cancer*. 2021;7:54.
59. Katagiri R, Goto A, Nakagawa T, et al. Increased levels of branched-chain amino acid associated with increased risk of pancreatic cancer in a prospective case-control study of a large cohort. *Gastroenterology*. 2018;155:1474−1482.e1.
60. Tobias DK, Hazra A, Lawler PR, et al. Circulating branched-chain amino acids and long-term risk of obesity-related cancers in women. *Sci Rep*. 2020;10:16534.
61. Hur B, Gupta VK, Huang H, et al. Plasma metabolomic profiling in patients with rheumatoid arthritis identifies biochemical features predictive of quantitative disease activity. *Arthritis Res Ther*. 2021;23:164.