Myopia genetics in genome-wide association and post-genome-wide association study era

Xuan Liao, Qing-Qing Tan, Chang-Jun Lan

Department of Ophthalmology, Affiliated Hospital of North Sichuan Medical College; Department of Ophthalmology and Optometry, North Sichuan Medical College, Nanchong 637000, Sichuan Province, China

Correspondence to: Xuan Liao and Chang-Jun Lan. Department of Ophthalmology, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, Sichuan Province, China. aleexand@163.com; eyelanchangjun@163.com

Received: 2018-10-05 Accepted: 2019-05-21

Abstract

- Genome-wide association studies (GWAS) of myopia and refractive error have generated exciting results and identified novel risk-associated loci. However, the interpretation of the findings of GWAS of complex diseases is not straightforward and has remained challenging. This review provides a brief summary of the main focus on the advantages and limitations of GWAS of myopia, with potential strategies that may contribute to further insight into the genetics of myopia in the post-GWAS or omics era.

- KEYWORDS: myopia; genetic variation; genome-wide association studies; omics

DOI:10.18240/ijo.2019.09.18

Citation: Liao X, Tan QQ, Lan CJ. Myopia genetics in genome-wide association and post-genome-wide association study era. Int J Ophthalmol 2019;12(9):1487-1492

INTRODUCTION

Myopia, also known as near-sightedness or short-sightedness, is characterized by which the images of distant objects fail to be properly focused on the retina plane but rather in the front of the retina. Myopia is accepted as a multifactor disorder that involves in genetic (nature) and non-genetic environmental or behavioral (nurture) risk factors plus their complex interaction, likely together with the effects of stochastic factors. Meanwhile, it is considered as a polygenic disease that involves a critical number of candidate genes joint action or more complex genetic mechanisms, rather than any of the simple Mendelian patterns of inheritance. There was a dramatic rise in myopia prevalence over the last 30y in many countries, especially among younger people in urban East Asia[1-3]. This phenomenon may be caused by increasing educational pressures or lifestyle changes and potentially gene-environment interactions, suggesting the role of environmental exposures in myopia susceptibility. Despite epidemiological heterogeneity, however, the genetic basis of myopia has been established based on the molecular genetics studies and genetic epidemiological evidences of myopia in the early stage. Heritability estimates from family and twin studies for myopia ranging between 50% and 90%, continue to play a significant role in enhancing the interpretability of genetic evidences.

The advent of the genome-wide association study (GWAS) era is accompanied with the revolution of molecular technology and information. The unbiased nature of genome-wide measurements coupled with the statistical power of association studies have yield new insights into myopic pathogenesis without any prior knowledge of function. There are several popular instances such as the Meta-analyses that could promote power to reveal more loci by pooling information from multiple GWAS: the imputation that could extend appraisal of associations across the genome by inferring frequency based on neighboring variant frequencies, the permutation that could build an empirical estimation of the null distribution by conservatively multiple corrections. On the basis of the advantages, GWAS has rapidly become one of powerful and affordable tool to discover common risk variants of the complex diseases, and also has been successfully applied in ophthalmic field. Rather than giving an exhaustive review of all reported findings for myopia, this brief review will focus on recent work on GWAS and farther strategies in the post-GWAS era of myopia.

IMPLIEDATIONS OF GWAS WITH REFERENCE TO MYOPIA

Initial Results and Further Findings  In the first place, Nakanishi et al[4] conducted a two-stage GWAS survey in Japanese by typing 411 777 single nucleotide polymorphism (SNP) markers, recruiting 830 pathological myopia cases and 1911 general population controls. This earliest GWAS of myopia reported the strongest but not genome-wide significant association at SNP rs577948 on chromosome 11q24.1 \((P=2.22\times10^{-7})\). It was speculated that this associated locus located upstream of \(BLID\) gene was involved in mitochondrial apoptosis and then prompted mitochondrial regulatory
mechanism of myopia. This initial finding, however, has failed to be replicated in follow-up studies. There are two possible reasons for this: on the one hand, replicated studies could demonstrate a false negative (type II error, reject qualified applicants) due to insufficient statistical power to detect small genetic effects; on the other hand, in most cases where confounder of population stratification and overestimation of effect size are involved, the initial finding likely represents a false positive (type I error, admit unqualified applicants) rather than true associations. In addition, GWAS requires very stringent significance levels, that is, \( P \)-values less than \( 5 \times 10^{-8} \) to remain significant after Bonferroni correction for the very large number of genetic markers tested. In the circumstance, a great sample size was needed in order to detect robustly modest genetic effects that are typical for complex disease.

The paucity of causal variants identified has motivated action to expand sample size through empowering organizationally a number of international multicenter collaborative efforts. Two subsequent large-scale GWAS for common refractive error were performed concurrently each in a total of more than 15 000 European populations. These GWAS identified over 150 SNPs associated with myopia. Note that GWAS significant variants seem generally to be of low frequency and/or small effect with the allelic odds ratio range from 1.10-1.20\(^{[10]}\). However, it should be borne in mind that estimated of odds ratio is only a surrogate for the relative risk rather than the true genetic effect. Meanwhile, small effects can still uncover novel relevant insights into pathogenic mechanisms in a complex disease, due to natural selection, pleiotropic mutation, genetic drift and population history in evolution.

**Significant Progresses** Two of the heretofore largest GWAS of myopic refractive error were conducted independently and published successively by CREAM\(^{[12]}\) and 23andMe company\(^{[12]}\). In addition to confirming previously reported loci\(^{[5-6]}\), both two studies provided compelling results of additional associated loci linked to myopia and refractive error. The CREAM conducted a genome-wide Meta-analysis comprised of 32 across ancestry cohorts, and discovered 16 novel quantitative trait loci (QTL) associated with refractive error in 37 382 individuals of European origin, of which 8 were shared with 8376 individuals of East Asians\(^{[11]}\). The 23andME group reported results from the largest genome-wide survival analysis in a European derived population. The Cox model survival analysis of 45 711 discover samples identified 20 new loci, of which 10 were replicated in a separate cohort of 8323 samples with early onset myopia before 10 years old\(^{[12]}\). The comparison of two investigations indicated that CREAM and 23andMe overlapped with each other to a startling extent, as well as associated loci had consistent direction of estimated effects. Not surprisingly, such robust results could be replicated again in a Japanese population\(^{[13]}\). These discoveries strengthened the existing viewpoint of signaling cascade from the retina triggered to the sclera remodeled and then ultimately leading to eye growth. More recently, Tedja et al\(^{[14]}\) also further revealed that a light-dependent retina-to-sclera signaling cascade acted on refractive error by a GWAS in 160 420 participants and replication in 95 505 participants.

These salient studies have provided additional information to explore the etiology and pathogenesis of myopia. As compared to CREAM conventional acquirement of phenotype information, 23andME utilized questionnaire data which may generate substantial misclassification errors because of lack of validation. Nevertheless, the Cox proportional hazard survival analysis produced an increasing statistical power, benefiting from distributional flexibility to study a wide variety of censored traits such as age of onset. Despite methodological biases, replication has a crucial role in showing associations that are identified reflect interesting biological processes. In addition, a linear relationship between hazard ratio of 23andME and effect size of CREAM was established, where locus specific hazard ratio for myopia onset age would predict the degree of refractive error throughout life\(^{[15-16]}\).
initial attempt to predict risk has a limited role primarily due to the relatively small effect size of the significantly associated variants. Hence, risk prediction may not be a good recommendation before a larger proportion of the risk variants underlying the myopia have been identified. It is envisioned that the development of risk prediction algorithms, incorporating massive genetic markers and biomarkers with risk factors, will facilitate a promising clinical application to determine the exact individual risk of developing myopia.

**The Related GWAS of Myopia** These GWAS also have demonstrated additional insights to shed light on the association of the two major determinants of refractive error, ocular axial length and corneal curvature, with myopia. A Meta-analysis in 12 531 Europeans and 8 216 Asians identified nine genome-wide significant loci for axial length\(^{[17]}\); of which five associated with refraction (LAMA2, GJD2, CD55, ALPPL2, and ZC3H11B\(^{[11-12]}\) were replicated, and differential gene expression was further observed in myopic animal experiments. Another confirmation that linked the attractive phenotypes trait with myopia was showed in genome-wide significant associated variants (PDGFRA, MTOR, CMPK1 and RBP3) for corneal curvature; particularly, a missense rs11204213 of RBP3 indicated larger effects on both corneal curvature and axial length compared with others\(^{[18]}\). Although these locus were not reported previously in myopia GWAS, homozygous nonsense mutations of the RBP3 gene was found to be associated with high myopia\(^{[19]}\). A large-scale GWAS (n=86 335) for corneal astigmatism identified four novel loci and one of which (NPLOC4) has previously demonstrated association with myopia, suggesting further support for the shared genetic susceptibility of myopia and astigmatism\(^{[20]}\).

What’s more, Simpson et al\(^{[21]}\) observed two genome-wide significant regions on 15q14 and 8q12 for hyperopia, which overlapped with previously reported loci of myopia age at onset, indicating GWAS also have provided evidences for myopia and hyperopia as dichotomous refractive error traits underlying the enmetropization mechaniam.

GWAS in human complex trait have already proven a resounding success, which have underpinned effectively the outcomes of genetic variants associated with myopia. This represents a key milestone in myopia genetics. A number of new loci have been implicated in myopia phenotype and, in sharp contrast with linkage and candidate studies, showing predominantly consensus among fellow-up studies. Albeit many variant loci proved robust, almost all of them did not capture causal associations but rather only tagged a causative event in a specific region of the human genome. Thus, those crucial events would still have to be elaborated. For the time being, GWAS discoveries have significantly broadened our knowledge of the genetic basis of common forms of myopia development but they have yet to demonstrate clinical implications. Given this, a great deal more work will be needed to further explore unexplained information and understand the underlying biologic mechanisms of these genetic variants.

**ROUTE TO FOLLOW IN POST-GWAS ERA**

**Seeking Heritability** Although GWAS have successfully proven in identifying multiple genetic variants that contribute to myopia phenotype, these variants together account for only a minority of the observed heritability\(^{[14]}\). The missing heritability may arise potentially from undiscovered common variants that are concealed by stringent significance thresholds of GWAS, rare variants that are ignored for GWAS approach on basic of common disease/common variant hypothesis, structural variations in the genome such as copy number variation (CNV) that escapes from current genotyping platforms. With the sequencing technologies advanced and cost decreased rapidly, it will be feasible to utilize whole genome sequencing in numerous populations to identify both common and rare variants with a modest effect underlying the myopia traits\(^{[22]}\). As an important source of human genome variability, CNV is being explored in the context of myopia. Yip et al\(^{[23]}\) adopted a systematic strategy to investigate the role of CNVs in high myopia, and identified 22 significant CNVs which still are needed to further explored. One animal study showed that the CNV of muscarinic acetylcholine receptor genes (CHRM), especially CHRM3, were significantly different between control and myopia, even among various degrees of myopia\(^{[24]}\). Metlapally et al\(^{[25]}\) reported that TEX28 gene CNV appeared to be associated with the MYP1 locus in X-linked high myopia phenotypes. Beleggia et al\(^{[26]}\) performed an CNV analysis in two affected individuals from MACOM syndrome with severe myopia, and identified CRIM1 CNV as an important factor in eye development. Although CNV has been speculatively involved in susceptibility to various complex diseases in human, the effect of CNV in missing heritability for myopia remains largely undefined.

While many variants surely remain to be found, phantom heritability may be another hypothesis caused by huge overestimation with no consideration for genetic interactions\(^{[27-28]}\). Such the absence of heritability could be partly attributed to gene-environment, gene-gene or more specifically variant-variant interactions. A series of epidemiological studies investigated the interactions between the myopia genetic variants and the main environmental factors, and demonstrated that educational attainment and genetic effects had strong interactions. Fan et al\(^{[29-30]}\) provides evidence of the interactions among education stratum and GWAS-associated loci such as ZMAT4 presented in both Asians and Europeans. Such education-environment interactions have also been implied for susceptibility variants in MMPs\(^{[31]}\). A joint Meta-
analysis based on gene-environment-wide interaction study (GEWIS) demonstrated that several genome-wide significant loci (AREG, GABRI and PDEI0A) interacted significantly with education by identifying SNP-education interaction effects on refraction error, and the interactions are more evident in Asians[32]. Tkatchenko et al[33] identified a low-frequency variant in APLP2 associated with the children exposed to large amounts of daily reading using a combination of GWAS, gene set enrichment analysis (GSEA) and functional analysis of animal model. An novel phenotypic-genotypic interaction between myopia and intelligence has been investigated recently[34]. However, despite the application of statistical and computational methods is conducive to identification of nonlinear epistatic interactions of genetic effects[35], it remains challenging to identify small-effect variants for complex traits and reduce the burden of multiple hypothesis-testing. The most promising route for identification of variants lies through combining biological functional evidence with statistical genetic evidence.

Epigenetic Studies Epigenetic modifications in the human genome serve as a genetic mechanism by which environmental exposures modulate disease risk, as well as play diverse roles in gene expression and function at different molecular levels[34]. The most well-known epigenetic modifications are the DNA methylations, histone modifications, and non-coding RNA activity so far. Methylation at cytosine-phosphate-guanine (CpG) sites was one of major repressive epigenetic modification. It was recently revealed that hypermethylation of CpG in the COLIA1 gene promoter may underlie the reduction of sclera collagen synthesis and then the development of myopia[37]. The expression of COLIA1 mRNA was decreased at the transcriptional level in myopic mice, corresponding to an increase in the frequency of CpG methylation. By means of large-scale MicroRNA (miRNA) expression profiling in a myopic mouse model, Tkatchenko et al[38] identified that a number of miRNAs were involved in the regulation of refractive eye development, and most of which were differentially upregulated in the myopic retina. Xie et al[39] reported that rs157907 polymorphism G allele of miR-29a targeted gene COLIA1 was significantly associated with myopia as a protective factor, and speculated that rs157907 might regulate miRNA expression and thereby affect collagen synthesis by binding specific mRNAs. On the other hand, the vast majority of GWAS associated variants were located on non-coding intergenic and intronic regions. The Encyclopedia of DNA Elements (ENCODE) and other projects have provided ample epigenomic data for functional annotation of non-coding variants, and discovered the majority of the GWAS associated SNPs in connection with epigenomic elements[40]. With these data, we thus tried to perform a functional annotation of index SNPs and proxy SNPs which aimed to prioritize potential regulatory variants and susceptibility genes[41]. Despite the challenges faced, these preliminary explorations will provide clues to data mining and integration toward further understanding of etiologies and treatments.

Integrative Pathway or Network Analysis GWAS approach typically focused on single SNP-based association test suffering from low power if each tested marker is incomplete linkage disequilibrium with undefined quantitative trait loci. Nevertheless, the polygenic basis of complex traits implicated that epistasis and pleiotropy appeared to be inherent properties of biomolecular networks rather than isolated occurrences. This has motivated the interest in multi-locus-based systemic approach to integrate GWAS data and other data modalities to yield additional insight within a biological context[42]. Actually, pathway analysis has previously been performed within GWAS. The CREAM identified several novel pathways involved in myopia by considering all the genes identified in the text and using the Ingenuity Pathways Analysis (IPA) database and Disease Association Protein-Protein Link Evaluator (DAPPLE)[31]. The Wnt receptor signaling pathway was identified in a recent GWAS result for axial length from CREAM effort, further reinforced that the signaling pathway plays a prominent role in vertebrate eye development[18]. Some studies have integrated visually significant genotype-phenotype associations with gene annotations databases to build pathways. The miRNA-mRNA interaction networks or functionally collaborative networks also have been conducted to identify the potential signaling pathways involved in form-deprivation myopia models. For example, Tkatchenko et al[40] found that nine signaling pathways were involved in regulation of neurogenesis; Mei et al[41] discovered that the regulation of transcription, axon guidance and TGF-β signaling pathways were significantly enriched. Meanwhile, it was suggested that miRNAs may serve as key regulators of the signaling cascades related to the development of myopia. Reconstruction models of regulatory network, constituted by binding events of transcription factors, might help understand and interpret the roles of genetics and epigenetics in myopic mechanism on the other hand. Despite of so much inaccurate and incomplete, the dynamic context-specific nature (distinct combinations of factors bind at specific genomic locations) of regulatory network is beginning to take its role in dissecting the genetics pathogenesis. Pathway analysis will next be extended to examining rare variants, other omics and interaction data. Through long-term exploration and unremitting efforts, a framework for unraveling the genetic basis of complex traits has just been initially established. For myopia genetics research, the present achievements are only the first step in this process and, ever larger studies would undoubtedly result
in more genetic discoveries but smaller effects. One challenge is how to tackle the fine mapping and functional dissection of already-identified GWAS loci. Furthermore, increasing emphasis will be placed on biological understanding and personalized discovery of diagnostics and therapeutics in clinical settings. Even so, its phenotypic predictability remains very low. New methodologies and perspectives will be needed to fully tackle related problems. The promising route for identification of missing low-frequency and small-effect variants lies through combining biological functional evidence with statistical genetic evidence. Identification of remaining trait variance will acquire additional discoveries, specially underlying rare variants and causal common variants and refined estimates of heritability. Functional validation, integrating the growing genetic and omics data, will produce omnibearing analysis of biological pathways, gene regulation networks and protein interaction maps. The improvement of molecular genetics combined with other methods is expected to become widespread medical application in humans in the end.

ACKNOWLEDGEMENTS

Foundation: Supported by Projects of Science & Technology Department of Sichuan Province (No.2019YJ0381).

Conflicts of Interest: Liao X, None; Tan QQ, None; Lan CJ, None.

REFERENCES

Myopia genetics in GWAS and post-GWAS era


42 Kao PY, Leung KH, Chan LW, Yip SP, Yap MK. Pathway analysis of complex diseases for GWAS, extending to consider rare variants, omics and interactions. *Biochim Biophys Acta Gen Subj* 2017;1861(2):335-353.