Integrated genome-wide analysis of human facial morphology

by

Dongjing Liu

BS, Beijing Normal University, 2014

MS, Peking University, 2017

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Dongjing Liu

It was defended on

September 21, 2020

and approved by

Seth M. Weinberg, PhD, Associate Professor, Department of Oral Biology School of Dental Medicine, Department of Human Genetics Graduate School of Public Health, University of Pittsburgh

Danial E. Weeks, PhD, Professor, Departments of Human Genetics & Biostatistics, Graduate School of Public Health, University of Pittsburgh

Jiebiao Wang, PhD, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Dissertation Director: John R. Shaffer, PhD, Assistant Professor, Department of Human Genetics Graduate School of Public Health, Department of Oral Biology School of Dental Medicine, University of Pittsburgh

Integrated genome-wide analysis of human facial morphology

Dongjing Liu, PhD

University of Pittsburgh, 2020

The human face is a highly multipartite structure resulting from the intricate coordination of multiple factors. The high heritability of facial morphology has long been appreciated, yet little is known about the contributions of specific genes. Knowledge of the genetic architecture of facial morphology is important for understanding craniofacial morphogenesis and how these processes contribute to craniofacial disorders. Studying facial genetics may also provide a basis for DNA-informed facial prediction, which has several real-world applications.

Genome-wide data on well-characterized human cohorts has great potential for generating novel insights in the post-Genome-Wide Association Study (GWAS) era. Moving beyond the conventional single variant-single trait association in GWAS, this study analyzed existing genome-wide data using three different approaches to glean insights into facial morphology, by leveraging state-of-the-art advances in 3D facial phenotype modeling and multivariate statistical approaches. Specifically, analysis in Aim 1 for the first time demonstrated the contribution of rare and low-frequency coding variants in facial variation, with eight genes being significant associated, one of which (*NECTIN1*) had known craniofacial function. Transcriptome-wide association analysis in Aim 2 extended a previous GWAS effort by refining potential causal genes for future functional characterization. Findings from this aim also pointed to several novel candidate genes. Finally, analysis in Aim 3 explored the role of Variance Quantitative Trait Loci and highlighted the importance of studying facial variability in addition to facial mean differences in gene discovery and mechanistic exploration. The variance prioritization strategy adopted in this aim also demonstrated its advantage in detecting gene by gene interactions involved in facial morphology. These results expanded our understanding of the genetic basis of normal-range facial morphology and will have important implications for future studies. The application of several recently developed statistical tools also helped to evaluate and generalize their utility to multivariate settings and identify their limitations.

This study had public health relevance. Our findings can help provide a roadmap for understanding the genetic underpinnings of craniofacial morphogenesis and birth defects, pave the way for advances in personalized prevention and therapeutics of related conditions, and inform DNA-based facial prediction for clinical and forensic application.

Table of Contents

List of Tables

List of Figures

Preface

I am grateful to all of those who have helped me along the way during my PhD journey. Without them, I would not have been able to complete this dissertation project and would not have made it through my PhD degree. I relish all I have learned and all the friends I made along.

I would like to express my deep appreciation and gratitude to my dissertation committee, Dr.John Shaffer, Dr.Daniel Weeks, Dr.Seth Weinberg and Dr.Jiebiao Wang. Their insightful guidance throughout have made undertaking this project an inspiring experience for me. As my chairman, Dr.Shaffer oversaw all aspects of this project and provided a great amount of invaluable feedback to guide my work. Dr.Weinberg has extensive domain knowledge into the subject matter - human craniofacial anatomy and development - which steered me through this research. Dr.Weeks is thoroughly knowledgeable in statistical genetics and has helped me substantially in ensuring the statistical aspects of my research to be solid and sound. Dr.Wang is a real expert in biostatistics and bioinformatics and has given me plenty of actionable suggestions to improve my work. Their wise guidance contributes a lot to my intellectual growth during my years of study here. Thank you to all my committee members, without your wise guidance this project would have not been the same. Beyond this project, each one of my committee members have provided me extensive personal and professional guidance, taught me a great deal about doing science, and have been extremely supportive of my career goals.

Special thanks to my supervisor, Dr.John Shaffer, for the patient guidance and mentorship he provided to me, all the way from when I first applied to the PhD program in the Human Genetics Department, through to the completion of this degree. Dr.Shaffer has taught me more than I could ever give him credit for here. His scientific brilliance and knowledge have immensely helped me in making progress in all the research projects of mine; he was always willing and enthusiastic to help me in any way he could. Thank you for being a truly amazing mentor.

Whole-heartedly I would like to say thank you to everyone at CCDG, the research center where I conducted most of my PhD research work at. Thank you to Dr.Mary Marazita and Dr.Eleanor Feingold, from whom I always see passion, enthusiasm and energy in genetic research, which motivated me to offer my full effort to every project. Thank you to all the staff members there for your collaborative effort during data collection, cleaning and management. Your work has provided me with the necessary resources that I needed in my research. I am truly fortunate to have had the opportunity to work in such a fabulous team and get to learn something from each one of its members. I am also thankful to my lovely colleagues in the Shaffer lab who together maintained a warm and supportive lab culture and have made my lab experience very enjoyable.

I am thankful to the Department of Human Genetics and all the professors and staff members for providing me with an excellent environment to learn, to research, to be curious, and to make achievements. I appreciate everyone being so caring about the students in the department. Also thank you to my fellow doctoral students for your feedback and friendship.

I cannot forget to thank my beloved family and friends for all the unconditional support in my endeavours. You earned this degree right along with me.

I have had great pleasure to work with so many talented people during this and other related projects. They have shown me by their own example what a good scientific researcher should be. It is sincerely appreciated that the role models they set up direct the path for me to keep making progress toward my goals.

This was an enriching journey and I relish all I have learned and all the friends I made along.

1.0 Overall Research Goal and Specific Aims

The human face is a highly multipartite structure resulting from the intrinsic complexity of facial morphogenesis and the intricate coordination of multiple factors that impact facial morphology over the lifespan. The strong genetic basis and high heritability of human facial morphology has been long appreciated, yet relatively little is known about the contributions of specific genes to facial features [1]. Knowledge of the genetics underlying normal-range facial variation is important for understanding the mechanisms of craniofacial morphogenesis and processes leading to aberrant craniofacial development, and will further provide the foundation for developing effective public health interventions. Genome-wide data of carefully phenotyped human cohorts have yielded substantial knowledge in genome-wide association studies (GWASs) of facial traits. Beyond the conventional single variant-single trait association analysis in GWAS, those genome-wide data can be analyzed in many other ways to gather additional insights into facial variation. This dissertation project focused on three such analyses to study facial genetics in the post-GWAS era.

The overall objective of the current project was to promote the genetic understanding of human facial variation. At their core, the strategies to achieve this goal were a series of genotype-phenotype association analyses that leveraged existing data resources and state-of-the-art advances in 3D facial phenotype modeling. Specifically, this project comprised of three aims to improve knowledge of the genetic architecture of facial morphology:

Aim 1: Interrogate the role of rare and low-frequency coding variants in facial morphology through an exome-wide gene-based association analysis

Aim 2: Identify putative functional genes underlying the SNP-face association via the integration of expression quantitative trait locus (eQTL) data with GWAS data, with specific analyses including transcriptome-wide association study (TWAS), conditional analysis and colocalization analysis

Aim 3: Examine the genetic effects on facial phenotypic variability by conducting a genome-wide scan for Variance Quantitative Trait Loci (vQTLs), and explore the underlying mech-

anisms for the identified vQTLs with a special focus on interaction effects

Studies of other human polygenic diseases and traits have demonstrated the usefulness of these analysis strategies [2–9]. By applying them for the first time to the field of facial genetics, this project represented critical steps to make full use of the genome-wide data on well-characterized human cohorts, the results of which greatly forwarded our understanding of the genetic architecture of facial variation. The application of various statistical genetic approaches provided opportunities to evaluate their usefulness for polygenic, multivariate, and morphological traits, which will help inform future study of other complex phenotypes. Knowledge generated from this study can help provide a roadmap for understanding the genetic underpinnings of craniofacial phenotypes and birth defects, pave the way for advances in personalized prevention and therapeutics of related conditions, and inform DNA-based facial prediction for clinical and forensic application.

2.0 Introduction

2.1 Face development

Human facial morphogenesis is the result of an intricate coordination of a series of precisely timed embryonic events. The process involves extensive interaction and movement of the derivates of all germ layers. The majority of anterior facial tissues, including cartilage, bone, and musculature are derived from the cranial neural crest cells (CNCC) which migrate ventrally into the first pharyngeal arch and frontonasal process [10]. Pharyngeal arches are the earliest primordia of face and appear at around 4th week of gestation. The first arch is of primary relevance to face and it later splits into upper maxillary and lower mandibular swellings (aka processes or prominences), which will give rise to the cartilage, bones, and connective tissue of the upper and lower jaw, respectively [11]. The maxillary and mandibular prominences are initially bilateral, and increasingly grow towards the midline as embryonic development continues. The frontonasal prominence, derived directly from CNCCs, give rise to the lateral and medial nasal processes which will later be separated from the maxillary prominence by deep furrows. The inward growth of the two maxillary prominences pushes the two medial nasal prominences together such that they fuse to form the midline of the nose and philtrum of the upper lip. The mandibular processes also grow together to form a single mandible. All facial prominences are fully integrated by the 10th week of human embryonic development. Dworkin et al. [12] devised a schematic diagram of human facial development (Figure 2.1) displaying the above processes. Table 2.1 relates the embryonic structures to the corresponding components of the adult face. Disruption of the development of these facial prominences can result in a variety of facial anomalies including orofacial cleft, cleft lip, frontonasal dysplasia, etc.

Many studies have contributed to the revelation of pathways and genes involved in craniofacial development. Some critical regulators include SHH [13, 14], BMP [15, 16] and FGF families [17, 18], and there is extensive co-regulation and interaction among them [19, 20]. A

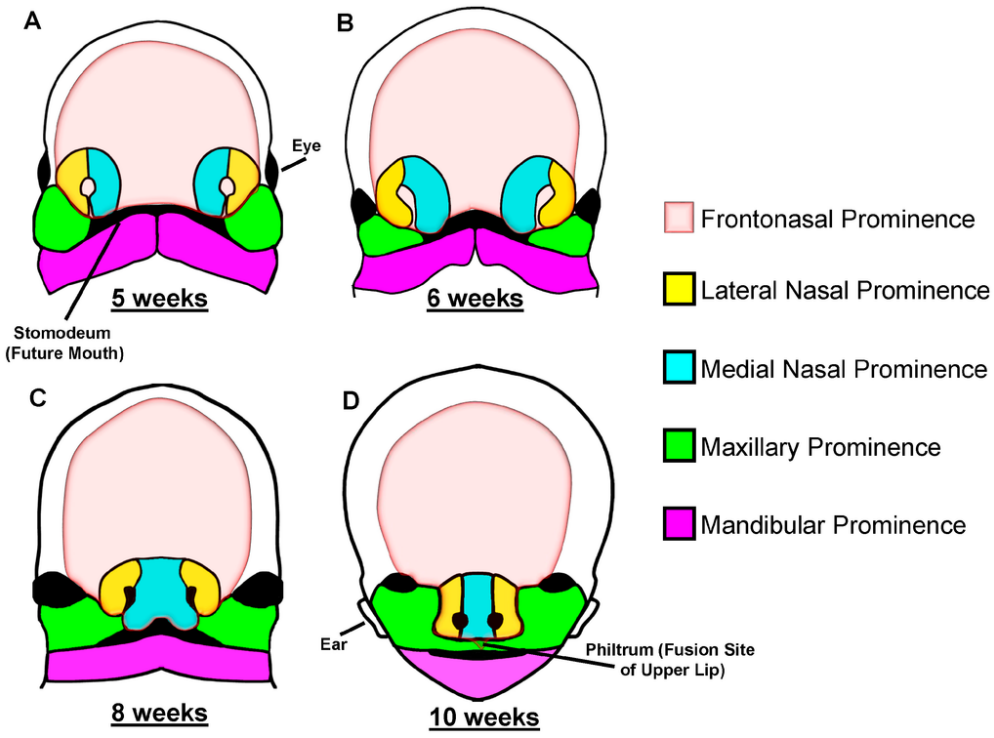


Figure 2.1: Human facial development. Reproduced from Dworkin et al. 2016 [12].

Table 2.1: Contribution of the embryonic prominences to the adult face

Prominence	Adult structure
Frontonasal	Forehead, nose bridge
Medial nasal	Nose midline, philtrum
Lateral nasal	Alae of nose
Maxillary	Cheeks, lateral upper lip
Mandibular	Lower lip, jaw

large number of genes are expressed and involved in the formation of the face [21]. Given the complexity and the multipartite nature of head and face, the molecular mechanisms crucial for the patterning and formation of craniofacial structures are not yet completely understood.

2.2 A review of human facial genetics study

2.2.1 GWAS of normal-range facial variation

Human facial morphology is a highly variable, complex trait with strong genetic influence as demonstrated by animal studies [22–24], human genetic syndromes [25], and heritability studies [26–29]. For more than a decade, significant progress has been made in elucidating the genetic basis of normal-range facial shape variation by the collective efforts from various research groups [1]. Table A1 summarizes the genes that have been identified in GWASs of facial morphology to date together with their putative biology functions. The initial successes largely featured GWASs on well-characterized cohorts with both 3D facial images and genomic data [30–38]. These studies derived a variety of facial biometrics from three-dimensional facial images or head magnetic resonance images, including facial landmark distances, angles, curvatures, and composite variables as a more comprehensive representation of shape. Faces studied came from several major populations around the world (European, African, Latin American and Asian) and span a wide age range, with the typical size of studies being a few hundreds to several thousands. In addition to cohorts of unrelated participants, some studies took efforts to phenotype relatives and focused primarily on highly heritable facial aspects in order to expedite gene mapping [32]. All of these efforts have together led to the identification and replication of a fair amount of common genetic variants influencing facial morphology. Some of the identified variants/genes are being further investigated in functional experiments to verify their role in craniofacial morphogenesis.

Despite the substantial progress, relatively little is known about how specific genetic

variants influence the phenotypic expression of many facial traits. Few of the GWAS hits have been rigorously pinpointed to a specific gene, and roles of these genes in craniofacial morphogenesis are yet to be experimentally validated. The highly complex nature of craniofacial morphogenesis and growth also indicates the existence of many undiscovered genetic factors. The infamous issue of missing heritability has motivated researchers studying other complex phenotypes to look beyond the marginal effects of common SNPs that are routinely studied in GWAS [39], yet so far little such effort has gone into facial genetics. Besides efforts toward larger cohorts and meta-analysis, strategies of tracking the genetic contribution missing from existing GWASs can be equally useful and worth pursuing in parallel.

2.2.2 Gene discovery was stymied by the suboptimal quantification of facial shape

Modeling three-dimensional geometry has long been a challenge in facial gene discovery. The traditional quantification of facial features gives simple linear distances or angles based on landmark points on facial surface. Such measures rely heavily on the placement of landmarks and do not represent well certain areas of the face, resulting in a great loss of geometric information. Moreover, they are chosen *a priori* based on convenience rather than biological relevance. Under the conventional phenotyping scheme, the entire facial shape was analyzed as a combination of sparse univariate "surrogates" that are far from complete and efficient. Early GWASs analyzing these distance/angle phenotypes individually usually only generated no more than a dozen significant associations, despite with tens of facial features tested [34,37,40]. In addition, there is no simple solution for balancing the trade-off between the number of traits to include and the burden of multiple testing in statistical analysis, which has hindered research progress.

2.2.3 A more efficient phenotyping approach

In order to improve upon the standing landmark-based strategy with more effective approaches, our group developed a comprehensive global-to-local modular phenotyping system

in 2018 [31]. This innovative approach was based on the modular organization of the face as a natural consequence of processes involved in craniofacial morphogenesis. The new representation of the facial morphology is a hierarchical framework that partitions the 3D surface into nested segments arranged from global to local, with the full face at the most global level and sets of confined facial segments arranged towards the most local level. This organization allows for an open-ended scan for genetic variants exhibiting a range of effects at different scales. The shape variation within individual facial segments was extracted through principal component analysis (PCA) and represented by multiple principal components (PCs), which enables subsequent statistical analysis to be done in a multivariate framework. These information-dense multivariate phenotypes were shown to be much more effective in gene mapping compared to the traditional sparse univariate traits. We were able to identify 38 loci in a cohort of 2,329 individuals by analyzing these multivariate facial phenotypes [31], demonstrating a greater success despite a smaller sample compared to prior works on univariate facial features [34, 37, 40].

2.3 Rationale for the specific aims

2.3.1 Aim 1: Rare and low-frequency coding variants analysis

Rare and low-frequency variants are commonly defined as variants with a minor allele frequency (MAF) below 1% and in between 1 and 5%, respectively. They have been suggested to account for part the missing heritability in both theoretical and empirical studies [41–44]. Conceptually, variants that heavily predispose individuals to diseases are under strong purifying selection pressure and thus are kept at a low frequency [45–47]. Evolution dictates an inverse effect size-population frequency relationship, which is widely observed and well appreciated as a common law for most complex traits, although not without exception (Figure 2.2). GWAS was not designed for studying these variants; fortunately, lessons on their effects are beginning to emerge from exome-wide and genome-wide sequencing studies. A number of studies have explored the contribution of rare and low-frequency variants in various hu-

man traits and disorders [3, 48–52]. For example, rare coding variants were estimated to account for nearly 5% of the heritability of multiple sclerosis risk [50]. An exome-wide study of human height identified 83 coding variants with frequencies below 5% and large effects of up to 2 centimeters per allele. About 1.7% of the heritable variation of adult height was attributed to rare and low-frequency coding variants. A more recent study with an exclusive focus on variants below 0.1% reported 64 statistically significant genes for a variety of human phenotypes in over 70,000 exomes, using electronic health records [48]. These studies serve as a proof-of-principle in support for the general involvement of low-frequency variants in complex human traits.

Studying rare and low-frequency variants also has great potential in unraveling additional core genes and functions which may lead to more direct insights into disease biology. Coding variants are more interpretable and experimentally tractable than non-coding variation. This feature complements well with GWAS, where identifying and characterising causal variants and mechanism presents remarkable challenges. As shown in Figure 2.2, detectable low-frequency variants are expected to have larger effect sizes compared to common SNPs, making the variants more amenable to experimental verification. The aforementioned height study followed their findings from the human association data and pursued *in vitro* functional validation for *STC2*, demonstrating how they were able to trace the molecular path from a single-base DNA substitution to changes in protein binding activity and eventually to an altered phenotype in the expected direction at the organismal level [2]. Another successful example was reported in Liu and colleagues' exome-wide study, where the statistical association between *A1CF* and blood triglyceride levels was strongly supported by their functional work in human cell lines and mouse models [51]. The authors observed reduced secreted APOB transcript levels in human hepatoma cells with induced deletion of *A1CF*, and further found that knock-in mice had 46% higher triglyceride than the wild-type mice. Altogether, these results illustrate the biological value of studying rare and low-frequency genetic variants.

Detecting variants that fall to the left in the allele frequency spectrum presents special challenges. Insufficient power and difficulty in replication are among the most prominent

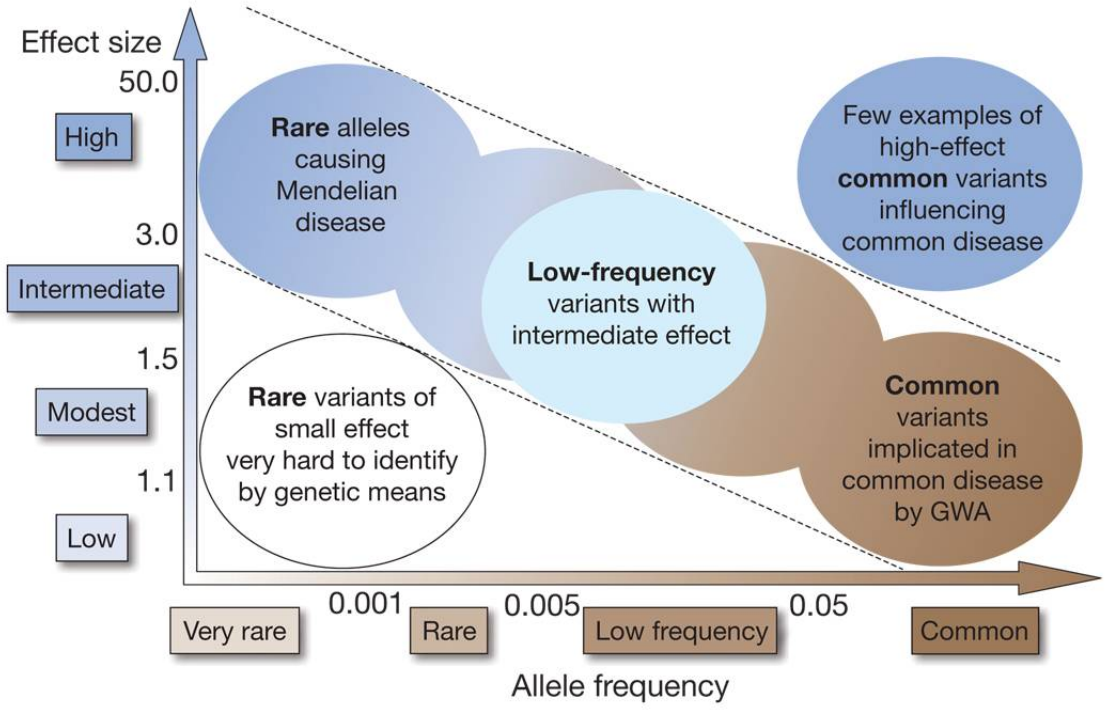


Figure 2.2: Variant effect size-population frequency relationship. Reproduced from Manolio et al 2009 with permission [39].

ones, both stemming from the scarcity of variant carriers in the populations. Classical single-variant tests would have at best modest power given the typical sample size for rare and low-frequency variant studies. A simple yet effective way to increase power is to use collapsing approaches [53]. The idea is to aggregate effects by binning qualifying variants into groups, usually combined with strategies to enrich risk effects by only focusing on a certain class of variation, such as nonsynonymous variants, loss of function variants or variants predicted to be deleterious. Units of association are usually selected to be biologically relevant regions such as genes, exons, enhancers, and pathways. This collapsing strategy has been widely adopted by many rare variant studies to date. The other challenge, replication difficulty, is a more intrinsic issue. By nature, rare and low-frequency variants tend to be population-specific and are likely to be absent in study cohorts different from the one where they were originally detected. When they do present in multiple cohorts, a meta-analysis can help. Tools specifically designed for meta-analyzing rare variants including RAREMETAL [54, 55]

and MetaSKAT [56] have demonstrated their usefulness in identifying additional variants and genes. Some studies managed the replication problem by taking another route - skipping statistical replication and seeking directly for functional validation [2], which may lend even stronger support although being demanding on resources.

There are a good number of statistical methods and associated software packages specifically configured for collapsing tests, most of which fall into two main categories: burden tests and variance-component tests [57, 58]. The burden test and its generalizations include CAST [59], CMC [60], WSS [61], aSum [62], VT [63], etc. These methods assume that all variants in a test unit influence the phenotype under study in the same direction. Variance-components type of tests, such as sequence kernel association test (SKAT) [64], SSU [65] and C-alpha [66], lift this assumption and are more powerful in the presence of both trait-increasing and trait-decreasing variants, or when not all variants are causal. When there is limited prior knowledge of the underlying genetic architecture, omnibus tests such as SKAT-O [67] and Mist [68] can be more effective by combining evidence from both burden and variance component tests. The success of these approaches in identifying causal genes for a variety of human traits has illustrated the applicability of collapsing analyses over diverse disease architectures, study designs and sample sizes [2, 69, 70].

Rare and low-frequency variants have not been studied for associations with human facial features. Based on findings from other polygenic morphological traits, it is reasonable to hypothesize that rare and low-frequency variants also contribute to facial morphology. This aim therefore asked several key questions concerning the genetic architecture of facial morphology: 1) Do rare and low-frequency variants have a role in facial morphology? 2) How big is their role? 3) Which genes are implicated? 4) Do these genes overlap with those already implicated by common variants in prior GWASs? 5) Do these genes also play a role in craniofacial anomalies, in particular orofacial clefts? Answering these questions will broaden the allele frequency spectrum of genetic variants ever studied and give a more comprehensive view of the genetic architecture of facial morphology.

2.3.2 Aim 2: Transcriptome-wide association study

Despite the substantial progress made by GWAS, top SNPs at GWAS loci often reside in non-coding regions and the identified trait-associated loci often harbor multiple genes, making it challenging to identify the causal ones [71]. This is in part due to the difficulty connecting regulatory variants to their target genes in the right cell types and right physiological contexts using existing data. This missing functionality issue represents one of the biggest challenges in GWAS and has drawn great research efforts worldwide. A major area of research with the intention to tackle this challenge involves the integration of functional genomics datasets with the genome-wide data [72]. Multi-omics data, such as genomics, transcriptomics, proteomics, epigenomics, metabolomics and so on, can be not only useful in establishing the molecular or statistical links across complex biological networks, but also in inferring the intermediate molecular profiles of individuals on whom not all of these data are collected. This effort is made possible by the rich data resources generated from consortia such as GTEx, ENCODE and Roadmap projects [73–75]. A wide range of statistical methodologies have been developed or extended based on this idea, which includs methods falling under the categories of fine-mapping approaches, transcriptome-wide association study (TWAS), colocalization analysis, Mendelian Randomization with the exposure of interest being molecular phenotypes, and methods for inferring disease tissues and cell types.

TWAS originates from the idea that non-coding risk variants can influence individuals' genetic propensity through their regulatory role on gene expression. TWAS tests for the association between a trait and the genetically regulated gene expression (GRex) predicted from a set of pre-computed gene expression weights and either the individual-level genotype data or summary-level association statistics [76]. The pre-computed weights, also referred to as expression prediction models, have to be generated from matched genome and transcriptome data on the same individual. GTEx [73] is one of the major ongoing projects collecting, housing and updating such data, and has greatly promoted the discovery of abundant cis expression quantitative trait loci (cis-eQTLs) with a small number of trans-eQTLs as well. Cis and trans refer to whether an eQTL is locally-acting or remotely-acting on its target gene.

Trans-eQTLs are especially difficult to pinpoint because they can be located anywhere in the genome and tend to have only modest effects on gene expression. Although they are known to explain important expression variation for many genes, most existing TWAS tools are not able to assess those distant genetic regulation effects due to the limited number of discovered trans-eQTLs. At the core of these TWAS tools are the expression prediction models, which were already provided by others and a user will only need individual level genome-wide data or summary level GWAS statistics to run a TWAS. Since its first introduction, TWAS has been widely applied and has shown its usefulness in studying diverse kinds of diseases and traits [5, 6, 77–79].

The TWAS design confers several strengths: (1) there is a reduced multiple testing burden compared to GWAS because the number of genes is substantially lower than the number of common SNPs; (2) by using genes as testing units, results from TWAS are more biologically interpretable; (3) predicting expression from genome data is much less costly than measuring expression by RNA-sequencing, which makes its application possible on biobank-scale sample size; (4) predicting gene expression from (loosely speaking) fixed genotype obviates the difficult argument on whether changes in expression level cause or follow from disease onset, whereas a direct mRNA-trait association analysis is inevitably prone to the issue of reverse causality; (5) directionality of the gene-trait association is apparent in TWAS, which can give hints for the mechanisms of the genetic effect.

The intermediate molecular trait does not have to be expression; a variety of other types of QTLs, such as methylation QTL, splicing QTL, and protein QTL can all be used in a similar way. They are less popular than eQTLs simply because that expression is the first step in gene regulation and the field has not gotten far enough beyond expression. Nevertheless, the body of research on various molecular phenotypes is growing fast, and in fact alternative splicing prediction has already been incorporated into some TWAS software [80]. Several implementations of TWAS are available. PrediXcan [81], the ever first TWAS software, initially trained elastic net expression prediction models and required individual level data on its application cohort. Later, the authors derived a mathematical formulation which achieved comparable results using only GWAS summary level data (S-PrediXcan) [82]. More

recently, PrediXcan further updated their models using a different strategy, Multivariate Adaptive Shrinkage in R (MASHR), for eQTLs on the GTEx v8 release data [80]. The MASHR-based model takes advantage of the correlation of gene expression between tissues and has a superior performance than those early versions trained by elastic net.

Another popular TWAS tool which also takes in summary statistics is FUSION [83]. FUSION adopts a Bayesian predictor to impute gene expression from genotypes, and incorporates a convenient function to run conditional analysis for examining how much of the GWAS signal remains after the gene-trait association is adjusted for. It also provides an omnibus test for effects across multiple reference panels. The application of FUSION in a large prostate cancer study identified 217 genes at 84 independent loci, several of which were novel, i.e. with no nearby GWAS signals [5]. The study also emphasized the importance of alternative splicing in prostate cancer risk by using splicing QTL data. The authors further extended their method to include chromatin marks, which was shown to be able to sort out candidate regulatory elements for the phenotype [6]. Their work represented a successful application of TWAS where it was able to discover a considerable number of novel genes that do not overlap with known GWAS loci.

More recently, a nonparametric Bayesian approach, Transcriptome-Integrated Genetic Association Resource (TIGAR), was developed specifically for modelling the complex genetic architecture of transcriptome [84]. The authors showed in the simulations as well as in the real-data analysis that TIGAR was able to fit transcriptomic imputation models for more genes and to achieve an improvement in expression imputation accuracy. Shortly after TIGAR was developed, the authors further extended the Bayesian TWAS approach to leverages both cis- and trans-eQTL information for TWAS [85]. By applying their new method to several Alzheimer's dementia datasets, they identified a novel risk gene that was completely driven by trans-eQTL.

There has not been TWAS conducted on human facial traits before this dissertation project. Given the rapid progress on the methodology and its successful application to other complex traits, TWAS will likely also benefit the gene mapping efforts for human facial morphology.

2.3.3 Aim 3: Variance quantitative trait locus analysis

Human genetic studies of quantitative traits so far have overwhelming focused on detecting variants whose genotype groups have differential mean values of the phenotype, and have been very successful in mapping QTLs and estimating their effects averaged in a population. On the other hand, the variability of human phenotypes across genotype groups has not received much attention despite it also being genetically regulated [86, 87]. Phenotypic variation being a phenotype on its own and the genetic control of it are indeed widely appreciated concepts in quantitative and evolutionary genetics, agriculture, animal breeding and plant genetics [88–90], but much less so in the field of human genetics. With the contribution of variability to overall phenotypic variation being ignored, traditional GWAS may potentially miss an important axis of genetic variation underlying individual differences [88].

Phenotypic variation may stem from developmental plasticity, a phenomenon wherein individual genotypes are able to produce distinct phenotypes in different genetic or environmental contexts [91]. The ability to produce plastic responses under heterogeneous conditions is a fundamental requirement of evolutionary processes, may confer adaptive advantages and thus be promoted by natural selection [92–94]. On the other hand, being able to produce a stable and robust phenotype under environmental and genetic perturbation is also an important feature for organisms to persist during evolution [95, 96]. The intricate balance between these two opposite forces is not yet fully understood [97]. Robustness and plasticity can be studied in model organisms but not in human populations using experimental approaches, which partly explains why human geneticists have not been as much involved.

The genetic basis of phenotypic variability can be examined by means of variance quantitative trait locus (vQTL). SNPs at vQTLs exhibit interindividual intragenotypic variability, i.e. one of the alleles is associated with a larger phenotypic variance compared to the other. vQTLs have been shown to be as common as quantitative trait loci (QTLs) in *Arabidopsis* [89], yet their occurrence and effects remain largely uncharacterized in humans. In traditional QTL studies where means of a trait are compared among genotype groups, any differences in trait variation are treated as a nuisance, or assumed to be absent. On the contrary, vQTL studies consider such differing variation as the primary target of investigation.

vQTLs may or may not also have an effect on the phenotypic mean, and those without QTL effects would be missing from conventional association studies. Phenotypic variability and vQTLs are topics that have remained largely unexplored for human facial morphology.

Differential phenotypic variance may arise from interaction effects [98, 99]. Interaction happens when trait-influencing factors crosstalk with each other as they operate, rather than act independently. Interaction between genetic factors is also known as epistasis. There is a consensus that the undetected interaction represents a critical source of missing heritability [100]. However, as a non-additive genetic effect, interaction is so far a much under-researched area due to low statistical power and the difficulties in measuring environmental fluctuations. Polygenic traits are influenced by numerous genetic and non-genetic factors, which in combination may give rise to vast and usually unknowable interaction map. For practical reasons, most association studies focus solely on the main effect of contributing factors, leaving interaction effects under-researched for many diseases/traits.

There are formal statistical tests for interaction, but a shortcut is to take advantage of vQTLs. Certain types of gene by gene interaction (G $\times$ G) and gene by environment interaction (G $\times$ E) lead to unequal variance of a quantitative trait across different genotype groups [98, 99, 101]. A graphical illustration can be seen in Figure 2.3, where the increased variance is the result of alleles at one SNP moderating the effects of alleles at a second SNP. The same principle applies to cases where the interaction involves a SNP and an environmental factor. This interaction-variance heterogeneity connection can therefore be used to prioritize SNPs for formal interaction tests. A number of studies have demonstrated the usefulness of this strategy. For example, a study found that a vQTL in *FTO* gene conferred a difference of 0.5 kilograms in standard deviation of weight between opposite homozygous groups, and this effect was consistent with the reported *FTO*-environment interactions for BMI [102]. The same approach was applied to 13 quantitative traits in UK Biobank samples, and a significant enrichment of G $\times$ E effects was observed among the detected vQTLs [8].

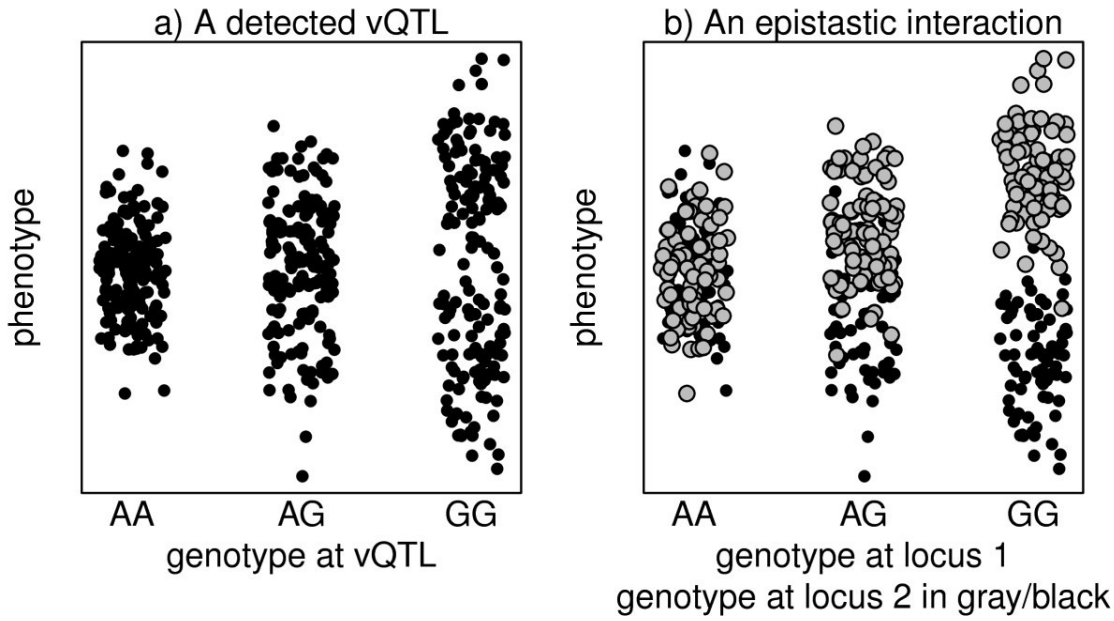


Figure 2.3: Relation between a vQTL and a G × G. Reproduced from Ronnegard et al 2012 [103]. (a) Phenotype values stratified by genotype at a hypothetical vQTL. (b) Variance heterogeneity could have arisen through a G × G with a second locus that segregates two genotypes (black and gray).

Molecular phenotypes such as gene expression, DNA methylation and protein levels have also been the subject of vQTL and/or interaction studies [86, 101, 104, 105]. For example, significant vQTLs were identified for two cardiovascular protein biomarkers in Women's Genome Health Study, where the observed variance heterogeneity was suggested to be explained by interacting effects between vQTLs and BMI/smoking [101]. It should be noted that interaction is one of the explanations of variance heterogeneity; other mechanisms can also manifest as differential variance. Although vQTLs and interaction effects are neither sufficient nor necessary to each other, a detected vQTL does signify the possible presence of unmodeled statistical interaction, which can be examined further.

Studying variance effects has important medical implications. Many diseases emerge as their underlying quantitative traits exceed phenotypic thresholds, such as obesity and hypertension, which are defined based on high BMI and blood pressure, respectively. For diseases without a simple quantitative diagnostic criterion, the liability threshold model of diseases proposes that there exists an underlying liability which summarizes all disease-related ex-

posures for an individual. This liability can be thought of as the analogous quantitative phenotype. At a vQTL, genotypes that produce a more variable phenotype will have a larger proportion of individuals exceeding the dividing threshold than other genotypes, leading to higher disease risk even when different genotype groups do not differ in the mean of the continuous phenotype or liability. Thus, alleles with extreme variability may be harmful to an individual and increase disease risk in the population.

2.4 Public health relevance

Evidence so far has suggested an overlapping genetic control between normal facial morphogenesis and abnormal facial development [106–108]. Studying the genetics underlying normal-range facial variation, therefore, can improve our knowledge of the complex relationship between genotype and phenotype in craniofacial syndromes and birth defects. These conditions are a source of major morbidity with both physical and psychosocial consequences imposing great public health burden [109, 110]. Knowledge on normal and abnormal craniofacial development will be essential for devising novel clinical approaches to prevent, treat and manage craniofacial conditions.

Another promising application of the genetic knowledge on facial morphology is DNA-based facial prediction and reconstruction, which has been accumulating enormous interests in clinical and forensic settings. Personalized planning of craniofacial surgery and orthodontics procedures, by leveraging the natural craniofacial growth informed by one's genotype, is already shifting from theory to reality [111]. Likewise, genetic facial prediction for forensic purposes is an area of interest to law enforcement [112]. The recovery of face from DNA remains challenging because the complex molecular and environmental interactions involved in facial development is still incompletely understood [113]. Research to better understand the genetic basis of human facial shape will provide the necessary scientific foundation for DNA-based facial profiling and ultimately inform such real-world applications.

3.0 Cohorts and Data

3.1 Cohorts and genotyping

This dissertation project involved three human cohorts consisting of unrelated, healthy individuals of different ancestries. (1) The PITT cohort included 2329 Europeans collected in Pittsburgh, Houston, Seattle and Iowa City, USA. The cohort ranged in age from three to 49 years and had a female proportion of 62%.(2) The Denver cohort had 664 Hispanic White individuals collected at Denver and San Francisco, USA. The cohort was consist of children ranging in age from three to 12 years and had a female proportion of 52% (3) The Tanzanian cohort consisted of 2595 African participants collected at several sites in Tanzania.The Tanzanian cohort ranged in age from three to 21 years and had a female proportion of 56%. For all cohorts, participants were eligible if they had not experienced facial trauma, major surgery, congenital facial anomalies that could potentially affect their natural facial structure. In the PITT and the Denver cohorts, 3D images of participant's resting face were captured using the 3dMD face camera system. The Tanzania participants were photographed using Creaform MegaCapturor II or Gemini white light 3D photogrammetry systems.

The PITT and the Denver cohorts were genotyped on the Illumina OmniExpress + Exome v1.2 array, which assays nearly 1 million SNPs including approximately 245,000 coding variants, and were fully imputed to the 1000 Genomes Project phase 3 reference panel. The Tanzania cohort was genotyped on the Illumina HumanOmni2.5+Exome-8v1A array covering over 2.5 million SNPs, which have been fully imputed to the 1000 Genomes Project phase 1 reference. Data cleaning and quality control procedures were performed using standard analysis pipelines [114]. Specifically, samples were interrogated for genetic sex, chromosomal aberrations, concordance between genetic and self-reported ancestry, biological relatedness among participants, missing call rate, and batch effects. Variant-level evaluation included missing call rate, discordance between duplicates, Mendelian errors, deviation from Hardy-Weinberg equilibrium, and sex differences in allele frequency and heterozygosity (for

autosomal and pseudo-autosomal SNPs). Imputed SNPs were filtered by SNP-level INFO score and genotype-per-participant-level genotype probability.

3.2 Phenotyping

The proposed project will take advantage of a novel approach, global-to-local hierarchical modular phenotyping, for modeling the multi-dimensional facial shape. This approach was developed by Peter Claes's group at KU Leuven, and a detailed description of the process can be found in Claes et al 2018 [31]. A flowchart of this hierarchical modular phenotyping approach can be seen in Figure 3.1. Briefly, approximately 10,000 points, also called quasi-landmarks, were automatically placed across the facial surface, by a non-rigid registration of a standard facial template onto each surface. The result is that each quasi-landmark represents the same facial position across all participants [115]. The configurations were then co-aligned to their mean with generalized Procrustes analysis. The landmarks were then clustered into groups of co-varying points in order to partition the full face into segments (also called modules). This was accomplished through spectral clustering applied to the pairwise correlation matrix of quasi-landmarks. Each partition generated two sub-segments, and process was repeated for a total of five iterations to generate a hierarchy of 63 facial modules comprising overlapping groups of quasi-landmarks. These modules formed successive levels representing the shift from more globally integrated to more locally focused morphology. The shape variation characterizing each module was represented by the 3D coordinates of all quasi-landmarks contained therein. Principal components analysis and parallel analysis were performed on the quasi-landmarks to form multi-dimensional phenotypes in which the shape variation is represented by principal component scores (PCs). Each module was represented by 5$\sim$50 PCs which were corrected for sex, age, height, weight, facial size and genetic ancestry. These 63 multivariate modules will serve as the phenotype to be analyzed individually in the statistical tests.

The hierarchical clustering of facial shape is shown in Figure 3.2 for the PITT cohort

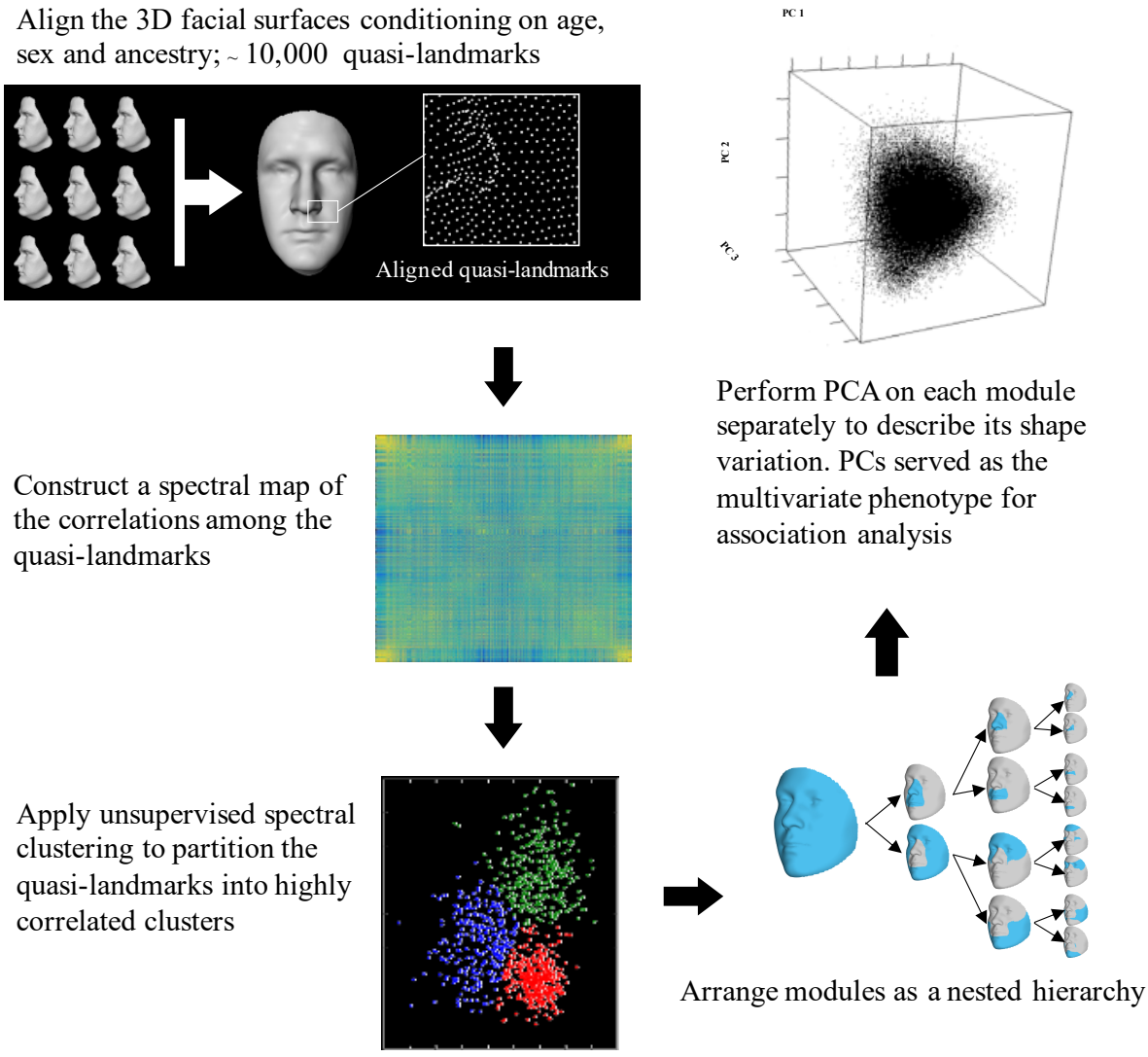


Figure 3.1: A flowchart of the hierarchical modular phenotyping approach developed by Claes et al [31]

and in Figure 3.3 for the Tanzania cohort. Segments are shown in several colors with a background face in gray. In the center is shown the most global segment representing the full face, which is split into segments 2 and 3 as shown in the innermost concentric circle. Moving outward across the five concentric circles, each facial segment is further partitioned into two more localized segments. These segments are also referred to as modules hereafter. Comparing Figure 3.2 and Figure 3.3, one can notice that the segmentation in the two cohorts was in general consistent with each other. The second tier of the hierarchy in both had four modules representing roughly the mouth/lip, upper face, lower face, and middle face/nose area. At a finer scale, individual facial modules in one cohort may or may not have an obvious counterpart in another. The forehead modules 30 in PITT and 20 in Tanzania, for instance, largely overlap each other, while the orbital module 45 in Tanzania was not present in PITT.

The Denver sample was projected onto each of the phenotypic traits (PCs in facial modules) learned in the PITT cohort because its sample size did not support a stable data-driven segmentation. Consequently, the facial traits were kept fixed and thus consistent between the PITT cohort and the Denver cohort, which facilitated using Denver as a replication dataset for the PITT cohort and the meta-analysis of the two.

This global-to-local segmentation framework confers several benefits. First, it achieves a compact and comprehensive representation of the 3D facial morphology. Shape variation is preserved to the maximum extent so that unlike univariate traits derived from a limited number of facial landmarks, little information is lost in our analysis. Meanwhile, PCA within each segment produces a manageable set of phenotypes and investigating all PCs simultaneously in a single multivariate framework would help reduce the multiple testing burden. Second, the organization allows for an open-ended interrogation for genes exhibiting a range of effects at different scales. By gradually zooming into more localized facial area, genes associated with a high level segment can be further refined to more specific facial regions. The phenotypic pattern of association signals can also help with inferring the possible biological mechanism and guiding functional experiments. Third, although the segmentation is unsupervised and learned exclusively from the data, the resulting modules

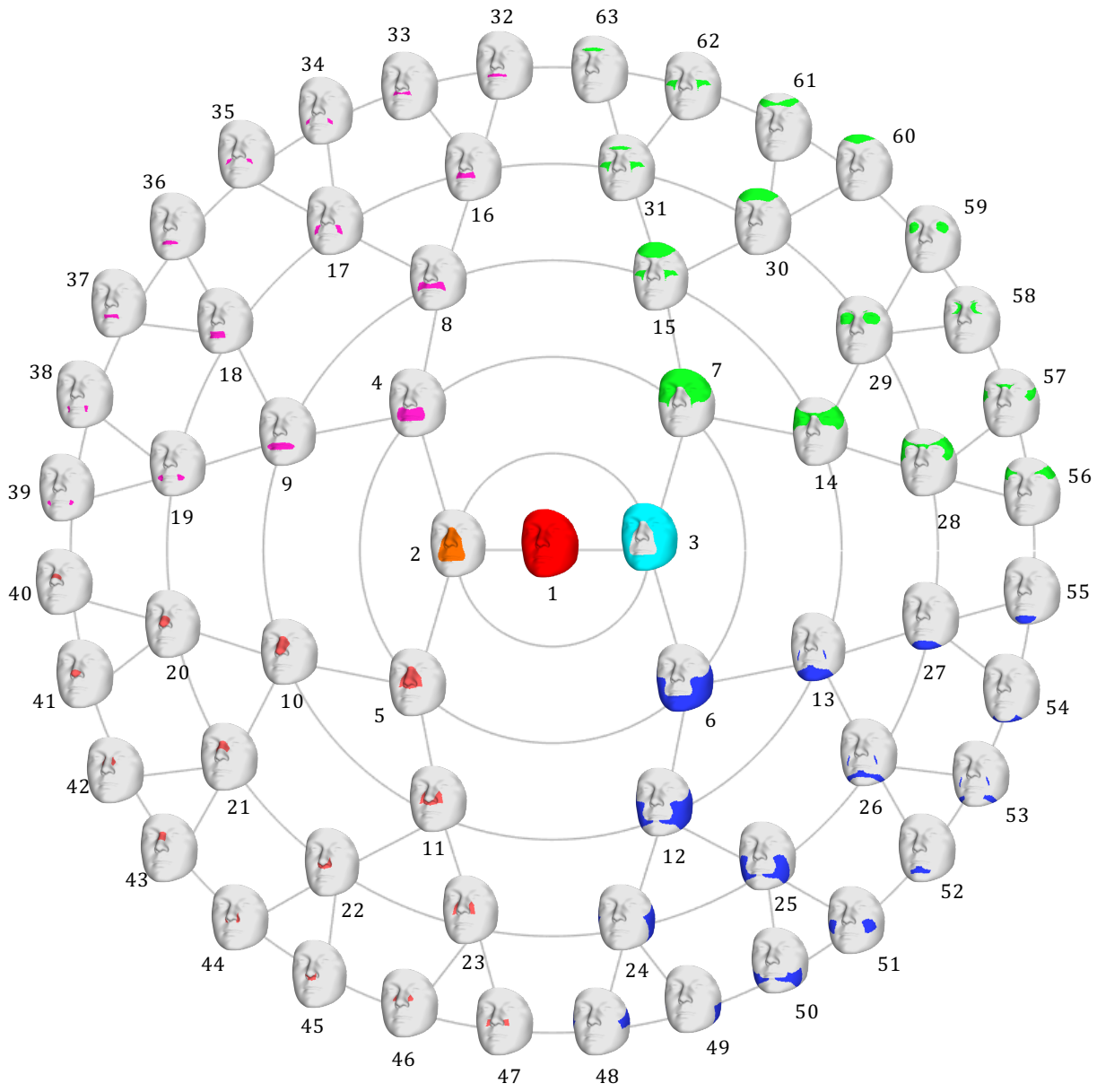


Figure 3.2: Global-to-local facial segmentation of the PITT cohort. Obtained using hierarchical spectral clustering. Modules are colored on a gray background face.

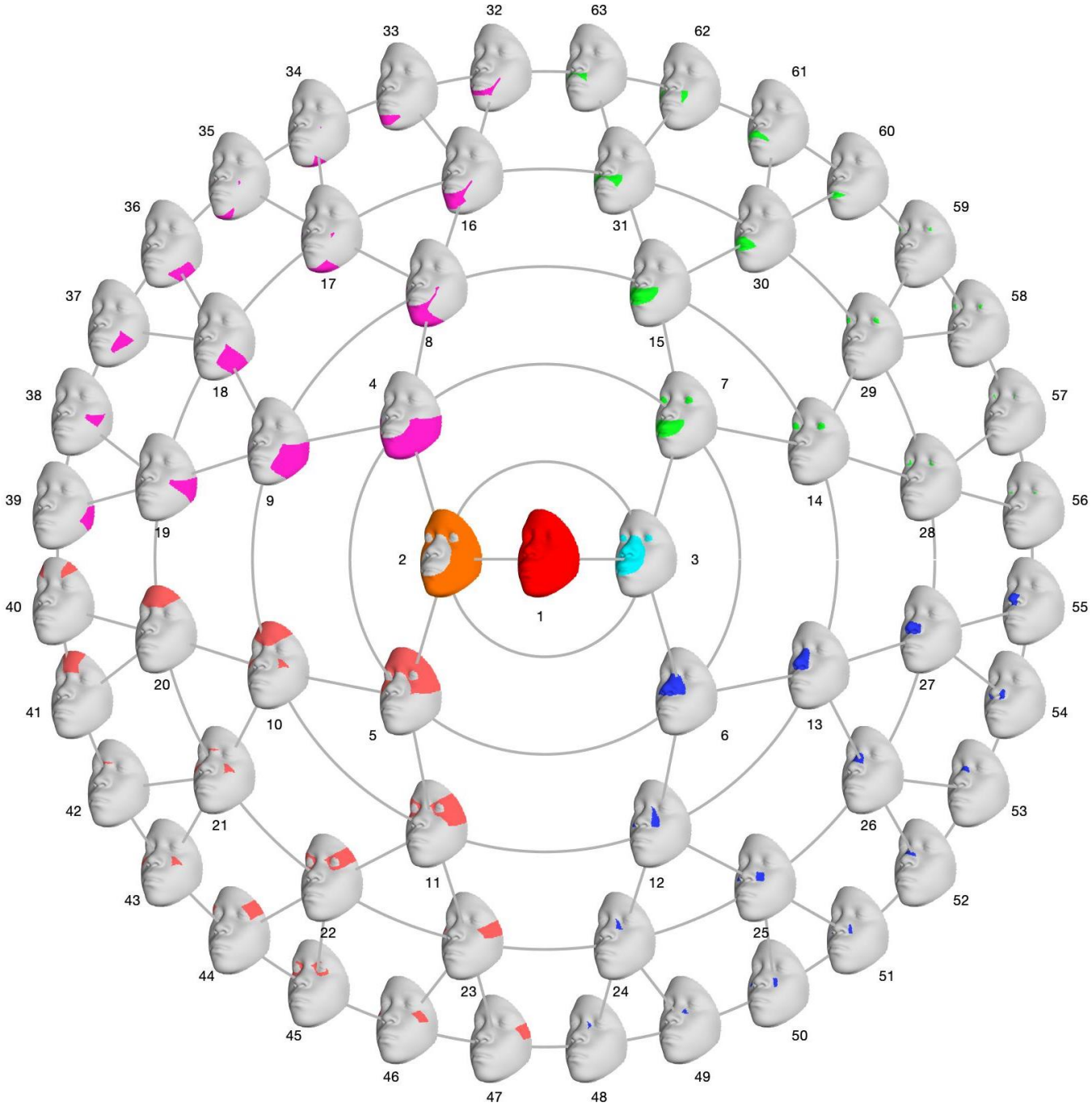


Figure 3.3: Global-to-local facial segmentation of the Tanzania cohort. Obtained using hierarchical spectral clustering. Modules are colored on a gray background face.

did show some degree of biological relevance. For example, module 4-7 at the second to the innermost concentric circle roughly echo the three embryonic facial prominences - frontonasal, maxillary and mandibular. Such biological relevance will facilitate the interpretation of gene mapping results. Fourth, the data-driven approach circumvents the inherent difficulty in defining proper facial measurements *a priori* and avoids the arbitrariness in the phenotype selection process.

3.3 Consideration of the ethnicity

Cohorts involved in this dissertation project were each of a different ancestry (European, Hispanic White, and African). The effect of certain genetic variants on individual level phenotypes and molecular traits are known to be different among populations/ethnic groups, and this heterogeneity complicates multi-ethnic studies where groups are to be analyzed as a whole. Two additional complicating factors need to be taken into account for this specific project. First, low-frequency genetic variants in Aim 1 would rarely be shared across ethnic groups and thus a joint analysis using individual level data, i.e. a mega-analysis, should not be expected to add much to the study. Second, the data-driven phenotyping approach generated different sets of facial modules for different cohorts. Some modules have counterparts (those overlapping a lot) across cohorts, whereas other do not. Operation of a meta-analysis on the summary-level data is much less clear as none of the modules were identical between the two cohorts. With the consideration of these complications, this project analyzed each cohort separately and except in Aim 1, did not pursue either a meta- or a mega-analysis.

4.0 Rare and Low-frequency Coding Variant Analysis

The analysis of the PITT cohort included in this chapter is under review as of this writing.

4.1 Introduction

Although the strong genetic control of facial features has long been appreciated, knowledge on specific genes underlying normal-range human facial variation is still poorly understood. Common SNPs identified by GWAS cumulatively explain only a small fraction of the heritable phenotypic variation. Based on large-scale genomic studies of other complex morphological traits, such as height [2, 3, 51], we hypothesized that functional variants at hundreds or perhaps thousands of loci have yet to be discovered. One proposed source of the missing heritability lies in variants with lower frequencies that have not been investigated in GWAS. We therefore aimed in this section to discover facial morphology-associated low-frequency variants, in a gene-based manner, first in each of the three study populations separately and then combing results across the two US cohorts by performing a meta-analysis.

4.2 Methods

4.2.1 MultiSKAT and Meta-MultiSKAT

This study used MultiSKAT and Meta-MultiSKAT [116,117] to carry out the exome-wide gene-based low-frequency coding variant scans. The two methods belong to a recently developed statistical framework accommodating both (1) the low statistical power associated with analyzing individual low-frequency genetic variation by implementing set-based tests, and (2) the multi-dimensional nature of the facial phenotype. More specifically, MultiSKAT extends

the mixed models-based kernel association test to a multivariate setting while adjusting for covariates and accounting for family relationships. It is flexible in relating multivariate phenotypes to a collection of genotypes through the use of different kernel options, and includes an omnibus test for integrating results across kernels in the absence of a priori knowledge. As a set-based test, MultiSKAT assigns larger weights to rarer variants when aggregating across variants within a testing unit. Meta-MultiSKAT further extends MultiSKAT for the purpose of meta-analyzing across datasets. Since MultiSKAT and Meta-MultiSKAT are brand-new methods, we closely monitored their behavior and performed quality control diagnostics in applying these methods to our dataset. To the best of our knowledge, there are not yet any alternatives, readily available tools for set-based association testing of multivariate traits.

MultiSKAT uses kernels to model one-to-many relationships: the phenotype kernel describes how one variant affects multiple traits and the genotype kernel specifies how multiple variants influence one trait. MultiSKAT provides two choices of the genotype kernel and four choices for the phenotype kernel, and suggests using multiple kernels to obviate the need of making upfront guesses of the unknown underlying genetic architecture. This study applied their omnibus test to combine across four kernel combinations, with either SKAT or burden as genotype kernel and either Hom or Het as phenotype kernel: (SKAT, Hom), (SKAT, Het), (burden, Hom), and (burden, Hom).The Hom kernel assumes the effect sizes of a variant on different phenotypes are homogeneous while the Het kernel assumes the effect sizes of a variant on different phenotypes are heterogeneous. Among all four available phenotype kernel options (Hom, Het, PhC, PC), three of them (Het, PhC, PC) are essentially the same for standardised variables (facial PCs). Therefore, only one among the three was selected for analysis, giving a total of two selected phenotype kernels. For the two available genotype kernels, we used both of them.

4.2.2 Quality control on facial phenotypes

Analyzing rare variants presents several challenges, one of which is that many widely applied statistical tests become more prone to inflated type I error when phenotypic normality does not hold [118]. In fact, we observed inflated MultiSKAT quantile-quantile (Q-Q) plots

at first and identified a need to re-examine the phenotypic distribution of the facial PCs. To accomplish this, we visualized the joint distribution of all PCs underlying each facial module in Q-Q plots where the chi-squared quantiles were plotted against robust squared Mahalanobis distances. Outliers that fell far apart from the rest of the sample can be identified by visual inspection.The Mahalanobis distance is a metric measuring how far an observation is from the center of the joint distribution, which can be thought of as the centroid in multivariate space. Facial images associated with outlier observations were revisited to confirm data validity and sample eligibility. After this procedure, one facial image was labelled as a multivariate outlier for PITT module 27 (Figure B1). Pairwise distributions of the 16 facial PCs in this module further revealed that this outlier had extreme values on four individual PCs (PC9, PC11, PC13 and PC14, Figure B2). Subsequent analysis involving module 27 therefore did not use data from this individual. Inspection of the 3D image suggested a potential artifact in the facial region represented by module 27. No outlying phenotypes were spotted in either the Tanzania cohort or the Denver cohort.

To minimize potential spurious associations, the outermost layer of the module rosette in Figure 3.2 and Figure 3.3 were excluded from low-frequency variant association test. These modules were shown to be less robust in the previous GWAS in that SNPs significantly associated with only the outermost modules but not more inner ones generally failed replication [31]. For variants with lower frequency and a brand-new statistical method that has yet had sufficient real-world application to prove its empirical performance, it is safer to restrict the analysis to better-behaved facial segments. The remaining 31 modules arranged as a top-down tree structure for the PITT cohort can be seen in Figure 4.1. The effective number of tests (number of independent modules) was estimated to be 19 in the PITT and the Denver cohort, and 20 in the Tanzania cohort according to Li and Ji's eigenvalue-based method [119]. These numbers were used for multiple testing correction.

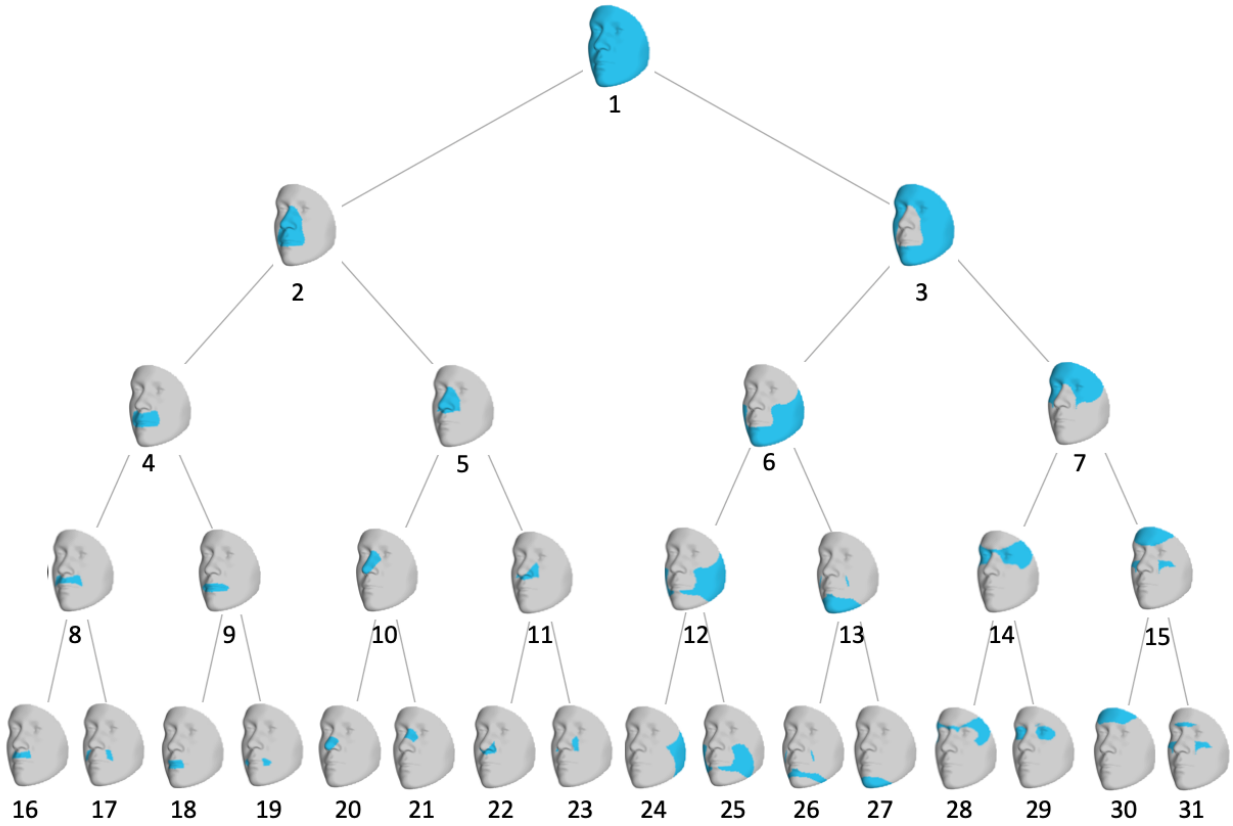


Figure 4.1: The 31 PITT facial modules included in low-frequency variant analysis. Modules are colored in blue.

4.2.3 Gene-level analysis

Imputed genotypes with a certainty above 0.9 were used to fill in any sporadic missingness among genotype calls of the directly genotyped variants. Wholly unobserved, imputed SNPs were not included in this analysis. Ancestry PCs based on common linkage disequilibrium (LD)-pruned SNPs were constructed and regressed out from the modular traits at the phenotyping stage. Variants were grouped into genes if they (1) overlap with hg19 exons so that only coding variants were included, (2) had a MAF below 1%, and (3) had at least 4 minor allele counts (MAC) for analysis in the PITT and Denver data, or at least 2 MAC for analysis in the Tanzania data. Genes with less than two qualified variants were excluded. The last criterion, a minimum MAC cutoff, was applied per suggestion from the original MultiSKAT paper and further customized for each dataset to include as many genes as possible

while maintaining a well-behaved statistic. The final numbers of variants included for the analysis in the PITT, Tanzania and Denver cohort were 31347, 20468, 10684, respectively, and the corresponding numbers of genes tested are 8091, 6235, 3660. The p-value thresholds for declaring significance (set for each individual cohort) were $0.05/(8091 \times 19) = 3.3\times10^{-7}$ in PITT, $0.05/(6235 \times 20) = 4.0\times10^{-7}$ in Tanzania, and $0.05/(3660 \times 19) = 7.2\times10^{-7}$ in Denver.

Genes and variants were annotated by Ensembl GRCh37. Several GWAS result annotation tools, including GREAT [120], FUMA [121] and ToppFun [122], were used to perform gene-set enrichment analysis with default parameter settings. We queried the GTEx database for gene expression levels in face-related tissues. Results from this gene-based low-frequency variant scan were compared with previous GWAS results in the same sample to explore the effects of variants with different allele frequencies.

4.2.4 Variant-level analysis

For genes identified by MultiSKAT, we scrutinized the quality of genotype calls by inspecting allele intensity cluster plots. We then performed association tests of individual variants with the corresponding facial segments using MultiPhen [123]. MultiPhen finds the linear combination of facial PCs with the strongest association with the genotypes at a SNP, and is especially appealing for this study due to its robust performance in scenarios where rare variants are tested against non-normal phenotypes.

Variant-level bioinformatic annotation was done using CADD [124]. CADD is a comprehensive metric that weights and integrates diverse sources of annotation, by contrasting variants that survived natural selection with simulated mutations. The scaled CADD score expresses the deleteriousness rank in terms of order of magnitude. A score of 10, for instance, is interpreted as ranking in the top 10% in terms of the damaging degree amongst reference SNPs, and a score of 20 refers to 1%, 30 to 0.1%, etc. Individual exonic variants were looked up in literature and by PhenoScanner [125] for existing genotype-phenotype associations.

The magnitude of phenotypic effect of individual low-frequency variants was quantified

by the difference between averaged faces of variant carriers and non-carriers. This difference was further compared with that of the significant common variants identified in our previous GWAS [31]. Specifically, the centroids of the multidimensional space defined by PCs in a certain module were computed separately for people carrying the variant and people who do not carry the variant. Then the Euclidean distance between the two centroids was calculated as a measure of variant effect size. The 95% confidence interval of the effect size was obtained by 5000 bootstraps. This work was done in collaboration with Harry Matthews from KU Leuven.

4.2.5 Lookups of significant genes in orofacial clefting GWAS

Our group has previously conducted a low-frequency variant association analysis in orofacial clefting (OFC) datasets [126]. Following our hypothesis that genes influencing typical facial presentation may also be involved in facial anomalies, we examined whether any significant genes in the current study were associated with non-syndromic cleft palate with or without cleft lip (NSCL/P). Summary statistics of the identified face-associated genes were retrieved from the OFC study, where two types of statistical tests (SKAT and burden test) were applied in four individual cohorts. The lookup p-values were compared to a Bonferroni adjusted threshold of p-value $< 0.05/(4 \times 2 \times$ number of genes looked up).

4.2.6 Replication and meta-analysis

The PITT and the Denver cohorts can serve as independent replication cohorts for each other. Both gene-level and single variant-level replication analysis were performed for genes identified in the discovery phase of individual cohorts. Meta-MultiSKAT was used to conduct a meta-analysis of the PITT and the Denver cohort. The Tanzania dataset was excluded from this effort because (1) the facial modules were not the same as those of the US cohorts, a result of the data-driven phenotyping process; (2) the participants were of a different ancestry, adding another layer of complexity, especially for low-frequency variants on which the effect of population structure is not yet well understood. On the other hand, the PITT

and Denver cohorts can be combined by meta techniques, free of the former concern and to a lesser extent of the latter. Genes common in both cohorts were meta analyzed by meta-MultiSKAT, assuming (1) the same and (2) different gene effects between cohorts, followed by a combination test of the two circumstances.

4.2.7 Low-frequency coding variant analysis for univariate facial traits

In addition to the multi-dimensional modular phenotypes, this project also analyzed a set of traditionally defined univariate facial traits. This analysis was conducted in the PITT cohort only, as the hand-placed landmarks used to generated linear distances were not available in the Denver or Tanzania datasets. A total of 24 preselected linear distance measures (Figure 4.2) were available for 2447 PITT subjects (mostly overlapping with the 2329 individuals included in the analysis of modular phenotypes). These univariate phenotypes represented different morphological information than the multivariate modules, shown by the low-to-medium Pearson correlations between them B3). Phenotypic residuals were generated by regressing out the effect of sex, age, age^2, height, weight and facial size on the 24 facial measures. The independent number of phenotypes estimated according to Li and Ji [119] was 17. We tested low-frequency variants on these residual phenotypes in a gene-based manner while adjusting for ancestry PCs. We used two complementary score-statistics based tests, SKAT and Combined Multivariate and Collapsing (CMC) methods as implemented in rvtest [127]. Coding variants with a MAF below 1% were grouped into genes. A total of 10725 genes with at least two qualified variants were included in the analysis, giving an exome-wide significance threshold of $0.05/(10725 \times 17) = 2.74\times10^{-7}$. For significantly associated genes, we scrutinized the quality of genotype calls by inspecting allele intensity cluster plots, and further performed single-variant association tests.

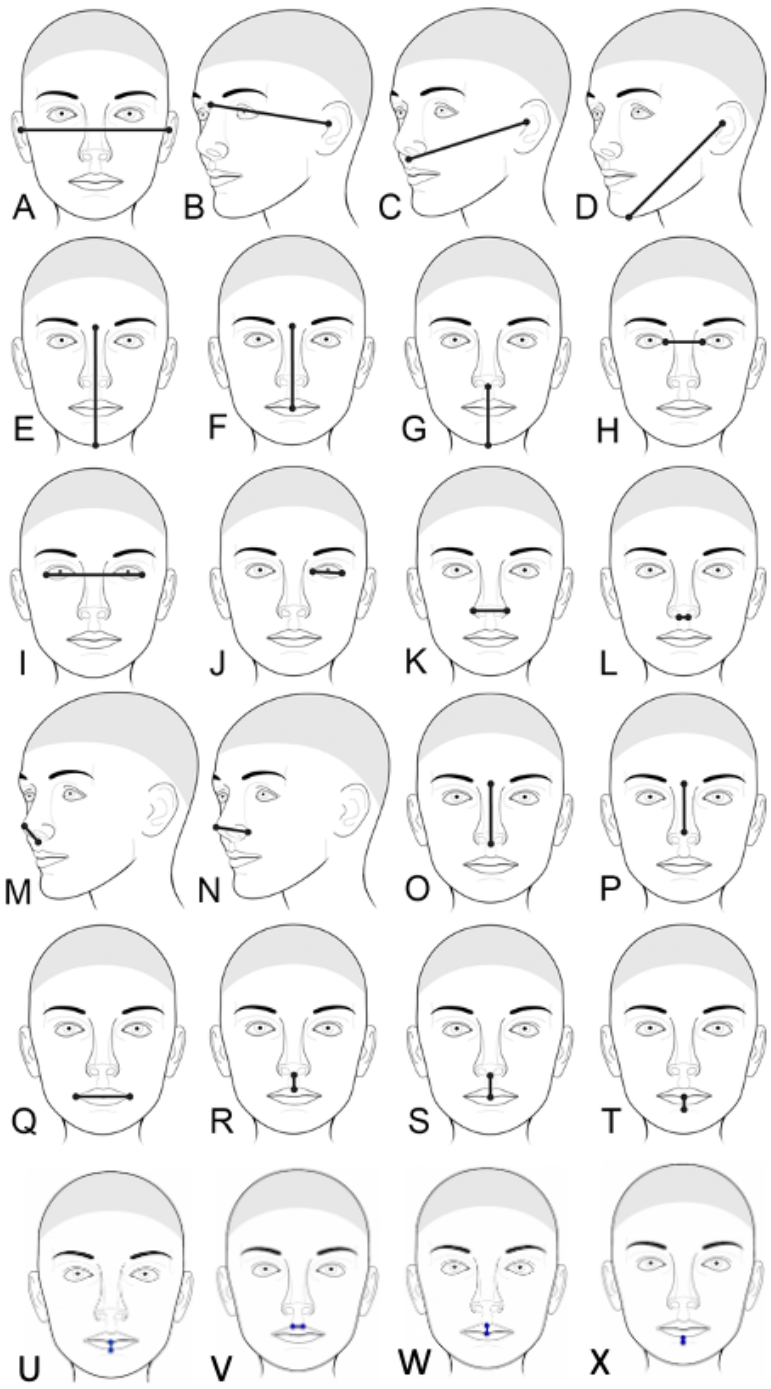


Figure 4.2: Univariate facial phenotypes. (A) Cranial base width, (B) Upper facial depth, (C) Middle facial depth, (D) Lower facial depth, (E) Morphological facial height, (F) Upper facial height, (G) Lower facial height, (H) intercanthal width, (I) Outercanthal width, (J) Palpebral fissure length, (K) Nasal width, (L) Subnasal width, (M) Nasal Protrusion, (N) Nasal ala length, (O) Nasal height, (P) Nasal Bridge Length, (Q) Labial fissure length, (R) Philtrum length, (S) Upper lip height, and (T) Lower lip height, (U) lower vermilion height, (V) Philtrum width, (W) Upper vermilion height,(X) Cutaneous Lower Lip Height

4.3 Results

4.3.1 Low-frequency variant association results in the PITT cohort gene-level analysis results

MultiSKAT identified seven genes (*HFE*, *NECTIN1*, *CARS2*, *LTB4R*, *TELO2*, *AR*, *FTSJ1*) significantly associated with at least one facial module. Figure 4.3 shows the composite Manhattan plot overlaying the results from all modules analyzed. Significant genes are linked to their associated facial segments together with the p-values. Four genes (*HFE*, *CARS2*, *LTB4R*, and *TELO2*) were associated with nose-related modules, and the others were associated with modules of the chin, mouth, and cheek. Three genes reached the significance level in more than one module. We observed well-calibrated test statistics and little evidence of inflation in the Q-Q plots (Figure B4).

A more comprehensive view of the genes' phenotypic effects, shown by the association p-values across the entire facial segmentation hierarchy, can be viewed in Figure 4.4. For each gene, its -$\log_{10}$(P-value) with 31 modules is shown as color shades ranging from the minimum to the maximum (within each gene). A top-down tracing of the shade degree gives an easy visualization of genes' effects propagating along the branching paths from the more global segments to the more local segments. *FTSJ1* had a broad range of effects involving facial regions from the full face to the most refined modules, while effects of other genes were more confined to local modules. Table 4.1 lists all facial segments and the corresponding morphological areas where the seven genes showed an association p-value $< 10^{-4}$.

We performed gene-set enrichment analysis to explore the functions associated with the seven identified genes. Enrichment was detected for a variety of biological processes (Figure 4.5), especially those related to ion, metabolism, transportation and regulation. Enriched molecular functions tended to be housekeeping and general processes, e.g. signaling receptor and protein binding activity. Two genes with relatively well characterized functions, i.e. *HFE* and *AR*, contributed a lot to the enrichment terms. In the GTEx database, these seven genes showed measurable expression level in adipose, skin and muscle-skeletal tissue (Figure B5), among which the strongest expression was seen for *NECTIN1* in skin.

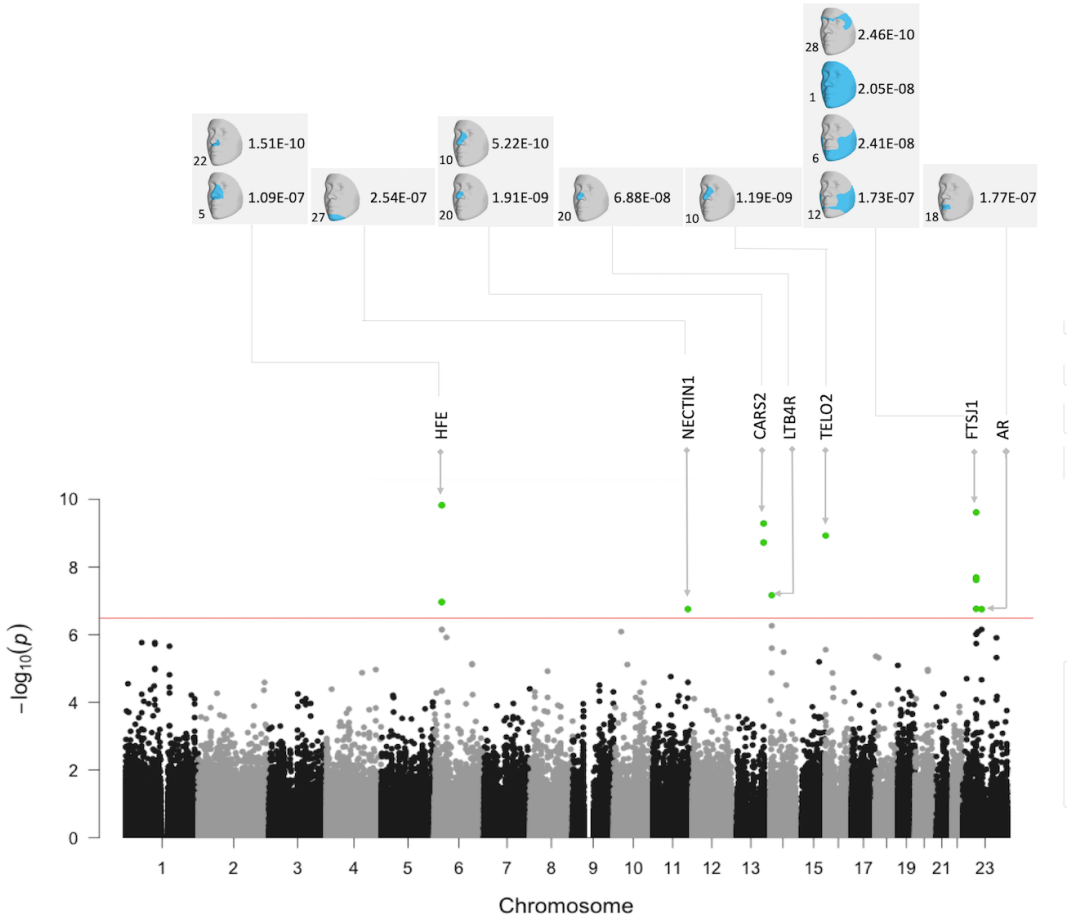


Figure 4.3: Composite Manhattan plot showing results across 31 facial modules in the analysis of the PITT cohort. The Manhattan plot shows the position of genes on the x-axis and MultiSKAT p-values on the y-axis. A total of 31 points are plotted for each gene, representing p-values obtained by testing their association with each of the 31 modules. The red horizontal line indicates the study-wide significance threshold. The associated facial modules and the corresponding p-value for each gene that surpassed the threshold are shown above the Manhattan plot.

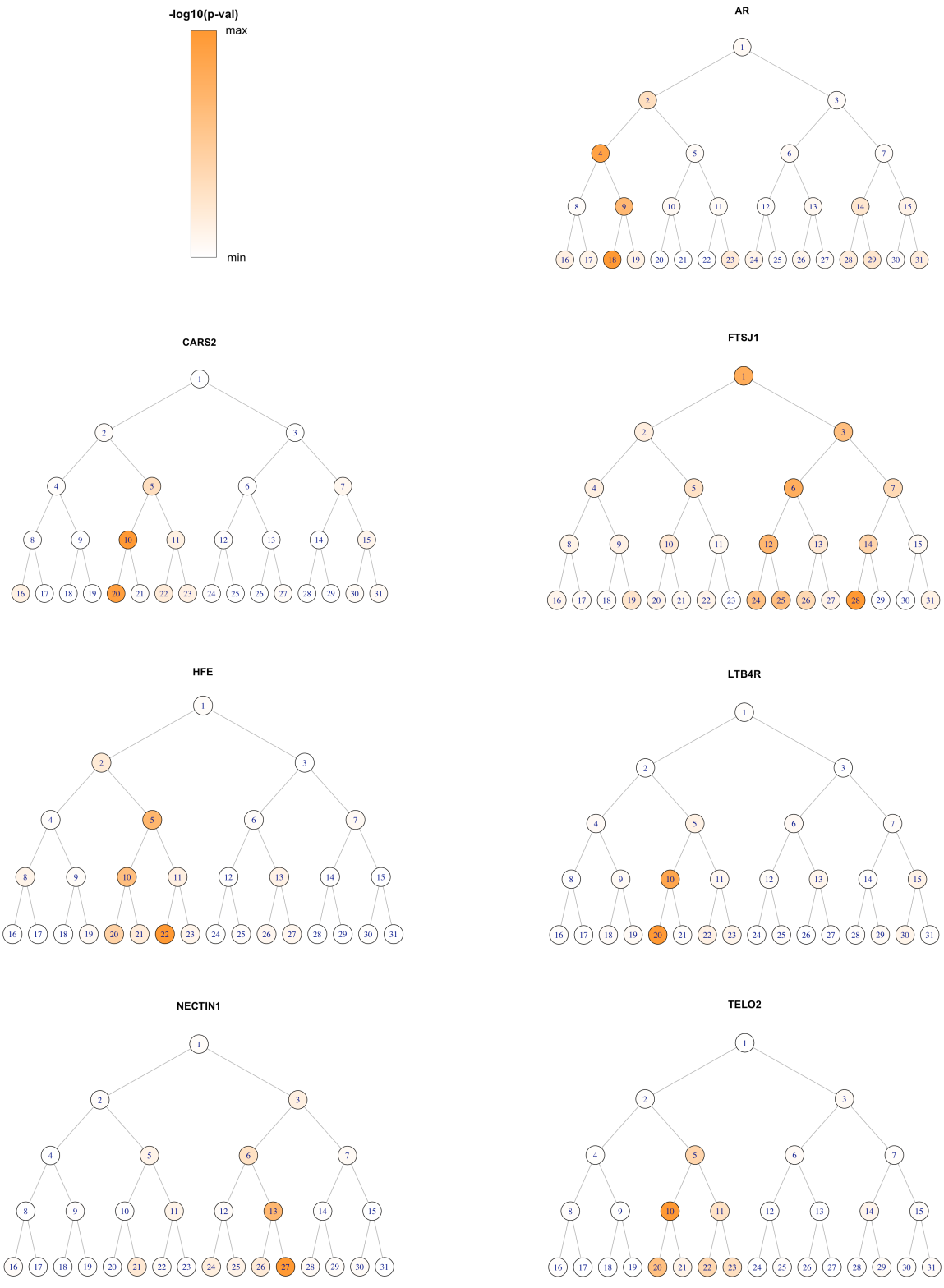


Figure 4.4: Module-wide association results for significant genes. For each gene, the –log10(p-value) is shown as color shades ranging from the minimum to the maximum, for 31 facial segments arranged the same way as in Figure 4.1.

Table 4.1: Module-wide association results of genes identified by MultiSKAT in the PITT cohort. Here we show associations with a p-value $< 10^{-4}$

Chr	Gene	Module	P-value	Facial region
6	*HFE*	22	1.51E-10	nose
		5	1.09E-07	
		10	7.11E-07	
		20	4.58E-05	
11	*NECTIN1*	27	2.54E-07	chin
		13	2.56E-05	
13	*CARS2*	10	5.22E-10	nose
		20	1.91E-09	
14	*LTB4R*	20	6.88E-08	nose
		10	5.47E-07	
16	*TELO2*	10	1.19E-09	nose
		20	2.79E-06	
23	*AR*	18	1.77E-07	lip, philtrum
		4	7.00E-07	
		9	2.15E-05	
23	*FTSJ1*	28	2.46E-10	full face, cheek
		1	2.05E-08	
		6	2.41E-08	
		12	1.73E-07	
		3	9.48E-07	
		25	9.84E-07	
		24	1.84E-06	
		14	8.26E-05	

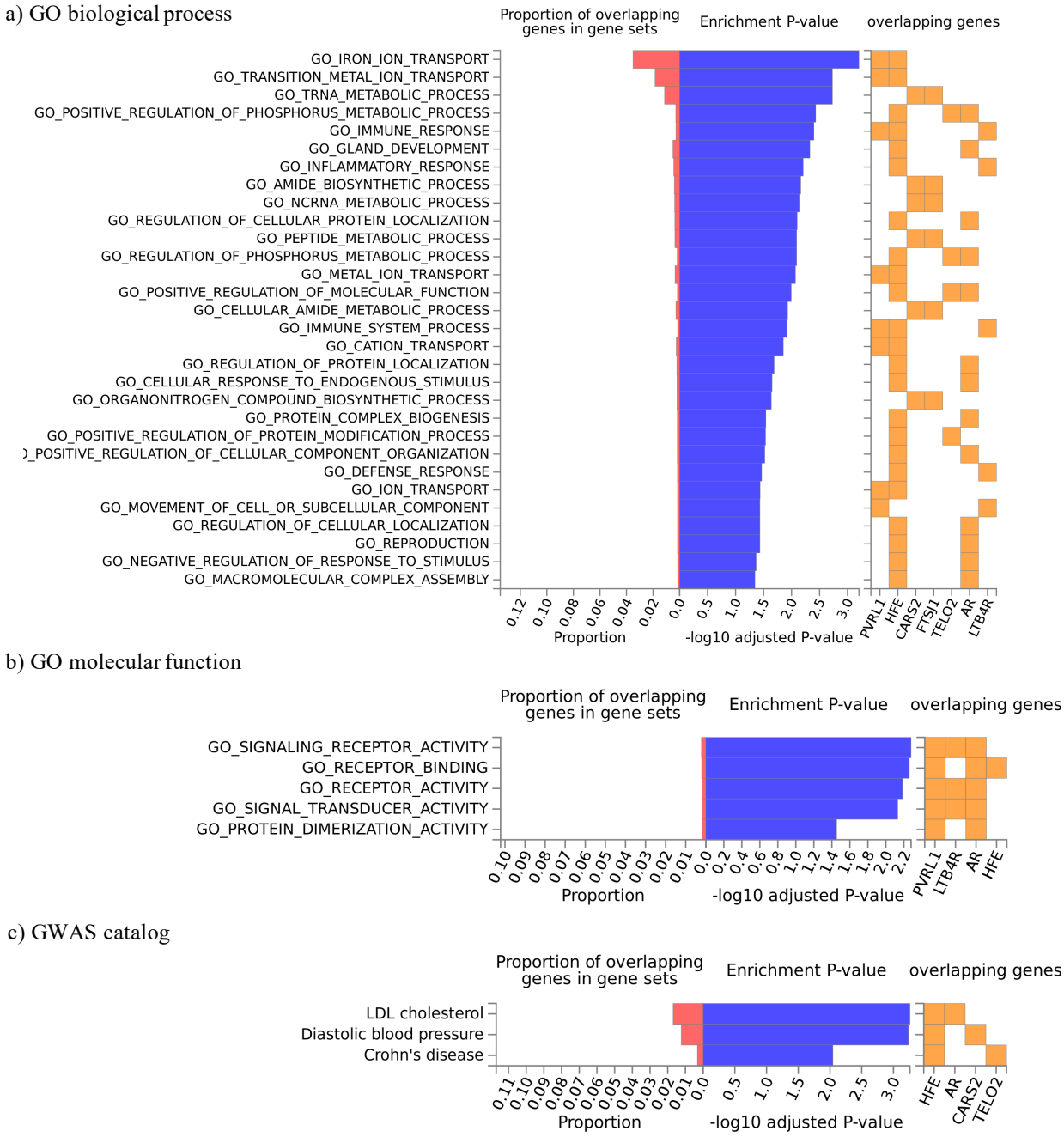


Figure 4.5: FUMA gene-set enrichment results for the seven MultiSKAT significant genes in the PITT cohort

To explore whether facial genes also affect the risk of orofacial clefts, we retrieved the summary statistics of five of the seven face-associated genes from our previous OFC study [126]. Two genes (*AR* and *FTSJ1*) were not available due to a lack of variant carriers. Table 4.2 displays the SKAT and CMC test p-values for the five genes in Europeans, Asians, Latino South Americans (LSA in Table 4.2) and combined populations. Two associations passed a Bonferroni corrected threshold for 40 tests (5 genes $\times$ 4 populations $\times$ 2 type of tests): *TELO2* with a CMC p-value = 6.5×10^{-4}, and *HFE* with a CMC p-value = 1.1×10^{-3}, both in the combined population of all ancestry groups.

4.3.2 Variant-level analysis results in the PITT cohort

Single variants in the seven significant genes were further tested individually for association with the corresponding facial module, the results of which are displayed in Table 4.3. Six SNPs from five genes yielded an association p-value < 0.05. SNP rs142932029 in *FTSJ1* was strongly associated with module 28 with p-value = 1.59×10^{-14}. Because no single variant in *HFE* or *LTB4R* showed nominal significance, the gene-level signals reflected the cumulative effects of multiple low-frequency variants and the increased power of gene-based test. Significance of other genes in the MultiSKAT test were to some degree driven by a specific SNP. Most of the individual variants appeared at frequencies much lower than 1%, and all are nonsynonymous variants except one splice site SNP in *FTSJ1* (this SNP is also an exon variant for a non-coding transcript of *FTSJ1*). Variants in *NECTIN1*, *CARS2* and *AR* were predicted to have a high probability of being deleterious with a CADD score greater than 20. PhenoScanner linked these coding variants with a variety of human traits/disorders in previous studies (Table B1), including height, vascular diseases, osteoporosis, neoplasms etc., suggesting that coding variants influencing facial shape may be pleiotropic and play roles in other biological processes. We compared the effect sizes of all individual low-frequency variants in the seven genes to that of the common SNPs identified in previous GWAS. Phenotypic differences between variant carriers and non-carriers were consistently larger for low-frequency variants than for common SNPs, as shown in Figure 4.6.

Table 4.2: SKAT and CMC test results of the association between the seven facial genes and NSCL/P in a multi-ethnic study

Gene	CHR	Number of variants	Study population	Sample size	SKAT P-value	CMC P-value
HFE	6	6	All pops	3621	3.68E-03	**1.10E-03**
HFE	6	1	Asia	528	5.07E-01	5.07E-01
HFE	6	3	Euro	1411	5.35E-01	7.08E-01
HFE	6	4	LSA	1676	2.23E-02	3.49E-02
NECTIN1	11	5	All pops	3621	1.43E-01	5.30E-02
NECTIN1	11	1	Asia	528	1.31E-01	1.31E-01
NECTIN1	11	3	Euro	1411	5.95E-01	3.86E-01
NECTIN1	11	5	LSA	1676	3.91E-03	1.15E-02
CARS2	13	17	All pops	3621	5.55E-01	8.83E-01
CARS2	13	2	Asia	528	5.11E-02	9.95E-02
CARS2	13	7	Euro	1411	4.00E-01	4.59E-01
CARS2	13	11	LSA	1676	2.03E-01	8.58E-01
LTB4R	14	7	All pops	3621	8.25E-01	4.64E-01
LTB4R	14	3	Asia	528	3.37E-01	4.67E-01
LTB4R	14	4	Euro	1411	7.36E-01	4.83E-01
LTB4R	14	6	LSA	1676	1.33E-01	6.46E-01
TELO2	16	20	All pops	3621	7.44E-03	**6.51E-04**
TELO2	16	3	Asia	528	7.63E-01	8.29E-01
TELO2	16	7	Euro	1411	1.49E-01	5.96E-01
TELO2	16	13	LSA	1676	7.78E-02	1.22E-01

4.3.3 Low-frequency variant association results in the Denver cohort and results for replication analysis

Two genes surpassed the exome-wide significance level in the analysis of the Denver cohort. *ZNF268* was associated with lip shape and *ARHGEF18* with forehead shape (Last two rows in Table 4.4). Both genes were significant in only one facial module. Table 4.5 shows the results from single variant analysis of these two genes as well as the variant-level replication in the PITT cohort. SNP rs140297736 in *ARHGEF18* had a p-value = 1.73×10^{-4} in a forehead module. Other variants did not show strong association by themselves. All five SNPs involved in the Denver cohort were also present in the PITT cohort with comparable MAF, yet none showed evidence of replication.

Variant-level replication for the PITT results in the Denver cohort was not pursued as all PITT variants were absent from the Denver cohort. This unavailability was a result of the much smaller size of the Denver sample. Gene-level replication was also compromised by limited sample in that only *LTB4R* was eligible in the Denver cohort, and it did not have a small replication p-value (Table 4.4). Genes originally discovered in the Denver cohort failed the replication as well.

4.3.4 Results for meta-analyzing the PITT and the Denver cohort

We included 3429 genes shared between the PITT and the Denver cohort in the meta-MultiSKAT analysis. Two genes passed the study-wide significant threshold (Table 4.6). Complement C6 was associated with nose shape, and Proprotein Convertase Subtilisin/Kexin Type 9 (*PCSK9*) was associated with philtrum shape. Neither were significant in either individual cohort. Three out of the nine genes identified in either PITT or Denver data were tested in the meta-analysis although none were significant, indicating that their effects were unique to the discovery cohort (Table 4.4).

Table 4.3: Single variant association and functional prediction for variants contributing to the gene-level significance in the PITT cohort

Chr	Gene	Gene-level association		Variant-level association							
		Module[a]	MultiSKAT P-value[a]	SNP	Pos (hg19)	Ref/Alt[b]	Function[c]	CADD score[d]	MAF	Module[e]	MultiPhen P-value[e]
6	*HFE*	5, **22**	1.51E-10	rs149342416	26087686	G/C	Arg6Ser	15.3	0.09%	22	7.0E-02
				rs143662783	26087718	C/G	Thr17Ile	13.4	0.09%	5	8.7E-01
11	*NECTIN1*	27	2.54E-07	rs142863092	119548369	G/A	Arg210His	25.2	0.09%	27	1.1E-03
				rs137991779	119549425	G/A	Gly44Ser	29.2	0.11%	27	1.5E-01
13	*CARS2*	**10**, 20	5.22E-10	rs151097801	111296817	C/T	Pro138Leu	22.4	0.09%	20	1.2E-01
				rs117788141	111357899	G/A	Val69Ile	28	0.09%	10	1.0E-02
14	*LTB4R*	20	6.88E-08	rs143666989	24780865	A/G	Gln332Arg	16.6	0.11%	20	1.1E-01
				rs148153989	24780915	A/T	Met349Leu	12.5	0.09%	20	5.9E-01
16	*TELO2*	10	1.19E-09	rs140903666	1544313	G/A	Ala11Thr	6.3	0.22%	10	8.2E-04
				rs144863771	1544314	C/A	Ala11Asp	10.7	0.22%	10	8.2E-04
				rs147858841	1555541	C/T	Ala132Val	9.4	0.11%	10	4.3E-01
23	*AR*	18	1.77E-07	rs142280455	66905875	A/G	Ser598Gly	22.4	0.13%	18	8.1E-01
				rs137852591	66941751	C/G	Gln267Glu	25	0.13%	18	3.9E-03
23	*FTSJ1*	1, 6, 12, **28**	2.46E-10	rs142932029	48341118	G/A	Ser161Asn	7.4	0.08%	28	1.6E-14
				rs201095751	48341414	C/T	Splice site	0.1	0.11%	12	1.0E-01

[a] For genes associated with multiple facial modules, the most significant module is in bold and only it's p-value is shown
[b] Alleles are listed as alternative/reference alleles on the forward strand of the reference genome
[c] For missense variant, amino acid substitution is given
[d] Bioinformatic prediction of variant effect, higher score indicates greater damaging effect
[e] Variants were tested against all module(s) with gene-level significance, and for genes associated with multiple modules, only the module yielding the smallest p-value in the variant-level test is shown

4.3.5 Results for univariate facial traits in the PITT cohort

In the analysis of facial linear distances, one association surpassed the exome-wide significance threshold. *GRAMD1B* was associated with the height of the lower vermilion with a SKAT p-value = 6.34×10^{-12}, a permutation SKAT p-value = 2.94×10^{-6} and a CMC p-value = 1.64×10^{-7}. See Figure 4.7 for the Manhattan plot of lower vermilion height where *GRAMD1B* was highlighted in both SKAT and CMC. Variant-level analysis identified a missense variant, rs191981781, whose minor allele was significantly associated with thicker vermilion (p-value=2.83×10^{-11}; Table 4.7). SIFT and PolyPhen predicted this variant to be benign. *GRAMD1B* is a highly conserved gene near a leukemia GWAS locus [128, 129]. Its molecular function has not been well characterized.

4.3.6 Low-frequency variant association results in the Tanzania cohort

MultiSKAT analysis of the Tanzania data did not identify significant genes. There was no overall inflation of the test statistics (Figure B6). We explored whether genes identified in the European cohorts also showed any evidence of association in this African cohort. Four out of the nine genes in Table 4.4 were available for lookup, none of which were associated with facial morphology in the Tanzania cohort (Table B2, threshold $0.05/4/20=6.25\times10^{-4}$).

4.4 Discussion

With rare and low-frequency variants being the primary study focus, this chapter presented a discovery effort to identify coding variants associated with normal-range human facial morphology by undertaking gene-based association tests and subsequent analyses using Exome chip genotyping data. Overall, we demonstrated that part of the morphological variation of facial shape is attributable to rare and low-frequency coding variants. Seven genes were detected in the analysis of modular phenotypes and one additional gene was found to be associated with traditional univariate facial measures. Notably, *NECTIN1* is

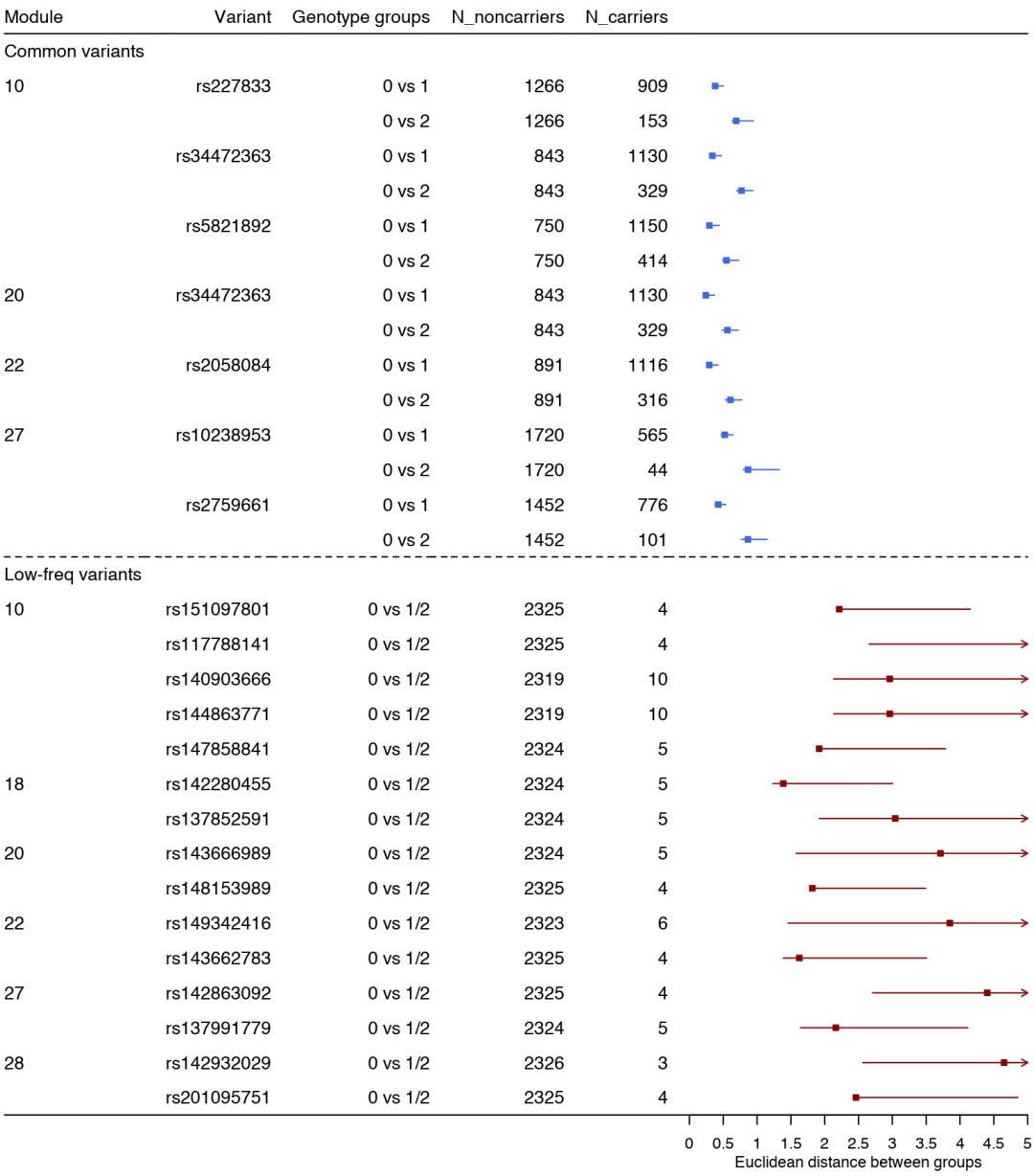


Figure 4.6: Magnitude of variant effects in the PITT cohort. Magnitudes were quantified by the Euclidean distance between averaged faces of different genotype groups. Horizontal lines in the rightmost column indicate the 95% confidence interval. The farther away the blue (common, significant GWAS SNPs) or red (low-freq variants) rectangular boxes are from x=0, the greater the effect. The genotype groups column indicates the two groups of people for whom the faces were averaged and the distance computed. 0 = major allele homozygotes, 1 = heterozygotes, 2 = minor allele homozygotes, and 1/2 = the combined group of heterozygotes and minor allele homozygotes. For example, 0 vs 1/2 means major allele homozygotes versus anyone else. The following two columns indicate sample sizes of the two groups in comparison.

Table 4.4: Replication and meta p-values for genes identified in either PITT or Denver

			Discovery			Replication		Meta	
CHR	Gene	Module	Cohort	Num Var	P-value	Num Var	P-value	Num Var[a]	P-value
6	*HFE*	5	PITT	2	1.09E-07	-[b]	-	-	-
6	*HFE*	22	PITT	2	1.51E-10	-	-	-	-
11	*NECTIN1*	27	PITT	2	2.54E-07	-	-	-	-
13	*CARS2*	10	PITT	2	5.22E-10	-	-	-	-
13	*CARS2*	20	PITT	2	1.91E-09	-	-	-	-
14	*LTB4R*	20	PITT	2	6.88E-08	3	0.96	0	4.39E-04
16	*TELO2*	10	PITT	3	1.19E-09	-	-	-	-
23	*AR*	18	PITT	2	1.77E-07	-	-	-	-
23	*FTSJ1*	1	PITT	2	2.05E-08	-	-	-	-
23	*FTSJ1*	6	PITT	2	2.41E-08	-	-	-	-
23	*FTSJ1*	12	PITT	2	1.73E-07	-	-	-	-
23	*FTSJ1*	28	PITT	2	2.46E-10	-	-	-	-
12	*ZNF268*	18	Denver	2	6.59E-07	4	0.20	2	1.00
19	*ARHGEF18*	30	Denver	3	6.89E-07	6	0.17	3	1.00

[a] Number of variants common in both cohorts
[b] These genes were not testable because they did not have at least two qualified variants in the replication cohorts. As this was a gene-level replication, we count any variants in a gene, rather than the particular variants tested in the discovery cohort.

Table 4.5: Discovery and replication results for individual variants in genes identified in the Denver cohort

								Discovery in Denver		Replication in PITT	
CHR	Gene	SNP	Position	Ref/Alt	Type	CADD score	Module	MAF	P-value	MAF	P-value
12	*ZNF268*	rs80217340	133778798	C/G	missense	9.5	18	0.30%	1.34E-02	0.13%	0.715
		rs200561453	133780754	T/C	stop gain	36.0	18	0.30%	7.36E-01	0.17%	0.125
19	*ARHGEF18*	rs140297736	7512016	G/C	missense	22.4	30	0.60%	1.73E-04	0.43%	0.175
		rs117824875	7533891	A/G	missense	12.2	30	0.46%	4.60E-01	0.61%	0.351
		rs199567237	7505163	A/G	missense	16.5	30	0.53%	4.85E-02	0.65%	0.053

Table 4.6: Significant genes in the meta-analysis of PITT and Denver

Gene	CHR	Module	Meta P-value	P-value in PITT	Num Var in PITT	P-value in Denver	Num Var in Denver	Num var in both[a]
C6	5	11	1.55E-07	7.24E-05	8	6.72E-03	4	3
PCSK9	1	16	5.09E-07	1.72E-06	3	6.09E-01	3	3

[a] Number of variants common in both cohorts

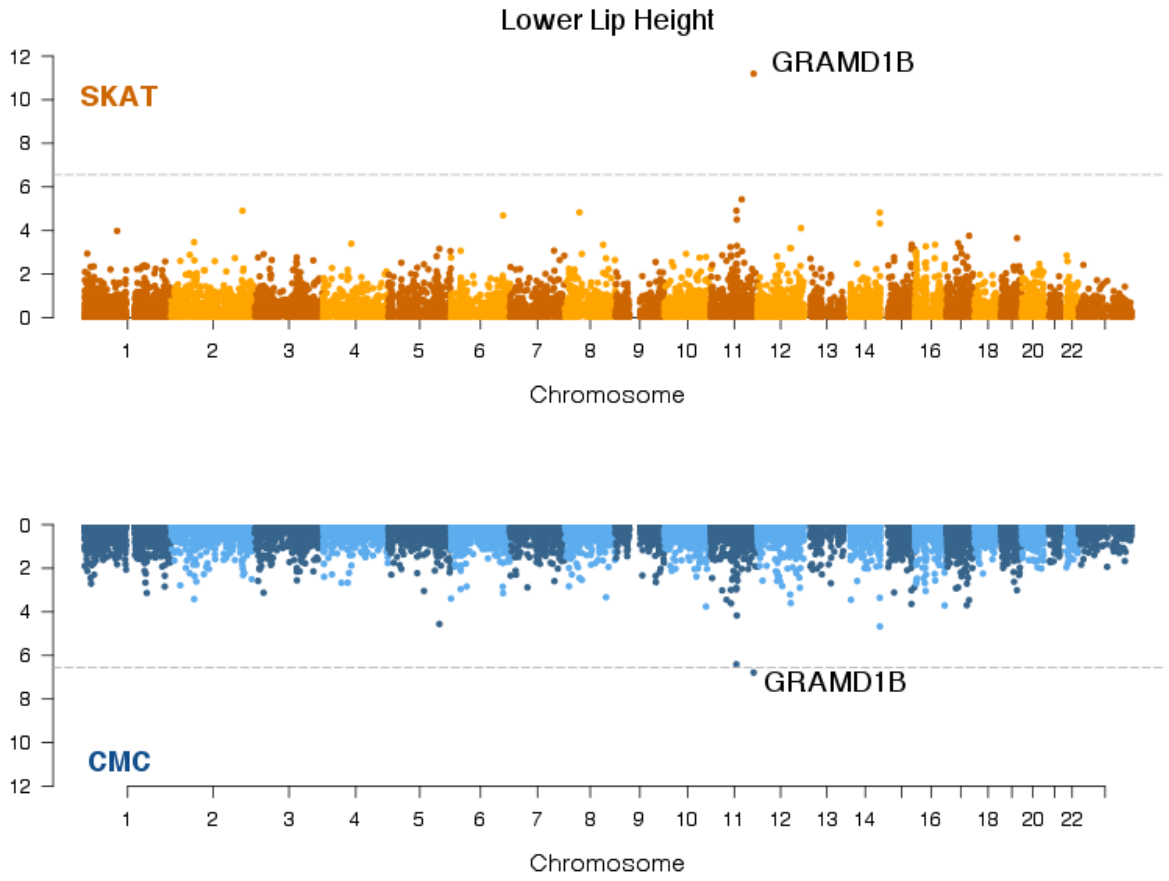


Figure 4.7: Manhattan plot for lower vermilion height highlighting *GRAMD1B* in the PITT cohort. The upper orange figure shows the -1og10-transformed SKAT P-values. The lower blue figure shows the -1og10-transformed CMC p-value, with y-axis in reverse orientation.

Table 4.7: Single variant association for *GRAMD1B* and univariate facial linear distances in the PITT cohort

Gene	Chr	Pheno	SNP	Position	Ref/Alt	MAF	Beta	SE	P-value
GRAMD1B	11	Lower	rs191981781	123477383	A/G	0.082%	6.17	0.92	2.8E-11
		Vermilion	rs7950063	123399604	C/T	0.021%	1.28	1.86	0.49
		Height	rs7935338	123473532	A/G	0.041%	-0.11	1.32	0.94

known to cause orofacial clefts, a craniofacial malformation that can be associated with alterations in facial shape. These results enhanced our understanding of the genetic architecture of human facial variation.

Among the identified genes, *NECTIN1* is a known important player in craniofacial morphogenesis and has been linked with both syndromic and isolated forms of orofacial clefting [130–132]. Individuals with cleft lip/palate-ectodermal dysplasia syndrome (OMIM:225060) have distinctive facial features including an underdeveloped lower jaw [133], which is consistent with the facial segment (chin) where *NECTIN1* association was observed in this study. NECTIN1 protein belongs to the subfamily of immunoglobulin-like adhesion molecules which are key components of cell adhesion junctions, playing essential roles in the fusion of palatal shelves during palatogenesis [134]. A handful of NECTIN1 mutations that potentially disrupt gene function have also been documented in non-syndromic cleft patients [135–137]. In the current study, two coding variants in *NECTIN1* were implicated, and both encode amino acid substitution and were predicted to be deleterious on protein function. Further experimental validation of *NECTIN1* was conducted in the laboratory of Eric C. Liao at Massachusetts General Hospital, where a strong expression of *nectin1a* was observed in zebrafish embryo craniofacial structures and mutants exhibited distorted palate structure. Altogether, our study showed strong supports for the involvement of *NECTIN1* in craniofacial traits.

Our results on *NECTIN1* were consistent with the proposed shared genetics for normal-range facial variation and orofacial clefting [108], and further suggested that rare and low-frequency variants underlie this connection. As a risk gene for facial deformity, *NECTIN1*

was implicated in normal-range facial variation for the first time by our study. There was some evidence from our previous study [126] that *NECTIN1* was associated with NSCL/P in the Latino South American population (Table 4.2), although it did not survive multiple testing correction. Two other genes, *TELO2* and *HFE*, passed the significance threshold in the analysis of combined populations. These results were in line with the existing hypothesis and supported the presence of common genes underlying both normal and abnormal facial development.

The current study is an important extension and complement of our previous GWAS [31] by focusing exclusively on coding variants with a MAF below 1%. Importantly, the two studies generated distinct, non-overlapping candidate genes. Common SNPs in or near (using a 500kb window on either side) the seven identified genes showed little evidence of association (p-value > 0.001 for all) with the same facial modules. Nonetheless, it is possible that there exist trans-acting common GWAS SNPs that regulate the expression of the seven newly identified genes during facial morphogenesis. Low-frequency variants had much more salient effects compared to common SNPs (Figure 4.6). It should be noted that this difference could be a result of the drastically smaller group of variant carriers, and we therefore chose to refrain from a deeper interpretation of the comparison.

Despite similar sample sizes, the African cohort was much less successful in identifying face-associated low-frequency variants than the European cohort. It is reasonable to expect the opposite to be true given that Africans are more genetically diverse than other populations due to their demographic history and possess more population-specific rare and low-frequency variants [138]. On the other hand, greater genetic variation may actually decrease the power to detect those variants due to not enough carriers of specific variants and less LD. We noted a 23% reduction of qualified genes in the analysis of the Tanzania cohort compared to that of the PITT cohort. As rarer variants tend to be younger and more population specific, causal variants in one population may not be detectable in one of a different ancestry. The smaller number of qualified genes in the Tanzania cohort could also be a result from the fact that the Exome SNP chip used (llumina HumanOmni2.5+Exome-8) was not specifically design to capture African-specific variants. Another factor that may contribute

to the difference in the results from the two cohorts was our data-driven facial phenotyping. Morphological variation extracted by PCA was cohort-specific, and can differ a lot in direction in the high dimensional space even in facial segments where the two cohorts have similar counterparts. Furthermore, rare variants are expected to have more unique effects than common SNPs, and therefore do not share genetic effects on different phenotypes.

Our success in gene discovery demonstrated the power of gene-based collective testing of rare and low-frequency variants which are usually untestable individually. While some significant genes harbor variants with a small p-value in our single-variant association tests, others would have been missed if not tested in aggregate. With a moderate sample size of a few thousand, it is highly desirable to collapse variants into biologically meaningful units and perform burden-style tests. In addition to a boost in power, another key benefit with analyzing rare coding variants collectively is the improved biological interpretability compared to GWASs. The gene-centered design of coding variants facilities much clearer biological implications and options for experimental follow-up than the GWAS design where coding and non-coding variants are tested individually. Our success in the functional validation of *NECTIN1* resonates well with this point.

The hierarchical facial segmentation enabled the discovery of genetic effects at different scales. For example, the effect of *FTSJ1* was observed both globally in the full face and locally in much more refined areas on the side of the face. By contrast, the effect of *NECTIN1* was confined to localized facial parts only. These patterns may help with understanding gene action during craniofacial morphogenesis and therefore guide future experiments. The data-driven phenotyping approach obviates the need of selecting traits with some level of arbitrariness, captures more of the variation of facial shape in 3D space and benefits from higher efficiency for gene mapping.

Replication of rare variant association signals presents unique challenges. The prominent barrier is the limited sample sizes of available replication datasets. The scarceness or even absence of the carriers in independent populations hindered the replication efforts of our findings. The PITT and Denver data served as independent replication cohorts for each other, yet only one gene identified in the PITT cohort was available in the Denver cohort.

This failure of replication should not be taken as an indication of spurious associations in the discovery analysis, but rather a lack of ability to seek replication, due to the lack of variant carriers. Similarly, the result that neither of the two genes discovered in the Denver dataset were replicated in the PITT cohort does not necessarily disapprove the signals but can be partly explained by different variants contributing to the gene-level association. We acknowledge this limitation and hope to offset this weakness by testing larger samples with better genome coverage in future studies.

Despite the whole point being specifically targeting rare and low-frequency variants, this study still fell short in covering variants in the lower extreme of the MAF spectrum. Given the limited sample size and the Exome Chip design, this study was not adequately powered to identify genes harboring extra rare variants which may also contribute to facial traits. Taking the PITT cohort as an example, variants with a MAF $< 0.08\%$ were not included in the analysis by our filtering criteria, . Although in general complex traits are not expected to have a large fraction of the heritability explained by rare and private variants, such variants may be influential, predictive, and actionable at the individual level. We look forward to whole-exome or whole-genome sequencing efforts of large samples in the near future, which will be able to give deeper insights on the contribution of rare genetic variations and private mutations.

Meta-analysis to combine results from multiple cohorts is a common practice in GWAS and an effective approach to increase power. It, however, poses special challenges for studies whose primary focus is on rare and low-frequency variants. In a strict sense, meta-analysis requires all cohorts to be polymorphic at the exact same variants in a gene. This "commonality" becomes less and less likely as the variants become rarer. With a sample size of a few thousand, sharing the same variants across cohorts is improbable for low-frequency variants. Even when the summary statistics from individual cohorts are combined at a gene level, and thus do not require the genes to have exact same variants, a fair number of genes would still not be testable in all cohorts. That being said, the very nature of the variants under study largely diminishes the benefit of a meta-analysis. Although this study did not require a gene to have identical variants in different cohorts in the meta-analysis, we still landed with a

substantial loss of testable genes. Only three out of the nine genes identified in individual cohorts were common in both; none reached significance level in the meta-analysis. Future research should seek to phenotype more individuals for a better characterization of the role of rare and low-frequency variants on facial morphology.

Our findings have implications for future studies on polygenic morphological trait and the prediction of such phenotypes in the context of precision medicine. An attention to rare and low-frequency genetic variants is strongly encouraged for future studies on human craniofacial traits, including both normal-range variation and deformity and malformation. Although less-frequent variants may not explain sizable heritable variation at the population level, they can have crucial effect for the individuals who carry them. Given that rare and low-frequency genetic variation can be highly specific to certain populations and facial shapes have distinctive ancestry features, future studies may benefit from extending the discovery of influential low-frequency variants to other ethnic groups. As sample sizes of carefully phenotyped human cohorts continue to increase, we expect that additional rare and low-frequency variants will be uncovered and they will further enhance our understanding of facial morphology.

5.0 Transcriptome-Wide Association Study

5.1 Background

5.1.1 Overview

The Aim 2 of this dissertation was to study facial morphology from a perspective of the regulatory role of genetic variants, by conducting TWAS and several follow-up analyses. The context under which this aim was proposed largely concerns the infamous issue of unknown function and the challenge of locating the causal gene that accounts for an association signal in GWASs. Specifically, our GWAS in a larger cohort which included the PITT sample has identified over 200 significant hits, most of which fall in non-coding intergenic regions and do not harbor genes with a known role in facial morphology [139]. Connecting 200 loci to their target genes and mechanism in functional experiments requires substantial amount of resources. As an alternative route for identifying putative causal genes, this aim was proposed to leverage the external molecular QTL data which can be used to impute gene expression in unmeasured samples.

Many existing TWAS applications have shown the dual benefits of TWAS design: (1) pinpointing putative causal genes at GWAS loci and (2) discovering novel genes not overlapping known GWAS loci [5, 6, 77–79]. The former is deemed as an especially appealing application of TWAS, as GWAS loci usually harbor multiple genes and in most cases it's remarkably challenging to make immediate mechanistic interpretations of any of them. The gain of power in the second benefit (compared to GWAS) comes from the cumulative genetic effects from multiple eQTLs as well as an eased burden of multiple testing. The current study benefited from both features of the TWAS design.

It is important to keep in mind that TWAS does not perform causal inference when using it as a fine-mapping approach to augment GWAS. One should seek for evidence supporting a causal role from multiple sources. One such source can be a colocalization of GWAS and eQTL signals. A high probability of colocalization adds to the amount of support that the

target gene of the eQTLs accounts for the respective GWAS signal, in particular through regulating gene expression. Another line of evidence can come from conditional analysis, where the predicted expression level of a TWAS gene is adjusted for as a covariate in the SNP-trait association test. The current study followed top TWAS genes using both of these strategies to augment the original TWAS results.

We acknowledge that although getting closer to the underlying mechanism is the primary goal, TWAS does not necessarily imply causality. The prioritized genes can still be non-causal and need further verification. Aim 2 was therefore positioned not as an aim to establish causality but rather an aim where the main focus is prioritization and hypothesis generation. Genes discovered here will need independent replication as well as functional validation in future work.

5.1.2 MetaXcan and COLOC software

A number of existing computational programs are available to carry out the analyses outlined in this aim. The exhaustive comparison and application of various methods was not a goal of this project; we chose to use PrediXcan/MetaXcan (referred to as MetaXcan hereafter), based on its improved performance and its acceptance of individual-level data as input. Features of MetaXcan that were relevant to this study are described below. MetaXcan released two different families of prediction models trained on GTEx v8: MASHR-based and Elastic Net (EN) -based [80]. MASHR stands for Multivariate Adaptive Shrinkage in R, which was used to smooth cis-eQTL effect size and standard error estimates by taking advantage of the correlation of gene expression regulation among different tissues. The resulting models are parsimonious, which means that only putative causal eQTLs are used as predictors. In short, MASHR-based models use fine-mapped variants and are biologically informed. By contrast, EN-based models use all cis-eQTLs of a gene without making further prioritization. It is a safe, robust alternative with decreased prediction performance, and is directly comparable to past versions of MetaXcan models. All models use locally-acting eQTLs (MAF $> 1\%$) located within 1 MB upstream/downstream of the target gene transcript start/end sites.

When signals from two different GWASs are located at the same loci, it is not clear whether the same variants are responsible for both or whether it is distinct causal variants close to each other. Intermediate cellular phenotypes, in particular gene expression, are commonly involved in such two-GWAS comparisons with the hypothesis being that gene expression mediates the disease signals. A number of GWAS–eQTL colocalization methods were developed to test this hypothesis, with popular ones including eCaviar [7], RCT [140], Sherlock [141], COLOC [142], enloc [143] and so on. No single best approach fits all situations; different methods are based on different underlying models and outperform others when its assumptions hold. COLOC [142] is a Bayesian procedure which gives intuitive posterior probabilities which are easily interpreted. It requires only p-values and SNPs' MAF, and its analytical solution enables genome-wide scale computation. For these reasons, we selected COLOC to test for colocalization between gene expression and facial morphology.

5.2 Methods

5.2.1 Workflow

All Aim 2 analyses were conducted in individual cohorts separately. A flowchart of the analysis can be viewed in Figure 5.1. Specifically, MetaXcan prediction models in the four chosen tissues were first used to obtain the genetically regulated gene expression (GRex), which were then tested for association against 63 multivariate facial modules by canonical correlation analysis. Genes with a p-value less than the predetermined threshold went into the follow-up stage where they were further examined by conditional and colocalization analysis. In parallel with the follow-up inspection, significant genes went through extensive annotation, literature search and bioinformatics analysis. Significant and suggestive genes collectively as a group were also annotated by gene-set enrichment tools. The following subsections describe each step in this workflow in more detail.

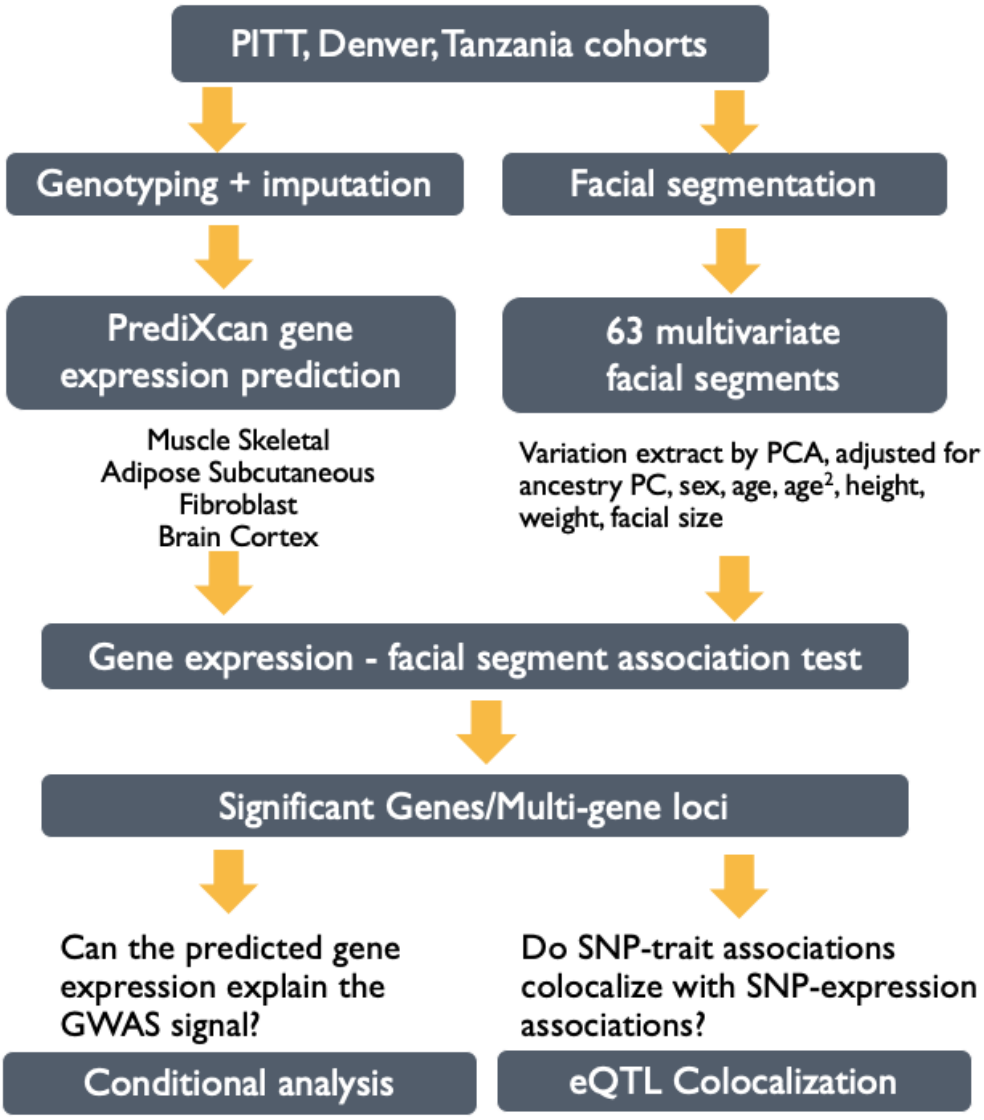


Figure 5.1: TWAS flowchart

5.2.2 Predicting gene expression by MetaXcan

MetaXcan provides 49 GTEx tissue types and requires a decision made by users regarding which tissue(s) to use. Due to the context-dependent nature of gene expression regulation, the ideal choice would be the mechanistically relevant tissue(s) for the phenotype under study. However, for morphological traits primarily formed during embryonic development, the optimal tissues and cells are rarely available. Currently there is no available eQTL data in embryonic craniofacial tissues and their developmental precursors. This study therefore selected four GTEx tissues that are highly related to facial morphology: muscle skeletal, subcutaneous adipose, cultured fibroblasts and brain cortex.

This study applied both MASHR-based and EN-based models in MetaXcan to complement the results from one another. Although EN-based models presumably have inferior performance compared to MASHR-based models, we found our analysis could benefit from them in two aspects. First, the sets of SNPs available in our data (either genotyped or imputed) were better matched to the set of SNPs in EN-based models. In the Tanzania cohort, 80% of the eQTLs in EN-based models were available; this number dropped down to 55% for MASHR models. With the smaller intersection, fewer genes would be predictable and testable. Second, the two types of models each have their unique genes although the majority overlap. Taking the union of the two sets will give a larger set of genes available for association test.

SNPs passing the quality control criteria in the original GWAS of each cohort [31] were used to predict gene expression. Specifically, on top of the basic cleaning steps, we further required SNPs to have a MAF >1% and imputation INFO score > 0.5 (for imputed SNPs) to be included.

5.2.3 Testing the association between the genetically regulated gene expression (GRex) and facial modules

Canonical correlation analysis (CCA) was used to test the association between GRex and 63 multivariate facial modules using the R package CCA [144]. CCA finds linear combi-

nations of the dependent variables (PCs) that maximally correlate with linear combinations of the independent variables (GRex) and gives the canonical correlation coefficients between the paired linear combinations. The canonical correlations were then subject to an asymptotic test to assign statistical significance, using the F-approximation of Wilks's Lambda as a test statistic. CCA was used in our previous GWAS in the same cohort on the same set of phenotypes [31]; applying the same statistical method would enable a straightforward comparison between the results of this TWAS and those of the previous GWAS, by which GWAS loci can be fine-mapped.

To be consistent with the multiple testing correction strategy in the previous GWAS for the PITT cohort, we considered two levels of significance thresholds in this study. The study-wide significance threshold was calculated as 0.05/(effective number of independent facial modules $\times$ number of genes tested in a tissue). The effective number of independent modules was computed according to the eigenvalue-based procedure by Li and Ji's [119], which was 39 for the PITT modules (which were also the Denver modules) and 40 for the Tanzania modules. With the intention not to miss any potential candidate genes, a relaxed, genome-wide significance threshold was also used to identify suggestive genes, which equaled 0.05/number of genes tested in a tissue. This study did not further account for the four tissue types and the two families of models, in consideration of the co-regulation of gene expression across tissues and the large number of overlapping genes between models.

5.2.4 Follow-up analysis of TWAS genes

TWAS does not necessarily imply causality. One of the most common causes is LD-contamination, which occurs when the expression predictor SNPs and phenotype-causing SNPs are different but in LD. Additional analysis is required to rule out spurious association due to LD-contamination. Such analysis can also help with a further prioritization of multi-gene TWAS loci. We conducted two types of follow-up analyses – conditional analysis for significant TWAS genes and colocalization analysis for multi-gene TWAS loci.

In a conditional analysis, the marginal SNP-trait association was first obtained by Multivariate Analysis of Variance (MANOVA), which is equivalent to CCA. Next, the GRex was

added as a covariate with all other terms being the same. This new model was evaluated by Multivariate Analysis of Covariance (MANCOVA). This analysis considered a 1 MB flanking window on either side of the gene.

eQTL Colocalization analysis was carried out for two multi-gene TWAS loci (the chromosome 1p12 and 3q21.3). COLOC [142] was used to estimate the posterior probability of shared or independent GWAS-eQTL signal. COLOC's Approximate Bayes Factor (ABF) analysis estimates the support for five hypotheses:
H0: neither trait has a genetic association in the region
H1: only trait 1 has a genetic association in the region
H2: only trait 2 has a genetic association in the region
H3: both traits are associated, but with different causal variants
H4: both traits are associated and share a single causal variant

A high posterior probability (PP) of H4 indicates evidence for colocalization. On the other hand, a low PP of H4 may not indicate evidence against colocalization but rather simply be the result of limited power, which is supported by a large sum of PP of H0, H1 and H2. By assessing all hypotheses simultaneously, distinct GWAS and eQTL signals can be distinguished from low power.

Summary statistics of the SNP-gene expression association analysis in the four tissues selected for TWAS were downloaded from GTEx v8 data repository. Summary statistics of the facial module GWAS were obtained from our previous publication [31]. Facial GWAS SNPs were first liftovered to assembly hg38 and then harmonized with the eQTL statistics. Following the criteria suggested by Barbeira et al. [82], we demanded a colocalized signal to simultaneously fulfill (1) PP of H4 > 0.5, (2) PP of H3 < 0.5, and (3) PP of H0+H1+H2 $<$ 0.3. To visualize the colocalization, the R package LocusCompareR [145] was used to make joint scatter plots for facial associations and gene expression associations.

5.2.5 Gene-set enrichment analysis

FUMA gene-set enrichment analyses were performed to investigate the potential functions of the identified genes. FUMA is a functional mapping and annotation tool for genetic associations [121], and has a GENE2FUNC function module for gene-set enrichment analysis. We provided two lists of genes as input: (1)The nine genes that passed our study-wide threshold adjusting for both number of phenotypes and number of genes in any of the three cohorts; and (2) an expanded list of 71 genes, further incorporating suggestive genes below the more liberal genome-wide p-value cutoff, discovered in any of the three cohorts. The parameters used for FUMA were FDR adjusted p-value (Benjamini-Hochberg procedure) $<$ 0.05 and a minimum of two overlapping genes.

5.3 Results

5.3.1 TWAS

Table 5.1 gives an overview of the TWAS results, listing the number of genes tested in each tissue type using each of the two models, the corresponding significant p-value thresholds, and the number of hits for PITT, Denver and Tanzania cohorts separately. Overall, TWAS in three cohorts identified nine genes at six loci (*WARS2/HAO2*, *RUVBL1/EEFSEC/SEC61A1*, *C1orf53*, *EYA4*, *CCDC91*, *SLK*) being associated with at least one facial segment with a p-value below the study-wide significance level. Table 5.2 details the analysis parameters and association statistics. All but *SLK* were found in the analysis of the PITT cohort; *SLK* was discovered in the Denver cohort; *EEFSEC* was the only significant gene in the Tanzania cohort analysis. *RUVBL1*, *EEFSEC*, *SEC61A1* and *EYA4* showed association with multiple facial segments. Except for *CCDC91*, all were located at previously identified face-associated GWAS loci [31], and the associated facial modules of the same gene in the GWAS and the TWAS were largely consistent. This implies an overall decent power and reliable performance of the TWAS method. TWAS statistics were well-behaved and there was no evidence of inflation or deflation (Figure C5 and C6).

Table 5.1: Overview of TWAS results

Cohort	Model Type	Tissue	Num Genes	Study-wide threshold	Num Significant Hits	Genome-wide threshold	Num Suggestive Hits
PITT	MASHR	Muscle Skeletal	11915	1.10E-07	3	4.20E-06	13
		Brain Cortex	12344	1.00E-07	3	4.10E-06	15
		Cells Cultured Fibroblasts	12534	1.00E-07	6	4.00E-06	19
		Adipose Subcutaneous	13029	9.80E-08	9	3.80E-06	22
	EN	Muscle Skeletal	7566	1.70E-07	1	6.60E-06	6
		Brain Cortex	5478	2.30E-07	7	9.10E-06	18
		Cells Cultured Fibroblasts	8911	1.40E-07	4	5.60E-06	14
		Adipose Subcutaneous	8617	1.50E-07	3	5.80E-06	10
Denver	MASHR	Muscle Skeletal	11863	1.10E-07	0	4.20E-06	4
		Brain Cortex	12283	1.00E-07	1	4.10E-06	2
		Cells Cultured Fibroblasts	12495	1.00E-07	0	4.00E-06	3
		Adipose Subcutaneous	13001	9.90E-08	0	3.80E-06	4
	EN	Muscle Skeletal	7566	1.70E-07	0	6.60E-06	2
		Brain Cortex	5477	2.30E-07	0	9.10E-06	0
		Cells Cultured Fibroblasts	8909	1.40E-07	0	5.60E-06	1
		Adipose Subcutaneous	8614	1.50E-07	0	5.80E-06	1
Tanzania	MASHR	Muscle Skeletal	9600	1.30E-07	0	5.20E-06	2
		Brain Cortex	9716	1.30E-07	1	5.20E-06	3
		Cells Cultured Fibroblasts	10325	1.20E-07	0	4.80E-06	4
		Adipose Subcutaneous	10668	1.20E-07	0	4.70E-06	4
	EN	Muscle Skeletal	7558	1.70E-07	0	6.60E-06	3
		Brain Cortex	5467	2.30E-07	1	9.10E-06	6
		Cells Cultured Fibroblasts	8894	1.40E-07	0	5.60E-06	5
		Adipose Subcutaneous	8614	1.50E-07	0	5.80E-06	6

Table 5.2: Study-wide significant TWAS genes

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
PITT	12p11.22	*CCDC91*	21	4.37E-08	Cells_Cultured_fibroblasts	MASHR	3
	1p12	*HAO2*	7	2.52E-08	Adipose_Subcutaneous	EN	13
	1p12	*WARS2*	7	2.03E-09	Muscle_Skeletal	MASHR	4
	1p12	*WARS2*	7	1.86E-09	Cells_Cultured_fibroblasts	MASHR	3
	1p12	*WARS2*	7	1.11E-08	Brain_Cortex	MASHR	2
	1q31.3	*C1orf53*	27	1.22E-07	Muscle_Skeletal	EN	2
	3q21.3	*EEFSEC*	2	2.07E-09	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	2	1.40E-07	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	4	6.30E-08	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	8	9.45E-11	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	8	1.26E-09	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	16	4.50E-09	Adipose_Subcutaneous	EN	50
	3q21.3	*EEFSEC*	16	1.22E-09	Cells_Cultured_fibroblasts	EN	51
	3q21.3	*EEFSEC*	16	3.33E-16	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	16	9.99E-16	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	16	4.32E-10	Cells_Cultured_fibroblasts	MASHR	2
	3q21.3	*EEFSEC*	16	6.92E-11	Brain_Cortex	MASHR	2
	3q21.3	*EEFSEC*	32	2.62E-09	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	32	1.13E-09	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	33	2.85E-10	Cells_Cultured_fibroblasts	EN	51
	3q21.3	*EEFSEC*	33	7.40E-08	Adipose_Subcutaneous	EN	50
	3q21.3	*EEFSEC*	33	4.06E-14	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	33	1.93E-11	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	33	9.07E-11	Cells_Cultured_fibroblasts	MASHR	2
	3q21.3	*EEFSEC*	33	2.77E-11	Brain_Cortex	MASHR	2
	3q21.3	*EEFSEC*	45	1.42E-08	Adipose_Subcutaneous	MASHR	2
	3q21.3	*EEFSEC*	45	7.25E-08	Brain_Cortex	EN	8
	3q21.3	*RUVBL1*	2	2.58E-09	Cells_Cultured_fibroblasts	MASHR	2
	3q21.3	*RUVBL1*	16	1.13E-08	Muscle_Skeletal	MASHR	2
	3q21.3	*RUVBL1*	16	1.06E-09	Cells_Cultured_fibroblasts	MASHR	2
	3q21.3	*RUVBL1*	33	9.25E-10	Muscle_Skeletal	MASHR	2
	3q21.3	*SEC61A1*	16	3.17E-09	Cells_Cultured_fibroblasts	EN	13
	3q21.3	*SEC61A1*	33	8.67E-08	Cells_Cultured_fibroblasts	EN	13
	6q23.2	*EYA4*	30	3.54E-09	Adipose_Subcutaneous	MASHR	1

Table 5.2: Continued from previous page

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
	6q23.2	*EYA4*	60	1.03E-09	Adipose_Subcutaneous	MASHR	1
	6q23.2	*EYA4*	61	1.29E-08	Adipose_Subcutaneous	MASHR	1
Denver	10q24.33	*SLK*	21	8.40E-08	Brain_Cortex	MASHR	2
Tanzania	3q21.3	*EEFSEC*	27	2.54E-08	Brain_Cortex	EN	8
	3q21.3	*EEFSEC*	27	1.14E-07	Brain_Cortex	MASHR	2

[a] Number of SNPs used to predict gene expression

Taking the results in the PITT cohort as an example, Figure 5.2 presents a composite Manhattan plot overlaying p-values for all modules and all tissues. Significant loci were pointed to their associated phenotypes (in blue). Propagation of signals in facial segments that are linked across different levels can be clearly visualized in the plot.

EEFSEC is the sole significant gene shared across cohorts. It is also a multi-effect gene, i.e. had associations with different traits spanning multiple facial quadrants. In the PITT hierarchical clustering rosette, the strongest association of *EEFSEC* was seen in the mouth and philtrum quadrant, with gene's effects started being detectable in the first level of the hierarchy (module 2) and went all the way to the most localized level (module 32 and 33, upper lip and philtrum). It also showed an effect in the nose quadrant, which has a close physical proximity to and shares the same upper level segment with lip and philtrum segments. This association pattern across multiple regions echoed the SNP-trait association at this locus in previous GWAS [31]. In the analysis of the Tanzania cohort, although the signal of *EEFSEC* was less strong, it did show association in the similar nose quadrant. Besides its robustness across cohorts, *EEFSEC* signals were also detected across all four tested tissues.

In addition to the nine study-wide significant genes, we also identified 62 suggestive genes that surpassed the genome-wide threshold although did not reach the study-wide cutoff (Table C1). Some of them were at or nearby facial GWAS association loci, such as *CRB1*, *WARS2* and *RUNX2*, yet most suggestive genes have not been implicated in craniofacial traits before. These genes showed associations with the shape of a variety of facial regions.

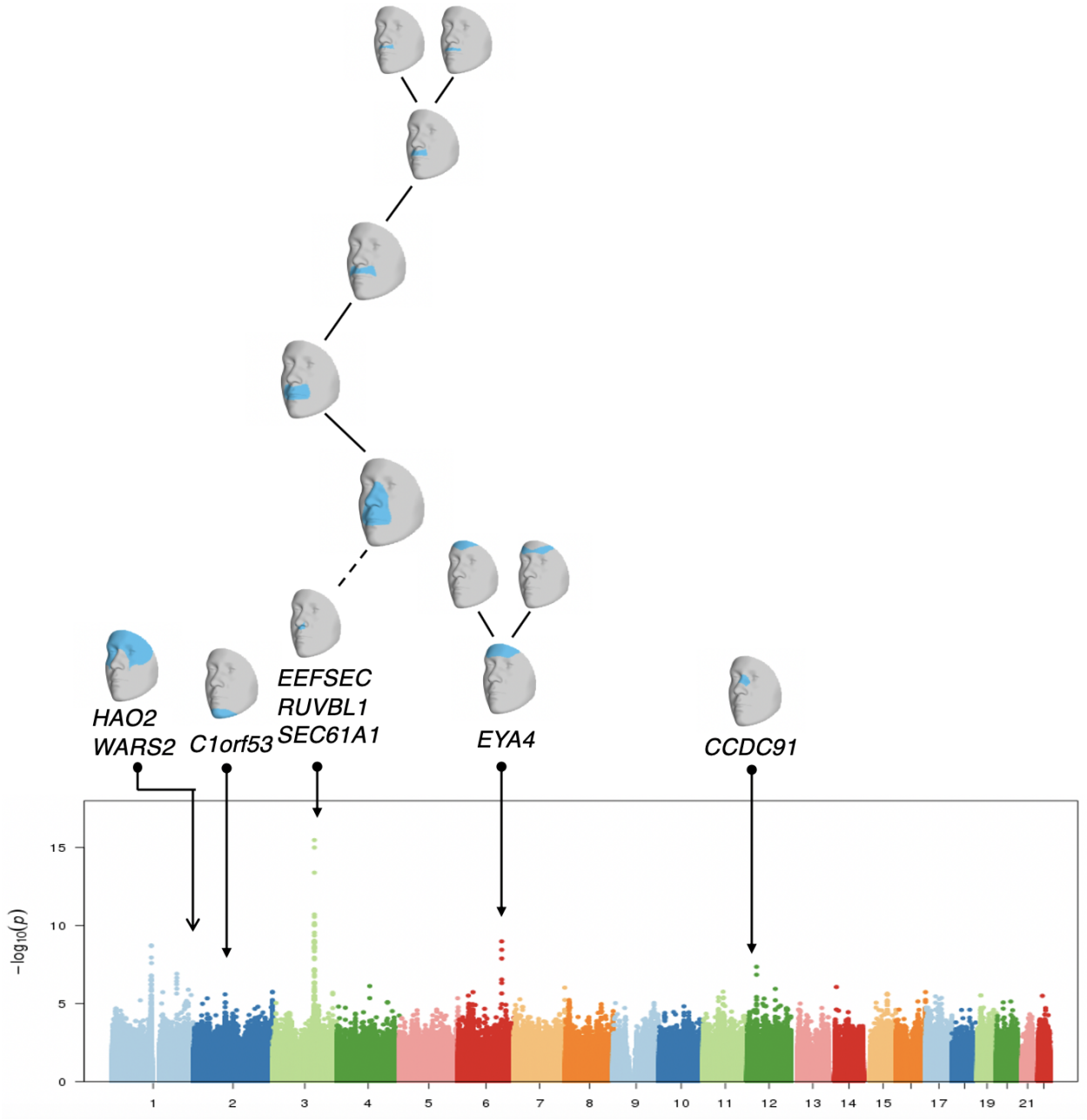


Figure 5.2: Composite Manhattan plot for TWAS results in the PITT cohort. The plot shows the position of genes on the x-axis and TWAS p-values on the y-axis. A total of 63×4 points are plotted for each gene, representing their p-values for 63 modules in the analysis of 4 tissues respectively. Horizontal lines for significance threshold were not drawn because the calculation was tissue-specific. The associated facial modules for each gene that surpassed the study-wide threshold are shown above the Manhattan plot.

5.3.2 Conditional analysis for significant TWAS genes

Results from the conditional analysis are summarized in Table 5.3. All genes but *CCDC91* were located at/near to a GWAS peak, as indicated by a minimum SNP-trait association p-value $< 5\times10^{-8}$ before conditioning. After conditioning on the predicted expression, GWAS signals at *CCDC91*, *EEFSEC*, *EYA4*, *HAO2*, and *SLK* were no longer significant (shown by a minimum p-value $> 5\times10^{-8}$ in the 5th column). On the contrary, although conditioning on *C1orf53* largely undermined the SNP-trait association (a 10-order-of-magnitude larger p-value), this locus remained to be a genome-wide significant one. For *RUVBL1*, *SEC61A1*, and *WARS2*, there were mixed results in different tissues and they were not able to fully explain the GWAS association.

Table 5.3: SNP-trait association p-values before and after conditioning on the predicted expression of TWAS significant genes

Cohort	Gene	Module	Minimum p-value before conditioning	Minimum p-value after conditioning	Tissue	Model
PITT	*C1orf53*	27	7.55E-19	8.88E-09	Muscle_Skeletal	EN
PITT	*CCDC91*	21	2.53E-06	1.19E-04	Cells_Cultured_fibroblasts	MASHR
PITT	*EEFSEC*	2	7.86E-10	8.22E-04	Adipose_Subcutaneous	MASHR
PITT	*EEFSEC*	2	7.86E-10	3.18E-05	Brain_Cortex	EN
PITT	*EEFSEC*	4	8.00E-09	4.71E-04	Brain_Cortex	EN
PITT	*EEFSEC*	8	3.05E-12	1.34E-03	Adipose_Subcutaneous	MASHR
PITT	*EEFSEC*	8	3.05E-12	1.45E-03	Brain_Cortex	EN
PITT	*EEFSEC*	16	1.43E-17	1.64E-07	Adipose_Subcutaneous	EN
PITT	*EEFSEC*	16	1.43E-17	7.14E-05	Adipose_Subcutaneous	MASHR
PITT	*EEFSEC*	16	1.43E-17	7.60E-05	Brain_Cortex	EN
PITT	*EEFSEC*	16	1.43E-17	2.33E-06	Brain_Cortex	MASHR
PITT	*EEFSEC*	16	1.43E-17	1.13E-07	Cells_Cultured_fibroblasts	EN
PITT	*EEFSEC*	16	1.43E-17	3.63E-07	Cells_Cultured_fibroblasts	MASHR
PITT	*EEFSEC*	32	4.44E-10	1.87E-03	Adipose_Subcutaneous	MASHR
PITT	*EEFSEC*	32	4.44E-10	2.04E-03	Brain_Cortex	EN
PITT	*EEFSEC*	33	3.63E-15	4.94E-07	Adipose_Subcutaneous	EN
PITT	*EEFSEC*	33	3.63E-15	1.72E-03	Adipose_Subcutaneous	MASHR

Table 5.3: Continued from previous page

Cohort	Gene	Module	Minimum p-value before conditioning	Minimum p-value after conditioning	Tissue	Model
PITT	*EEFSEC*	33	3.63E-15	2.29E-03	Brain_Cortex	EN
PITT	*EEFSEC*	33	3.63E-15	1.27E-03	Brain_Cortex	MASHR
PITT	*EEFSEC*	33	3.63E-15	1.90E-04	Cells_Cultured_fibroblasts	EN
PITT	*EEFSEC*	33	3.63E-15	4.83E-04	Cells_Cultured_fibroblasts	MASHR
PITT	*EEFSEC*	45	3.20E-08	4.32E-04	Adipose_Subcutaneous	MASHR
PITT	*EEFSEC*	45	3.20E-08	4.09E-04	Brain_Cortex	EN
PITT	*EYA4*	30	6.14E-12	1.43E-04	Adipose_Subcutaneous	MASHR
PITT	*EYA4*	60	1.86E-11	3.49E-03	Adipose_Subcutaneous	MASHR
PITT	*EYA4*	61	3.34E-10	1.22E-03	Adipose_Subcutaneous	MASHR
PITT	*HAO2*	7	5.75E-14	1.20E-06	Adipose_Subcutaneous	EN
PITT	*RUVBL1*	2	7.86E-10	6.89E-04	Cells_Cultured_fibroblasts	MASHR
PITT	*RUVBL1*	16	1.43E-17	1.65E-08	Cells_Cultured_fibroblasts	MASHR
PITT	*RUVBL1*	16	1.43E-17	1.15E-08	Muscle_Skeletal	MASHR
PITT	*RUVBL1*	33	3.63E-15	6.75E-05	Muscle_Skeletal	MASHR
PITT	*SEC61A1*	16	1.43E-17	4.94E-07	Cells_Cultured_fibroblasts	EN
PITT	*SEC61A1*	33	3.63E-15	5.51E-07	Cells_Cultured_fibroblasts	EN
PITT	*WARS2*	7	5.75E-14	1.78E-08	Brain_Cortex	MASHR
PITT	*WARS2*	7	5.75E-14	6.83E-07	Cells_Cultured_fibroblasts	MASHR
PITT	*WARS2*	7	5.75E-14	1.55E-08	Muscle_Skeletal	MASHR
Denver	*SLK*	21	1.14E-08	1.87E-03	Brain_Cortex	MASHR
Tanzania	*EEFSEC*	27	1.55E-09	1.72E-05	Brain_Cortex	EN
Tanzania	*EEFSEC*	27	3.20E-10	2.78E-06	Brain_Cortex	MASHR

The chromosome 3q21.3, a multi-gene TWAS locus, harbors three genes (*EEFSEC*, *SEC61A1*, *RUVBL1*) which were significant in at least one TWAS test, and one suggestive gene (*RPN1*). Taking results from PITT module 16 in adipose as an example, Figure 5.3 shows the SNP-trait association p-values before (in orange) and after (in black) conditioning on the GRex of *EEFSEC*. By completely flattening the GWAS peak, the conditional model showed strong evidence that *EEFSEC* is the putative mediator at this locus.

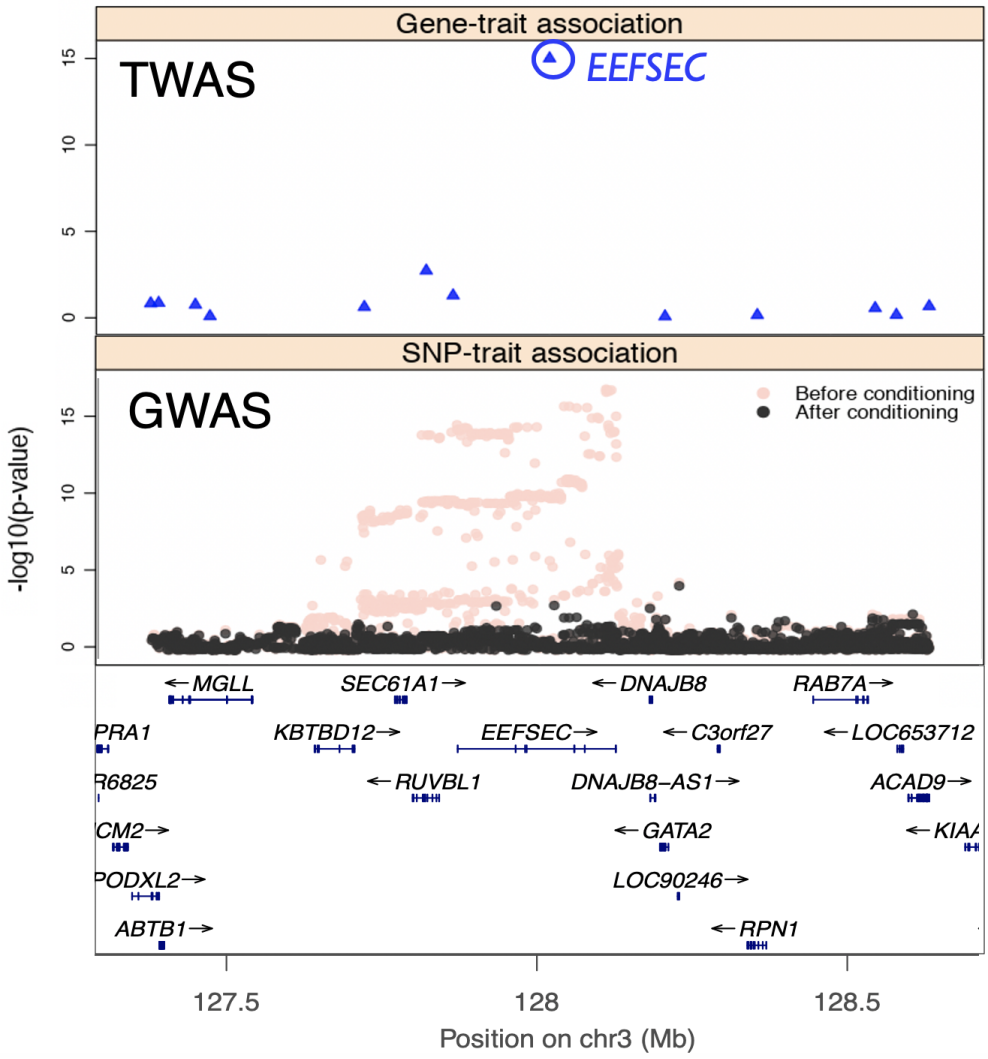


Figure 5.3: SNP-philtrum associations conditioning on *EEFSEC* at the 3q21.3 locus. Top TWAS panel displays the GRex-module 16 (philtrum) association p-values for genes at 3q21.3 locus, obtained in subcutaneous adipose tissue. The GWAS panel displays the SNP association p-values at the same locus, with orange and black dots showing the unconditioned and conditioned results, respectively. Conditioning on the GRex of *EEFSEC* significantly diminished the GWAS signal, in support of a mediating role of *EEFSEC*.

5.3.3 COLOC analysis of multi-gene TWAS loci

We performed phenotype-eQTL colocalization analysis for the two multi-gene TWAS loci, 3q21.3 and 1p12. For 3q21.3, GWAS summary statistics of *EEFSEC*, *SEC61A1*, *RUVBL1*, and *RPN1* with PITT module 16 and 33 and their eQTL signals in four GTEx tissues were used as COLOC input data. For 1p12, we used GWAS summary statistics of *WARS2* and *HAO2* with PITT module 7. Table 5.4 lists signals that met our criteria for a shared causal SNP. At the 3q21.3, COLOC pointed to *EEFSEC* in brain cortex and adipose tissue, and to *SEC61A1* in muscle skeletal tissue. Figure 5.4 is a graphic illustration for the first row in Table 5.4. The colocalized signal is supported by the "diagonal pattern" in the left panel–p-values from the two plotted GWASs aligned along the diagonal given the local LD pattern. At the 1p12 locus, *HAO2* was highlighted in adipose tissue (Figure 5.5) whereas *WARS2* did not colocalize. A complete list of COLOC results (including those tested but did not meet the criteria for a colocalized signal) can be seen in Table C2.

Table 5.4: COLOC analysis of multi-gene TWAS loci (only show the colocalized signals)

Locus	Module	Gene	Tissue	PP.H0+H1+H2[a]	PP.H3[b]	PP.H4[c]
3q21.3	33	*EEFSEC*	Brain_Cortex	1.14E-03	0.34	0.66
	33	*SEC61A1*	Muscle_Skeletal	1.30E-02	0.39	0.60
	33	*EEFSEC*	Adipose_Subcutaneous	3.99E-07	0.45	0.55
1p12	7	*HAO2*	Adipose_Subcutaneous	9.83E-04	0.47	0.53

[a] A small sum of the PP of H0, H1 and H2 indicates a high power for the colocalization analysis

[b] Posterior probability of different causal variants

[c] Posterior probability of shared causal variant

5.3.4 Gene-set enrichment analysis

Several GWAS catalog phenotypes were enriched among the nine study-wide significant TWAS genes, including gestational age at birth, chronic obstructive pulmonary disease, anthropometric measurements and cancers (Figure 5.6a). Most of these terms remained

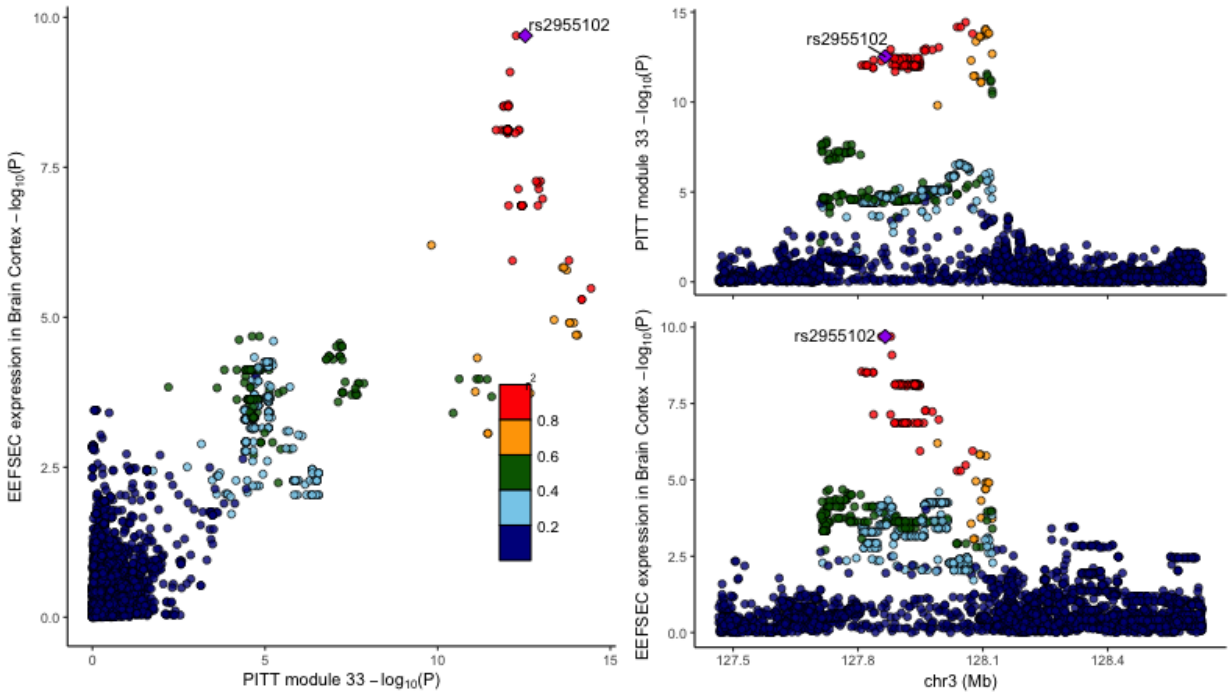


Figure 5.4: Locuscompare plot for *EEFSEC*. Top right: regional scatter plot for PITT module 33 GWAS. Bottom right: regional scatter plot for *EEFSEC* expression GWAS in brain cortex. Left: joint distribution of the p-values from the two GWASs. Color indicates LD with the lead SNP.

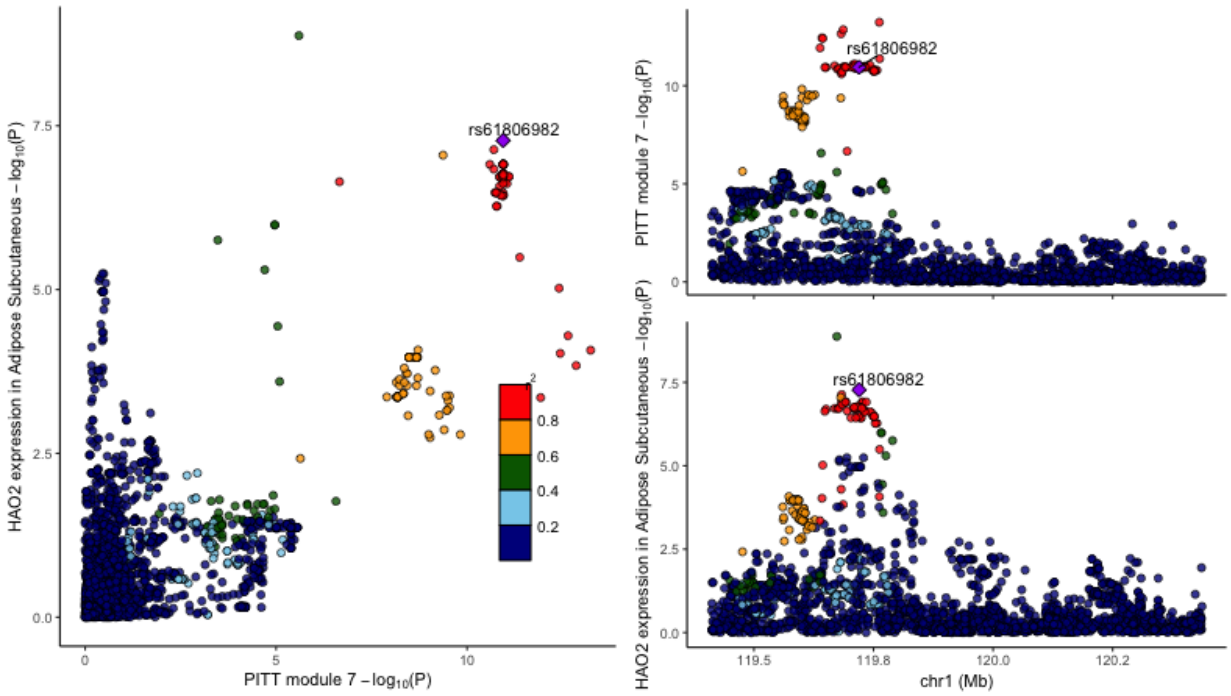
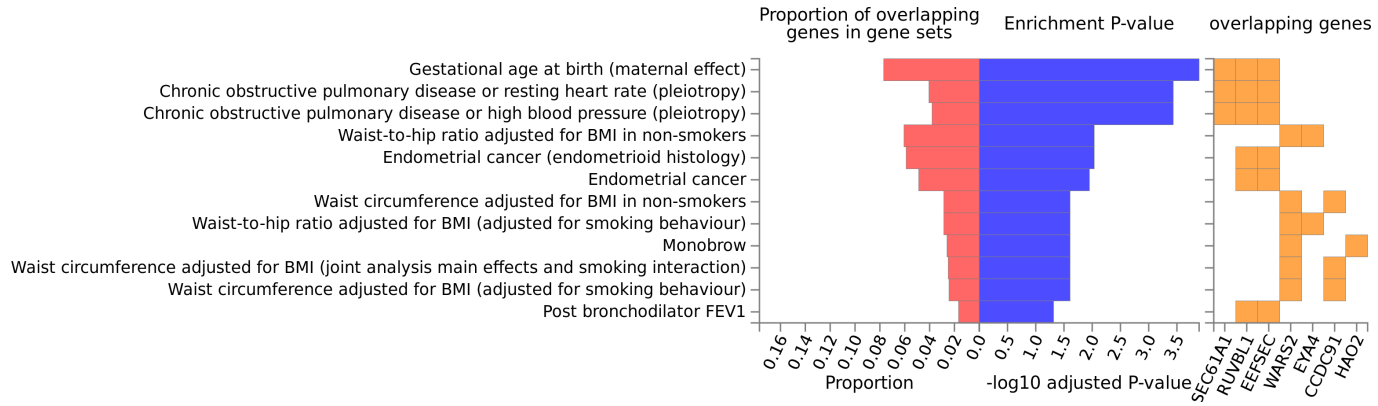
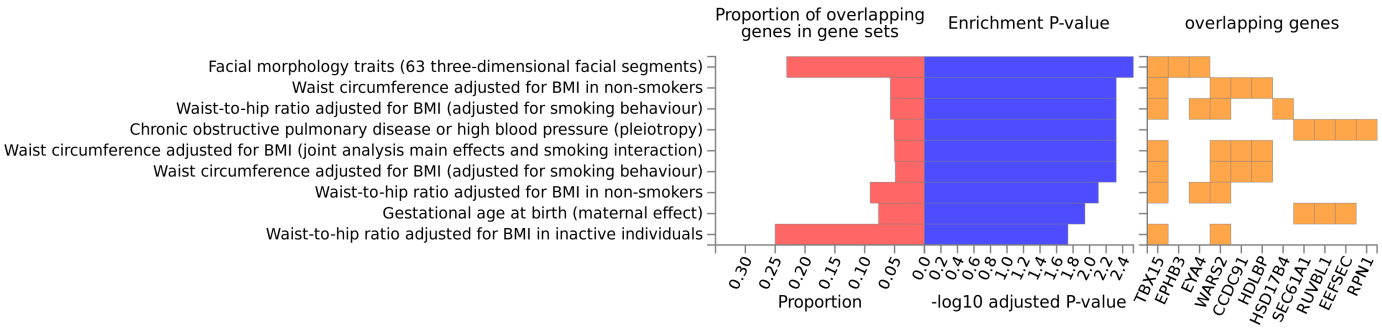


Figure 5.5: Locuscompare plot for *HAO2*. Top right: regional scatter plot for PITT module 7 GWAS. Bottom right: regional scatter plot for *HAO2* expression GWAS in adipose subcutaneous. Left: Joint distribution of the p-values from the two GWASs. Color indicates LD with the lead SNP.

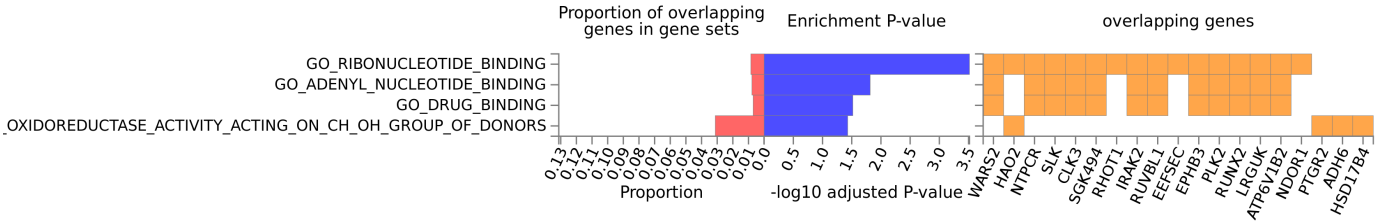
significant when the extended list comprised of 71 significant and suggestive genes was used as input (Figure 5.6b). The top enriched gene ontology (GO) molecular functions were ribonucleotide binding and adenyl nucleotide binding (Figure 5.6c), suggesting regulatory roles of the involved genes.



(a) GWAS catalog terms enriched among the nine significant TWAS genes



(b) GWAS catalog terms enriched among 71 significant and suggestive TWAS genes



(c) GO molecular function terms enriched among 71 significant and suggestive TWAS genes

Figure 5.6: Significant enrichment terms for TWAS results by FUMA

5.3.5 TWAS identified novel genes not revealed by previous GWAS

CCDC91 was a novel facial gene that has not been discovered before, including in our previous GWAS where SNPs nearby had small p-values but did not reach genome-wide significance (Figure C1 bottom panel). In this TWAS, *CCDC91* was significantly associated with nose shape morphology in fibroblasts and was suggestive in adipose tissue. Other genes in the vicinity did not show association.

Among genes with a TWAS p-value $< 10^{-6}$ in at least one module, we tested SNP-face associations for all eQTLs used in predicting gene expression in the TWAS. Table C3 shows that four genes (*CCDC91*, *LTB4R*, *ADH6*, and *LINC01006*) had no eQTL SNPs which yielded a p-value $< 10^{-6}$. A visualization of each gene's region can be viewed in Figure C1-C4. In fact, the eQTLs of the latter three all showed p-values $> 10^{-4}$, two orders of magnitude greater than that of the gene-level association. These genes represent cases where TWAS was able to confer greater power for gene discovery than GWAS due to the collective effect of multiple eQTL SNPs.

5.4 Discussion

In this chapter, we conducted TWAS in three independent cohorts with a goal of identifying potentially functional mediators of the SNP-trait associations. Our results emphasized the important role of gene expression regulation in the genetic control of facial shape variation. Nine genes reached the study-wide significance level in four relevant tissues; an additional set of 62 genes showed suggestive association with facial shape. Results from conditional analysis sorted out putative mediating genes from genes that were not able to account for the GWAS signal. At multi-gene TWAS loci, a single gene can be prioritized based on eQTL colocalization. Our results highlighted *EEEFSEC* being a promising candidate for future studies pursuing mechanistic understanding of facial genetics. In addition to genes already implicated in previous facial GWAS, this study also identified three novel candidates (*CCDC91*, *LTB4R*, *ADH6*), some of which have relevant craniofacial function. Taken

together, these findings facilitated gene prioritization at facial GWAS loci and demonstrated the benefits of TWAS in mapping facial genes.

Gene-set enrichment analysis showed several interesting connections between facial morphology and other complex traits/diseases. The *EEFSEC/RUVBL1/SEC61A1* locus was found to be associated with gestational duration in previous GWAS [146], which raises interesting questions such as whether gestational age at birth influences facial shape and how far into later life this effect persists. Previous studies to some extend provided support for an effect of gestational age on certain dimensions of facial features. For example, a study reported differential orbital and ocular measures of babies born at different gestational age [147]; Another group found gestational age associated with linear distances between several facial landmarks [148]. However, the long-term effect of gestational age on facial morphology into adulthood may be another interesting question, on which there has not been studied previously. Whether there is a genetic component underlying this connection is a natural following step awaiting further study.

It is not surprising that face-associated genes were enriched for other anthropometric traits such as BMI and waist-to-hip ratio, given that adiposity and fat distribution influence both faces and anthropometrics. Our results also suggested a potential correlation of facial morphology and lung conditions, which has not been reported before. When further adding suggestive genes to the input list, the ribonucleotide binding term in the GO molecular function analysis showed a significant enrichment, suggesting regulatory activity being an important mechanism through which genes affect facial morphology. Studies dissecting the correlation between normal-range facial variation and other complex traits from a genetic perspective are just starting, and findings from this study generated interesting hypotheses to be tested in future research.

Results from TWAS, conditional, and colocalization analysis all featured *EEFSEC* being the most likely functional gene mediating the SNP-trait association at 3q21.3. *EEFSEC* is also a multi-effect facial gene with association signals seen in different facial quadrants. Possible mechanisms for multi-effect genes include (1) variants in the same gene may have effects across multiple cell populations; (2) variants act early in facial development, presum-

ably before or not long after the initial appearance of facial prominences, which are then carried along to different facial regions. *EEFSEC* is a translation factor necessary for the synthesis of selenoprotein. Selenoprotein is the major active form of micronutrient selenium in human body and supports various cellular and organismal activities, including mammalian development [149]. So far there is little known role of *EEFSEC* in facial morphogenesis; its face-related function remains to be explored.

LTB4R represents an interesting novel gene that was not observed in GWAS. Two rare missense variants in this gene were found to influence the shape of nose in the PITT cohort (Results from Chapter 4). In the current study, the effect of *LTB4R* was detected in the mandible segment. It is another multi-effect gene affecting facial regions spanning different quadrants, and even more intriguingly, with SNPs located at the opposite ends of the allele frequency spectrum both contributing their effects. Future study into the mechanism of *LTB4R* effects would be particularly insightful. It is not unreasonable that quantitative changes in gene expression level affects traits in a somewhat different but still related way than qualitative alternation in protein function (caused by changes in amino acid sequence). It is also possible that different facial regions are sensitive to different cues during morphogenesis, with some respond in a dose-dependent manner while others do not display such relationship. *LTB4R* is a receptor for leukotriene B4, a potent chemoattractant involved in inflammation and immune response. So far there has been no study on *LTB4R* in the context of facial development; future research is needed to demonstrate the biological relevance of this gene.

In addition to *LTB4R*, this study also nominated *CCDC91* and *ADH6* as novel facial candidates. *CCDC91* was shown to have a critical role in promoting the transport of carrier vesicles between the Golgi and lysosomes [150], and has been linked to skeletal problems such as hyperostosis and ossification. *ADH6* belongs to alcohol dehydrogenase family which encode enzymes involved in alcohol and drug metabolism. Polymorphisms of *ADH6* were found to be associated with alcohol consumption, alcohol dependence, and schizophrenia [151–153]. What roles these genes play in craniofacial development remain an open question be in investigated in future studies.

Similar to what we observed in Chapter 4, TWAS in the Tanzania population was less fruitful than that in the PITT cohort, despite comparable sample sizes. This in fact is generally true for many studies involving both African and European populations. Explanations may include (1) historically, African genomes have undergone more recombination events and therefore have less LD compared to the Europeans; (2) the external eQTL resources and gene expression imputation models were originally built using data of majority Whites, and extrapolation to another population, especially one of a different ancestry, would inevitably introduce extra uncertainty; (3) compared to the Europeans, fewer genes were tested in the Tanzania cohort due to a smaller overlap between African SNPs (genotyped and imputed) and European eQTLs. The cohort-specific nature of the multivariate facial phenotypes complicates a meta-analysis of the two cohorts, which we did not end up pursuing in this project.

The unavailability of craniofacial-specific omics data is a notable challenge for any study of craniofacial phenotypes, including the current one. Most of the cartilage, bone and connective tissues that make up the developing head are derived from craniofacial neural crest cells formed early in embryonic development. The nature and timing of the processes involved make it extremely difficult to collect the most relevant tissues and cell types in their native contexts. Such data are not available for us to build up our own craniofacial-specific eQTL resource and gene expression prediction model. We compromised by analyzing four adult tissue types that are closely related to facial structure, and these tissues represented the best options for the purpose of this study. Suboptimal though as compared to craniofacial neural crest cells from an embryonic development point of view, they are still essential components of the head and have great impacts on the morphology of the facial surface. Our tissue choices were intended to approximate the gene expression regulation in (1) facial bones and the attached muscles, (2) fibroblasts and adipocytes, which are the principal active cells of the connective tissues forming the base of the skull and face, and (3) frontal bone, which forms the forehead and interacts closely with brain cortex. The relevance of the selected tissue and cell types was also supported by the largest-so-far GWAS of normal-range facial variation from our group, where a subset of the identified GWAS peaks were enriched for enhancer activity in in vitro-derived osteoblasts, chondrocytes, differentiating skeletal muscle myoblasts, fibroblasts, and keratinocytes [139]. Even though the best avail-

able eQTL data were originally generated from steady-state bulk postmortem tissues and the expression regulation of them can be very different from that happening earlier in embryonic development, they still bear useful information and reflect biological processes progressively modifying facial morphology later in life. By analyzing the selected tissues, our study did achieve a fair amount of discoveries including genes previously implicated in face-related traits, demonstrating the usefulness of these four tissues in studying facial variation.

Given the limited tissue availability, the unknown aspects about facial growth, and a high level of tissue- and developmental-stage-specificity of expression regulation, this study refrained from comparing across tissues to not over-interpreting the results. That said, the current data was not able to answer questions such as which tissue contributes the most and should be considered as the most relevant for studying craniofacial conditions.

An inherent limitation of TWAS is that gene expression prediction can only be as good as the current knowledge on gene regulation, which is far from complete. Theoretically, prediction accuracy is capped by the heritability mediated by the genetic component of gene expression levels. Indeed, that cap is even stronger because the prediction only makes use of the cis subset with detectable effects coming from common variants. A recent study on UK Biobank molecular data estimated this upper boundary to be $11\,\% \pm 2\,\%$ across 42 common traits and 48 GTEx tissues [154], suggesting at best a modest capability in gene expression imputation of the cis-eQTL data. Another illuminating finding from their study was that the majority of the phenotype heritability mediated by gene expression is explained by genes with weak QTL effect sizes. Solving the problem of unknown functions will need more studies focusing on weak eQTLs, tissue- and context-specific eQTLs, trans-eQTLs without cis mediators, and rare eQTLs. Growing knowledge on the genetic architecture of gene expression will eventually improve the effectiveness of future TWAS.

Interpreting TWAS results presents a specific challenge – it is not always possible to pinpoint a single putative functional gene among all genes at a certain locus. This issue represents another intrinsic limitation of TWAS design and is a result of complexities at several layers. First, attrition of predictable genes occurs during the gene expression model training and testing processes where many genes were lost due to unsatisfactory prediction

accuracy. The typical numbers of predictable genes range from a few thousands to around 15 thousand in different tissues, which accounts for $10\% \sim 50\%$ of all human genes. Because gene prioritization relies on comparison among genes, when some genes are not available for testing, it is dangerous to prioritize the best available one to be the most likely causal one. Second, attrition of predictable genes also occurs when the intersection of variants in the GWAS and eQTLs in the prediction models does not contain enough SNPs. Third, there can be cases where multiple adjacent genes are significant in TWAS. A true disease-causing biology for all genes is possible, but most of time this is due to shared eQTLs or distinct eQTLs in high LD. Such a scenario was indeed observed in this study for *EEFSEC* and *RUVBL1* in the PITT cohort. As shown in figure 5.7, the two genes are next to each other and both had significant association with the philtrum shape. A further examination revealed that they shared one of the eQTLs. Fourth, interpreting TWAS results can be challenging when genes shared eQTLs or have distinct eQTLs in high LD. In these cases, conditional analysis may not be able to sort out the causal gene because SNP-trait association signals will diminish after conditioning on the GRex of both genes. Figure 5.8a and 5.8b display the marginal SNP-trait association statistics before and after conditioning on the TWAS-predicted expression of *EFFSEC* and *RUVBL1*, respectively. Both accounted for the GWAS signal to a similar degree, posing an apparent challenge in designating a single gene for follow-up experimentation. However, this difficulty in assigning a causal gene is not to abrogate the value of a TWAS. Indeed, no single fine-mapping approach can do a perfect job and the convergence of evidence from various sources should suffice as strong support for certain genes being functional. We therefore enhanced TWAS associations with results from conditional and colocalization analyses to make the best rigorous interpretation with bearing the limitations in mind.

In summary, analyses in this chapter demonstrated how eQTL data can be efficiently utilized to aid with GWAS loci fine-mapping as well as to identify novel candidate genes for normal-range facial variation. Our findings have made an important step forward from the earlier GWAS effort and paved the way for future functional experiments. These results will ultimately help with the construction and refinement of our mechanistic understanding of processes governing human facial morphology.

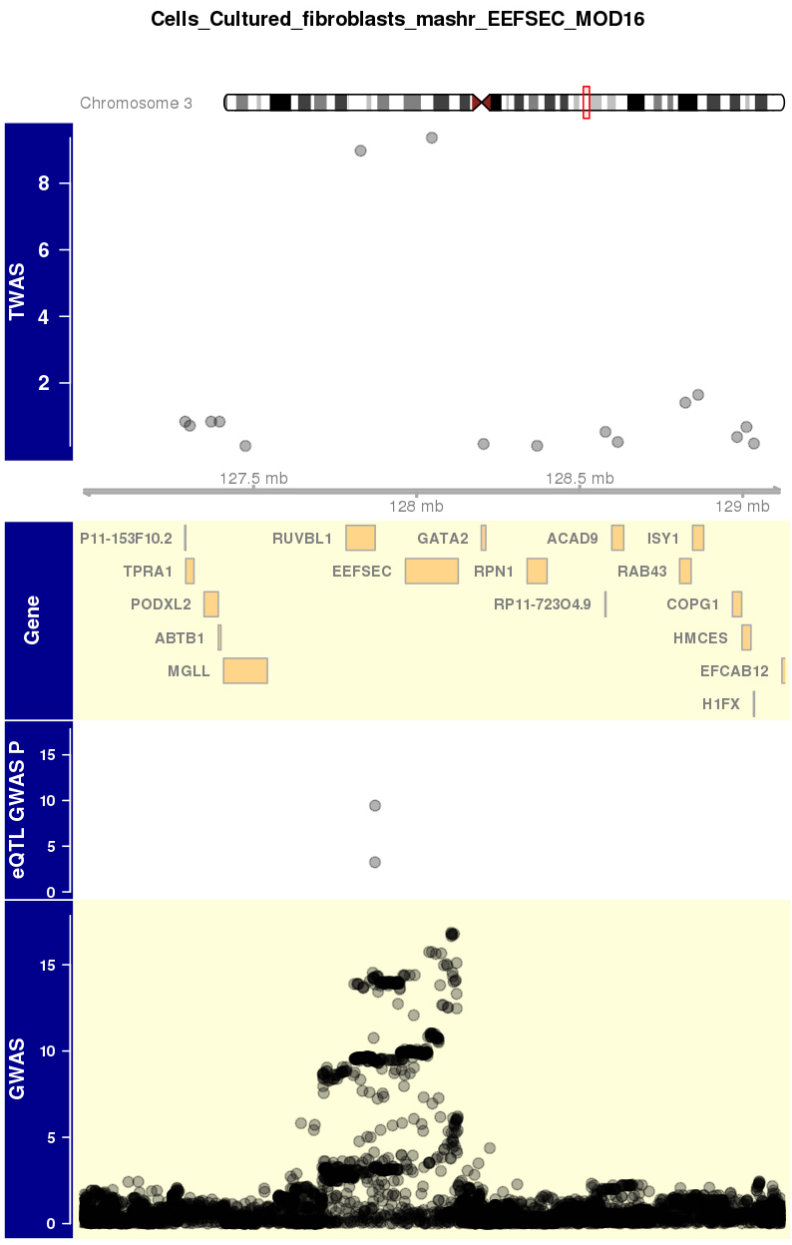
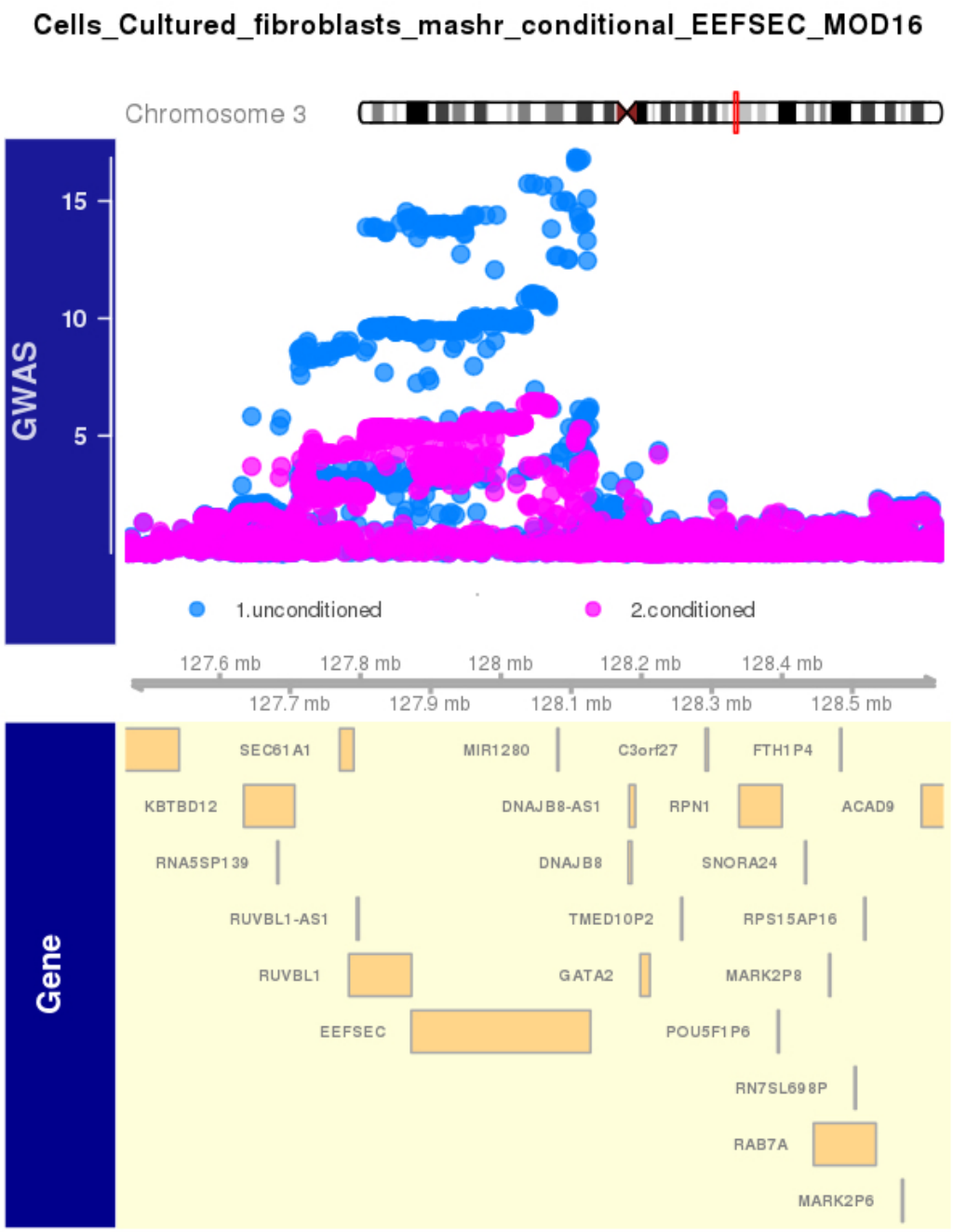
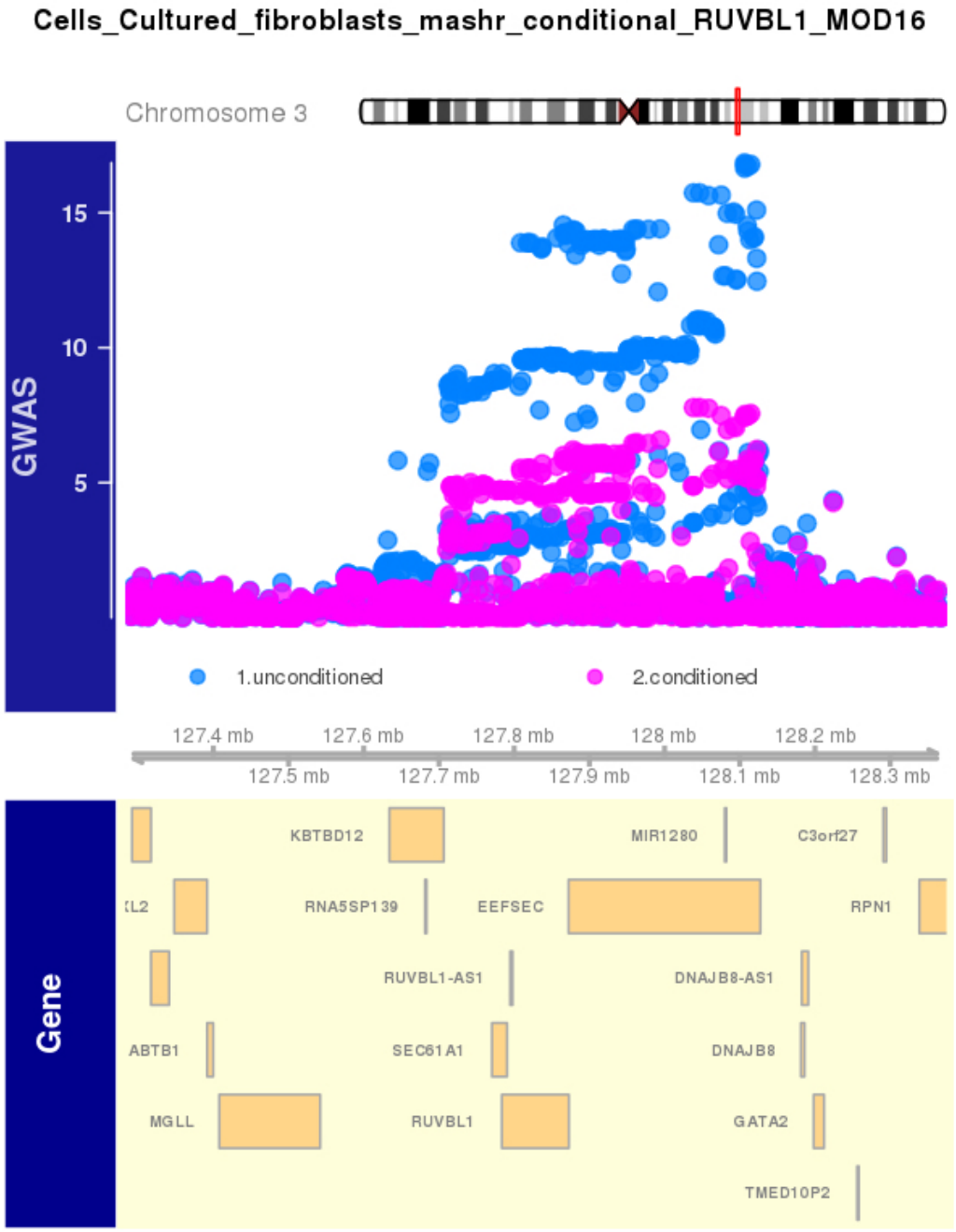


Figure 5.7: An example TWAS locus where two genes had significant association. At the 3q21.3 locus, both *RUVBL1* and *EEFSEC* were associated with philtrum shape in fibroblast cells. The top TWAS panel displays the gene-module association -$\log_{10}$(p-value) for genes annotated in the following Gene panel. The eQTL GWAS P panel highlights the eQTL SNPs used for predicting *EEFSEC* expression, with BP coordinates on the x-axis and GWAS -$\log_{10}$(p-value) on the y-axis. The bottom GWAS panel displays the GWAS -$\log_{10}$(p-value) for SNPs in this region.



(a) Conditioning on *EEFSEC*

Figure 5.8: Conditional analysis may fall short in sorting out the putative causal gene when nearby genes share eQTLs. Considering the example in Figure 5.7, conditioning on both (a) *EEFSEC* and (b) *RUVBL1* was able to partially reduce the association signal. The GWAS panel shows the SNP-module association p-values before (blue) and after (magenta) conditioning on gene expression.



(b) Conditioning on *RUVBL1*

Figure 5.8: Conditional analysis may fall short in sorting out the putative causal gene when nearby genes share eQTLs (cont.)

6.0 Variance Quantitative Trait Locus (vQTL) Analysis

6.1 Background

This chapter focuses on studying the genetic control of facial shape variability as well as detecting gene by gene (G × G) and gene by environment (G × E) interaction effects by taking advantage of the strategy of "variance prioritization" [101]. The main analysis performed was a genome-wide screen for loci associated with the variances of both multivariate facial modules and univariate facial distances, followed by a series of inspections into the underlying mechanism of the identified vQTLs.

6.1.1 Mechanisms for a SNP variance effect

Several underlying genetic mechanisms can manifest themselves as phenotypic variance heterogeneity across genotype groups [86]. Explanations for an observed variance effect include (1) the presence of a genuine interaction effect; (2) confounding by a mean-variance relationship for a non-normal phenotypic distribution; (3) induction by a nearby causal SNP with a mean effect; (4) induction by surrounding (multiple) rare causal variants. Sorting out mechanisms of potential biological interest from mechanisms simply reflecting statistical artifacts will provide valuable insights into the genetic architecture of the trait under study.

6.1.2 Variance prioritization applications

Answering the question of which mechanism(s) underlie an observed variance effect is a challenging task and would demand a lot of data. However, for the purpose of variance prioritization as described by Pare and colleague [101], such effort is not necessarily needed. The idea is to perform pre-screening and gather a set of SNPs enriched in interaction effects, and then search for the interacting factors using more direct approaches. The enriched set is comprised of SNPs conferring genetic variance effects, which can be relatively easily

identified. This strategy has been shown useful by Wang and colleagues who conducted a large-scale G × E scan for 13 quantitative traits in UK Biobank [8]. They identified 75 vQTLs associated with nine traits at an experiment-wise significance level, among which obesity-related phenotypes were disproportionally represented. By directly testing for SNP interactions with sex, age, smoking, physical activity and sedentary behavior, the study reported 16 out of the 75 showing significant G × E effects. When compared with a randomly chosen set of SNPs with QTL effects only, the set of 75 identified vQTLs was much more enriched for G × E effects. Findings from their paper also supported a polygenic model of phenotypic variance, indicating the existence of a large number of yet-to-be-discovered vQTLs with small effects. The BMI study by Yang et al [102] presents another exemplar application. The authors reported a vQTL in the *FTO* gene conferring a difference of 0.5 kilograms in standard deviation of weight between the opposite homozygous groups. They further showed that this variance effect was consistent with the reported *FTO* - environment factor interactions for BMI.

The variance prioritization strategy has also been applied to molecular traits such as gene expression, DNA methylation and protein concentration [86,101,104,105]. For instance, significant vQTLs were identified for two cardiovascular protein biomarkers in the Women's Genome Health Study, and the observed variance heterogeneity was suggested to be explained by an interaction between vQTL SNPs and BMI/smoking [101].

6.1.3 A review of variance homogeneity test

Several classical tests for homogeneity of variance exist, including Bartlett's test [155], Levene's test [156] and the Fligner-Killen test [157]. These tests are routinely conducted to check the assumption of variance homoscedasticity of many widely used statistical tests, such as Analysis of Variance (ANOVA) and Student's T-test. Bartlett's test requires normally-distributed data and has a significantly elevated false positive rate when this assumption is violated. Levene's test is more robust to departures from normality, and a more robust version of Levene's test can be achieved by substituting the group mean by the group median. Fligner-Killen test is also robust to the normality assumption. Newer parametric methods,

such as double generalized linear model [158] and likelihood-based test [159] were later developed to complement the traditional tests. More recently, Young and colleague [9] proposed their Heteroskedastic Linear Mixed Model, which is a statistical technique built upon double generalized linear models but goes beyond that by introducing a test for a dispersion effect independent of the mean-variance relationship of the phenotypic distribution. There is no universal single best method; these test statistics have different properties and may perform better than one another under different scenarios.

All of the aforementioned tests were initially proposed and mostly applied for testing the variance of a single variable. There has not been as much attention paid to variance/covariance in a multivariate setting, i.e. statistical tests for assessing the equality of two or more covariance matrices. Meanwhile, statistical theories of the variance lag far behind that of the mean. As a consequence, methods for variance are a lot less mature than methods for mean. As of the completion of this dissertation research, only Box's M test is readily applicable in multivariate settings to compare two or more covariance matrices [160]. However, Box's M test is well-known for being extremely sensitive to the violation of normality and the presence of outliers. Although bootstrapping was suggested to help with maintaining correct type I error rate, such a procedure can be prohibitively computationally intensive, preventing its application to the whole genome. There is a need for a robust, computationally tractable approach for testing variance heterogeneity of multivariate phenotypes.

6.2 Methods

All analyses in this chapter were conducted in the PITT and the Tanzania datasets separately. The smaller Denver cohort was not included due to insufficient power. In brief, a multivariate generalization of Levene's test was first implemented for a genome-wide search for vQTLs. Lead SNPs at detected vQTLs were then subjected to a series of result look-ups from Chapter 4 and 5 as well as from previous GWAS to sort out which could not be

explained by statistical artifacts. SNPs prioritized from the above procedure were further explored in interaction analysis where G × G and G × sex effects were formally tested in a regression framework. Finally, stratified analysis was performed to verify that the detected interaction effect of a vQTL truly accounts for its variance effect. As in Chapter 4, the PITT cohort was analyzed for both multivariate facial modules and univariate facial distances.

6.2.1 Levene's test with median and its multivariate generalization

This study used Levene's test with median and its multivariate generalization to identify SNPs associated with the variability of facial shape.The conventional Levene's test is equivalent to a test of group differences in mean deviation, with the deviation calculated as the absolute difference between the phenotypic value and its group-specific mean. This property implies that in parallel with the univariate Levene - ANOVA relationship, multivariate Levene's test can be operationalized by Multivariate Analysis of Variance (MANOVA) [161]. Specifically, the multivariate test was done by first taking the absolute deviation of each observation from its genotype group mean along each dimension (facial module PC), followed by a MANOVA with Pillai's multivariate test statistic on all dimensions in one facial module together. The statistical analysis was conducted using the Anova function (which also runs MANOVA) in the R package car.

6.2.2 Multiple testing correction

Detecting differences in group variances demands larger sample sizes than detecting differences in group means. Given that the sample of this study is smaller than the typical sample size of previous vQTL studies, we raised the MAF minimum cutoff for the sake of maintaining statistical power. On top of the quality control filters imposed in the TWAS in Chapter 5, SNPs with a MAF < 0.2 were further excluded from the vQTL analysis, giving a total of 2,818,796 qualified SNPs (genotyped and imputed) in the PITT cohort and 2,929,496 in the Tanzania cohort. Given this customized MAF filter, the conventional genome-wide significance threshold of 5×10^{-8} may no longer be appropriate for multiple testing correction.

We therefore used Genetic Type 1 Error Calculator [162] to recalibrate the effective number of tests (M_{eff}) and the corresponding p-value threshold in each cohort separately. A p-value $< 0.05/M_{eff}$ was considered as genome-wide significant, and a p-value $< 0.05/M_{eff}$/number of independent facial modules (or facial distances) was considered study-wide significant.

6.2.3 Power calculation

This project is the first of its kind in studying variance effect for facial traits and we do not have a justified speculation of the genetic effect sizes *in prior*. We therefore carried out a post hoc power calculation for MANOVA in the G*Power software [163]. A wide range of parameter choices were explored, and for illustrative purposes we chose the following based on the real phenotype and genetic data: three comparison groups (corresponds to the three genotype groups for any SNP tested); 11 dependent variables (for one of modules where significant vQTLs were detected); $\alpha = 1\times10^{-7}$; effect size being the maximum of the empirical effect size. The actual α level used in this study (see Results) was not in the acceptable range of G*Power; the closest α allowed was 1×10^{-7}.

6.2.4 Exploration of the mechanisms of variance heterogeneity for lead SNPs at vQTL

Several underlying genetic mechanisms can manifest themselves as phenotypic variance heterogeneity across groups. For (1) variance heterogeneity induced by a mean-variance relationship, we checked whether such a relationship exists by plotting absolute deviation versus group mean. For (2) variance heterogeneity induced by a nearby causal SNP with mean effect, we obtained the GWAS statistics of SNPs residing within 500 kb on either side of the lead vQTL SNP in the corresponding facial modules where the vQTLs were identified. For (3) variance heterogeneity induced by surrounding (multiple) rare causal variants, we obtained the gene-based association test p-values of genes overlapping with the same region as above, in the corresponding facial modules where the vQTLs were identified (from Chapter 4 results). For (4) variance heterogeneity as a result of interaction effects, we conducted tests

for vQTL × all other genome-wide SNPs and vQTL × sex, with more details described in the next sections.

6.2.5 G × E and G × G for lead SNPs at vQTL

We carried out formal interaction tests for the lead SNP at genome-wide significant vQTLs using a two-way MANOVA. Two-way MANOVA extends the traditional one-way MANOVA by including a second independent variable and an interaction term between the two factors. In the MANOVA model for a particular facial module, all of the PCs within this module would be the dependent variables, and the independent variables include vQTL, a second factor, and an interaction term vQTL × the second factor. The second factor can be any of the SNPs that were included in the vQTL analysis (MAF > 0.2) for G × G, and for G × E we restricted this second factor to be sex of the participants. For the sake of preserving power and not missing potential signals, we used 5×10^{-8} to declare a significant interaction effect without further adjusting for the number of vQTLs tested.

6.2.6 vQTL analysis for univariate facial traits

The 24 univariate facial traits were also analyzed in the PITT cohort. In Chapter 4 we showed that there were 17 independent traits among all 24, and the multivariate analysis in the current chapter have estimated the effective number of SNPs to be 448,915. The threshold used for declaring significance was therefore 6.55×10^{-9} (0.05/448,915/17). We used conventional Levene's test with median implemented in the OSCA software package (http://cnsgenomics.com/software/osca) [164] for an efficient genome-wide search. Top vQTLs were examined for the aforementioned four possible mechanisms in a similar manner. The G × G test was done by PLINK command –epistasis and the G × E test was done by fitting a linear regression model with the interaction term using PLINK command –linear. Colocalization plots for vQTL summary statistics and G × G summary statistics were generated by R package LocusCompareR [145].

As described in the Results section below, we detected a suggestive vQTL signal at the

PRICKLE1 locus and a significant G × G effect between *PRICKLE1* and *FOCAD*. With the assistance from Seongwon Cha's group at the Korea Institute of Oriental Medicine, we sought to validate these signals in an independent set of 5643 Korean individuals. A description of the cohort and genotype and phenotype data can be found in Cha et al 2018 [40]. The exact facial feature (cranial base width) was not available; zygion-to-zygion distance (zyR-zyL) was used as an substitute. Both are measures of facial base width and are highly similar to each other. The landmarks for calculating zyR-zyL can be seen in Fig 1C of Cha et al 2018 [40]. The same test in OSCA, Levene's test of homogeneity with median, was applied to SNPs in and near *PRICKLE1* using a 500 kb flanking window on either side, while adjusting for age, sex, height and weight. All SNPs located ± 500 kb of *PRICKLE1* were tested for interaction with all SNPs located ± 500 kb of *FOCAD* (the gene that *PRICKLE1* was interacting with in the PITT cohort) using the same method as in the discovery analysis, i.e. –epistasis function in PLINK. Initially, all SNPs with a MAF > 0.01 were analyzed. We then implemented a post hoc MAF filter, the lower boundary of which was set based on the minimum cutoff value which was able to give a well-behaved Q-Q plot. This cutoff was 0.2 for the vQTL test (consistent with the discovery analysis) and 0.1 for the G × G test. The independent number of SNPs was computed according to the eigenvalue-based procedure by Li and Ji's [119], using the 1000 Genome EAS data. For the vQTL test at the *PRICKLE1* locus, a p-value below $0.05/53 = 9.4\times10^{-4}$ was considered a significant replication. For the G × G analysis, we used the conventional genome-wide threshold of 5×10^{-8}. Whenever a specific SNP identified in the PITT cohort was not available in the Korean cohort, proxy SNPs were searched based on LD in the CEU population and if available, used as substitutes.

6.3 Results

6.3.1 Genome-wide vQTL scans for the multivariate facial modules

The M_{eff} was calculated to be 448,915 for the PITT data and 852,329 for the Tanzania data. We therefore used a genome-wide p-value threshold of 1.11×10^{-7} (0.05/448,915) and

a study-wide threshold of 2.86×10^{-9} (0.05/448,915/39, with 39 independent facial modules for the PITT sample) for the analysis in the PITT cohort, and a genome-wide threshold of 5.87×10^{-8} and a study-wide threshold of 1.47×10^{-9} for the analysis in the Tanzania cohort.

We did not identify any SNP at the study-wide significance level. However, eight loci yielded suggestive signals at the genome-wide level. Four of them were detected in the PITT cohort, and the affected facial regions included nose, lip, cheek and mandible (Table 6.1). Another four were identified in the Tanzania cohort and showed association with the shape variability of nose and zygoma area (Table 6.2). Multivariate Levene's test for the involved facial modules were well-behaved and there was no sign of inflation or deflation (Figure D1).

Zooming into the eight suggestive loci, SNPs' p-values and genes in vicinity can be viewed in Figure 6.1 - 6.8 panel (a). Most identified vQTLs did not overlap with coding regions, although there were two loci falling inside known genes. The chromosome 17q21.32 locus overlapped with a long non-coding RNA gene, and the lead SNP of 11q14.2 was an intronic SNP of gene *PRSS23*. The absence of an association "tower" at some loci could be a result of correlated SNPs being excluded if their MAF was below 20%. None of these loci harbor or lie close to a GWAS peak (Figure 6.1–6.8 panel (b)), nor were they reported in any previous facial morphology studies. A couple vQTLs harbor genes with some evidence of craniofacial relevance from past studies of cell lines, animal models, and human syndromes (see Discussion).

6.3.2 A post hoc power calculation for multivariate Levene's test

Figure 6.9 shows the power of detecting a global difference in a MANOVA as a function of total sample size, for an example module (module 37, where genome-wide significant vQTLs were identified). At an α level of 1×10^{-7}, the current sample size (N=2329) was able to detect a variance effect equal to or larger than the observed maximum effect size with a power greater than 90%. The power calculation assumed a balanced group size, i.e. same-sized groups of three genotypes, which was not true for most of the SNPs. With other parameters held constant, one should expect a lower power given this imbalance.

Table 6.1: Genome-wide significant vQTLs in the PITT cohort

Locus	SNP	BP	MAF	Module	P-value
8q23.3	rs13278165	117467575	0.420	35	8.88E-08
	rs35954699	117467839	0.420	35	8.88E-08
	rs34447117	117467891	0.419	35	9.74E-08
10p11.22	rs806812	32357760	0.355	37	4.26E-08
	rs73477	32400396	0.357	37	7.06E-08
	rs211428	32401087	0.357	37	7.06E-08
	rs2370703	32402827	0.357	37	7.06E-08
	rs405040	32406709	0.359	37	2.79E-08
10q25.3	rs7083764	115557964	0.410	2	5.25E-08
	rs4918871	115559190	0.410	2	3.95E-08
	rs4918872	115559311	0.404	2	1.02E-07
	rs4918873	115559458	0.404	2	1.02E-07
	rs9783229	115560438	0.410	2	3.21E-08
	rs9783196	115560876	0.410	2	3.49E-08
	rs10885508	115560970	0.410	2	3.49E-08
	rs7098623	115561346	0.410	2	2.86E-08
	rs10787503	115562202	0.411	2	1.53E-08
	rs4917669	115566844	0.409	2	2.18E-08
	rs7899612	115578872	0.409	2	2.27E-08
17q21.32	rs3744772	46713532	0.285	6	1.08E-07
	rs4793963	46715759	0.344	6	7.29E-08
	rs4793596	46717306	0.285	6	1.01E-07

Table 6.2: Genome-wide significant vQTLs in the Tanzania cohort

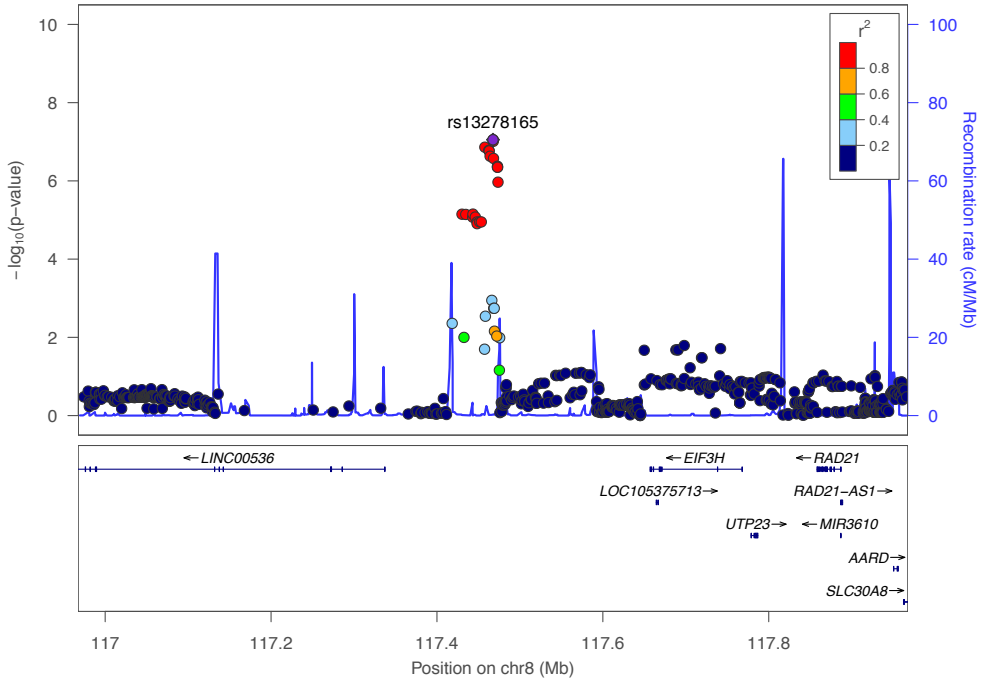
Locus	SNP	BP	MAF	Module	P-value
2q22.3	rs1919443	148398301	0.249	51	4.40E-08
	rs73011961	148404905	0.247	51	2.86E-08
4q13.1	rs935719	62035605	0.437	3	9.24E-09
11q14.2	rs57622507	86597905	0.253	47	3.16E-08
13q21.22	rs9542643	71864264	0.241	50	4.90E-08
	rs9529870	71864348	0.241	50	3.87E-08

6.3.3 Little overlap of identified vQTLs between the two cohorts

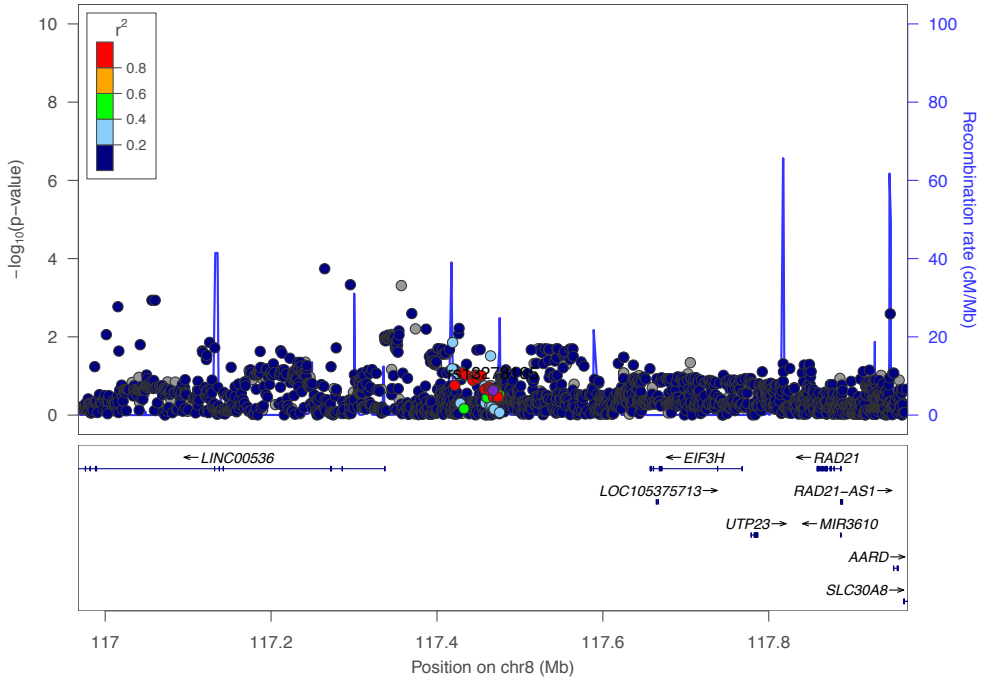
SNPs that passed the genome-wide threshold in either cohorts were looked up in the other for evidence of replication. Due to the cohort-specific facial segmentation procedure, all 63 facial modules were considered regardless of which module was originally involved in the vQTL. This led to 4 loci $\times$ 40 independent modules for looking up PITT vQTL SNPs in the Tanzania cohort and 4 loci $\times$ 39 independent modules for going the other direction. The full lookup results can be viewed in the affiliated supplementary file of this dissertation (Table S2); Table D1 presents results in the best module where a SNP obtained the minimum p-value in the lookup cohort. Most of the SNPs did not show evidence of replication. The only one that had a significant p-value after Bonferroni correction was rs806812 at the chromosome 10p11.22. It was identified in PITT upper lip module and showed the best association with a Tanzania cheek module (p-value = 2.6×10^{-4}).

6.3.4 Mechanisms of the observed variance heterogeneity

Four common mechanisms for SNP variance effects were scrutinized for the lead SNPs at the eight detected vQTLs. This section shows that the three mechanisms reflecting statistical artifacts were unlikely in our data, and in the next section we detail how G $\times$ G effects could possibly explain the observed variance heterogeneity at some vQTLs.

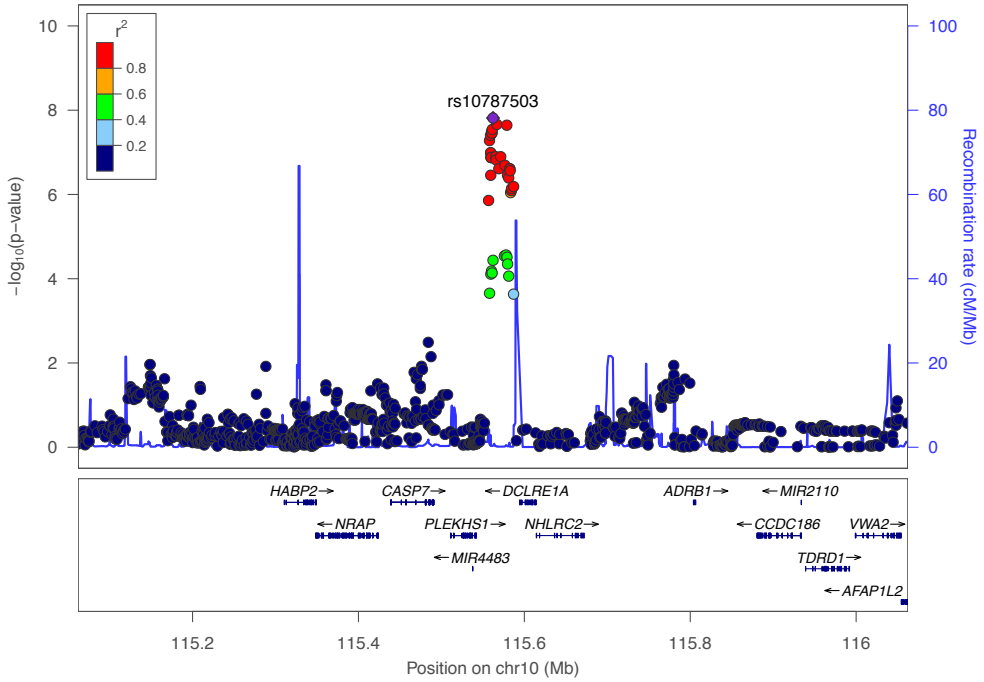


(a) vQTL analysis

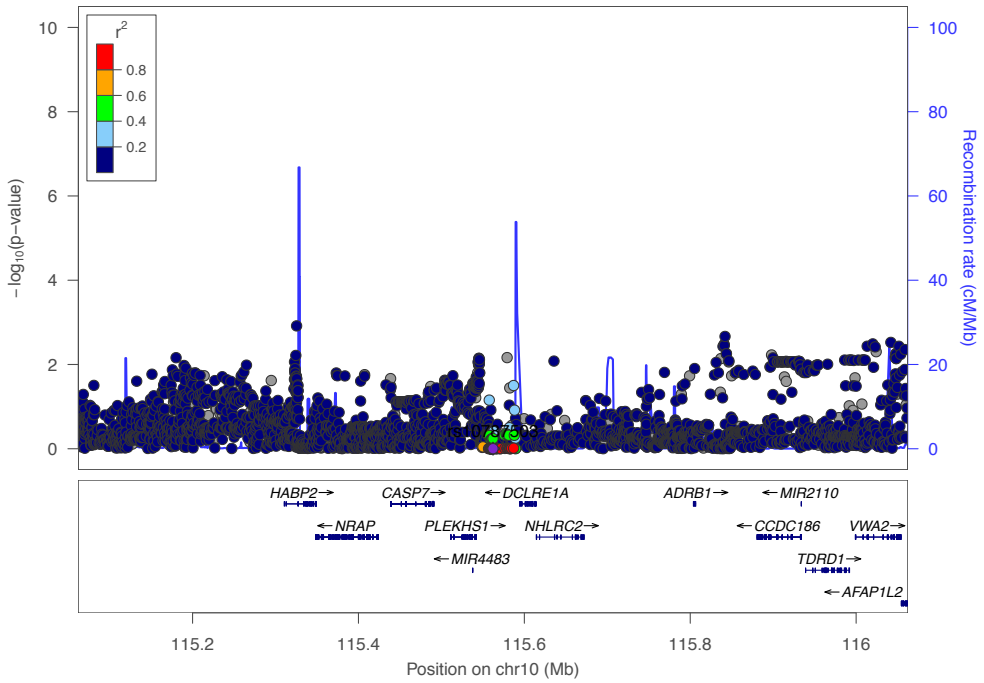


(b) GWAS

Figure 6.1: PITT vQTL: Regional plot of the 8q23.3 locus in module 35 (nasolabial)

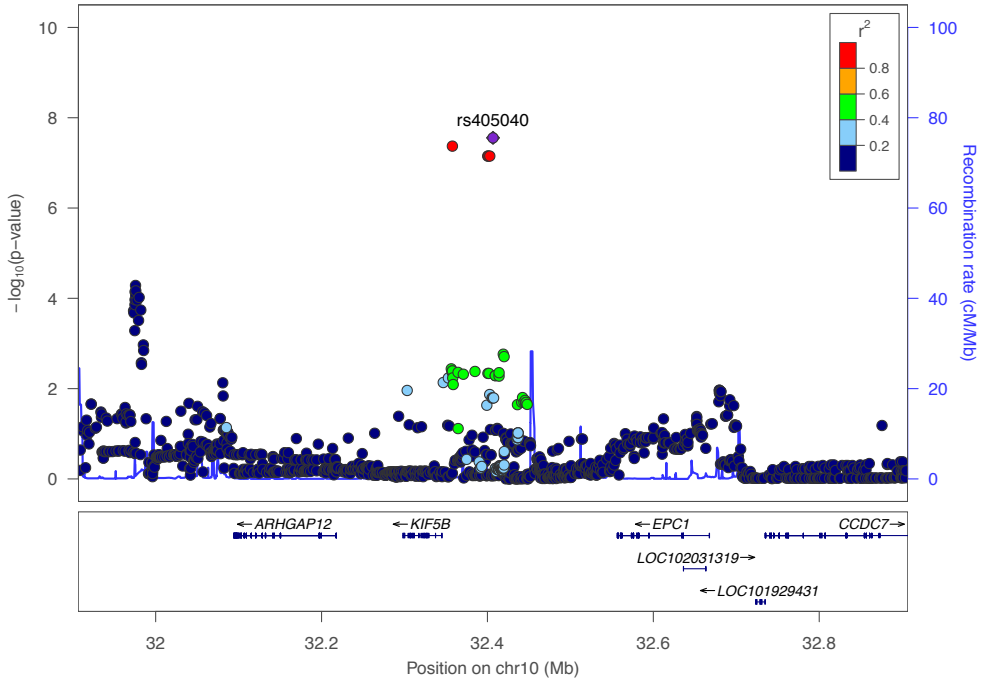


(a) vQTL analysis

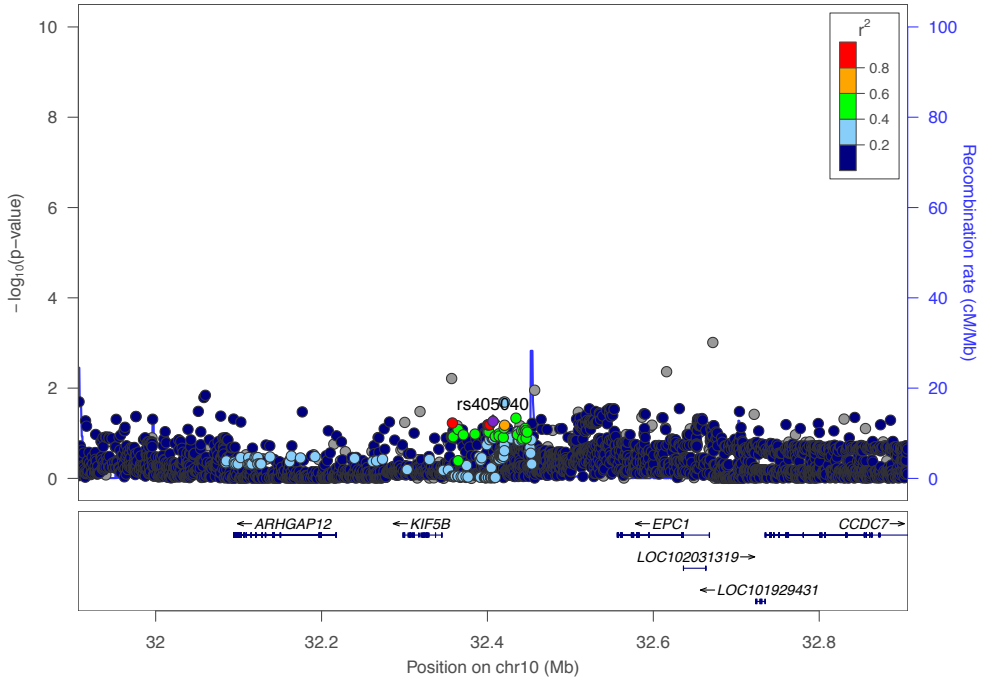


(b) GWAS

Figure 6.2: PITT vQTL: Regional plot of the 10q25.3 locus in module 2 (nose, mouth)

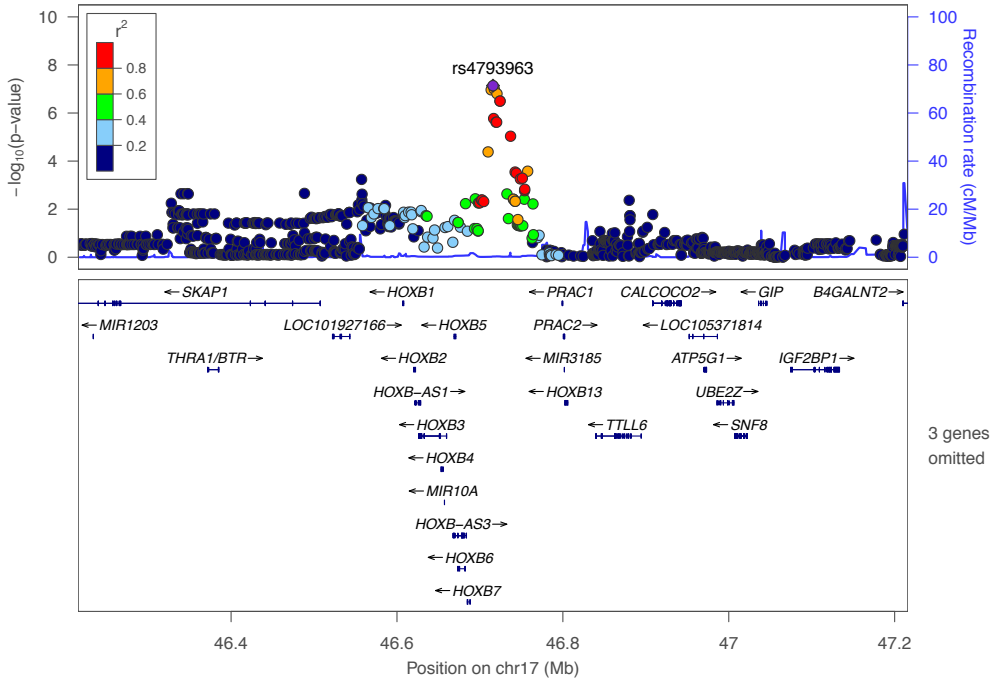


(a) vQTL analysis

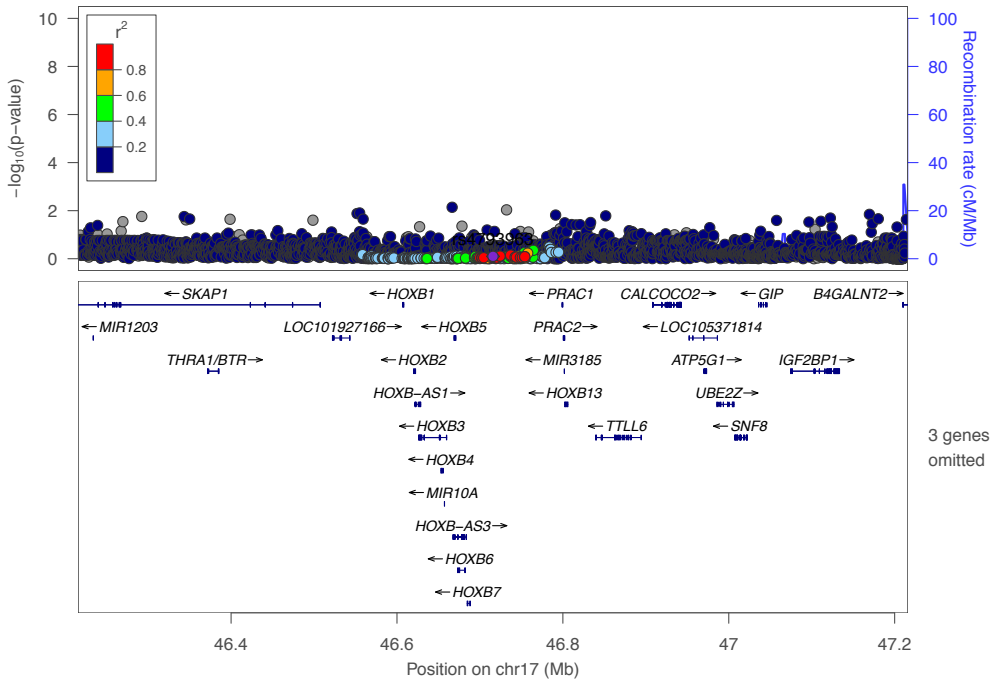


(b) GWAS

Figure 6.3: PITT vQTL: Regional plot of the 10p11.22 locus in module 37 (upper lip)

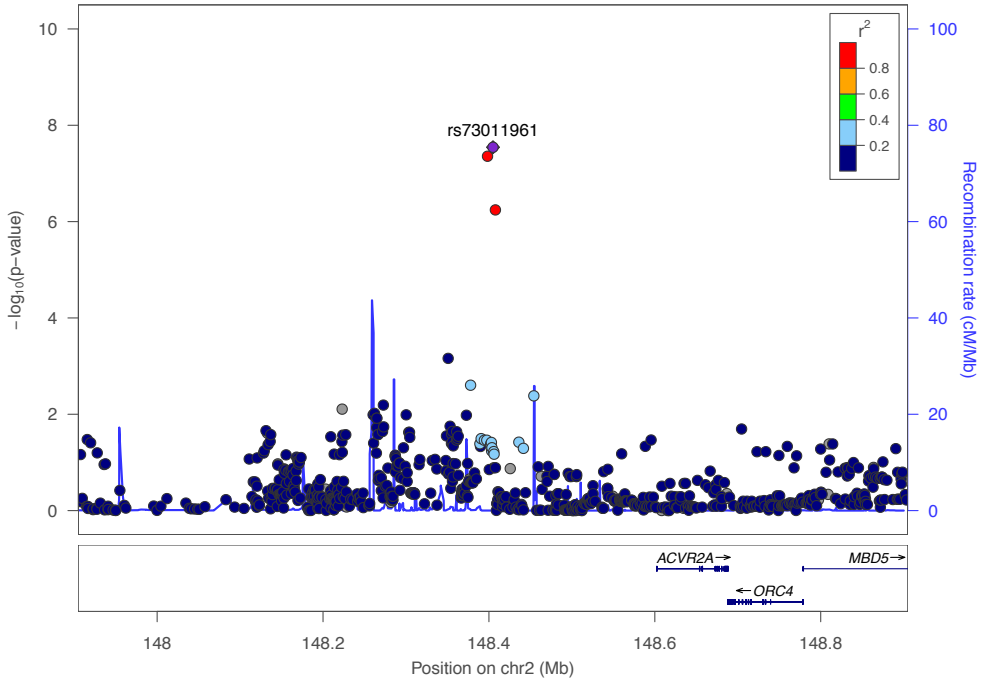


(a) vQTL analysis

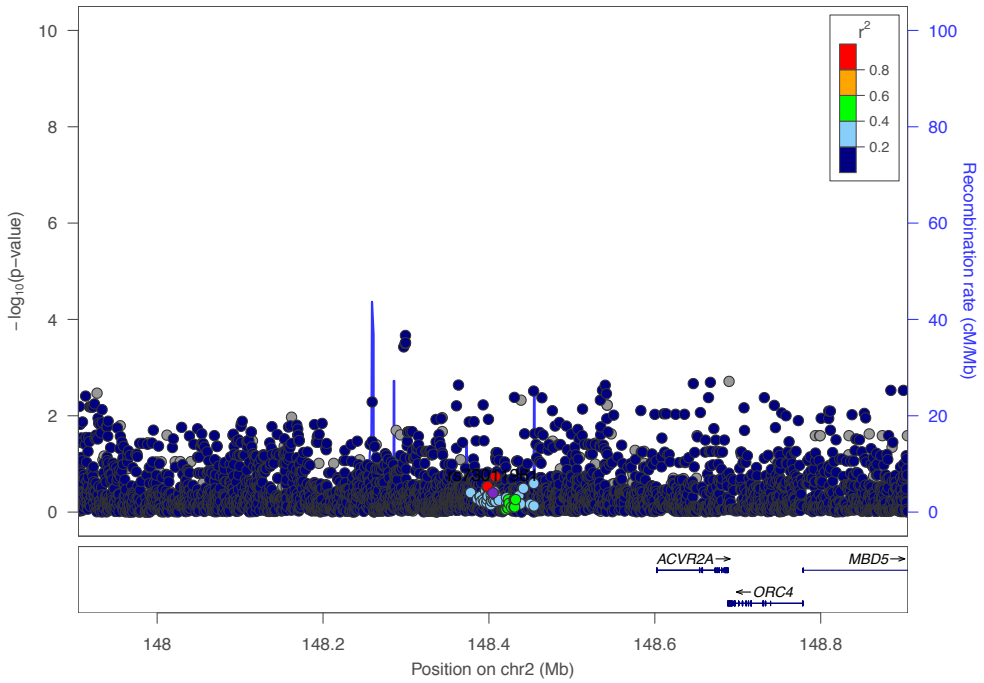


(b) GWAS

Figure 6.4: PITT vQTL: Regional plot of the 17q21.32 locus in module 6 (lower face)

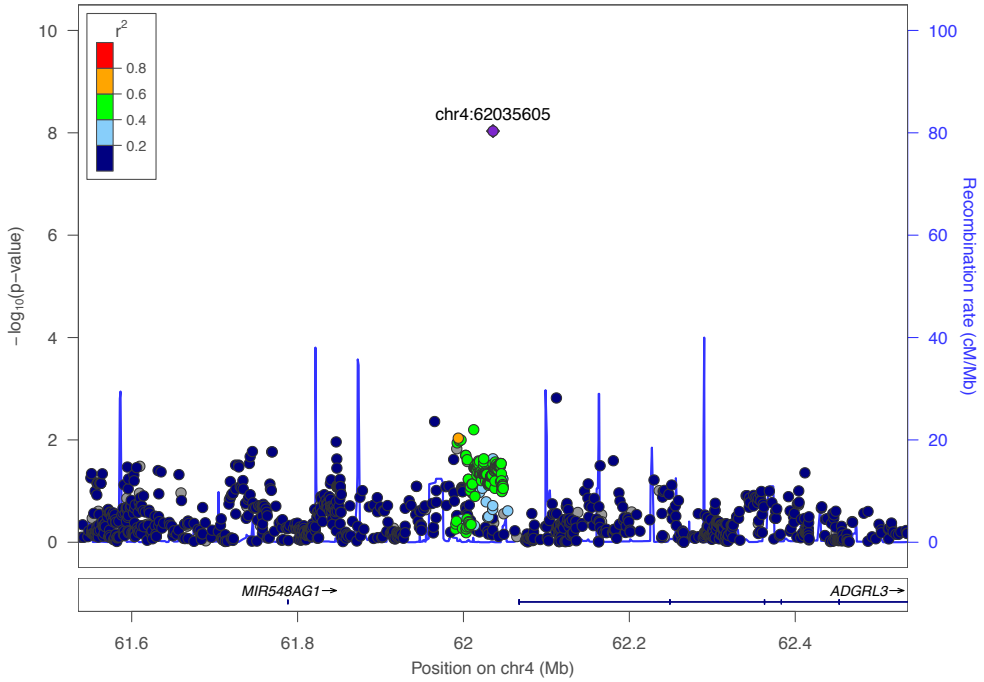


(a) vQTL analysis

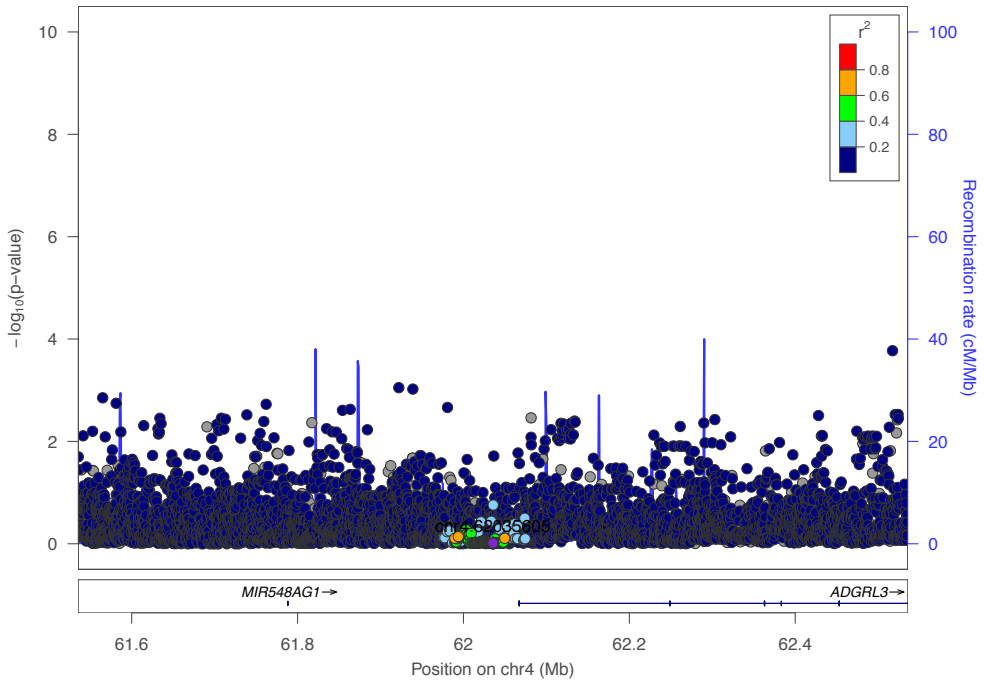


(b) GWAS

Figure 6.5: Tanzania vQTL: Regional plot of the 2q22.3 locus in module 51 (nose)

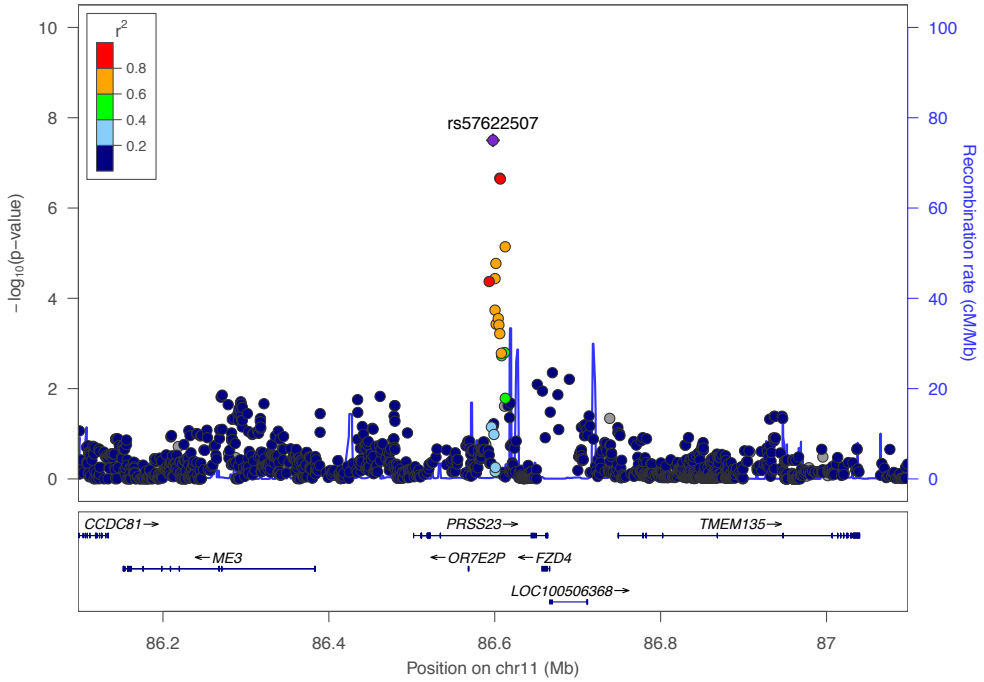


(a) vQTL analysis

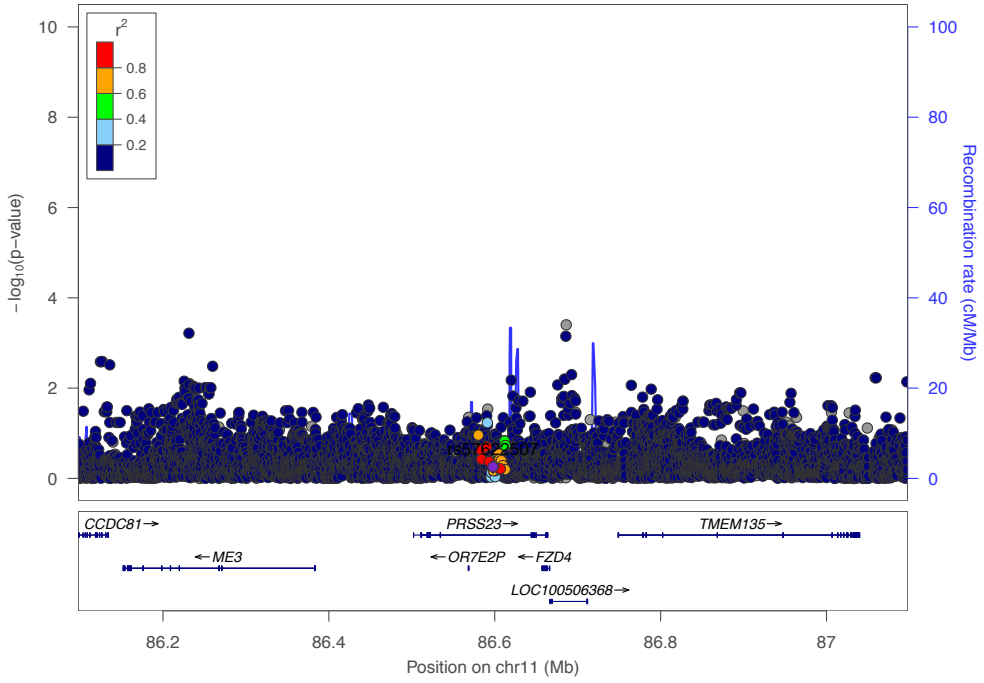


(b) GWAS

Figure 6.6: Tanzania vQTL: Regional plot of the 4q13.1 locus in module 3 (nose, eye)

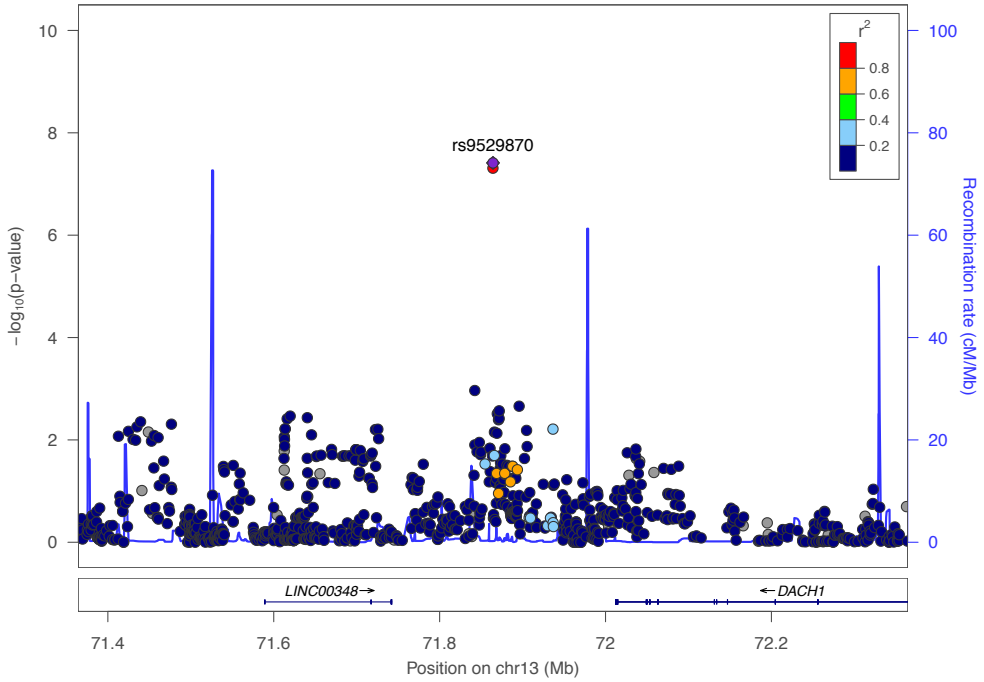


(a) vQTL analysis

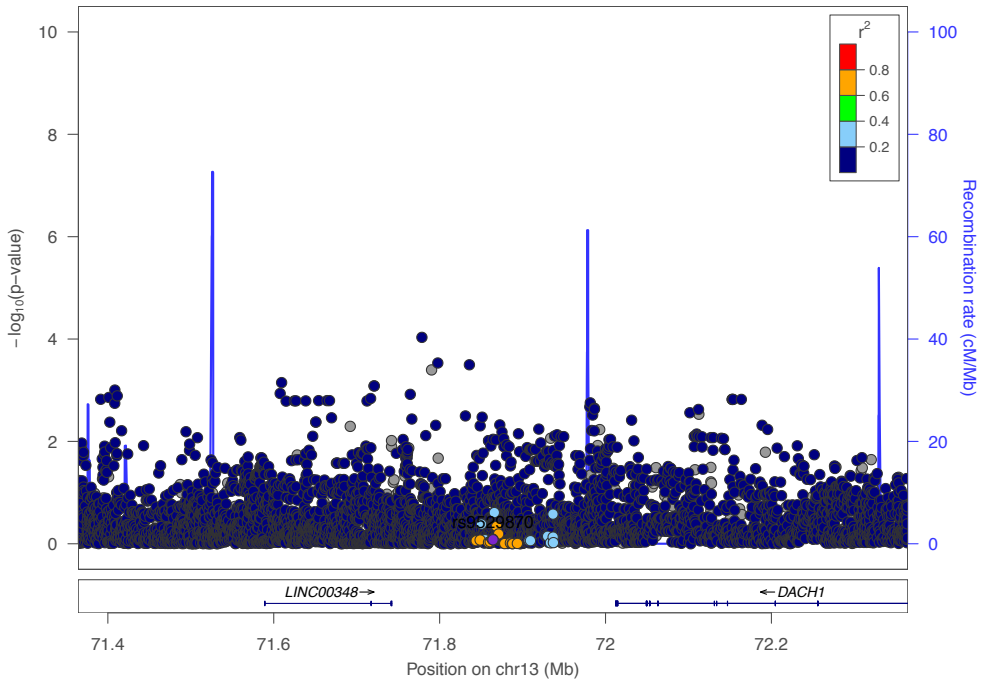


(b) GWAS

Figure 6.7: Tanzania vQTL: Regional plot of the 11q14.2 locus in module 47 (zygoma)



(a) vQTL analysis



(b) GWAS

Figure 6.8: Tanzania vQTL: Regional plot of the 13q21.22 locus in module 50 (nose)

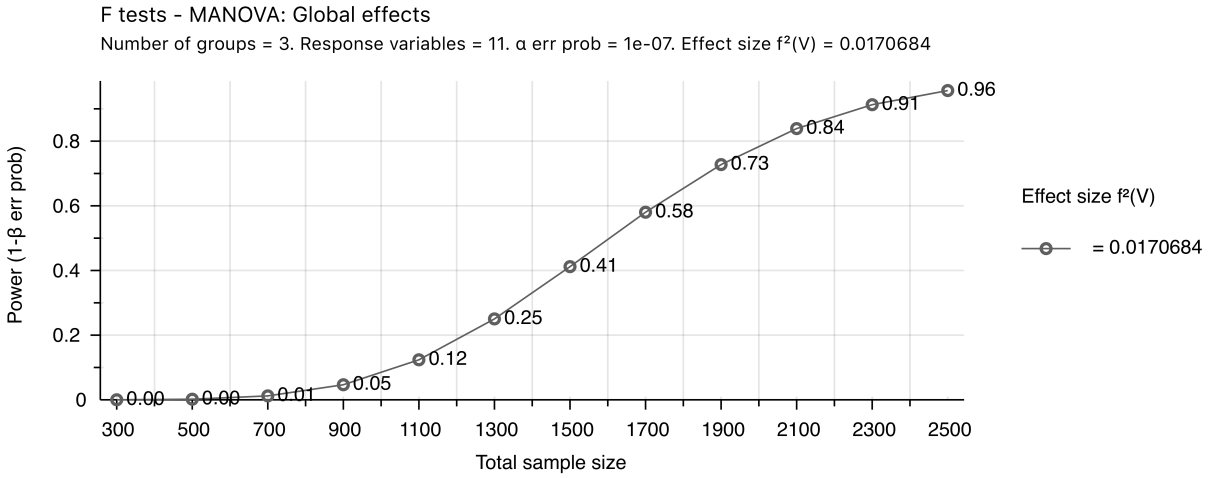


Figure 6.9: Post hoc power calculation for the vQTL analysis

(1) Variance heterogeneity induced by a mean-variance relationship. For this to be true, the SNP has to have a mean effect on the phenotype. As already mentioned, our previous GWAS did not identify any SNPs at the eight vQTLs to have mean effects. Therefore, the mean-variance relationship cannot be the underlying causal of the detected vQTLs.

(2) Variance heterogeneity induced by a nearby causal SNP with mean effect. Figure 6.1–6.8 panel (b) show that none of the eight identified loci overlap with a GWAS signal in a 1 MB window, which means that the presence of a nearby common SNP with mean effects was highly unlikely.

(3) Variance heterogeneity induced by surrounding (multiple) rare causal variants. In Chapter 4 we tested the effects of rare and low-frequency variants by aggregating coding SNPs into genes. Summary statistics of genes located in a 1 MB window centered at the lead SNP of each vQTL were retrieved from Chapter 4 results and displayed in Table 6.3. There was no evidence for any of these genes to have a significant effect on the corresponding facial module.

Table 6.3: Results of gene-based rare and low-frequency variant association test at vQTLs

Cohort	vQTL	Module	Gene	MultiSKAT p-value	Number of variants
PITT	10p11.22	2	*CASP7*	0.30	2
			CCDC186	0.04	2
			DCLRE1A	0.44	8
			HABP2	0.50	6
			NHLRC2	0.39	3
			NRAP	0.08	7
			PLEKHS1	0.69	2
			TDRD1	0.79	4
			VWA2	0.51	4
	17q21.32	6	*HOXB2*	0.78	2
			HOXB4	0.87	2
			HOXB9	0.79	2
			TTLL6	0.42	3
	10p11.22	37	*CCDC7*	0.45	7
			EPC1	0.43	2
			KIF5B	0.96	2
Tanzania	11q14.2	47	*CCDC81*	0.09	5
			FZD4	0.43	7
			TMEM135	1.00	7
	13q21.22	50	*DACH1*	0.03	2
	2q22.3	51	*ACVR2A*	0.06	2
			ORC4	0.66	6

6.3.5 G × Sex and G × G

Table 6.4 shows G × sex test p-values by fitting a MANOVA model. None of the lead SNPs showed evidence of G × sex effect.

Table 6.4: G × sex results for the lead SNP at the eight vQTLs

Cohort	Locus	Lead SNP	BP	Module	G × sex p-value
PITT	8q23.3	rs13278165	117467575	35 (nasolabial)	0.90
	10p11.22	rs405040	32406709	37 (upper lip)	0.45
	10q25.3	rs10787503	115562202	2 (nose, mouth)	0.35
	17q21.32	rs4793963	46715759	6 (lower face)	0.27
Tanzania	2q22.3	rs73011961	148404905	51 (nose)	0.78
	4q13.1	rs935719	62035605	3 (nose, eye)	0.94
	11q14.2	rs57622507	86597905	47 (zygoma)	0.44
	13q21.22	rs9529870	71864348	50 (nose)	0.97

The G × G tests did not yield signals at a significance level of $5\times10^{-8}/8$ (8 lead SNPs tested). One PITT vQTL and one Tanzania vQTL showed suggestive SNP × SNP interaction at a relaxed threshold of 5×10^{-8} (Table 6.5). SNP rs73011961 was detected to interact with the *GPC6* locus in the Tanzania data (Figure D2c). The PITT vQTL 10p11.22 showed a possible interaction with *CPO* and another locus near *EPHA4* (Figure D2). Q-Q plots and Manhattan plots of the genome-wide interaction analysis for the two involved facial modules can be viewed in Figure D3.

Table 6.5: Suggestive G × G for lead vQTL SNPs

Cohort	Module	vQTL Locus	vQTL lead SNP	Interacting locus	Interacting lead SNP	Interaction p-value
Tanzania	51 (nose)	2q22.3	rs73011961	13q31.3	rs7333716	2.81E-08
PITT	37 (upper lip)	10p11.22	rs405040	2q33.3	rs2084940	2.73E-08
				2q36.1	rs11885148	1.10E-08

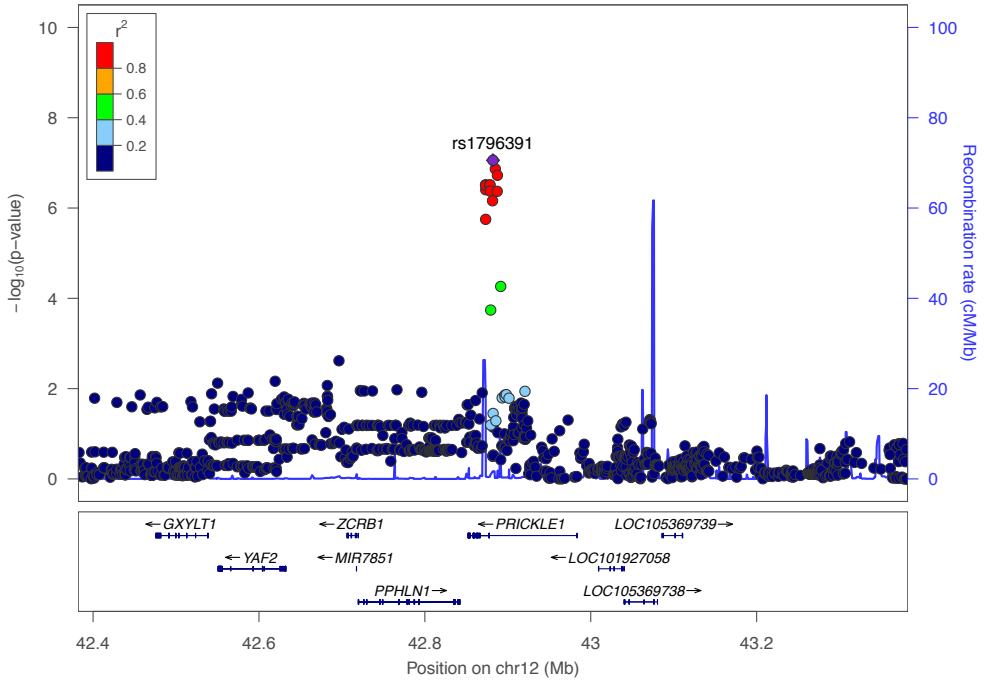
6.3.6 vQTL analysis for univariate facial traits

Although no study-wide significant vQTL was identified for any of the 24 facial distances, a few loci had p-values close to the threshold and were considered as suggestive hits. Table D2 shows the top associated SNP for each of the 24 traits. After examining into these suggestive signals, we identified the *PRICKLE1*-cranial base width association as particularly interesting and therefore focused on this association in the remaining of the Results section.

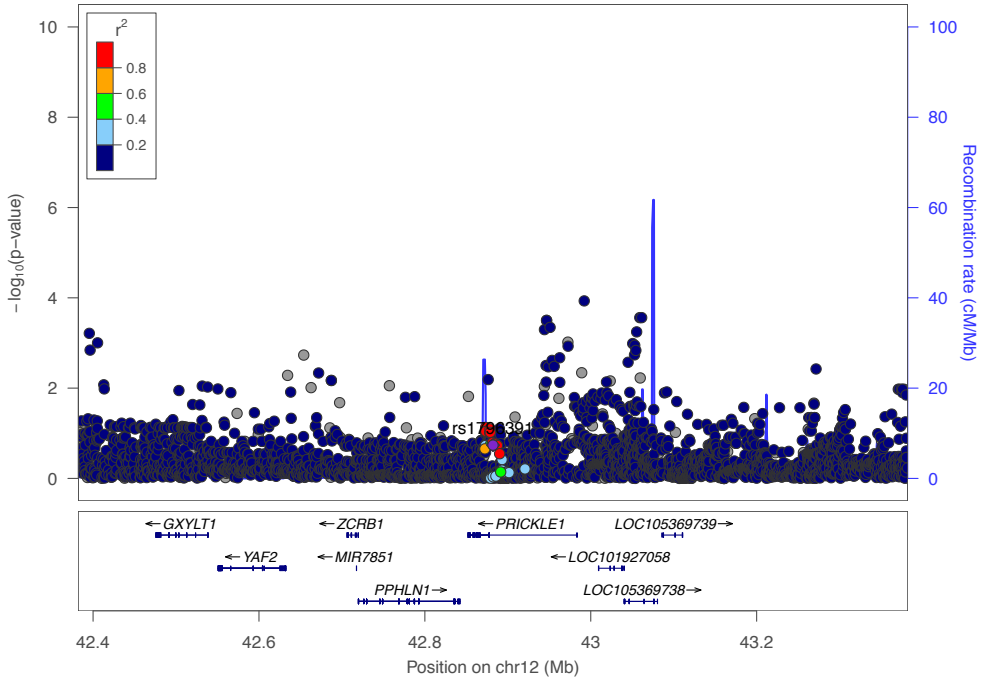
The top associated SNP for the variance of cranial base width was rs1796391 (MAF = 0.25), which falls in one of the introns of *PRICKLE1* (Figure 6.10a). It yielded a p-value=8.8×10^{-8} in the Levene's homogeneity test, and its minor allele was associated larger phenotypic variance. Levene's test statistics for cranial base width were well-behaved and there was no sign of inflation nor deflation (Figure D4).

Similar to how we followed up on the vQTLs of facial modules, we assessed three mechanisms as potential sources for the variance heterogeneity at rs1796391. SNP rs1796391 was not associated with the mean of cranial base width (GWAS p-value=0.18), nor did any of the SNPs located within 500 kb from rs1796391 show significant mean effect (Figure 6.10b). A linear regression model including the rs1796391 $\times$ sex term was fitted to test for a possible G $\times$ E effect, and neither the 1df test for the interaction effect only nor the 2df test for both the main and the interaction effects yield significant results (1df test p-value=0.69, 2df test p-value=0.45). We looked up gene-based association summary statistics from Aim1, and *PRICKLE1* had a CMC p-value=0.32 and a SKAT p-value=0.83. These numbers suggested that three of the possible mechanisms under consideration were unlikely to explain the variance effect of the *PRICKLE1* locus in cranial base width.

In the genome-wide search of SNPs interacting with *PRICKLE1* SNP rs1796391, we detected a significant G $\times$ G with the second locus being *FOCAD* on chromosome 9. The SNP pair consisting of rs1796391 (*PRICKLE1*) and rs10511683 (*FOCAD*) gave an interaction p-value=8.1×10^{-9}; no other SNPs at the *FOCAD* locus had an interaction p-value below 5×10^{-8} (Figure 6.11a). SNP rs10511683 was intronic and did not show a mean effect on cranial base width (Figure 6.11b). Figure 6.12 presents a visualization of the SNP $\times$ SNP interaction effect, where the non-parallel pattern of the red, green, and purple line indicated



(a) vQTL analysis



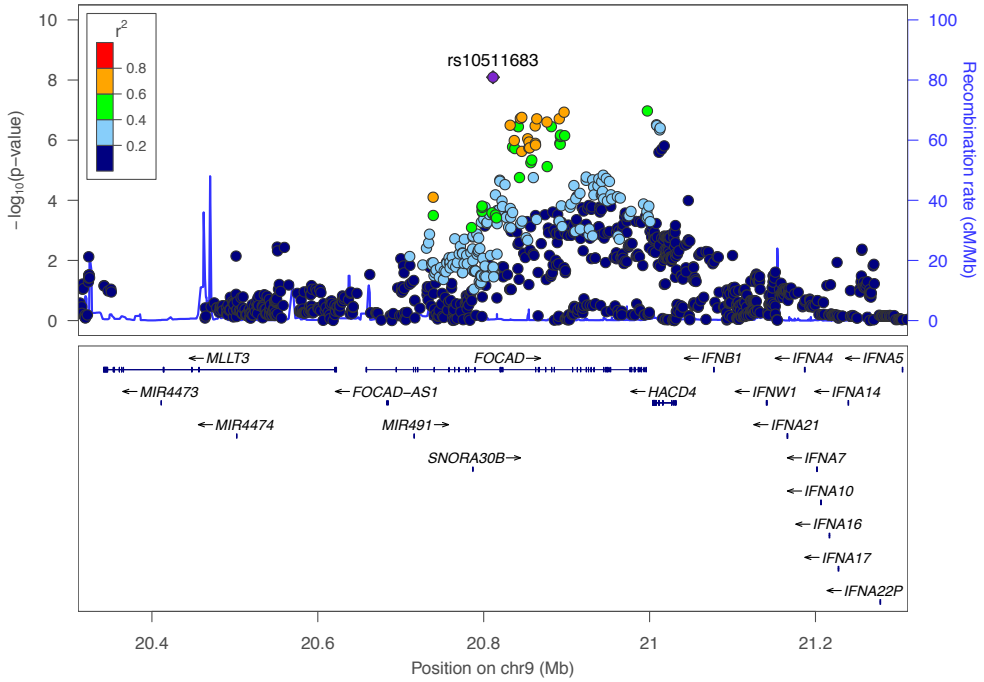
(b) GWAS

Figure 6.10: Regional plot of the *PRICKLE1* locus in cranial base width in the PITT cohort

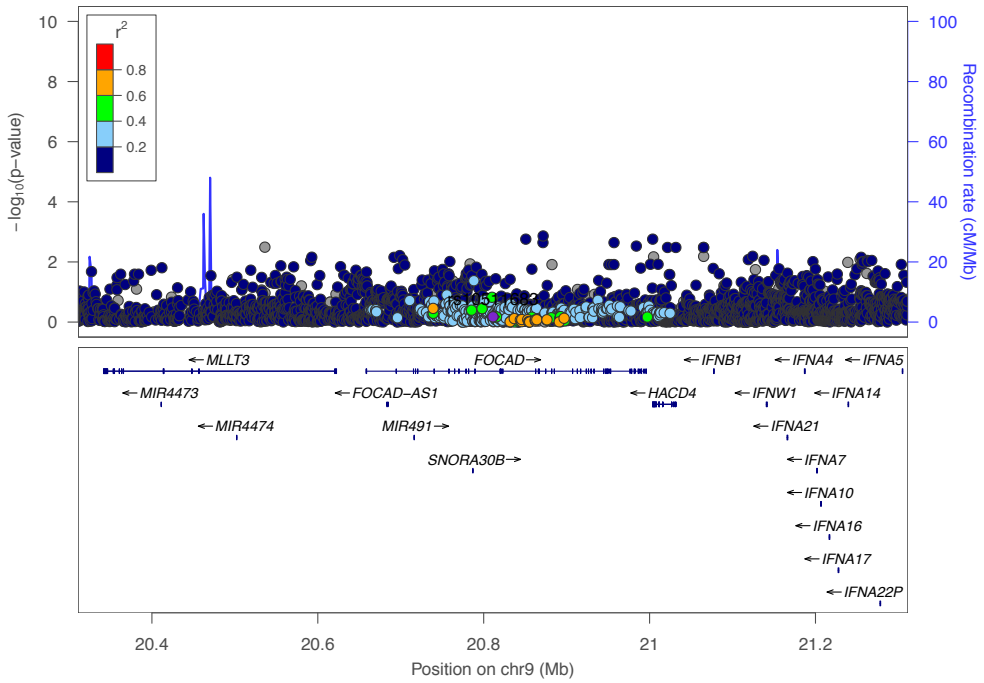
the differential conditional effects of rs10511683 depending on the genotype at rs1796391. The table to the right in Figure 6.11 displays association results of rs10511683 before and after the stratified analysis. The minor allele of rs10511683 in *FOCAD* was associated with wider cranial base in people carrying the AA genotype at rs1796391 (*PRICKLE1*), whereas it had an opposite effect direction in the GG group. The p-values in both AA and GG group were small, despite not meeting the genome-wide threshold. These opposite effects offset each other when the interacting locus was not taken into consideration, shown by the non-significant p-value in the first row of the table.

We performed two analyses to further explore the relationship between the detected vQTL signal and the G × G signal. First, focusing on the *PRICKLE1* locus, we compared the p-values from the Levene's tests with the p-values from the G × G tests in Figure 6.13. The upper right cluster in the left panel supported a colocalization of the signals from the two analyses while taking into account the local LD structure. Although this does not prove that the interaction was indeed the underlying cause for the variance heterogeneity, the colocalization at minimum demonstrates a possibility of such explanation. Second, we repeated the Levene's test for rs1796391 (*PRICKLE1*) after stratifying the entire sample by individual's genotype at rs10511683 (*FOCAD*). If the variance heterogeneity at rs1796391 (*PRICKLE1*) was induced by its interaction with rs10511683 (*FOCAD*), we would expect to see weakened signals in each stratum. In accordance with this expectation, the stratum-specific vQTL signals were substantially less significant compared with the unstratified p-value = 8.8×10^{-8} (Figure 6.14). We caution that the stratification inevitably led to a reduced power and a tendency for higher variance in smaller groups, which complicated the interpretation of this result. Nonetheless, results from these two analyses suggest that the G × G effect can potentially account for the observed variance heterogeneity at rs1796391 (*PRICKLE1*).

From figure 6.12, we noted one individual in the minor allele homozygous group appeared to be an outlier of cranial base width. We examined the influence of this outlier on the vQTL and G × G results by performing a sensitivity analysis with this individual removed. Levene's test gave a p-value = 9.03×10^{-7}, and the same SNP pair yielded a G × G p-value = 4.65×10^{-8}. Figure 6.15, which is the same to figure 6.12 except that the outlying observation

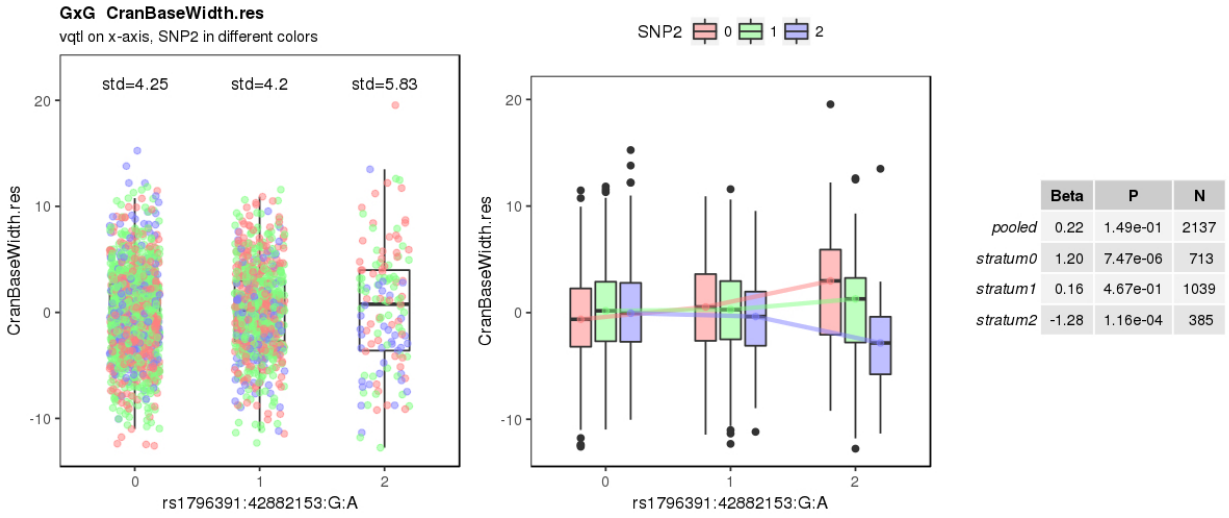


(a) G × G between rs1796391 in *PRICKLE1* and SNPs at the *FOCAD* locus



(b) GWAS

Figure 6.11: Regional plot of the *FOCAD* locus in cranial base width in the PITT cohort



	Beta	P	N
pooled	0.22	1.49e-01	2137
stratum0	1.20	7.47e-06	713
stratum1	0.16	4.67e-01	1039
stratum2	-1.28	1.16e-04	385

Figure 6.12: Interaction effect between rs1796391 (*PRICKLE1*) and rs10511683 (*FOCAD*) in cranial base width. Left: boxplots show the residual value of cranial base width (after regressing out age, age^2, sex, height, weight and facial size) in three genotype groups of the vQTL rs1796391 (*PRICKLE1*), colored by the genotype at the interacting SNP rs10511683 in *FOCAD* (SNP2 in the figure). Standard deviations were shown above each box. Middle: X-axis and Y-axis are the same as in the boxplots to the left, but further stratified and colored by rs10511683 (*FOCAD*). The non-parallel pattern is the hallmark of statistical interaction. Right: Association (with the mean of cranial base width) beta coefficients, p-values and sample sizes of rs10511683 (*FOCAD*) in the entire sample (top row) and in three strata defined by rs1796391 (*PRICKLE1*) (following rows).

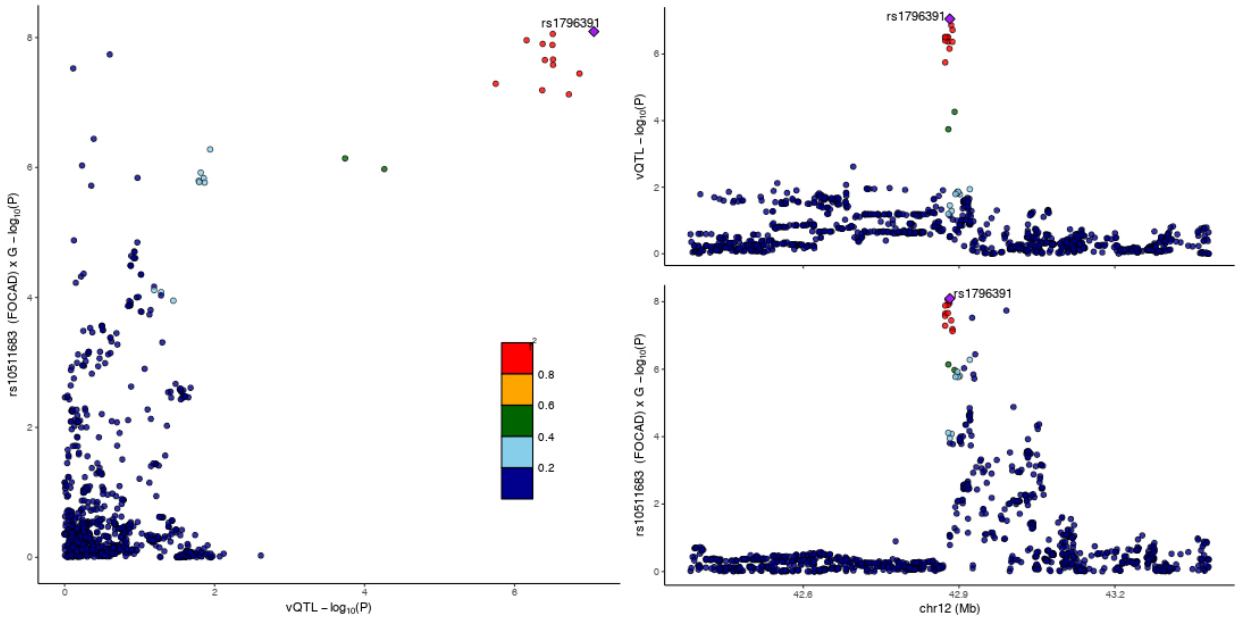


Figure 6.13: Locuscompare plot for *PRICKLE1*. Top right: regional scatter plot for the vQTL scan of cranial base width. Bottom right: regional scatter plot for the G × G test between rs10511683 (*FOCAD*) and SNPs at the *PRICKLE1* locus. Left: joint distribution of the p-values from the two analyses. Colored by LD with rs1796391.

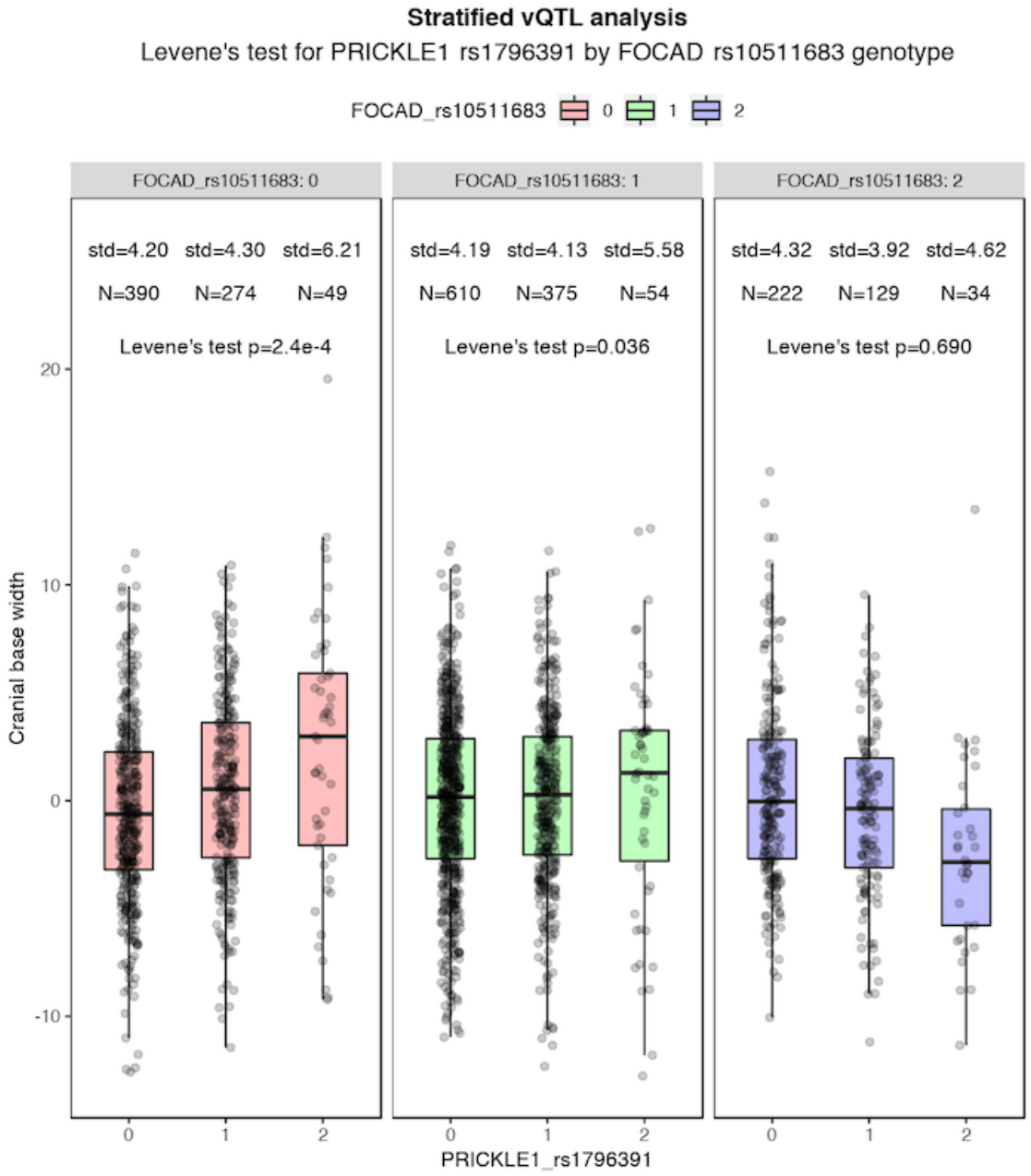


Figure 6.14: Stratified vQTL analysis of *PRICKLE1*. Plots show the residual value of cranial base width (after regressing out age, age^2, sex, height, weight and facial size) in three genotype groups of the vQTL rs1796391 (*PRICKLE1*). The left, middle, and right panels correspond to the AA, AG, and GG genotype at rs10511683 (*FOCAD*), respectively. Shown above each boxplot are the standard deviations of the phenotype, group sizes, and the p-values from the Levene's test in that genotype group of rs10511683. The vQTL signals in each stratum were less significant than that in the unstratified analysis.

was excluded, demonstrates the presence of variance heterogeneity and the interaction effect in the remaining samples. As one would expect, The signals were slightly weakened compared to the original analyses. These results suggest that the detected *PRICKLE1* vQTL and its interaction effect were not materially affected by the phenotypic outlier, which is expected given the robustness of Levene's test with median.

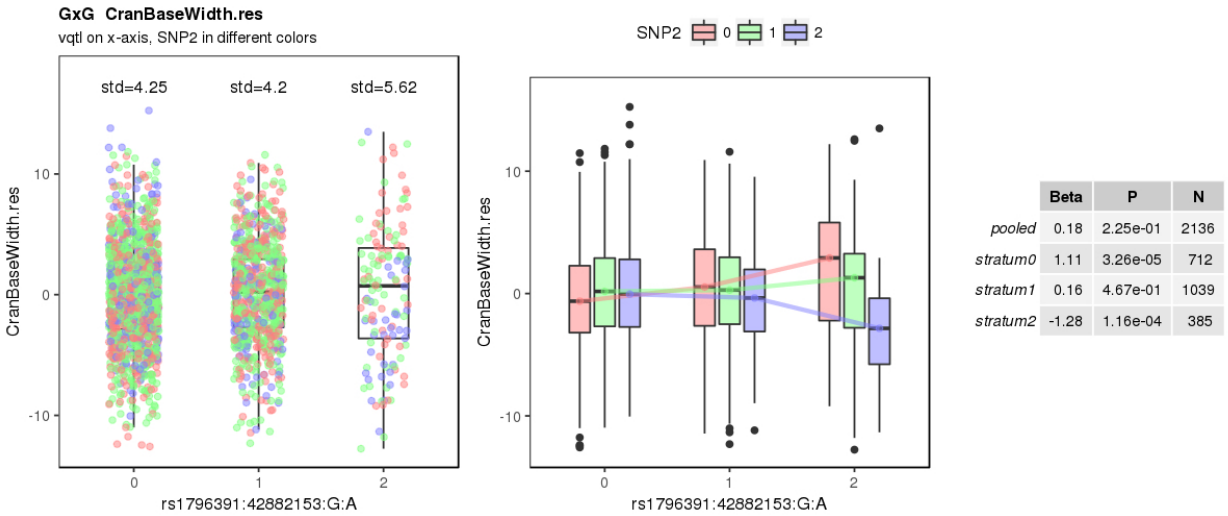


Figure 6.15: Sensitivity analysis of the *PRICKLE1* vQTL. This figure shows the same analyses as in figure 6.12, except that the phenotypic outlier in the homozygous group was removed. Left: boxplots show the residual value of cranial base width (after regressing out age, age^2, sex, height, weight and facial size) in three genotype groups of the vQTL rs1796391 (*PRICKLE1*), colored by the genotype at the interacting SNP rs10511683 in *FOCAD* (SNP2 in the figure). Standard deviations were shown above each box. Middle: X-axis and Y-axis are the same as in the boxplots to the left, but further stratified and colored by rs10511683 (*FOCAD*). The non-parallel pattern is the hallmark of statistical interaction. Right: Association (with the mean of cranial base width) beta coefficients, p-values and sample sizes of rs10511683 (*FOCAD*) in the entire sample (top row) and in three strata defined by rs1796391 (*PRICKLE1*) (following rows).

We sought to validate the vQTL and G × G signal of *PRICKLE1* in an independent set of Korean individuals. A minimum MAF cutoff of 0.2 (as used in the discovery analysis) was able to generate well-behaved test statistics on the Q-Q plot (Figure D5). The effective number of tests among the 663 qualified SNPs at the *PRICKLE1* was 53. No SNPs survived

the Bonferroni correction ($0.05/53=9.4\times10^{-4}$, Figure 6.16). The lead SNP at *PRICKLE1* from the PITT analysis were not available in the replication cohort. Four proxy SNPs which had a $r^2=1$ with the lead SNP rs1796391 were identified, yet none of them showed evidence of influencing phenotypic variance (p-value > 0.1 for all; Table 6.6 and 6.7).

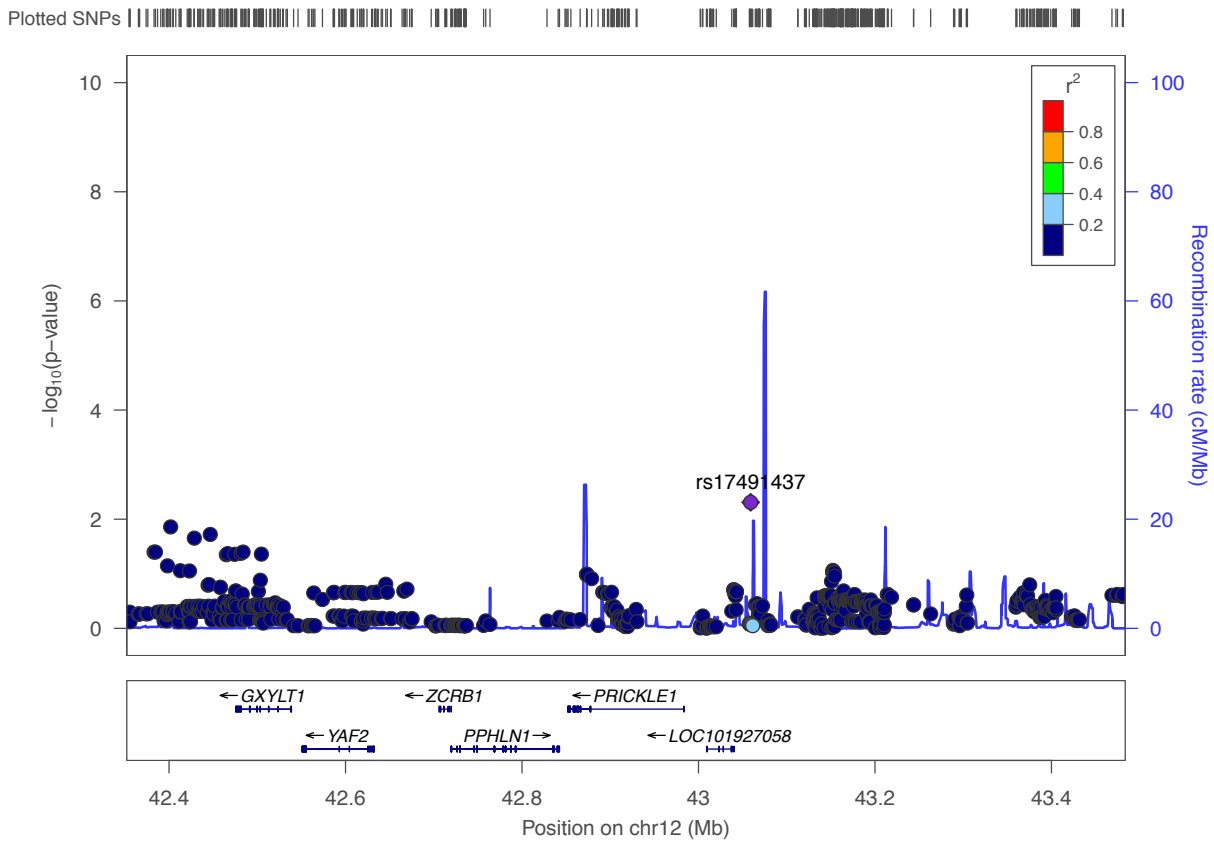


Figure 6.16: Results of the vQTL test for *PRICKLE1* locus in the replication cohort. There is little evidence of replication.

Since the lead interacting SNPs from the analysis of the PITT cohort were not available in the replication cohort, we looked for proxy SNPs based on LD. Four proxies were identified for *FOCAD* interacting SNP rs10511683 with r^2 ranging from 0.63 to 0.77 (Table 6.6). These four SNPs were tested for pairwise interaction effects with each of the aforementioned four proxies for *PRICKLE1* SNP rs1796391, the results of which were summarized in Table 6.8. There was no evidence of the presence of significant interaction effect when only these 16 SNP pairs were taken into account.

We further expanded the test region under consideration by analyzing all SNPs located in

Table 6.6: Proxy SNPs identified for rs1796391 (*PRICKLE1*) and rs10511683 (*FOCAD*) in the replication cohort

Gene	Proxy SNP	CHR:BP	MAF[a]	r2[b]
PRICKLE1	rs1796362	12:42873184	0.221	1
	rs1796361	12:42873402	0.221	1
	rs2708068	12:42873478	0.224	1
	rs1669916	12:42878879	0.209	1
FOCAD	rs4978017	9:20739231	0.333	0.631
	rs10757149	9:20831915	0.348	0.767
	rs2026994	9:20834837	0.348	0.647
	rs2151001	9:20836862	0.293	0.697

[a] minor allele frequency in the replication cohort

[b] r^2 with either rs1796391 or rs10511683 in the 1000 Genome CEU population

Table 6.7: vQTL replication analysis results for the proxy SNPs of the discovery lead SNP rs1796391

SNP	Beta	P-value
rs1796362	-0.039	0.104
rs1796361	-0.039	0.102
rs2708068	-0.039	0.102
rs1669916	-0.038	0.123

Table 6.8: G × G test p-values for the proxy SNP pairs in the replication cohort

		PRICKLE1 proxy SNP			
		rs1669916	rs1796361	rs1796362	rs2708068
FOCAD proxy SNP	rs10757149	0.257	0.248	0.247	0.249
	rs2026994	0.268	0.261	0.260	0.260
	rs2151001	0.146	0.151	0.150	0.166
	rs4978017	0.264	0.242	0.239	0.227

the 1 Mb window centered around *PRICKLE1* and *FOCAD*. Using a MAF filter of 0.1, 906 *PRICKLE1* SNPs and 796 *FOCAD* SNPs were tested for pairwise G × G. Figure D6 shows that the distribution of the resulting 721176 p-values was consistent with the null expectation, except for p-values at the tail. Two pairs of SNPs passed the conventional genome-wide threshold: rs10964862 (*FOCAD*) × rs11181736 (*PRICKLE1*) with a p-value=4.53×10^{-9} and rs10964862 (*FOCAD*) × rs11181735 (*PRICKLE1*) with a p-value=9.33×10^{-9}. The top interacting SNPs in the replication cohort were 340 kb and 380 kb away from that in the PITT cohort. SNP rs10964862 (*FOCAD*) showed significant or nearly significant interaction with multiple SNPs at the *PRICKLE1* locus (Figure 6.17). This SNP does not overlap with known coding sequence.

6.4 Discussion

With a shift in focus from facial shape mean to facial shape variability, this study for the first time examined the genetic basis of normal-range facial morphology from a variance perspective and demonstrated the existence and research value of vQTL. The extension of a conventional variance homogeneity test to multivariate outcomes helped with prototyping the statistical pipeline for application in future cross-trait settings. The result that none of the discovered vQTLs overlapped with known GWAS loci highlights the unique value of

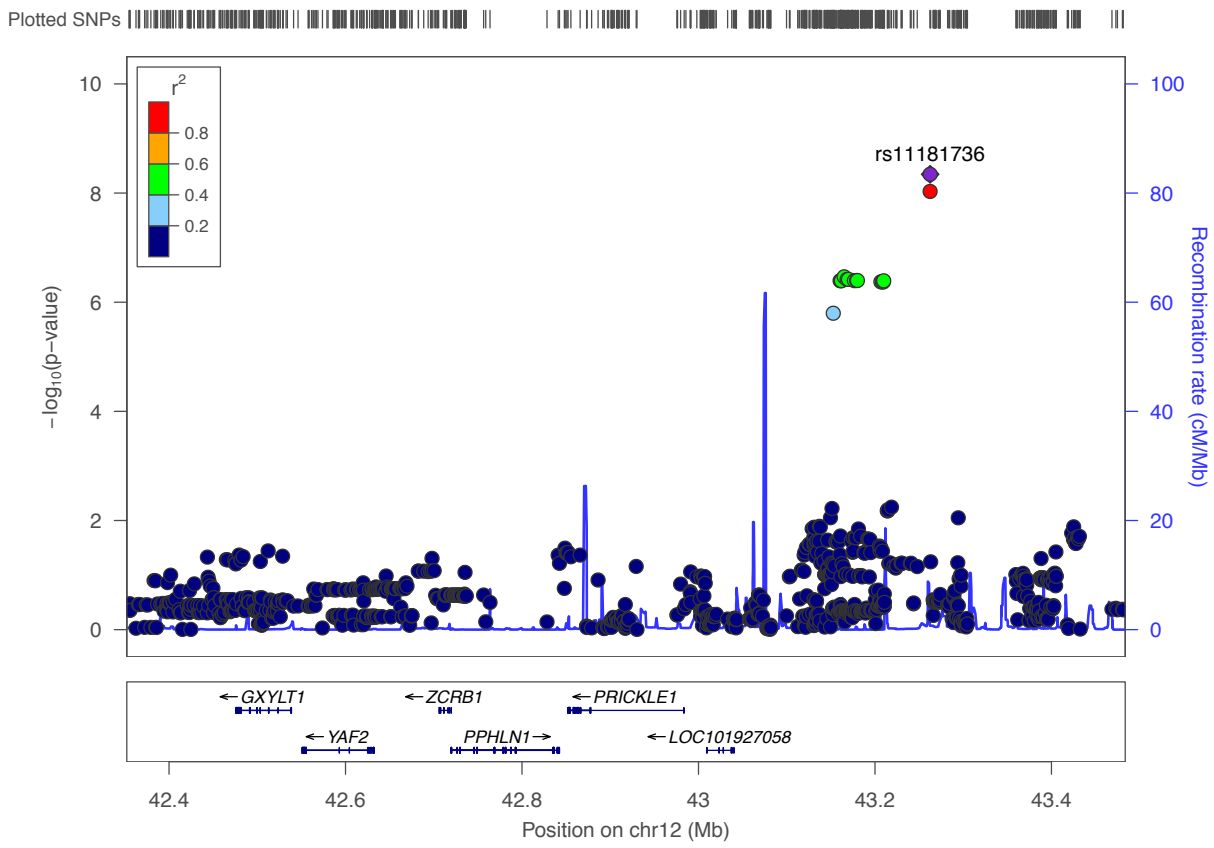


Figure 6.17: Replication of *PRICKLE1* × *FOCAD* interaction. Regional plot showing the interaction p-values between rs10964862 (*FOCAD*) and SNPs at the *PRICKLE1* locus

vQTL studies. Such loci can only be uncovered via a study focusing on phenotypic variance rather than on phenotypic means; the genes they harbor were completely missing by previous GWAS. By carefully scrutinizing the possible mechanisms for the discovered vQTLs, we uncovered SNP by SNP interactions pointing to biologically relevant genes. These findings expanded our understanding of the genetic control of human facial morphology and have great potential in guiding future studies to further forward field knowledge.

Our analysis of the multivariate facial modules identified novel candidate loci and genes, a few of which were reported to have relevant function in bone, limb, and craniofacial structures in previous studies. For example, the Tanzania vQTL signal at 2q22.3 is near to three genes with some previous evidence: *ACVR2A* has a role in cranial neural crest cell patterning and *acvr2a*-depleted zebrafish exhibited defects in cartilage, bone, and pharyngeal tooth structures [165]; mutations in *ORC4* can cause Meier-Gorlin syndrome 2 whose clinical hallmarks include flat philtrum, micrognathia and mandibular hypoplasia [166]; *MBD5* represents another gene which underlies human craniofacial syndromes [167]. Loci where interaction signals were detected also had promising candidate genes. For instance, one of the identified interaction loci in the Tanzania data harbors gene *GPC6* (Figure D2c). In a previous study, *GPC6* was able to impair ossification and cause omodysplasia (OMIM 604404) when disrupted by mutations [168]. Omodysplasia patients displayed short limbs, short stature, posteriorly rotated ears and mild micrognathia. Expression of mouse *Gpc6* was detected in proliferative chondrocytes at growth plate. Though supportive, these lines of evidences do not prove the functional role of these genes at the vQTLs. Like any other GWASs, fine-mapping is needed to prioritize potential causal genes.

Our vQTL analysis of the traditional facial distance phenotypes also generated new insights. *PRICKLE1* has been shown to play crucial roles in craniofacial morphogenesis in animal experiments and rare human syndromes but has never been reported being significant in human population studies, regardless of whether a candidate strategy or a genome-wide search was performed. By studying variance effects, our analysis of cranial base width for the first time highlighted *PRICKLE1* in a human association study. Although this signal did not surpass the study-wide significance threshold, we still consider it as being a

promising finding given its relevant biology and relatively small size of the study cohort. Several factors may contribute to the failure of replication of the *PRICKLE1* vQTL. First, the signal was discovered in European population while the replication cohort was of Asian ancestry. Differences in LD structure, the nonavailability of the original top SNPs, and/or differences in genetic and environmental background may all have a role in shaping the distribution of the variance effect itself and/or our ability to detect variance effects. Second, the facial phenotypes were slightly different in the two cohorts. Despite a lack of variance effects, we did replicate the G $\times$ G signal in the Korean dataset if we consider a window surrounding the original signal. The LD difference in the two populations may explain the shift in the location of the interaction signal. SNPs involved in the replication signal did not overlap with known regulatory elements of *PRICKLE1*/*FOCAD*. Their role in the context of craniofacial morphogenesis needs to be examined in future studies.

PRICKLE1 serves as a signaling factor in the noncanonical Wnt pathway, the disruption of which is known to cause cleft palate and stunt limb growth [169–171]. A recently study characterized the functions of Prickle1a and Prickle1b in zebrafish cranial neural crest cell development during epithelial-to-mesenchymal transition and migration [172]. Studies in mouse also support its essential role in craniofacial development. Prickle1 missense allele mutant mice were microcephalic and displayed several craniofacial defects including a cleft lip, incompletely penetrant cleft palate, and a shorter proximal-distal axis of the head [173]. These phenotypes were a result of the abnormal migration and differentiation of osteoblast precursors in the frontal bone in the absence of functioning Prickle1 protein. Sequencing studies of craniofacial syndrome patients implicated both rare and common *PRICKLE1* variants [174]. Our findings suggest that *PRICKLE1* influences facial width and disease risk not by shifting the average level of the phenotype but instead via its control over how variable the phenotype can be. We further demonstrated the G $\times$ G effect between *PRICKLE1* and *FOCAD*. There is so far little known about *FOCAD* except it being a potential tumor suppressor gene. The biological relevance of this *PRICKLE1* $\times$ *FOCAD* signal is currently being studied at one of our collaborators, Heather Szabo-Rogers's laboratory using *Prickle1* mutant mice, and the preliminary results are promising.

Consistent with findings from previous studies, our results suggested that vQTLs are much less common than QTLs (loci affecting phenotypic mean) for complex human traits. Studies on both more proximal molecular phenotypes and more distal disease status have all indicated a lesser contribution of vQTLs than QTLs to trait heritability [86, 101, 104]. This may also be a simple consequence of not having enough power to detect vQTLs, given the complexity and method immaturity in studying variance itself. Findings from existing vQTL studies are not at odds with the possibility of a large amount of vQTLs, each with a modest effect size or moderate-to-large effect size that only manifest under certain conditions. More studies in this area are needed to give further insights.

As aforementioned, variance heterogeneity can emerge from several different underlying mechanisms. How common these mechanisms are relative to one another is not known. This brings about the inherent limitation of the "variance prioritization" strategy - one cannot make immediate interpretation of the vQTL effect. We caution that although our data did not support mechanisms other than the interaction explanation, we cannot completely rule out other possibilities because the real causal SNPs may not be genotyped or imputed well and therefore their untagged causal effect may still confound a vQTL analysis. Additionally, we were not able to examine non-genetic factors other than sex in the G $\times$ E test, leaving the possibility that the detected vQTLs may interact with other unmeasured environmental factors in shaping facial morphology. Nonetheless, we have carefully scrutinized the mechanisms of the identified vQTLs within the capacity of available data and made meaningful justification for the most plausible one.

One factor that may or may not have affected our analysis was the complicated mathematical transformation that the multivariate facial phenotypes have gone through. Phenotype transformation has been suggested to have an influence on the power and/or false-positive rate of vQTL analysis [175]. The PC phenotypes used were standardized to have a mean of 0 and a variance of 1 in the entire sample, which may have added a layer of complexity to the variance test as well as brought extra vulnerability. However, imaging phenotypes do not have a natural "as-is" value like those directly quantifiable ones, such as height or bone mineral density, and they have to be transformed in some way. Our PC phenotypes have

been shown to be successful for gene mapping purpose in the previous GWAS [31]. Because the effects of phenotypic transformation is largely understudied for variance equality tests, it is also possible that such transformation may in fact be beneficial. The standardization process pulled the data closer to a normal distribution, and the pre-adjustment of covariates can lead to a minor improvement in statistical power [8]. One future direction from here is to evaluate the consequence of various types of variable transformation on variance homogeneity tests.

We acknowledge the limited power for detecting variance heterogeneity and suggest that there likely are more vQTLs for human facial shape that were undetected given the current sample size. Likewise, interactions involving the vQTLs could also have remained undetected due to insufficient power. This power issue is consistent with the fact that we failed to identify any study-wide significant signals. Our genome-wide significant vQTLs and G $\times$ G effects should not be considered genuine biological effects before successful replication in independent samples and verification by functional experiments. We were lucky to have access to an independent replication cohort via an external collaboration. Yet experimental validation was not a manageable task given the scope of the current dissertation work, and will continue to remain challenging even beyond the current project. Nevertheless, findings from this chapter raise interesting hypotheses which have great potential in generating novel insights by future investigations.

This vQTL aim is particularly innovative in that (1) vQTL is a highly understudied topic in current human genetics and has never been a theme of any kind in studies of human facial morphology; and (2) to the best of our knowledge, there has not been a single vQTL study on multivariate phenotypes in the literature, indicating a lack of attention on such phenotypes and insufficient research on statistical approaches suitable for multivariate settings. The multivariate analysis pipeline developed in this project provided an exemplar upon which studies of other multivariate phenotypes can build. As the first real case of vQTL study of multivariate phenotypes, this chapter demonstrated how a conventional variance equality test can be generalized and applied to 3D facial features, highlighted the potential of such a design to reveal interaction effects, and underscored the need to go beyond marginal effects

in future facial GWAS. Insights gained from this aim are also instructive for other complex, polygenic human traits and disorders. As biobank-scale cohorts with genome, phenome and exposome data are becoming increasingly available, the power of vQTL studies will eventually rise to a level where researchers will become much more confident in working on variance effects than they are today. We believe those studies will be able to add great value to our understanding of human genetics.

7.0 Conclusion

7.1 Summary

By leveraging existing genome-wide data and state-of-the-art advances in 3D facial phenotype modeling, this dissertation study addressed several gaps in current research and made important findings on human facial morphology. In Chapter 4 we performed an Exome-wide investigation of the role of rare and low-frequency genetic variants in a gene-based manner. Our findings extended the spectrum of genetic factors involved in facial morphology from common to rare and low-frequency variants. We highlighted novel candidate genes, some of which have shown promising evidence of verification in follow up experiments at a collaborating lab. The TWAS analysis in Chapter 5 represented a meaningful extension of our previous GWAS effort and refined potential mediating genes for establishing mechanistic connection. This work also pointed to novel genes where the individual SNPs never reached the genome-wide threshold in previous GWASs. Finally, the vQTL chapter featured the importance of studying variability as a phenotype in advancing the knowledge of genotype-phenotype relationships in complex traits. The variance prioritization strategy was shown to be useful in the discovery of G $\times$ G effects for facial morphology. These results expanded our understanding of the genetic basis of normal-range facial variation and will have important implications for future studies.

7.2 Significance

This dissertation project contributed to the genetic knowledge of human facial morphology in several aspects: (1) showing for the first time the role of rare and low-frequency variants and identifying promising genes in preliminary functional experiments; (2) refining GWAS loci for potential mediating genes as well as discovering novel candidate genes; and

(3) exploring the genetic factors underlying facial shape variability and uncovering gene by gene interaction effects. Findings from this study forwarded our understanding of the genetic architecture of facial shape variation, which will inform further exploration into the biology of craniofacial morphogenesis and the pathogenesis of craniofacial anomalies. The implementation of several recently developed statistical tools helped to evaluate and generalize their utility as well as identify their limitations. In the long-term, knowledge gained from this study will help lay the necessary scientific foundation for real-world applications such as DNA-informed facial prediction for clinical and forensic purposes.

Appendix A

Supplementary materials for Chapter 2

Table A1: Genes identified in normal-range facial variation GWAS

Gene	Function	Associated trait	Reference[a]
ACAD9, RAB7A	Regulates vesicle traffic	Philtrum	3
ALX3		Intercanthal width	9
ASPM	Mitotic spindle	Chin	3
BC039327	ncRNA	Nose prominence	3
CACNA2D3		Pronasale to left alare	8
C5orf64		Pronasale to left alare	8
C5orf50		Eyes to nasion, Zygions to nasion	7
CDH18		Nose bridge	3
COL17A1		Eyes to nasion	7
DCHS2	Likely function in cell adhesion, cartilage differentiation	Columella inclination, nose protrusion, nose tip angle	1
DHX35		Alae width	2
		Columella/nose tip	3
DLX6, DYNC1L1	Homeobox transcription factor	Chin	3
EDAR	Tumor necrosis factor receptor	Chin protrusion	1
EPHB3, DVL3	Ephrin	Nose bridge	3

Table A1: Continued from previous page

Gene	Function	Associated trait	Reference[a]
EYA4, RPS12	Ribosomal protein, may function in eye development	Forehead	3
FREM1	Expressed in the nasal midline	Central upper lip height	6
GL13	SHH signaling pathway	Nose wing breadth	1
HDAC8		Orbital spacing	6
		Intercanthal width	9
HNRNPR		Upper facial height, midfacial width	4
HOXD cluster	Morphogenesis	Mouth, philtrum, nose width	3
		Curvature of eyelid	2
KCTD15		Nose tip	3
MAFB		Cranial base width	9
MBTPS1	Craniofacial patterning	Upper facial profile height	5
NHP2, ZNF345A		Chin	3
OSR1-WDR35		Right facial angle of en-ex-go	2
PABP1, C1L2A, HADC8	Histone deacetylase	Intercanthal distance	9
PARK2	Proteasomal degradation	Upper facial height	6
PAX1	Chondrocyte differentiation	Nose wing breadth	1
		Nasal width	9
		Nasal width	3
PAX3	Active transcription factor in neural crest cells	Prominence and vertical position of nasion	8
		Nasion position	1

Table A1: Continued from previous page

Gene	Function	Associated trait	Reference[a]
		Nose quadrant	3
		Eyeballs to nasion	7
		Eyeballs to nasion	9
PAX9	Dental and craniofacial development	Cranial base width	9
PCDH15		Upper facial profile prominence	5
PDE8A	Expressed in ectoderm	Allometry, inner canthal distance	4
PKDCC		Mandible	3
PRDM16		Nasal width and height	7
		Nasal width and height	9
SCHIP1		Centroid size, face height and width	4
SOX9	Acts during chondrocyte differentiation	Nose prominence, nose width	3
SUPT3H, RUNX2	Bone development	Nasal width	1
		Nasal width	3
TBX15, WARS2		Upper face quadrant	3
TMEM163	Cadherin family	Eye height width and depth	5
TMTC2		Right endocanthion	8
TP63	Development signaling, epithelial morphogenesis	Distance between eyeballs	7
TRPC6	Cation channel subunit	Upper facial depth	9
WDR27		Eye tail length	2
ZNF219	Transcription partner of SOX9	Nasal ala length	9

[a] 1-Adhikari et al., 2016; 2-Cha et al., 2018; 3-Claes et al., 2018; 4-Cole et al., 2016; 5-Crouch et al., 2018; 6-Lee et al., 2017; 7-Liu et al., 2012; 8-Paternoster et al., 2012; 9-Shaffer et al., 2016

Appendix B

Supplementary materials for Chapter 4

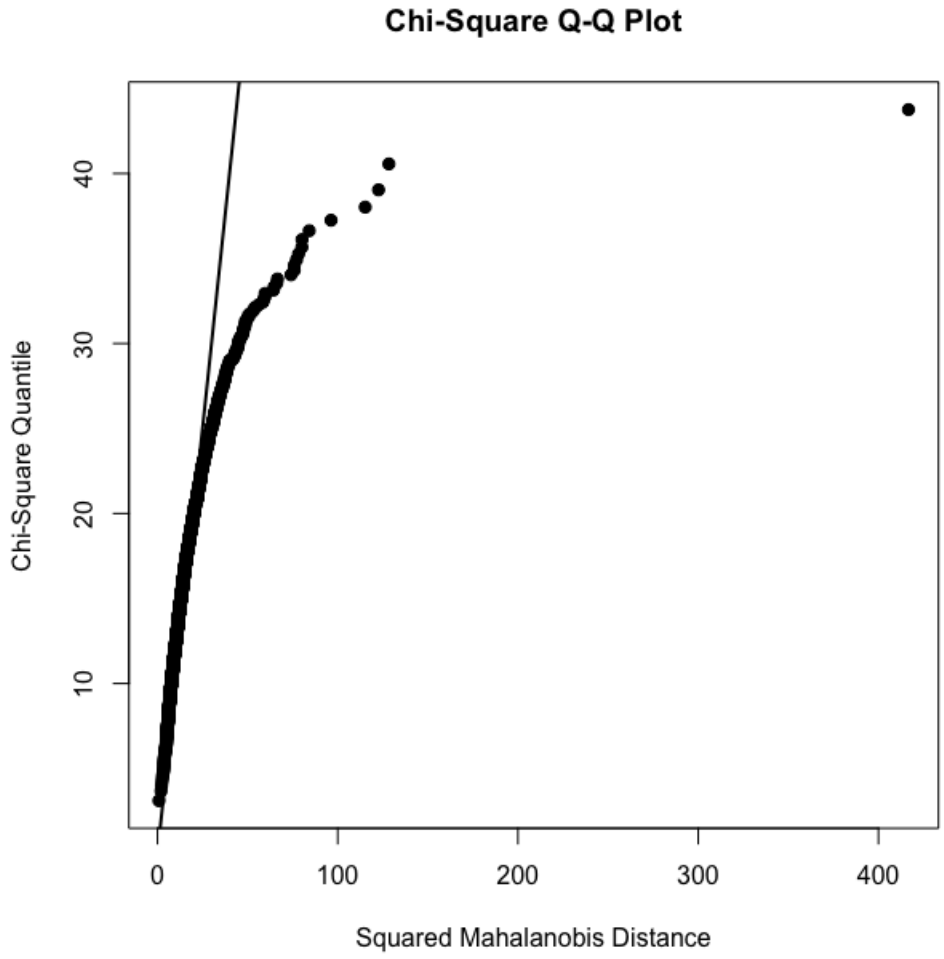


Figure B1: Multivariate outlier for PITT module 27. The outlier in the upper right fell far apart from the remaining data.

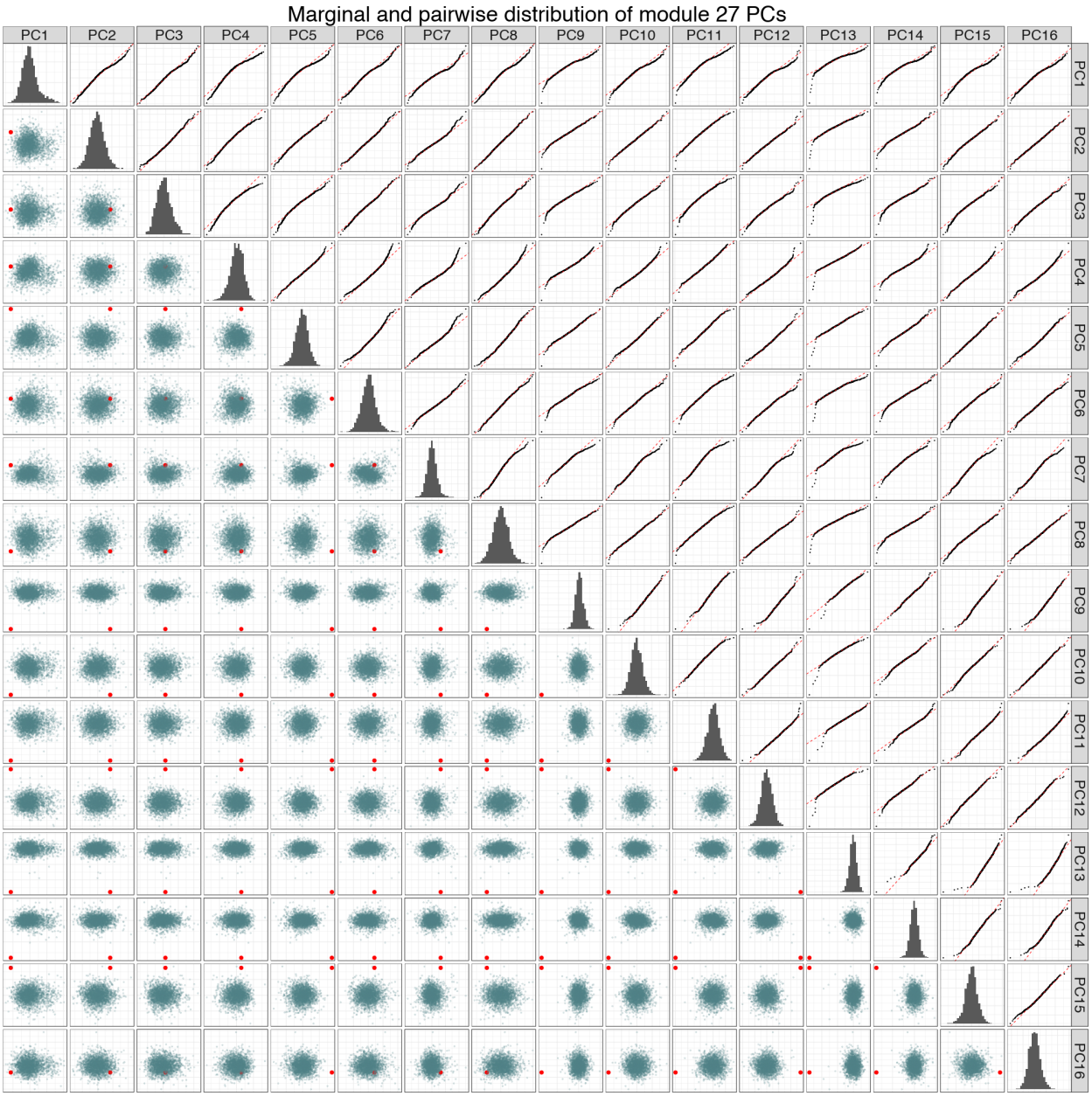


Figure B2: Distribution of PCs in PITT module 27. Panels in upper right triangle display the pairwise Q-Q plots. Panels in Bottom left triangle display the pairwise scatter plots. Histograms of individual PCs are shown along the diagonal. The PC values of the outlier identified in Figure B1 are highlighted in red.

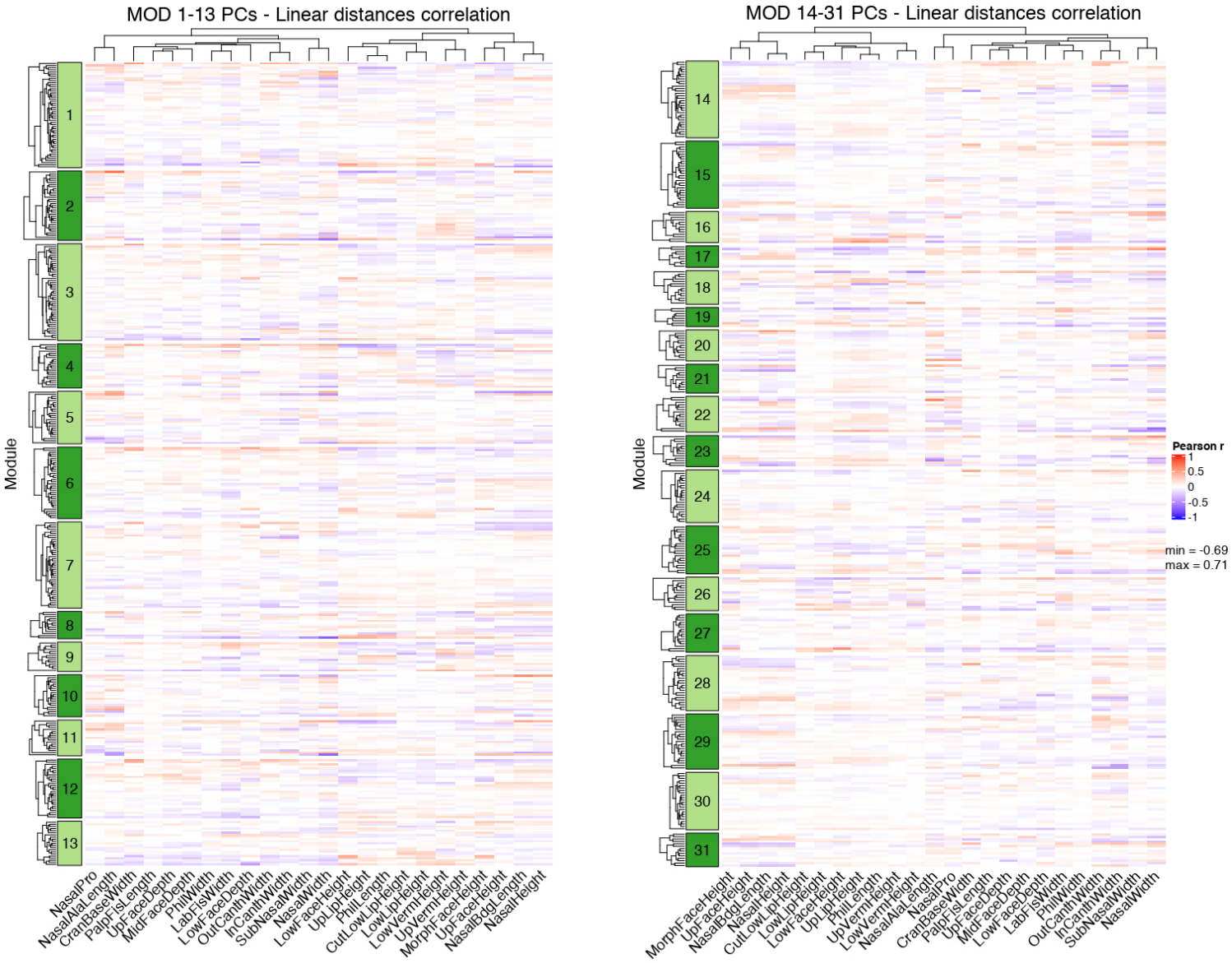


Figure B3: Pairwise Pearson correlation between individual PCs in 31 modules and 24 facial distances. Left: module 1-13; right: module 14-31. Rows are PCs and columns are facial distances. PCs are grouped into modules and within each module the rows are clustered by hierarchical clustering. Columns are also clustered. Color indicates Pearson correlation. The two sets of facial phenotypes were largely uncorrelated.

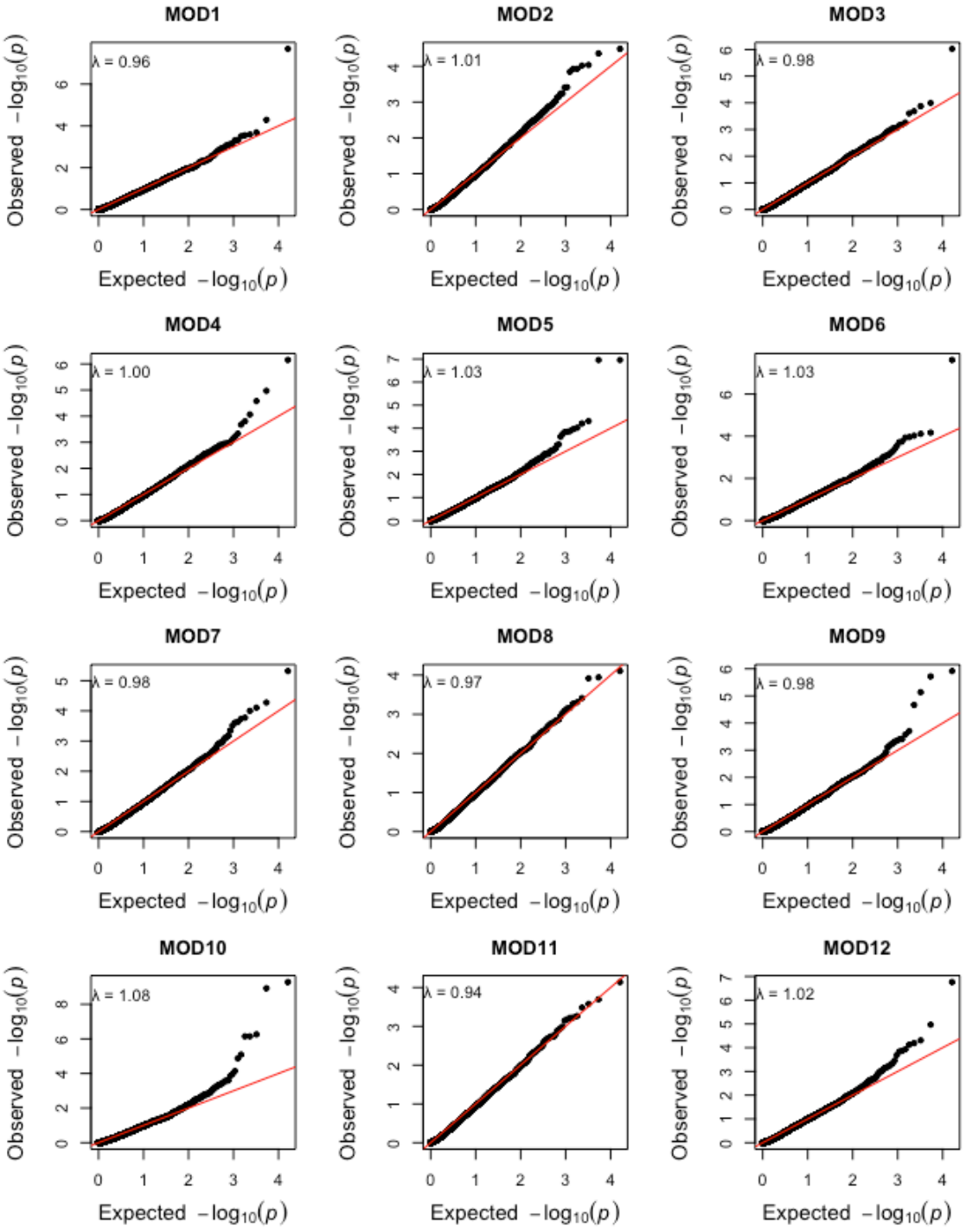


Figure B4: Q-Q plots for MultiSKAT analysis of 31 facial modules in the PITT cohort. Genomic inflation factor λ is shown in the top left corner of each sub figure.

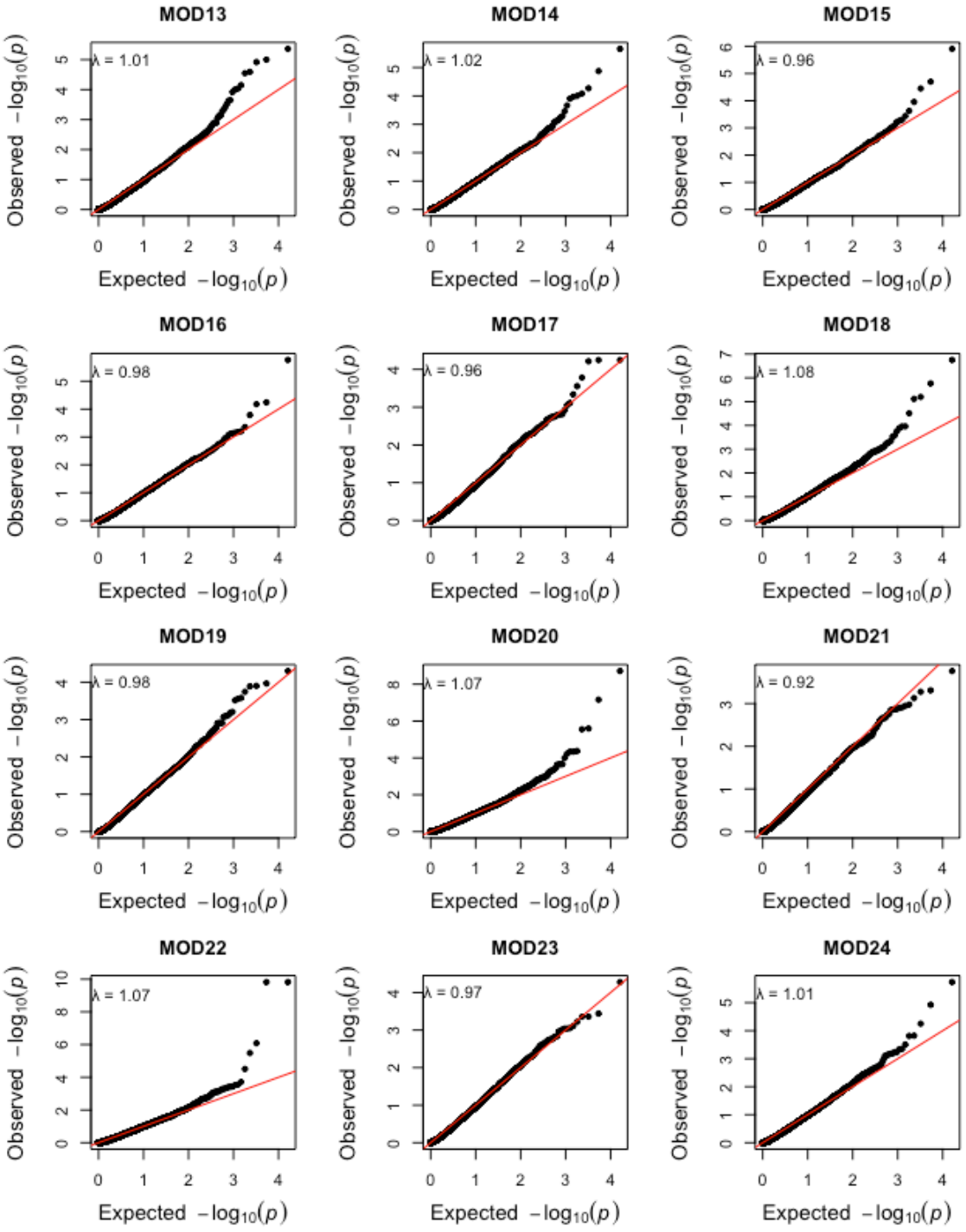


Figure B4: Q-Q plots for MultiSKAT analysis of 31 facial modules in the PITT cohort(cont.). Genomic inflation factor λ is shown in the top left corner of each sub figure.

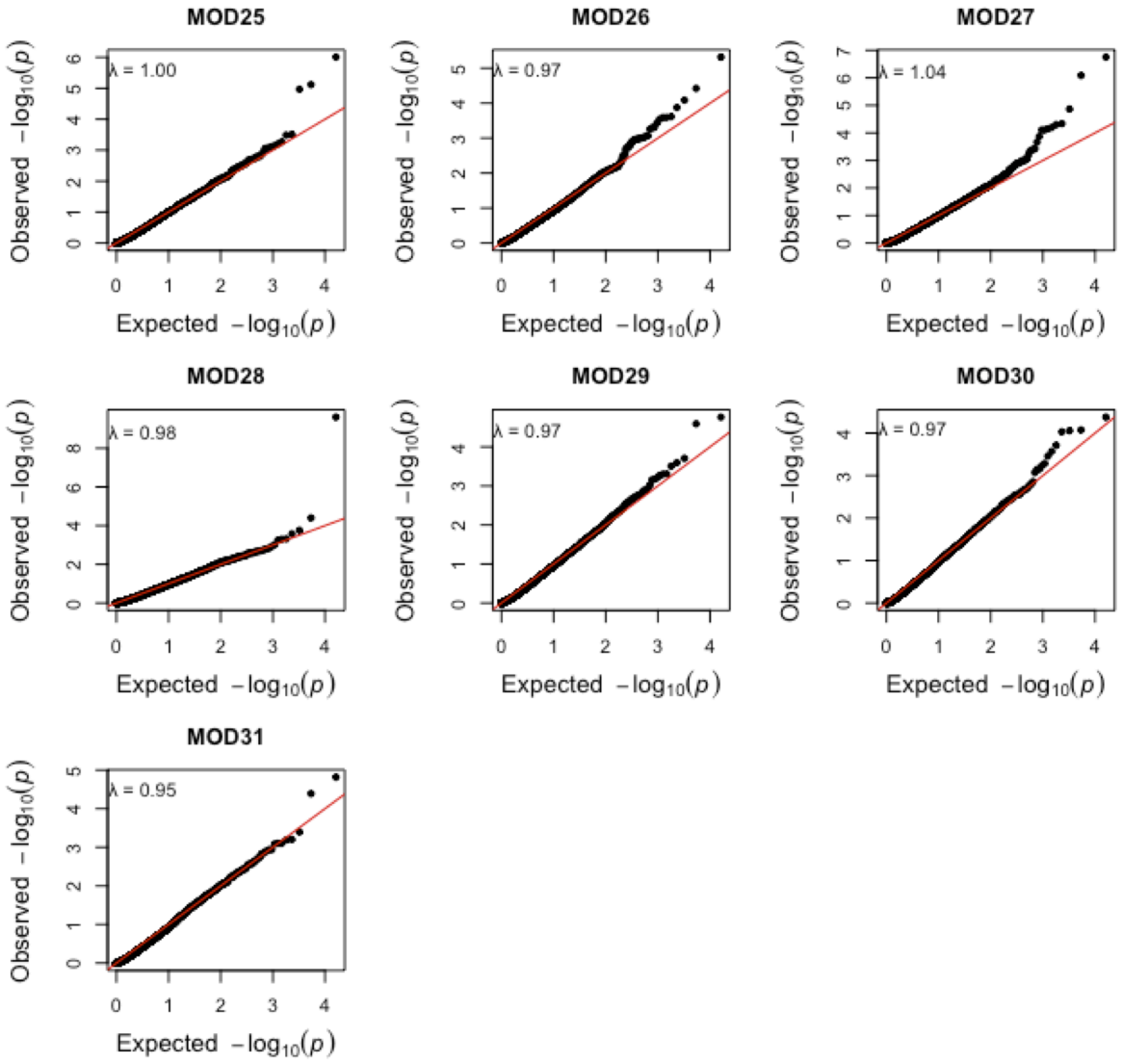


Figure B4: Q-Q plots for MultiSKAT analysis of 31 facial modules in the PITT cohort(cont.). Genomic inflation factor λ is shown in the top left corner of each sub figure.

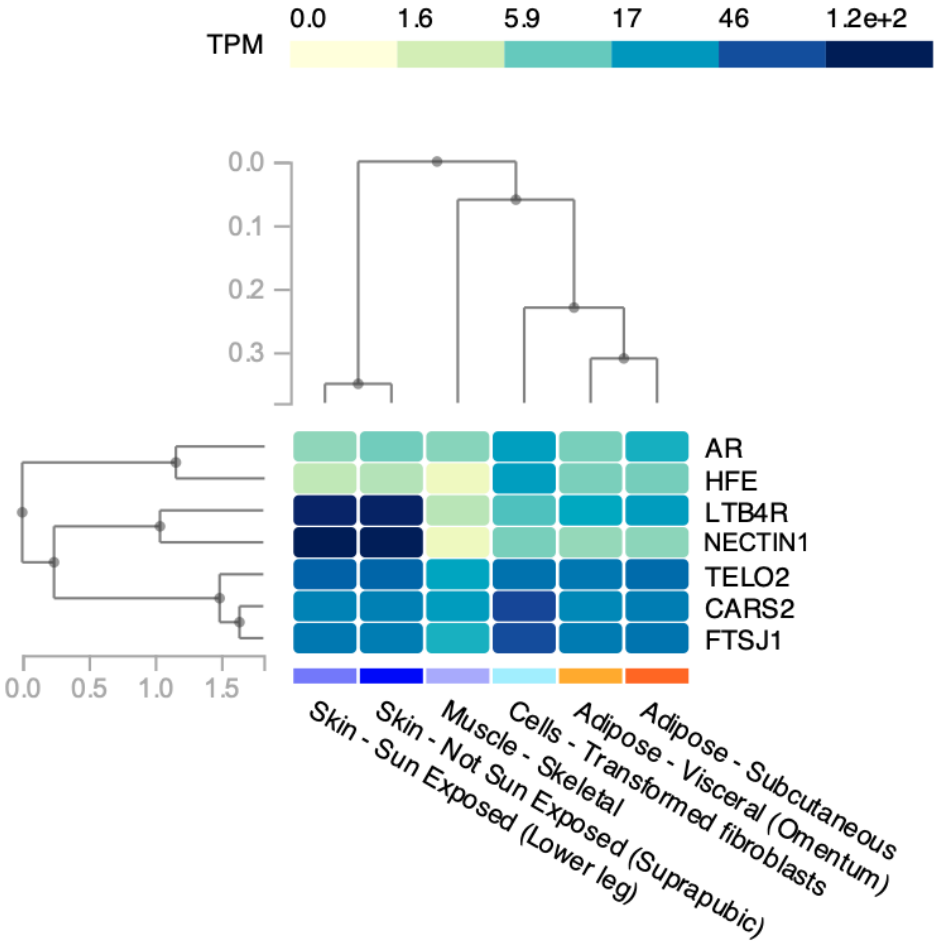


Figure B5: Expression of MultiSKAT significant genes in GTEx tissues relevant to facial morphology. Genes are arranged in rows and tissues in columns. Dendrogram shows similarity in expression levels. TPM, transcripts per million.

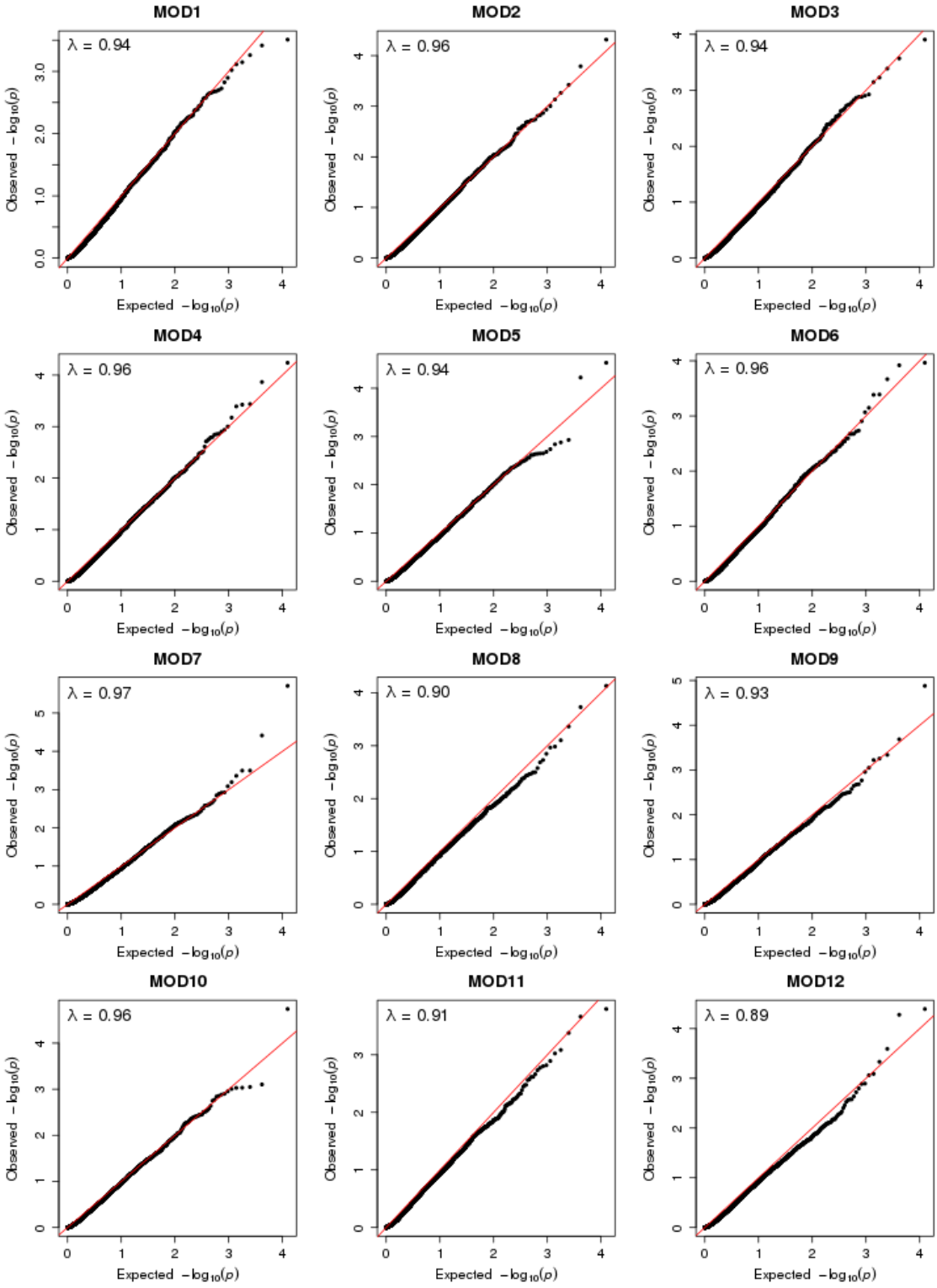


Figure B6: Q-Q plots for MultiSKAT analysis of 31 facial modules in the Tanzania cohort. Genomic inflation factor λ is shown in the top left corner of each sub figure.

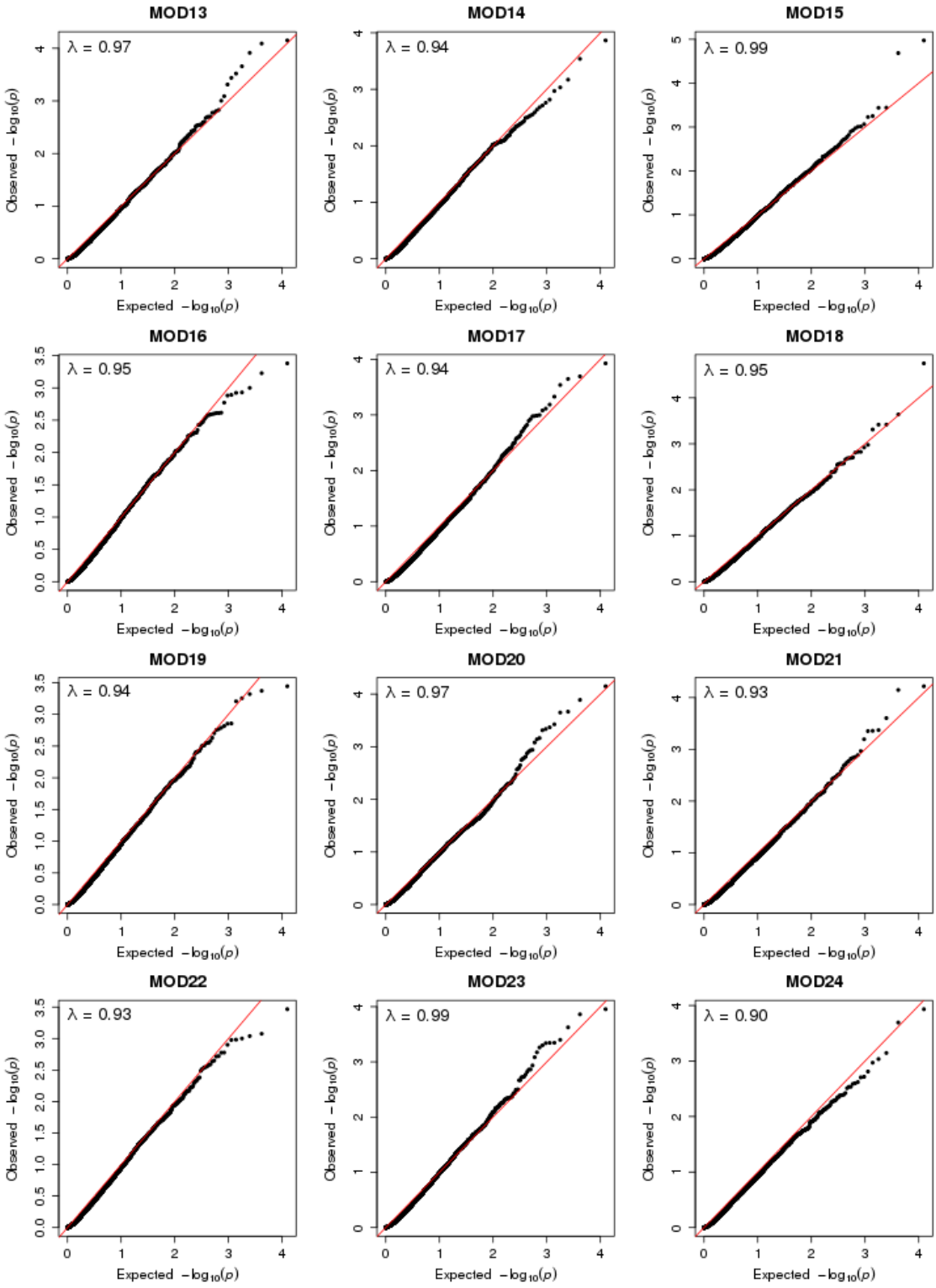


Figure B6: Q-Q plots for MultiSKAT analysis of 31 facial modules in the Tanzania cohort (cont.). Genomic inflation factor λ is shown in the top left corner of each sub figure.

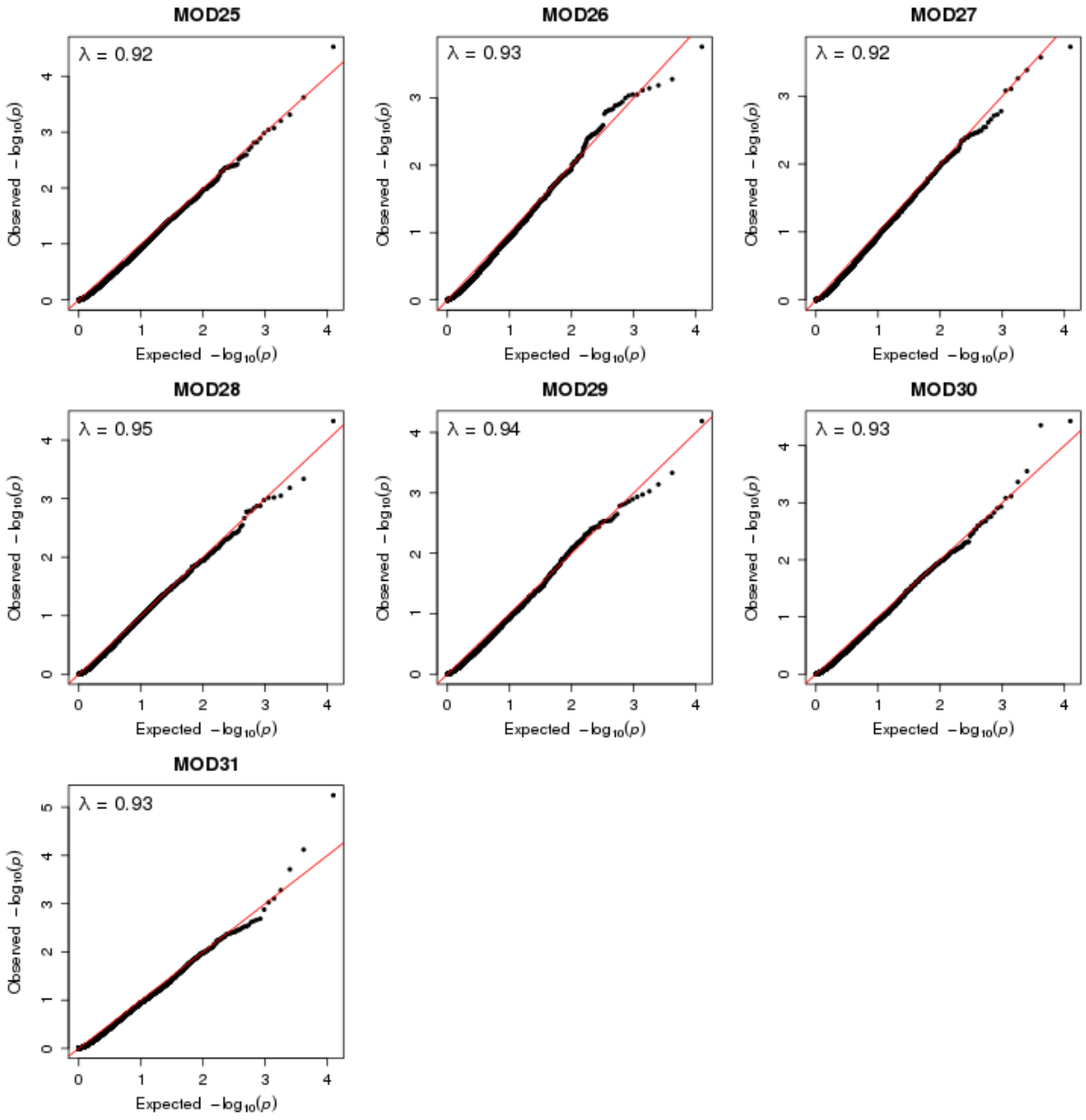


Figure B6: Q-Q plots for MultiSKAT analysis of 31 facial modules in the Tanzania cohort (cont.). Genomic inflation factor λ is shown in the top left corner of each sub figure.

Table B1: Previous genotype-phenotype associations for variants in the seven MultiSKAT significant genes. Only shows those with a p-value $< 10^{-4}$ in the PhenoScanner database

SNP	hg19 coordinates	Trait	Study	PMID	Ancestry	Year
rs117788141	chr13:111357899	Cause of death: thoracic aortic aneurysm, ruptured	Neale B	UKBB	European	2017
		Cause of death: myelodysplastic syndrome, unspecified	Neale B	UKBB	European	2017
		Intracranial injury	Neale B	UKBB	European	2017
		Treatment with indometacin	Neale B	UKBB	European	2017
		Treatment with movelat cream	Neale B	UKBB	European	2017
rs137852591	chrX:66941751	Height	GIANT	28146470	Mixed	2017
		Height	GIANT	28146470	European	2017
rs142863092	chr11:119548369	Cause of death: vascular dementia, unspecified	Neale B	UKBB	European	2017
		Treatment with progynova 1mg tablet	Neale B	UKBB	European	2017
		Osteoporosis with pathological fracture	Neale B	UKBB	European	2017
		Deafness	Neale B	UKBB	European	2017
		Treatment with lamotrigine	Neale B	UKBB	European	2017
		Self-reported tennis elbow or lateral epicondylitis	Neale B	UKBB	European	2017
		Cause of death: multiple myeloma	Neale B	UKBB	European	2017
		Self-reported uterine or endometrial cancer	Neale B	UKBB	European	2017
rs151097801	chr13:111296817	Cause of death: organ-limited amyloidosis	Neale B	UKBB	European	2017
		Cause of death: myelodysplastic syndrome, unspecified	Neale B	UKBB	European	2017
		Home area population density: postcode not linkable	Neale B	UKBB	European	2017
		Cause of death: malignant neoplasms of independent multiple sites	Neale B	UKBB	European	2017
		Treatment with clonazepam	Neale B	UKBB	European	2017
		Other and unspecified injuries of lower leg	Neale B	UKBB	European	2017
		Home area population density: Scotland large urban area	Neale B	UKBB	European	2017
		Other demyelinating diseases of central nervous system	Neale B	UKBB	European	2017

Table B1: Previous genotype-phenotype associations for variants in the seven MultiSKAT significant genes (cont.)

SNP	Beta	Se	P-value	Direction	N	N_cases	N_controls	N_studies	Unit	Dataset
rs117788141	-0.061	0.0075	4.87E-16	-	7637	7	7630	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.061	0.009	1.21E-11	-	7637	10	7627	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.006	0.0012	4.86E-06	-	337199	440	336759	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.004	0.0009	1.34E-05	-	337159	232	336927	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.002	0.0005	3.86E-05	-	337159	73	337086	1	risk diff	Neale-B_UKBB_EUR_2017
rs137852591	0.11	0.024	4.1E-06	+	458927	0	458927	147	IVNT	GIANT_Height_Mixed_2017
	0.11	0.024	7.7E-06	+	381625	0	381625	106	IVNT	GIANT_Height_EUR_2017
rs142863092	-0.057	0.0092	5.93E-10	-	7637	11	7626	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.004	0.0008	5.6E-07	-	337159	141	337018	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.003	0.0006	3.65E-06	-	337199	76	337123	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.003	0.0006	5.59E-06	-	323978	78	323900	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.006	0.0013	6.51E-06	-	337159	408	336751	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.004	0.0009	1.96E-05	-	337159	179	336980	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.103	0.0258	6.18E-05	-	7637	87	7550	1	risk diff	Neale-B_UKBB_EUR_2017
	-0.007	0.0019	7.49E-05	-	337159	843	336316	1	risk diff	Neale-B_UKBB_EUR_2017
rs151097801	0.0615	0.0075	4.06E-16	+	7637	7	7630	1	risk diff	Neale-B_UKBB_EUR_2017
	0.061	0.009	1.58E-11	+	7637	10	7627	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0012	0.0002	1.3E-08	+	333997	12	333985	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0597	0.0128	3.03E-06	+	7637	20	7617	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0044	0.0009	3.26E-06	+	337159	237	336922	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0023	0.0005	6.09E-06	+	337199	70	337129	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0339	0.0076	8.95E-06	+	333997	18867	315130	1	risk diff	Neale-B_UKBB_EUR_2017
	0.0033	0.0008	2.43E-05	+	337199	163	337036	1	risk diff	Neale-B_UKBB_EUR_2017

Table B2: MultiSKAT results of *ARHGEF18*, *CARS2*, *NECTIN1*, and *TELO2* in the Tanzania cohort

Module	MultiSKAT p-value of genes			
	ARHGEF18	*CARS2*	*NECTIN1*	*TELO2*
1	0.921	0.896	0.979	0.141
2	0.801	0.577	0.476	0.326
3	0.618	0.103	0.254	0.175
4	0.988	0.779	0.931	0.594
5	0.775	0.161	0.177	0.093
6	0.742	0.617	0.536	0.097
7	0.400	0.300	0.751	0.944
8	0.939	0.650	0.953	0.173
9	0.896	0.460	0.663	0.975
10	0.879	0.287	0.411	0.008
11	0.679	0.038	0.584	0.307
12	0.795	0.310	0.840	0.265
13	0.470	0.019	0.114	0.049
14	0.297	0.131	0.665	0.973
15	0.632	0.156	0.952	0.224
16	0.989	0.659	0.177	0.250
17	0.091	0.988	0.110	0.045
18	0.603	0.077	0.302	0.992
19	0.551	0.594	0.144	0.739
20	0.982	0.127	0.950	0.157
21	0.198	0.033	0.962	0.050
22	0.007	0.040	0.406	0.079
23	0.251	0.067	0.580	0.272
24	0.983	0.191	0.627	0.658
25	0.300	0.547	0.436	0.307
26	0.760	0.748	0.077	0.251
27	0.982	0.180	0.503	0.029
28	0.321	0.131	0.576	0.976
29	0.652	0.353	0.697	0.969
30	0.999	0.418	0.957	0.732
31	0.591	0.043	0.734	0.053

Appendix C

Supplementary materials for Chapter 5

Table C1: Suggestive TWAS genes

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
PITT	1q31.3	*C1orf53*	13	2.23E-07	Muscle_Skeletal	EN	2
PITT	1q31.3	*C1orf53*	55	1.11E-06	Muscle_Skeletal	EN	2
PITT	1q43	*CHML*	35	2.82E-06	Brain_Cortex	EN	108
PITT	1q31.3	*CRB1*	13	5.80E-07	Muscle_Skeletal	MASHR	1
PITT	1q31.3	*CRB1*	27	3.66E-07	Muscle_Skeletal	MASHR	1
PITT	1q31.3	*CRB1*	55	2.26E-06	Muscle_Skeletal	MASHR	1
PITT	1q42.2	*NTPCR*	59	1.29E-06	Adipose_Subcutaneous	MASHR	4
PITT	1p12	*TBX15*	1	5.73E-07	Cells_Cultured_fibroblasts	MASHR	1
PITT	1p12	*TBX15*	1	5.73E-07	Brain_Cortex	MASHR	1
PITT	1p12	*TBX15*	3	3.03E-06	Cells_Cultured_fibroblasts	MASHR	1
PITT	1p12	*TBX15*	3	3.03E-06	Brain_Cortex	MASHR	1
PITT	1p12	*TBX15*	25	1.62E-07	Cells_Cultured_fibroblasts	MASHR	1
PITT	1p12	*TBX15*	25	1.62E-07	Brain_Cortex	MASHR	1
PITT	1p12	*WARS2*	7	3.14E-07	Muscle_Skeletal	EN	49
PITT	1p12	*WARS2*	7	6.44E-07	Adipose_Subcutaneous	EN	23
PITT	1p12	*WARS2*	7	3.37E-07	Adipose_Subcutaneous	MASHR	3
PITT	1p12	*WARS2*	7	2.51E-07	Cells_Cultured_fibroblasts	EN	18
PITT	1p12	*WARS2*	7	3.07E-07	Brain_Cortex	EN	34
PITT	1p12	*WARS2*	15	1.34E-06	Muscle_Skeletal	MASHR	4
PITT	1p12	*WARS2*	15	3.19E-06	Cells_Cultured_fibroblasts	MASHR	3
PITT	1p12	*WARS2*	15	5.15E-06	Brain_Cortex	EN	34
PITT	1p12	*WARS2*	15	1.58E-06	Brain_Cortex	MASHR	2
PITT	1p12	*WARS2*	61	3.09E-07	Muscle_Skeletal	MASHR	4
PITT	1p12	*WARS2*	61	1.38E-06	Cells_Cultured_fibroblasts	MASHR	3
PITT	1p12	*WARS2*	61	1.00E-06	Brain_Cortex	MASHR	2
PITT	1q21.3	*ZBTB7B*	29	1.85E-06	Adipose_Subcutaneous	MASHR	1

Table C1: Continued from previous page

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
PITT	1q42.13	*ZNF678*	8	6.70E-06	Brain_Cortex	EN	61
PITT	2q11.2	*CIAO1*	62	2.55E-06	Brain_Cortex	EN	26
PITT	2q37.3	*HDLBP*	6	1.79E-06	Cells_Cultured_fibroblasts	EN	22
PITT	2q37.3	*HDLBP*	12	1.84E-06	Cells_Cultured_fibroblasts	EN	22
PITT	2p21	*LINC01914*	52	4.57E-06	Adipose_Subcutaneous	EN	13
PITT	3q21.3	*EEFSEC*	4	4.39E-07	Adipose_Subcutaneous	MASHR	2
PITT	3q21.3	*EEFSEC*	4	5.55E-06	Cells_Cultured_fibroblasts	EN	51
PITT	3q21.3	*EEFSEC*	4	3.16E-06	Brain_Cortex	MASHR	2
PITT	3q21.3	*EEFSEC*	8	8.36E-07	Adipose_Subcutaneous	EN	50
PITT	3q21.3	*EEFSEC*	8	1.14E-06	Cells_Cultured_fibroblasts	EN	51
PITT	3q21.3	*EEFSEC*	8	5.36E-07	Cells_Cultured_fibroblasts	MASHR	2
PITT	3q21.3	*EEFSEC*	8	1.87E-07	Brain_Cortex	MASHR	2
PITT	3q21.3	*EEFSEC*	22	4.58E-07	Adipose_Subcutaneous	MASHR	2
PITT	3q21.3	*EEFSEC*	32	4.57E-07	Adipose_Subcutaneous	EN	50
PITT	3q21.3	*EEFSEC*	33	3.17E-06	Muscle_Skeletal	EN	25
PITT	3q21.3	*EEFSEC*	37	4.21E-06	Brain_Cortex	EN	8
PITT	3q27.1	*EPHB3*	4	2.01E-06	Brain_Cortex	MASHR	2
PITT	3q27.1	*EPHB3*	17	2.55E-06	Brain_Cortex	MASHR	2
PITT	3p25.3	*IRAK2*	20	9.07E-06	Brain_Cortex	EN	6
PITT	3q21.3	*RUVBL1*	5	3.10E-06	Cells_Cultured_fibroblasts	MASHR	2
PITT	3q21.3	*RUVBL1*	8	3.64E-06	Muscle_Skeletal	MASHR	2
PITT	3q21.3	*RUVBL1*	16	5.01E-06	Cells_Cultured_fibroblasts	EN	55
PITT	3q21.3	*RUVBL1*	45	1.74E-06	Cells_Cultured_fibroblasts	MASHR	2
PITT	3q21.3	*SEC61A1*	32	1.23E-06	Cells_Cultured_fibroblasts	EN	13
PITT	3q21.3	*SEC61A1*	33	1.65E-06	Muscle_Skeletal	MASHR	2
PITT	4q23	*ADH6*	21	7.55E-07	Cells_Cultured_fibroblasts	EN	64
PITT	4q23	*ADH6*	43	4.45E-06	Cells_Cultured_fibroblasts	EN	64
PITT	6p21.33	*CCHCR1*	57	3.08E-06	Brain_Cortex	EN	14
PITT	6q23.2	*EYA4*	7	4.78E-07	Adipose_Subcutaneous	MASHR	1
PITT	6q23.2	*EYA4*	15	2.23E-06	Adipose_Subcutaneous	MASHR	1
PITT	6q23.2	*EYA4*	63	2.86E-07	Adipose_Subcutaneous	MASHR	1
PITT	6p21.1	*RUNX2*	10	1.84E-06	Adipose_Subcutaneous	MASHR	1
PITT	7q36.3	*LINC01006*	14	9.49E-07	Adipose_Subcutaneous	MASHR	2
PITT	11p11.2	*ATG13*	33	3.92E-06	Cells_Cultured_fibroblasts	MASHR	1

Table C1: Continued from previous page

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
PITT	11q12.2	*CCDC86*	49	1.73E-06	Brain_Cortex	EN	27
PITT	11q12.2	*CCDC86*	51	4.56E-06	Brain_Cortex	EN	27
PITT	12p11.22	*CCDC91*	21	5.70E-06	Adipose_Subcutaneous	EN	80
PITT	12p11.22	*CCDC91*	21	3.73E-06	Cells_Cultured_fibroblasts	EN	20
PITT	12p11.22	*CCDC91*	43	1.41E-07	Cells_Cultured_fibroblasts	MASHR	3
PITT	12q21.31	*NTS*	58	1.14E-06	Adipose_Subcutaneous	EN	17
PITT	14q12	*LTB4R*	54	8.58E-07	Muscle_Skeletal	EN	16
PITT	15q24.1	*CLK3*	11	2.42E-06	Muscle_Skeletal	MASHR	1
PITT	15q24.1	*CLK3*	11	2.42E-06	Adipose_Subcutaneous	MASHR	1
PITT	15q24.1	*CLK3*	11	2.42E-06	Cells_Cultured_fibroblasts	MASHR	1
PITT	15q24.1	*CLK3*	11	2.42E-06	Brain_Cortex	MASHR	1
PITT	15q24.1	*CLK3*	25	2.59E-06	Muscle_Skeletal	MASHR	1
PITT	15q24.1	*CLK3*	25	2.59E-06	Adipose_Subcutaneous	MASHR	1
PITT	15q24.1	*CLK3*	25	2.59E-06	Cells_Cultured_fibroblasts	MASHR	1
PITT	15q24.1	*CLK3*	25	2.59E-06	Brain_Cortex	MASHR	1
PITT	16q24.3	*DBNDD1*	35	1.82E-06	Adipose_Subcutaneous	MASHR	5
PITT	17q21.33	*MYCBPAP*	9	4.28E-06	Adipose_Subcutaneous	EN	45
PITT	17q11.2	*RHOT1*	17	3.42E-06	Muscle_Skeletal	MASHR	1
PITT	19p13.2	*ICAM5*	21	2.96E-06	Cells_Cultured_fibroblasts	MASHR	1
PITT	19p13.2	*ICAM5*	21	2.96E-06	Brain_Cortex	MASHR	1
PITT	22q12.1	*C22orf31*	51	3.17E-06	Brain_Cortex	EN	90
Denver	3q13.31	*ZDHHC23*	61	5.13E-07	Muscle_Skeletal	MASHR	1
Denver	3q13.31	*ZDHHC23*	61	5.13E-07	Adipose_Subcutaneous	MASHR	1
Denver	3q13.31	*ZDHHC23*	61	5.13E-07	Cells_Cultured_fibroblasts	MASHR	1
Denver	5q11.2	*PLK2*	63	3.29E-07	Muscle_Skeletal	MASHR	2
Denver	6p21.2	*MTCH1*	30	2.32E-07	Cells_Cultured_fibroblasts	MASHR	1
Denver	9q34.3	*NDOR1*	36	1.10E-06	Adipose_Subcutaneous	MASHR	2
Denver	9q34.3	*NDOR1*	50	4.23E-06	Adipose_Subcutaneous	EN	11
Denver	10q22.1	*KIF1BP*	3	1.99E-06	Adipose_Subcutaneous	MASHR	1
Denver	10q24.33	*SLK*	21	2.14E-06	Cells_Cultured_fibroblasts	MASHR	2
Denver	11q23.3	*USP2*	26	3.84E-06	Brain_Cortex	MASHR	2
Denver	15q24.3	*HMG20A*	41	4.13E-06	Muscle_Skeletal	EN	5
Denver	16q13	*RSPRY1*	59	1.20E-06	Muscle_Skeletal	MASHR	2
Denver	19p13.3	*GZMM*	4	2.30E-06	Adipose_Subcutaneous	MASHR	1

Table C1: Continued from previous page

Cohort	Locus	Gene	Module	P-value	Tissue	Model	Num SNPs[a]
Denver	19q13.33	*TSKS*	56	5.13E-06	Muscle_Skeletal	EN	2
Denver	20q13.33	*STMN3*	30	3.88E-06	Cells_Cultured_fibroblasts	EN	5
Denver	20q13.33	*TNFRSF6B*	30	3.33E-06	Muscle_Skeletal	MASHR	2
Tanzania	1p13.3	*C1orf194*	1	6.61E-07	Brain_Cortex	MASHR	2
Tanzania	1p21.2	*CDC14A*	4	4.62E-06	Brain_Cortex	MASHR	1
Tanzania	3q21.3	*RPN1*	18	8.01E-06	Brain_Cortex	EN	103
Tanzania	3q21.3	*RPN1*	56	6.42E-06	Brain_Cortex	EN	103
Tanzania	5p15.33	*CTD-2012J19.3*	27	1.81E-06	Adipose_Subcutaneous	EN	14
Tanzania	5q23.1	HSD17B4	31	5.69E-06	Adipose_Subcutaneous	EN	51
Tanzania	5q31.1	*VDAC1*	60	4.20E-06	Adipose_Subcutaneous	MASHR	1
Tanzania	5q31.3	*PCDHA3*	13	7.14E-06	Brain_Cortex	EN	21
Tanzania	7q11.23	*FGL2*	4	4.88E-06	Adipose_Subcutaneous	EN	7
Tanzania	7q11.23	*FGL2*	4	4.02E-06	Cells_Cultured_fibroblasts	EN	12
Tanzania	7q33	*LRGUK*	8	9.21E-07	Adipose_Subcutaneous	EN	39
Tanzania	8p21.3	*ATP6V1B2*	7	3.53E-06	Adipose_Subcutaneous	MASHR	1
Tanzania	8q24.3	*CPSF1*	62	2.48E-06	Brain_Cortex	EN	34
Tanzania	12p12.1	*ST8SIA1*	45	3.19E-06	Muscle_Skeletal	EN	9
Tanzania	14q11.2	*CMTM5*	22	6.24E-07	Muscle_Skeletal	MASHR	1
Tanzania	14q23.3	*PLEKHG3*	21	3.99E-06	Cells_Cultured_fibroblasts	MASHR	2
Tanzania	14q24.3	*PTGR2*	10	2.47E-06	Adipose_Subcutaneous	MASHR	5
Tanzania	14q24.3	*PTGR2*	10	2.41E-06	Cells_Cultured_fibroblasts	MASHR	6
Tanzania	15q15.2	*CCNDBP1*	24	4.72E-06	Brain_Cortex	EN	7
Tanzania	15q26.3	*CERS3*	49	3.06E-06	Adipose_Subcutaneous	EN	18
Tanzania	16q12.1	*SIAH1*	42	1.78E-06	Adipose_Subcutaneous	EN	13
Tanzania	17p11.2	*FBXW10*	53	6.18E-06	Muscle_Skeletal	EN	15
Tanzania	17p11.2	*PRPSAP2*	42	6.67E-06	Brain_Cortex	EN	25
Tanzania	17q11.2	*SGK494*	63	3.89E-06	Cells_Cultured_fibroblasts	EN	24
Tanzania	19q13.2	*AC006129.1*	60	1.21E-06	Cells_Cultured_fibroblasts	EN	18
Tanzania	19p13.11	*YJEFN3*	17	1.06E-06	Muscle_Skeletal	MASHR	3
Tanzania	19q13.43	*ZNF471*	60	5.37E-06	Muscle_Skeletal	EN	41
Tanzania	19q13.43	*ZNF667-AS1*	60	2.41E-06	Cells_Cultured_fibroblasts	MASHR	2

[a] Number of SNPs used to predict gene expression

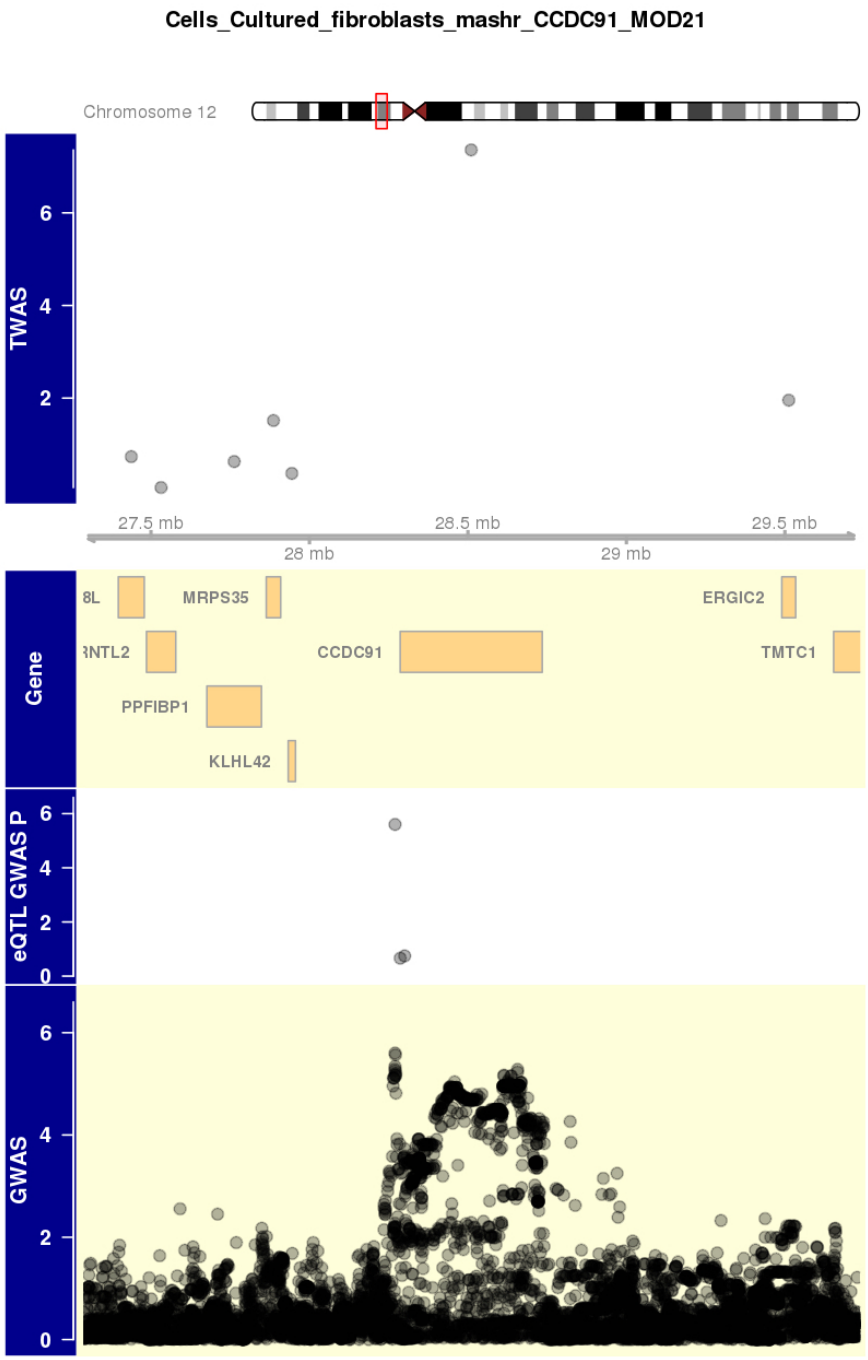


Figure C1: TWAS regional plot of *CCDC91*. TWAS of PITT module 21 (nose bridge) in fibroblasts using the MASHR model identified *CCDC91*. The top TWAS panel displays the gene-module association -$\log_{10}$(p-value) for genes annotated in the following Gene panel. The eQTL GWAS P panel highlights the eQTL SNPs used for predicting *CCDC91* expression, with BP coordinates on the x-axis and GWAS -$\log_{10}$(p-value) on the y-axis. The bottom GWAS panel displays the GWAS -$\log_{10}$(p-value) for all SNPs located within 1 MB on either side of *CCDC91*.

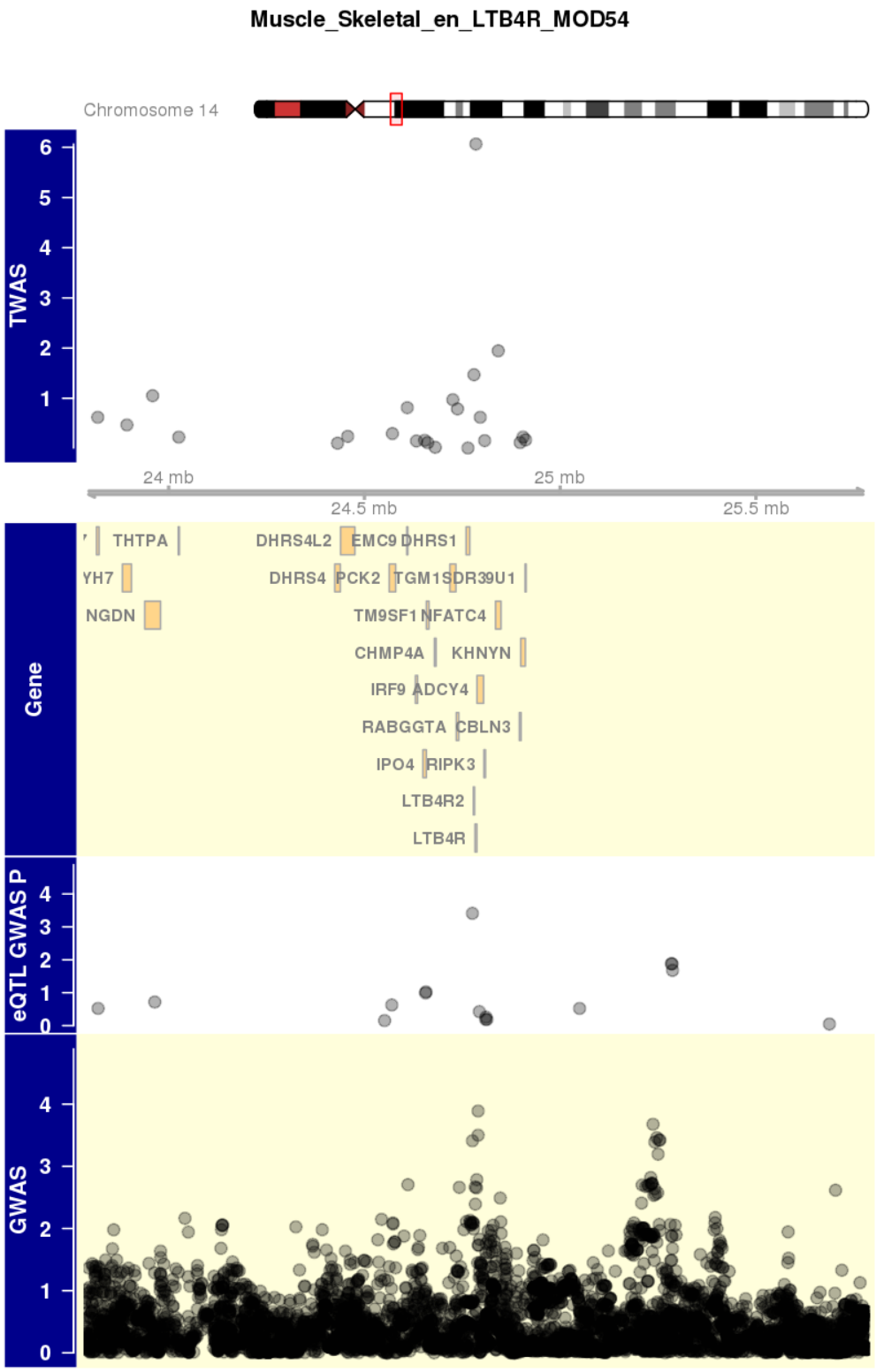


Figure C2: TWAS regional plot of *LTB4R*. TWAS of PITT module 54 (mandible) in Muscle Skeletal using the EN model identified *LTB4R*. Figure layout same as above.

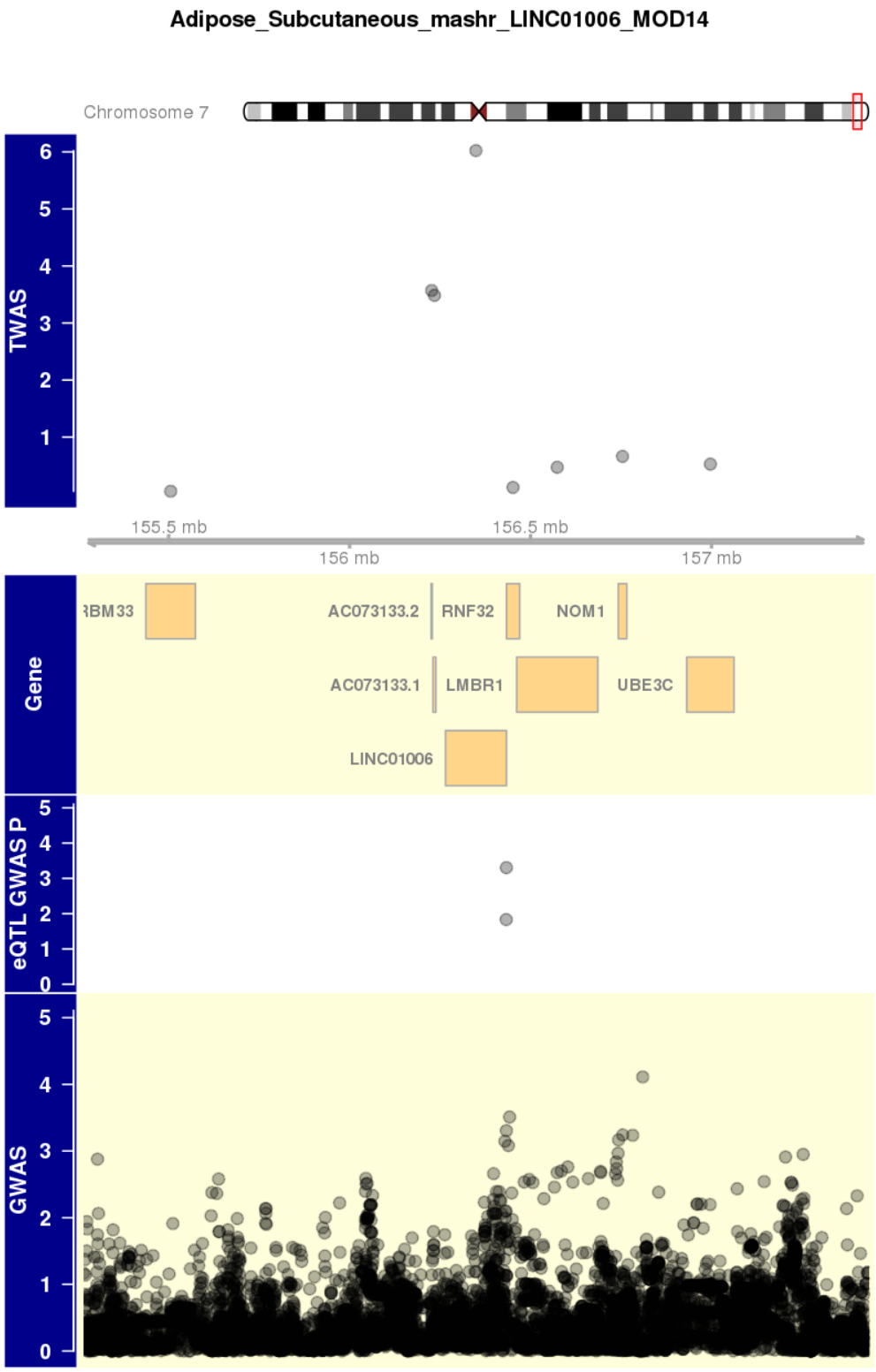


Figure C3: TWAS regional plot of *LINC01006*. TWAS of PITT module 14 (eye area) in Adipose Subcutaneous using the MASHR model identified *LINC01006*. Figure layout same as above.

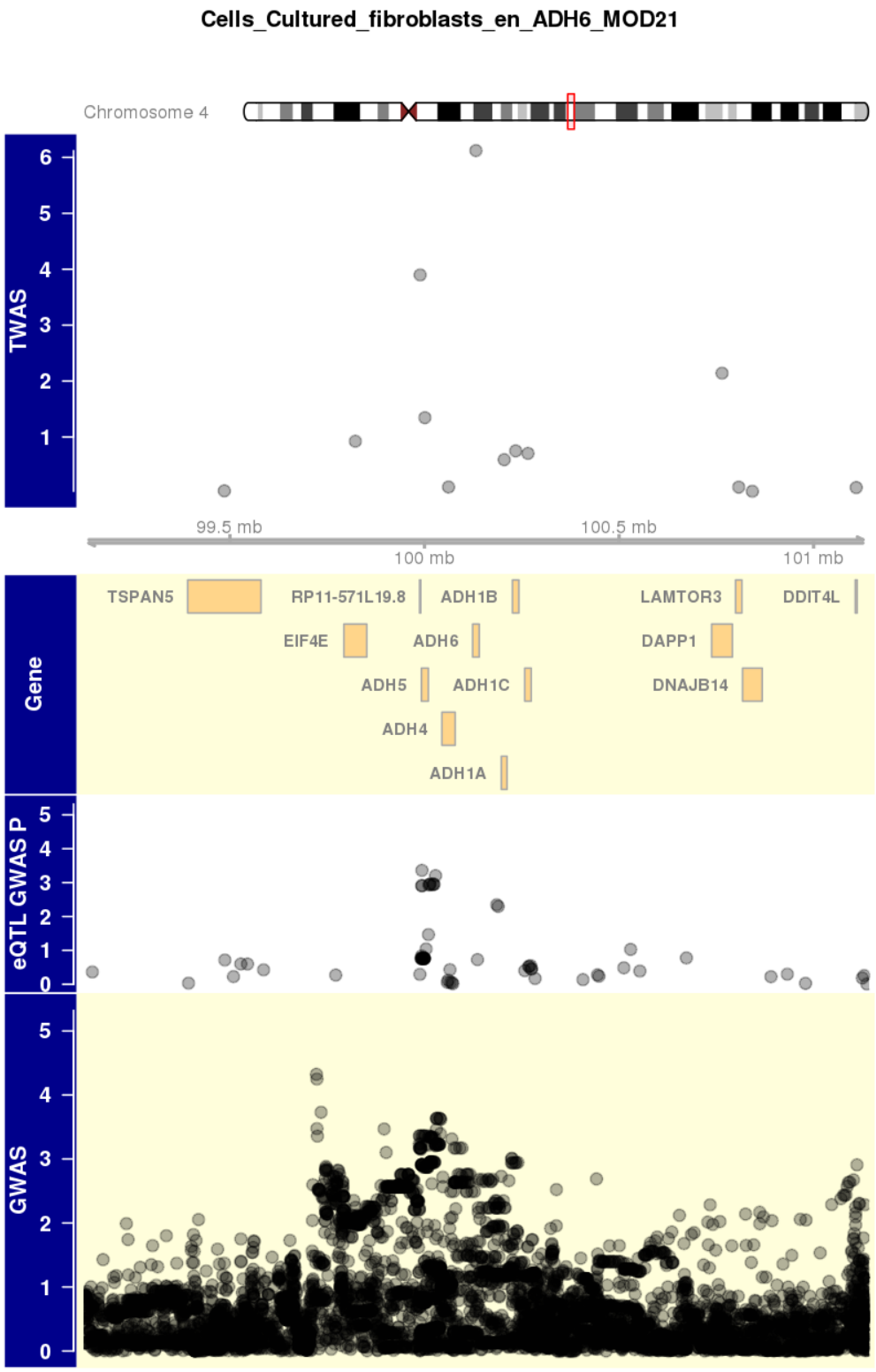


Figure C4: TWAS regional plot of *ADH6*. TWAS of PITT module 21 (nose bridge) in fibroblasts using the EN model identified *ADH6*. Figure layout same as above.

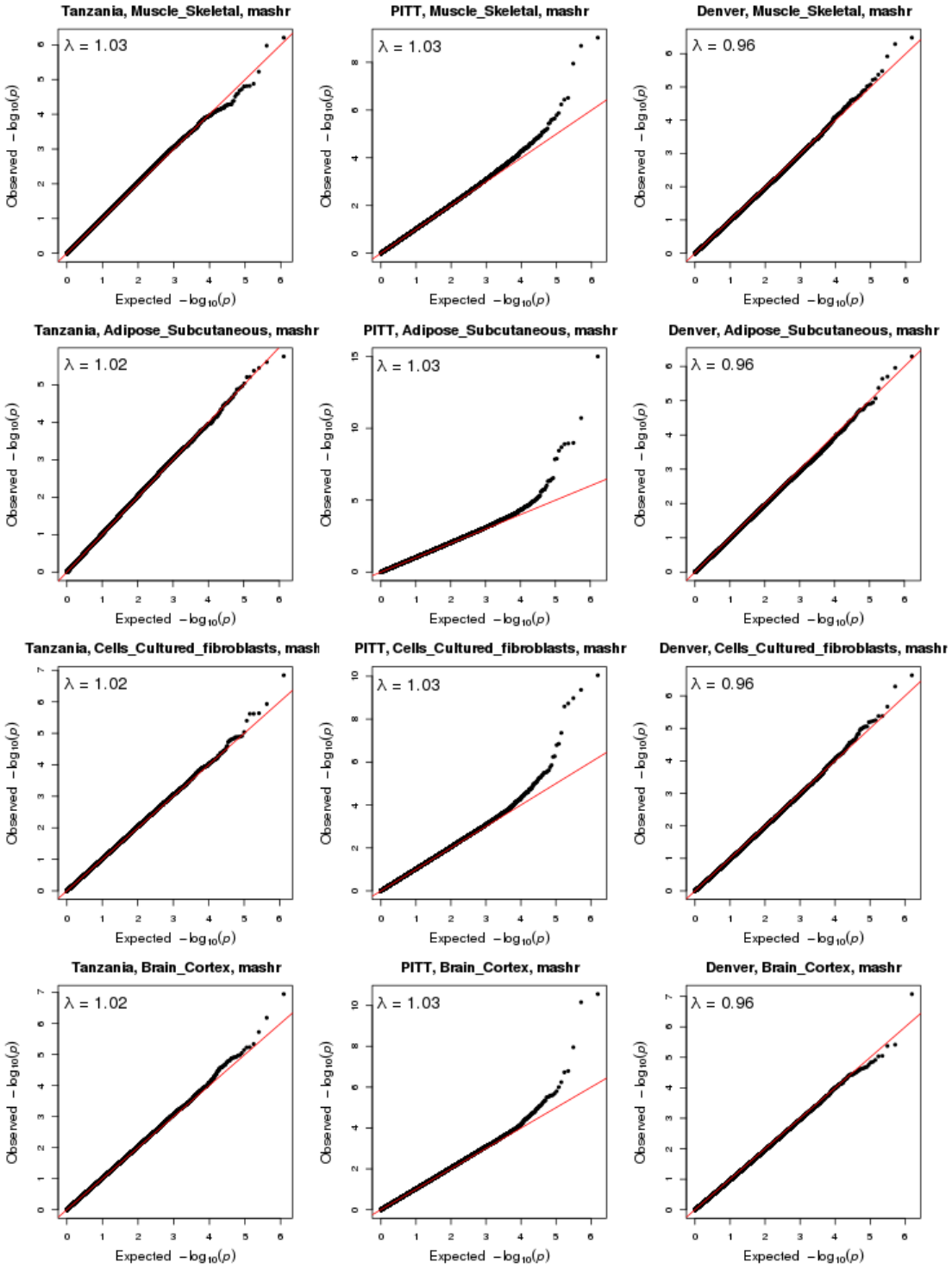


Figure C5: Q-Q plots for TWAS using the MASHR model. For simplicity, each figure contains p-values for analyzing all 63 facial modules. Figure titles indicate the cohort, tissue and type of prediction model used. Genomic inflation factor λ is shown in the top left corner of each sub figure.

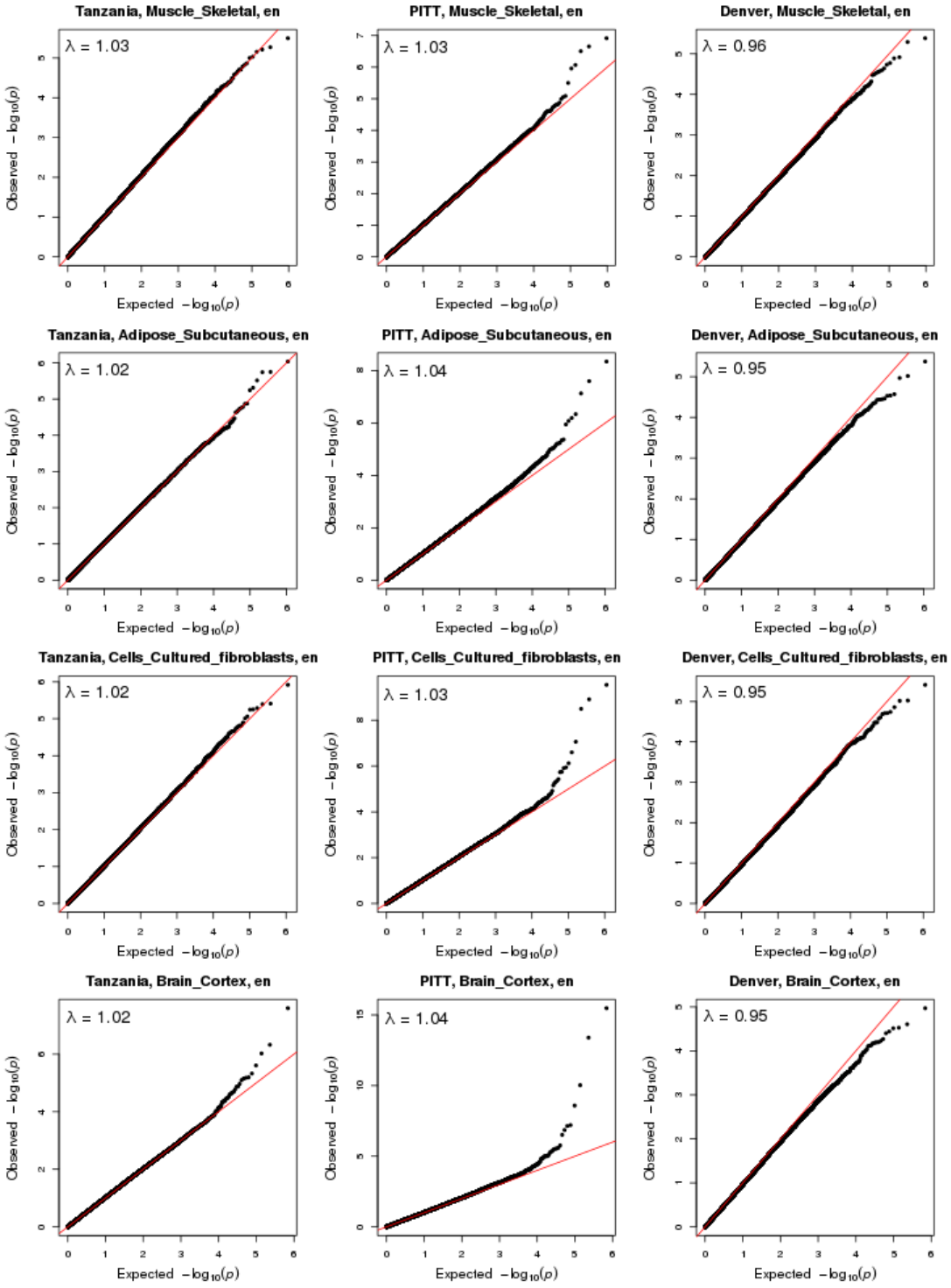


Figure C6: Q-Q plots for TWAS using the EN model. For simplicity, each figure contains p-values for analyzing all 63 facial modules. Figure titles indicate the cohort, tissue and type of prediction model used. Genomic inflation factor λ is shown in the top left corner of each sub figure.

Table C2: COLOC results at the 3q21.3 and the 1p12 locus

Locus	Module	Gene	Tissue	PP.H0+H1+H2[a]	PP.H3[b]	PP.H4[c]	Met colocalization criteria[d]
3q21.3	33	*EEFSEC*	Brain_Cortex	1.14E-03	0.34	0.66	Y
	33	*SEC61A1*	Muscle_Skeletal	1.30E-02	0.39	0.60	Y
	33	*EEFSEC*	Adipose_Subcutaneous	3.99E-07	0.45	0.55	Y
	33	*EEFSEC*	Cells_Cultured_fibroblasts	6.18E-06	0.55	0.45	N
	16	*EEFSEC*	Brain_Cortex	2.12E-03	0.63	0.37	N
	16	*EEFSEC*	Adipose_Subcutaneous	6.65E-07	0.74	0.26	N
	16	*EEFSEC*	Cells_Cultured_fibroblasts	9.15E-06	0.81	0.19	N
	16	*SEC61A1*	Muscle_Skeletal	2.66E-02	0.80	0.18	N
	33	*SEC61A1*	Brain_Cortex	0.59	0.25	0.17	N
	16	*RUVBL1*	Brain_Cortex	0.70	0.18	0.13	N
	33	*RUVBL1*	Brain_Cortex	0.72	0.18	0.10	N
	16	*SEC61A1*	Brain_Cortex	0.64	0.27	8.90E-02	N
	16	*RPN1*	Brain_Cortex	0.33	0.63	4.56E-02	N
	33	*RUVBL1*	Muscle_Skeletal	0.80	0.16	4.02E-02	N
	33	*RPN1*	Brain_Cortex	0.33	0.63	3.59E-02	N
	16	*RUVBL1*	Adipose_Subcutaneous	0.76	0.21	3.50E-02	N
	33	*RUVBL1*	Adipose_Subcutaneous	0.76	0.21	3.30E-02	N
	16	*RUVBL1*	Muscle_Skeletal	0.80	0.16	3.24E-02	N
	33	*SEC61A1*	Adipose_Subcutaneous	0.31	0.68	1.23E-02	N
	16	*SEC61A1*	Adipose_Subcutaneous	0.31	0.68	1.23E-02	N

Table C2: COLOC results at the 3q21.3 and the 1p12 locus (cont.)

Locus	Module	Gene	Tissue	PP.H0+H1+H2[a]	PP.H3[b]	PP.H4[c]	Met colocalization criteria[d]
3q21.3	33	*RPN1*	Muscle_Skeletal	2.50E-03	1.00	9.70E-05	N
	16	*RPN1*	Muscle_Skeletal	2.50E-03	1.00	9.69E-05	N
	33	*SEC61A1*	Cells_Cultured_fibroblasts	1.32E-09	1.00	3.15E-05	N
	16	*RUVBL1*	Cells_Cultured_fibroblasts	4.95E-07	1.00	3.01E-05	N
	33	*RUVBL1*	Cells_Cultured_fibroblasts	4.96E-07	1.00	9.25E-06	N
	16	*SEC61A1*	Cells_Cultured_fibroblasts	7.36E-12	1.00	5.56E-06	N
	33	*EEFSEC*	Muscle_Skeletal	1.73E-09	1.00	2.51E-06	N
	16	*EEFSEC*	Muscle_Skeletal	4.19E-10	1.00	7.04E-07	N
	33	*RPN1*	Cells_Cultured_fibroblasts	2.14E-09	1.00	7.79E-10	N
	16	*RPN1*	Cells_Cultured_fibroblasts	8.18E-10	1.00	2.75E-10	N
	33	*RPN1*	Adipose_Subcutaneous	1.33E-09	1.00	3.36E-11	N
	16	*RPN1*	Adipose_Subcutaneous	7.52E-12	1.00	1.98E-13	N
1p12	7	*HAO2*	Adipose_Subcutaneous	9.83E-04	0.47	0.53	Y
	7	*WARS2*	Brain_Cortex	4.08E-08	1.00	7.26E-04	N
	7	*WARS2*	Muscle_Skeletal	4.09E-08	1.00	6.10E-06	N
	7	*WARS2*	Adipose_Subcutaneous	4.09E-08	1.00	4.23E-06	N
	7	*WARS2*	Cells_Cultured_fibroblasts	4.09E-08	1.00	4.04E-06	N

[a] A small sum of the PP of H0, H1 and H2 indicates a high power for the colocalization analysis

[b] Posterior probability of different causal variants

[c] Posterior probability of shared causal variant

[d] Criteria include (1) PP.H4 > 0.5, (2) PP.H3 < 0.5, and (3) PP.H0+H1+H2 < 0.3. Y=yes, N=no

Table C3: Genes yielding stronger gene-level association in TWAS than SNP-level association in GWAS

TWAS gene-level association				GWAS SNP-level association		
Gene	Module	P-value	Model	eQTL SNP	eQTL BP	P-value
LTB4R	54	8.60E-07	EN	rs17090828	23819813	0.300
			Muscle Skeletal	rs2180852	23964516	0.192
				rs2281703	24551567	0.705
				rs2759407	24569947	0.236
				rs2295978	24656251	0.102
				rs2295977	24657226	0.095
				rs2281472	24775846	3.9E-04
				rs6573722	24793647	0.376
				rs3759625	24809683	0.648
				rs3759630	24810413	0.550
				rs8014581	24812899	0.648
				rs7143637	25049603	0.301
				rs898766	25284005	0.013
				rs12890808	25285800	0.013
				rs8007588	25287011	0.021
				rs7143241	25687817	0.898
ADH6	21	7.60E-07	EN	rs6532729	99146436	0.433
			Fibroblasts	rs1045655	99393671	0.929
				rs4699349	99486957	0.193
				rs4699353	99509133	0.594
				rs17027853	99528088	0.252
				rs13148540	99545196	0.253
				rs10470954	99586508	0.375
				rs13145463	99771854	0.536
				rs2924584	99988208	0.510

Table C3: Continued from previous page

TWAS gene-level association				GWAS SNP-level association		
Gene	Module	P-value	Model	eQTL SNP	eQTL BP	P-value
				rs7662987	99991642	0.173
				rs7684986	99991676	0.173
				rs11547772	99992793	0.146
				rs1061187	99992853	0.173
				rs6827292	99992994	0.001
				rs1803037	99993151	0.173
				rs28730646	99993478	4.4E-04
				rs28730644	99993686	0.001
				rs17595424	99993869	0.173
				rs7683802	99995138	0.173
				rs4699699	99997179	0.173
				rs4699701	99998447	0.173
				rs13118409	99998736	0.173
				rs13145727	99998866	0.173
				rs13119035	99999130	0.173
				rs13146416	99999160	0.173
				rs13125919	99999848	0.173
				rs7683704	100004226	0.091
				rs17216887	100010273	0.034
				rs7667261	100011300	0.001
				rs7681427	100013959	0.001
				rs7687322	100014342	0.001
				rs5003497	100021242	0.001
				rs6822742	100023959	0.001
				rs6823388	100023999	0.001
				rs6819724	100028451	0.001
				rs10026860	100058770	0.875

Table C3: Continued from previous page

TWAS gene-level association				GWAS SNP-level association		
Gene	Module	P-value	Model	eQTL SNP	eQTL BP	P-value
				rs13148577	100060836	0.761
				rs1800759	100065509	0.372
				rs4148884	100066287	0.877
				rs12649136	100070816	0.887
				rs2156731	100072567	0.962
				rs9307238	100136182	0.187
				rs17028758	100185697	0.005
				rs17028770	100189327	0.005
				rs2298753	100257907	0.398
				rs3762896	100268131	0.298
				rs4147542	100268553	0.298
				rs11936869	100273173	0.284
				rs17586163	100274571	0.355
				rs11499823	100274749	0.350
				rs7661978	100284199	0.677
				rs1072626	100407267	0.729
				rs13126513	100444684	0.526
				rs7665289	100448167	0.580
				rs11734413	100512350	0.328
				rs13306568	100529777	0.094
				rs10516449	100553510	0.408
				rs2162385	100673132	0.167
				rs6813978	100890487	0.598
				rs3886172	100932401	0.503
				rs4699769	100979170	0.938
				rs6853633	101125207	0.653
				rs2866207	101129818	0.554

Table C3: Continued from previous page

TWAS gene-level association				GWAS SNP-level association		
Gene	Module	P-value	Model	eQTL SNP	eQTL BP	P-value
				rs6839368	101136908	0.976
LINC01006	14	9.49E-07	MASHR Adipose Subcutaneous	rs76141520	156432793	0.015
				rs7788200	156433243	5.0E-04
CCDC91	21	4.37E-08	MASHR Fibroblasts	rs10843100	28270034	2.5E-06
				rs34742695	28286355	0.218
				rs139966291	28300347	0.180
CCDC91	43	1.41E-07	MASHR Fibroblasts	rs10843100	28270034	4.0E-06
				rs34742695	28286355	0.108
				rs139966291	28300347	0.222

Appendix D

Supplementary materials for Chapter 6

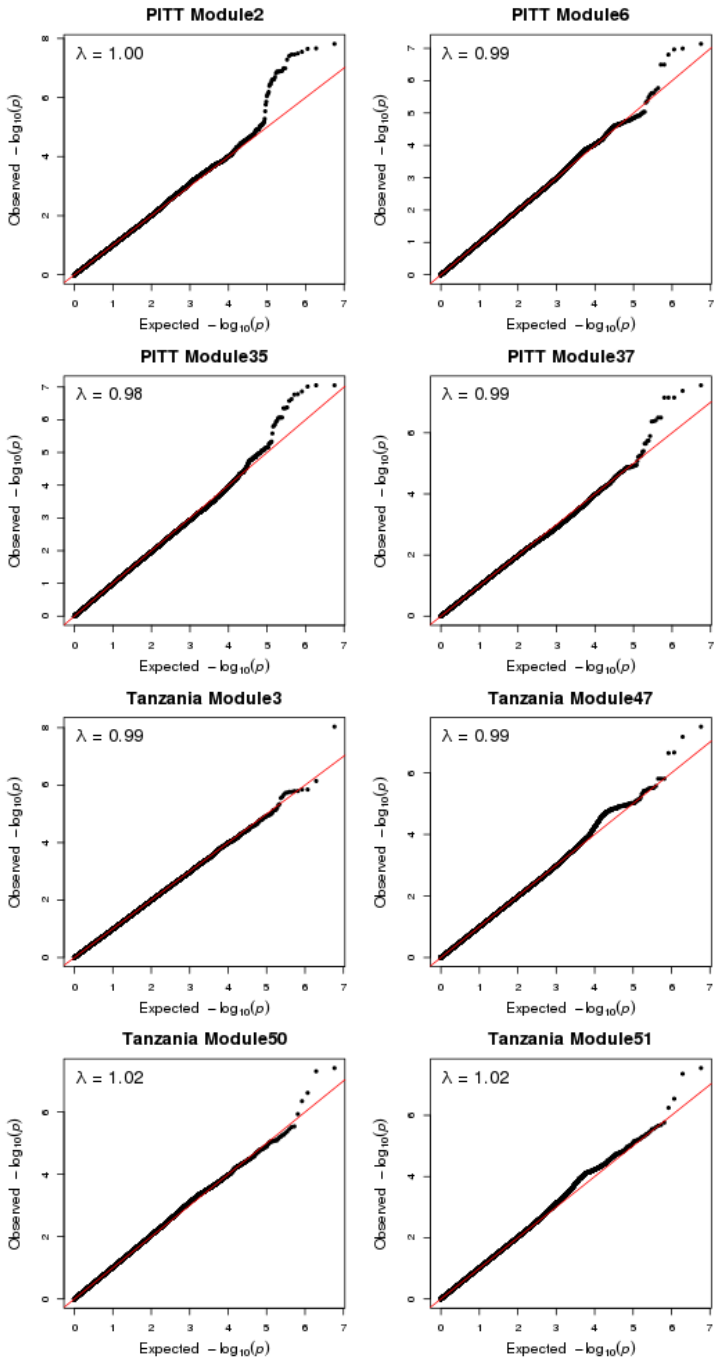
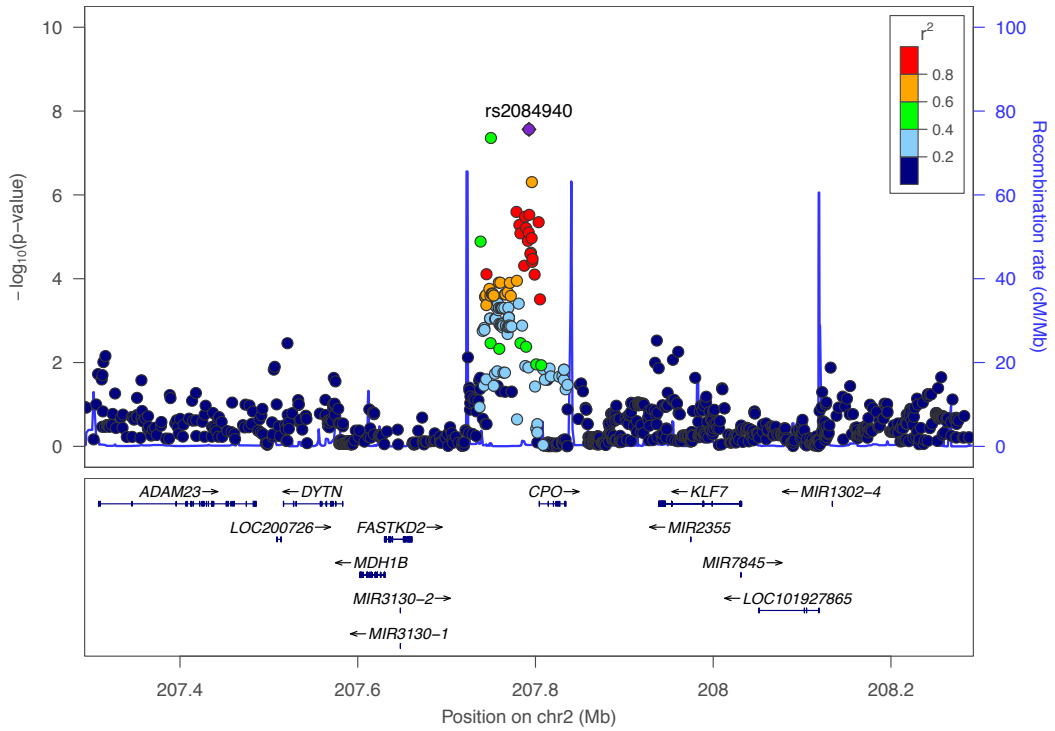
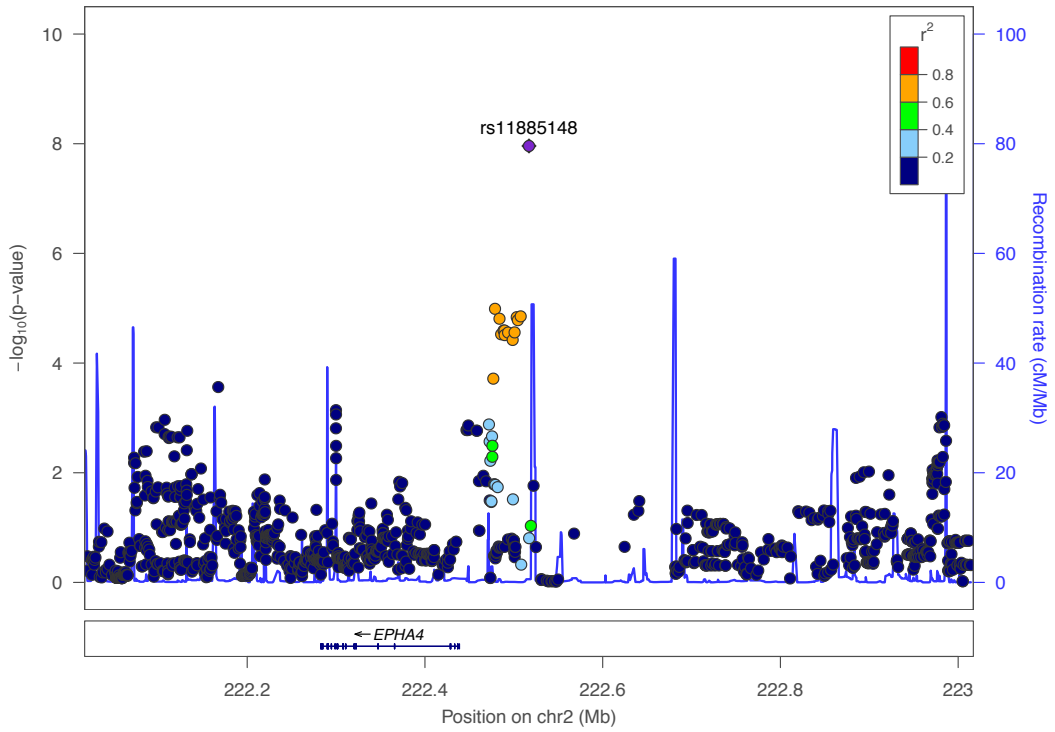


Figure D1: Q-Q plots for the vQTL analysis. Only facial modules where suggestive vQTLs were identified are shown for simplicity. See sub figure titles for cohort and facial module. Genomic inflation factor λ is shown in the top left corner of each sub figure. Statistics were well behaved for all analyses shown here.

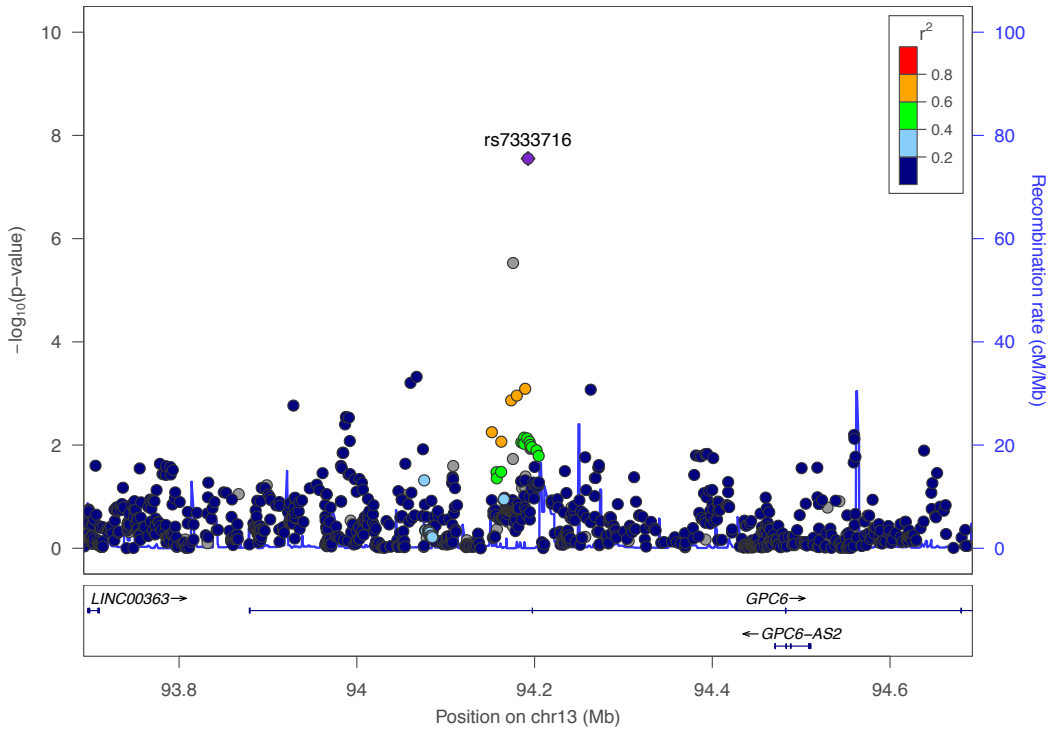


(a) The PITT vQTL rs405040 × 2q33.3 in module 37 (upper lip)



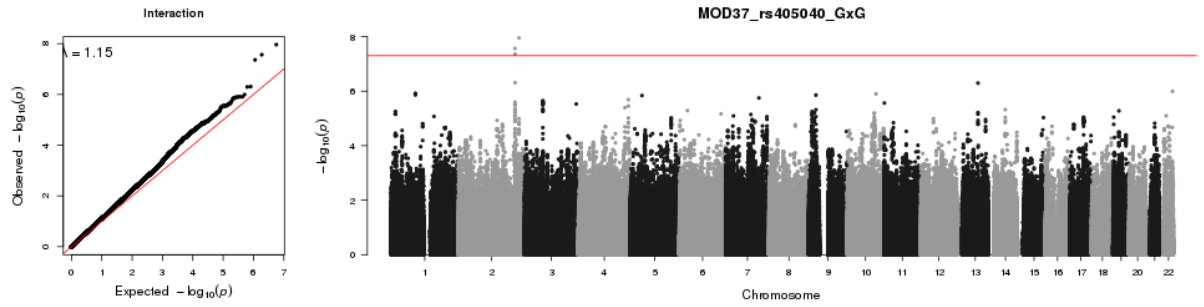
(b) The PITT vQTL rs405040 × 2q36.1 in module 37 (upper lip)

Figure D2: Locuszoom plot for suggestive G × G involving lead vQTL SNPs

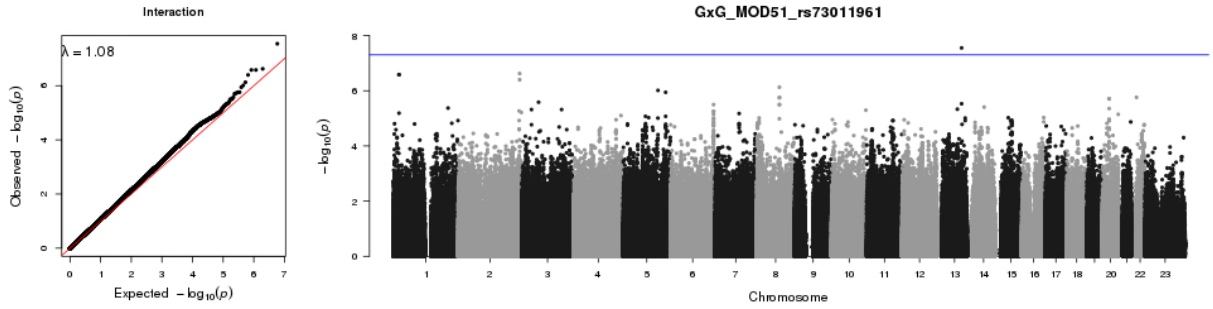


(c) The tanzania vQTL rs73011961 × 13q31.3 in module 51 (nose)

Figure D2: Locuszoom plot for suggestive G × G involving lead vQTL SNPs (cont.)



(a) The PITT vQTL rs405040 in module 37 (upper lip)



(b) The Tanzania vQTL rs73011961 in module 51 (nose)

Figure D3: Plots for the genome-wide search of G × G. The fixed interacting factors were the lead vQTL SNPs indicated in the sub figure titles. The red horizontal line in (a) and the blue horizontal line in (b) denote genome-wide significance level (5×10^{-8}).

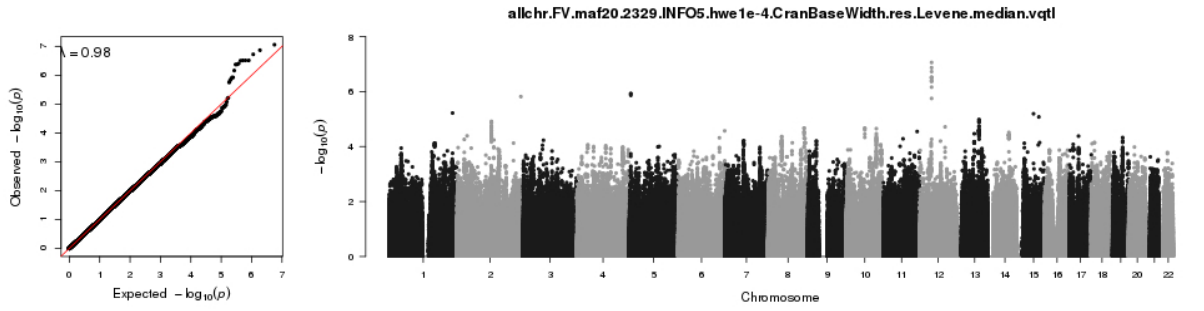


Figure D4: Q-Q and Manhattan plot for the vQTL test of cranial base width in the PITT cohort

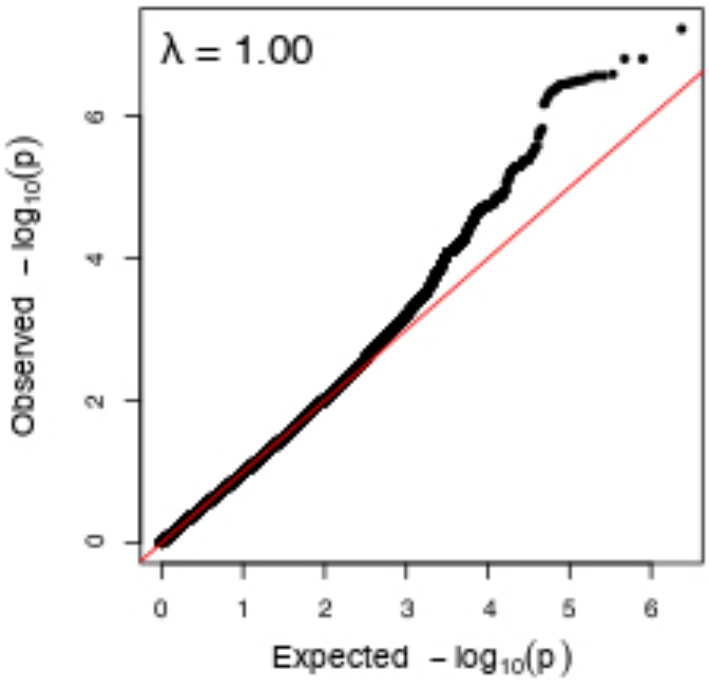


Figure D5: Q-Q plot for the vQTL test of zygion-to-zygion distance in the Korean replication cohort. We ran the test for genome-wide SNPs only to determine the minimum MAF requirement sufficient for removing systematic inflation. This plot shows the results from including SNPs with a MAF > 0.2, where the genomic inflation factor λ was satisfactory. In the subsequent replication analysis, only *PRICKLE1* SNPs above this MAF cutoff were considered.

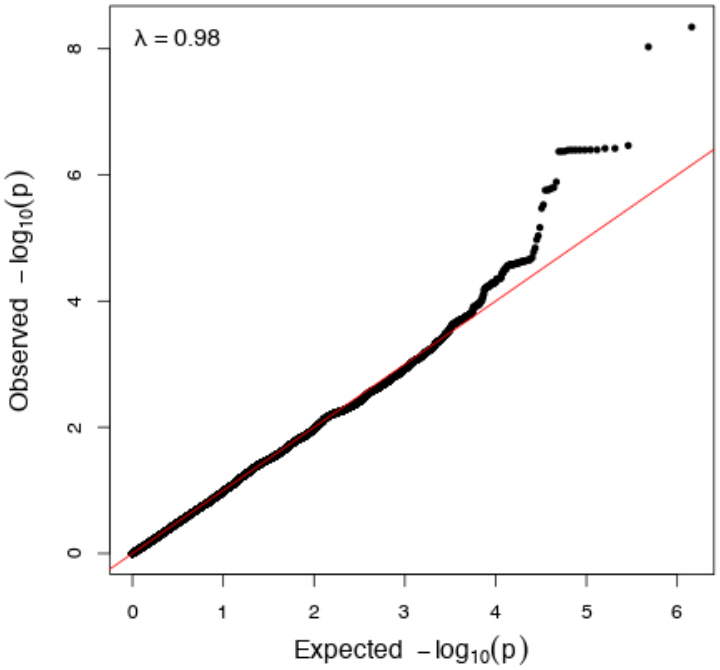


Figure D6: Q-Q plot for the G × G test of zygion-to-zygion distance in the replication cohort. SNPs with a MAF > 0.1 were included.

Table D1: Lookups of vQTLs in the PITT and the Tanzania cohorts

Discovery				Lookup best results	
Cohort	Loci	Module	SNP	P-value	Module
PITT	8q23.3	35	rs13278165	0.0200	29
		nasolabial	rs35954699	0.0233	29
			rs34447117	0.0202	29
	10p11.22	37	rs806812	0.0003	19
		upper lip	rs73477	0.0435	6
			rs211428	0.0320	21
			rs2370703	0.0299	21
			rs405040	0.0346	57
	10q25.3	2	rs4918871	0.0222	34
		nose	rs4918872	0.0199	34
		mouth	rs4918873	0.0199	34
			rs9783229	0.0199	34
			rs9783196	0.0219	34
			rs7098623	0.0170	57
			rs10787503	0.0306	3
			rs4917669	0.0164	34
			rs7899612	0.0121	59
	17q21.32	6	rs3744772	0.0452	14
		lower face	rs4793963	0.0310	14
			rs4793596	0.0464	14
Tanzania	4q13.1	3	rs935719	0.0051	13
		nose,eyes			
	13q21.22	50	rs9529870	0.0088	2
		nose	rs9542643	0.0093	2

Table D2: Top vQTL for each of the 24 univariate facial distances in the PITT cohort

Trait	Chr	BP	SNP	MAF	Beta	P-value	NMISS
LowFaceDepth	3	104414583	rs2947779	0.38	0.16	8.35E-08	2210
CranBaseWidth	12	42882153	rs1796391	0.25	0.19	8.76E-08	2200
NasalAlaLength	10	130048037	rs12256165	0.20	-0.19	1.00E-07	2300
CutLowLipHeight	20	9742096	rs4816172	0.21	0.19	1.07E-07	2305
LabFisWidth	13	59685916	rs9538303	0.30	0.17	1.83E-07	2296
NasalBdgLength	7	128158603	rs55638891	0.20	-0.19	1.89E-07	2272
MorphFaceHeight	3	185217189	rs2140287	0.46	-0.15	3.06E-07	2236
LowFaceHeight	10	73716143	rs4746102	0.30	0.16	5.37E-07	2259
NasalHeight	12	63691758	rs1146106	0.33	-0.16	6.06E-07	2320
LowLipHeight	21	37816927	rs11088349	0.25	0.17	6.20E-07	2285
UpLipHeight	1	34988954	rs10493058	0.49	0.14	8.76E-07	2318
OutCanthWidth	6	118655020	rs72967533	0.47	0.14	1.06E-06	2277
PhilLength	11	44893024	rs835760	0.23	0.17	1.09E-06	2255
UpFaceDepth	4	17170581	rs1522074	0.38	-0.15	1.19E-06	2226
LowVermHeight	1	208805883	rs2404677	0.21	0.17	1.25E-06	2308
UpVermHeight	1	42634101	rs6701382	0.23	0.17	1.30E-06	2212
NasalWidth	9	139929080	rs41317014	0.25	-0.16	1.48E-06	2299
SubNasalWidth	7	138358456	rs17160449	0.39	0.14	1.62E-06	2315
InCanthWidth	11	1380289	rs6421028	0.32	0.15	1.80E-06	2302
MidFaceDepth	13	55082227	rs6561782	0.24	0.17	2.50E-06	2164
PalpFisLength	1	187855391	rs147654748	0.27	-0.16	2.69E-06	2258
PhilWidth	4	114108991	rs4834321	0.49	0.14	2.95E-06	2318
NasalPro	5	42963864	rs11954543	0.47	0.13	4.55E-06	2304
UpFaceHeight	12	63691758	rs1146106	0.33	0.14	6.30E-06	2318

Appendix E

List of Abbreviations

3D - three-dimensional

ABF - approximate bayes factor

ANOVA - analysis of variance

CCA - canonical correlation analysis

CMC - combined multivariate and collapsing

CNCC - cranial neural crest cell

eQTL - expression quantitative trait locus

EN - elastic net

G $\times$ E - gene by environment interaction

G $\times$ G - gene by gene interaction

GO - gene ontology

GRex - genetically regulated gene expression

GTEx - gene-tissue expression

GWAS - genome-wide association study

LD - linkage disequilibrium

M_{eff} - effective number of tests

MAC - minor allele count

MAF - minor allele frequency

MANCOVA - multivariate analysis of covariance

MANOVA - multivariate analysis of variance

NSCL/P - nonsyndromic cleft lip with or without cleft palate

OFC - orofacial clefts

PC - principal component

PCA - principal component analysis

PP - posterior probability

QTL - quantitative trait locus

SKAT - sequence kernel association test

SNP - single nucleotide polymorphism

TIGAR - transcriptome-integrated genetic association resource

TWAS - transcriptome-wide association study

vQTL - variance quantitative trait locus

Bibliography

[1] Seth M Weinberg, Jasmien Roosenboom, John R Shaffer, Mark D Shriver, Joanna Wysocka, and Peter Claes. Hunting for genes that shape human faces: Initial successes and challenges for the future. *Orthodontics & craniofacial research*, 22 Suppl 1(S1):207–212, May 2019.

[2] Eirini Marouli, Mariaelisa Graff, Carolina Medina-Gomez, Ken Sin Lo, Andrew R Wood, Troels R Kjaer, Rebecca S Fine, Yingchang Lu, Claudia Schurmann, Heather M Highland, Sina Rüeger, Gudmar Thorleifsson, Anne E Justice, David Lamparter, Kathleen E Stirrups, Valérie Turcot, Kristin L Young, Thomas W Winkler, Tonu Esko, Tugce Karaderi, Adam E Locke, Nicholas G D Masca, Maggie C Y Ng, Poorva Mudgal, Manuel A Rivas, Sailaja Vedantam, Anubha Mahajan, Xiuqing Guo, Goncalo Abecasis, Katja K Aben, Linda S Adair, Dewan S Alam, Eva Albrecht, Kristine H Allin, Matthew Allison, Philippe Amouyel, Emil V Appel, Dominique Arveiler, Folkert W Asselbergs, Paul L Auer, Beverley Balkau, Bernhard Banas, Lia E Bang, Marianne Benn, Sven Bergmann, Lawrence F Bielak, Matthias Blüher, Heiner Boeing, Eric Boerwinkle, Carsten A Böger, Lori L Bonnycastle, Jette Bork-Jensen, Michiel L Bots, Erwin P Bottinger, Donald W Bowden, Ivan Brandslund, Gerome Breen, Murray H Brilliant, Linda Broer, Amber A Burt, Adam S Butterworth, David J Carey, Mark J Caulfield, John C Chambers, Daniel I Chasman, Yii-Der Ida Chen, Rajiv Chowdhury, Cramer Christensen, Audrey Y Chu, Massimiliano Cocca, Francis S Collins, James P Cook, Janie Corley, Jordi Corominas Galbany, Amanda J Cox, Gabriel Cuellar-Partida, John Danesh, Gail Davies, Paul I W de Bakker, Gert J de Borst, Simon de Denus, Mark C H de Groot, Renée de Mutsert, Ian J Deary, George Dedoussis, Ellen W Demerath, Anneke I den Hollander, Joe G Dennis, Emanuele Di Angelantonio, Fotios Drenos, Mengmeng Du, Alison M Dunning, Douglas F Easton, Tapani Ebeling, Todd L Edwards, Patrick T Ellinor, Paul Elliott, Evangelos Evangelou, Aliki-Eleni Farmaki, Jessica D Faul, Mary F Feitosa, Shuang Feng, Ele Ferrannini, Marco M Ferrario, Jean Ferrières, Jose C Florez, Ian Ford, Myriam Fornage, Paul W Franks, Ruth Frikke-Schmidt, Tessel E Galesloot, Wei Gan, Ilaria Gandin, Paolo Gasparini, Vilmantas Giedraitis, Ayush Giri, Giorgia Girotto, Scott D Gordon, Penny Gordon-Larsen, Mathias Gorski, Niels Grarup, Megan L Grove, Vilmundur Gudnason, Stefan Gustafsson, Torben Hansen, Kathleen Mullan Harris, Tamara B Harris, Andrew T Hattersley, Caroline Hayward, Liang He, Iris M Heid, Kauko Heikkilä, Øyvind Helgeland, Jussi Hernesniemi, Alex W Hewitt, Lynne J Hocking, Mette Hollensted, Oddgeir L Holmen, G Kees Hovingh, Joanna M M Howson, Carel B Hoyng, Paul L Huang, Kristian Hveem, M Arfan Ikram, Erik Ingelsson, Anne U Jackson, Jan-Håkan Jansson, Gail P Jarvik, Gorm B Jensen, Min A Jhun, Yucheng Jia, Xuejuan Jiang, Stefan Johansson, Marit E Jørgensen, Torben Jørgensen,

Pekka Jousilahti, J Wouter Jukema, Bratati Kahali, René S Kahn, Mika Kähönen, Pia R Kamstrup, Stavroula Kanoni, Jaakko Kaprio, Maria Karaleftheri, Sharon L R Kardia, Fredrik Karpe, Frank Kee, Renske Keeman, Lambertus A Kiemeney, Hidetoshi Kitajima, Kirsten B Kluivers, Thomas Kocher, Pirjo Komulainen, Jukka Kontto, Jaspal S Kooner, Charles Kooperberg, Peter Kovacs, Jennifer Kriebel, Helena Kuivaniemi, Sébastien Küry, Johanna Kuusisto, Martina La Bianca, Markku Laakso, Timo A Lakka, Ethan M Lange, Leslie A Lange, Carl D Langefeld, Claudia Langenberg, Eric B Larson, I-Te Lee, Terho Lehtimäki, Cora E Lewis, Huaixing Li, Jin Li, Ruifang Li-Gao, Honghuang Lin, Li-An Lin, Xu Lin, Lars Lind, Jaana Lindström, Allan Linneberg, Yeheng Liu, Yongmei Liu, Artitaya Lophatananon, Jian'an Luan, Steven A Lubitz, Leo-Pekka Lyytikäinen, David A Mackey, Pamela A F Madden, Alisa K Manning, Satu Männistö, Gaëlle Marenne, Jonathan Marten, Nicholas G Martin, Angela L Mazul, Karina Meidtner, Andres Metspalu, Paul Mitchell, Karen L Mohlke, Dennis O Mook-Kanamori, Anna Morgan, Andrew D Morris, Andrew P Morris, Martina Müller-Nurasyid, Patricia B Munroe, Mike A Nalls, and M... Nauck. Rare and low-frequency coding variants alter human adult height. *Nature*, 542(7640):186–190, February 2017.

[3] Xiangfeng Lu, Gina M Peloso, Dajiang J Liu, Ying Wu, He Zhang, Wei Zhou, Jun Li, Clara Sze-man Tang, Rajkumar Dorajoo, Huaixing Li, Jirong Long, Xiuqing Guo, Ming Xu, Cassandra N Spracklen, Yang Chen, Xuezhen Liu, Yan Zhang, Chiea Chuen Khor, Jianjun Liu, Liang Sun, Laiyuan Wang, Yu-Tang Gao, Yao Hu, Kuai Yu, Yiqin Wang, Chloe Yu Yan Cheung, Feijie Wang, Jianfeng Huang, Qiao Fan, Qiuyin Cai, Shufeng Chen, Jinxiu Shi, Xueli Yang, Wanting Zhao, Wayne H-H Sheu, Stacey Shawn Cherny, Meian He, Alan B Feranil, Linda S Adair, Penny Gordon-Larsen, Shufa Du, Rohit Varma, Yii-Der Ida Chen, Xiao-Ou Shu, Karen Siu Ling Lam, Tien-Yin Wong, Santhi K Ganesh, Zengnan Mo, Kristian Hveem, Lars G Fritsche, Jonas Bille Nielsen, Hung-fat Tse, Yong Huo, Ching-Yu Cheng, Y Eugene Chen, Wei Zheng, E Shyong Tai, Wei Gao, Xu Lin, Wei Huang, Goncalo Abecasis, Sekar Kathiresan, Karen L Mohlke, Tangchun Wu, Pak Chung Sham, Dongfeng Gu, and Cristen J Willer. Exome chip meta-analysis identifies novel loci and East Asian–specific coding variants that contribute to lipid levels and coronary artery disease. *Nature Genetics*, 49(12):1722–1730, October 2017.

[4] Steven Gazal, Po-Ru Loh, Hilary K Finucane, Andrea Ganna, Armin Schoech, Shamil Sunyaev, and Alkes L Price. Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations. *Nature Genetics*, 50(11):1600–1607, November 2018.

[5] Nicholas Mancuso, Simon Gayther, Alexander Gusev, Wei Zheng, Kathryn L Penney, Zsofia Kote-Jarai, Rosalind Eeles, Matthew Freedman, Christopher Haiman, Bogdan Pasaniuc, and PRACTICAL consortium. Large-scale transcriptome-wide association

study identifies new prostate cancer risk regions. *Nature Communications*, 9(1):4079–11, October 2018.

[6] Alexander Gusev, Nicholas Mancuso, Hyejung Won, Maria Kousi, Hilary K Finucane, Yakir Reshef, Lingyun Song, Alexias Safi, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Steven McCarroll, Benjamin M Neale, Roel A Ophoff, Michael C O'Donovan, Gregory E Crawford, Daniel H Geschwind, Nicholas Katsanis, Patrick F Sullivan, Bogdan Pasaniuc, and Alkes L Price. Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nature Genetics*, 50(4):538–548, April 2018.

[7] Farhad Hormozdiari, Martijn van de Bunt, Ayellet V Segrè, Xiao Li, Jong Wha J Joo, Michael Bilow, Jae Hoon Sul, Sriram Sankararaman, Bogdan Pasaniuc, and Eleazar Eskin. Colocalization of GWAS and eQTL Signals Detects Target Genes. *American journal of human genetics*, 99(6):1245–1260, December 2016.

[8] Huanwei Wang, Futao Zhang, Jian Zeng, Yang Wu, Kathryn E Kemper, Angli Xue, Min Zhang, Joseph E Powell, Michael E Goddard, Naomi R Wray, Peter M Visscher, Allan F McRae, and Jian Yang. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. *Science advances*, 5(8):eaaw3538, August 2019.

[9] Alexander I Young, Fabian L Wauthier, and Peter Donnelly. Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. *Nature Genetics*, 50(11):1608–1614, 2018.

[10] Dwight R Cordero, Samantha Brugmann, Yvonne Chu, Ruchi Bajpai, Maryam Jame, and Jill A Helms. Cranial neural crest cells on the move: their roles in craniofacial development. *American journal of medical genetics. Part A*, 155A(2):270–279, February 2011.

[11] Yang Chai and Robert E Maxson. Recent advances in craniofacial morphogenesis. *Developmental dynamics: an official publication of the American Association of Anatomists*, 235(9):2353–2375, September 2006.

[12] Sebastian Dworkin, Yeliz Boglev, Harley Owens, and Stephen J Goldie. The Role of Sonic Hedgehog in Craniofacial Patterning, Morphogenesis and Cranial Neural Crest Survival. *Journal of developmental biology*, 4(3), August 2016.

[13] José M Brito, Marie-Aimée Teillet, and Nicole M Le Douarin. Induction of mirror-image supernumerary jaws in chicken mandibular mesenchyme by Sonic Hedgehog-producing cells. *Development*, 135(13):2311–2319, July 2008.

[14] José M Brito, Marie-Aimée Teillet, and Nicole M Le Douarin. An early role for sonic hedgehog from foregut endoderm in jaw development: ensuring neural crest cell survival. *Proceedings of the National Academy of Sciences*, 103(31):11607–11612, August 2006.

[15] Amy E Merrill, B Frank Eames, Scott J Weston, Thayer Heath, and Richard A Schneider. Mesenchyme-dependent BMP signaling directs the timing of mandibular osteogenesis. *Development*, 135(7):1223–1234, April 2008.

[16] Xuguang Nie, Keijo Luukko, and Paivi Kettunen. BMP signalling in craniofacial development. *The International journal of developmental biology*, 50(6):511–521, 2006.

[17] Shahida Moosa and Bernd Wollnik. Altered FGF signalling in congenital craniofacial and skeletal disorders. *Seminars in cell & developmental biology*, 53:115–125, May 2016.

[18] X Nie, K Luukko, and P Kettunen. FGF signalling in craniofacial development and developmental disorders. *Oral Diseases*, 12(2):102–111, March 2006.

[19] Diane Hu, Nathan M Young, Xin Li, Yanhua Xu, Benedikt Hallgrimsson, and Ralph S Marcucio. A dynamic Shh expression pattern, regulated by SHH and BMP signaling, coordinates fusion of primordia in the amniote face. *Development*, 142(3):567–574, February 2015.

[20] Yu Lan and Rulang Jiang. Sonic hedgehog signaling regulates reciprocal epithelial-mesenchymal interactions controlling palatal outgrowth. *Development*, 136(8):1387–1396, April 2009.

[21] Joan E Hooper, Weiguo Feng, Hong Li, Sonia M Leach, Tzulip Phang, Charlotte Siska, Kenneth L Jones, Richard A Spritz, Lawrence E Hunter, and Trevor Williams. Systems biology of facial development: contributions of ectoderm and mesenchyme. *Developmental Biology*, 426(1):97–114, June 2017.

[22] Eric Van Otterloo, Trevor Williams, and Kristin Bruk Artinger. The old and new face of craniofacial research: How animal models inform human craniofacial genetic and clinical data. *Developmental Biology*, 415(2):171–187, July 2016.

[23] Trish E Parsons, Charlene M Downey, Frank R Jirik, Benedikt Hallgrimsson, and Heather A Jamniczky. Mind the gap: genetic manipulation of basicranial growth within synchondroses modulates calvarial and facial shape in mice through epigenetic interactions. *PLoS ONE*, 10(2):e0118355, 2015.

[24] Neus Martínez-Abadías, Philipp Mitteroecker, Trish E Parsons, Mireia Esparza, Torstein Sjøvold, Campbell Rolian, Joan T Richtsmeier, and Benedikt Hallgrimsson. The Developmental Basis of Quantitative Craniofacial Variation in Humans and Mice. *Evolutionary biology*, 39(4):554–567, December 2012.

[25] Fiorella Gurrieri, Brunella Franco, Helga Toriello, and Giovanni Neri. Oral-facial-digital syndromes: review and diagnostic guidelines. *American Journal of Medical Genetics Part A*, 143A(24):3314–3323, December 2007.

[26] Hanne Hoskens, Jiarui Li, Karlijne Indencleef, Dorothy Gors, Maarten H D Larmuseau, Stephen Richmond, Alexei I Zhurov, Greet Hens, Hilde Peeters, and Peter Claes. Spatially Dense 3D Facial Heritability and Modules of Co-heritability in a Father-Offspring Design. *Frontiers in genetics*, 9:554, 2018.

[27] Joanne B Cole, Mange Manyama, Jacinda R Larson, Denise K Liberton, Tracey M Ferrara, Sheri L Riccardi, Mao Li, Washington Mio, Ophir D Klein, Stephanie A Santorico, Benedikt Hallgrimsson, and Richard A Spritz. Human Facial Shape and Size Heritability and Genetic Correlations. *Genetics*, 205(2):967–978, February 2017.

[28] Jelena Djordjevic, Alexei I Zhurov, Stephen Richmond, and Visigen Consortium. Genetic and Environmental Contributions to Facial Morphological Variation: A 3D Population-Based Twin Study. *PLoS ONE*, 11(9):e0162250, 2016.

[29] Maja Šešelj, Dana L Duren, and Richard J Sherwood. Heritability of the Human Craniofacial Complex. *Anatomical record (Hoboken, N.J.: 2007)*, 298(9):1535–1547, September 2015.

[30] Seongwon Cha, Ji Eun Lim, Ah Yeon Park, Jun-Hyeong Do, Si Woo Lee, Chol Shin, Nam Han Cho, Ji-One Kang, Jeong Min Nam, Jong-Sik Kim, Kwang-Man Woo, Seung-Hwan Lee, Jong Yeol Kim, and Bermseok Oh. Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC GENOMICS*, 19(1):481, June 2018.

[31] Peter Claes, Jasmien Roosenboom, Julie D White, Tomek Swigut, Dzemila Sero, Jiarui Li, Myoung Keun Lee, Arslan Zaidi, Brooke C Mattern, Corey Liebowitz, Laurel

Pearson, Tomamp x000E1 s Gonzamp x000E1 lez, Elizabeth J Leslie, Jenna C Carlson, Ekaterina Orlova, Paul Suetens, Dirk Vandermeulen, Eleanor Feingold, Mary L Marazita, John R Shaffer, Joanna Wysocka, Mark D Shriver, and Seth M Weinberg. Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nature Genetics*, 50(3):1–16, February 2018.

[32] Daniel J M Crouch, Bruce Winney, Willem P Koppen, William J Christmas, Katarzyna Hutnik, Tammy Day, Devendra Meena, Abdelhamid Boumertit, Pirro Hysi, Ayrun Nessa, Tim D Spector, Josef Kittler, and Walter F Bodmer. Genetics of the human face: Identification of large-effect single gene variants. *Proceedings of the National Academy of Sciences of the United States of America*, 115(4):E676–E685, January 2018.

[33] Myoung Keun Lee, John R Shaffer, Elizabeth J Leslie, Ekaterina Orlova, Jenna C Carlson, Eleanor Feingold, Mary L Marazita, and Seth M Weinberg. Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. *PLoS ONE*, 12(4):e0176566–13, April 2017.

[34] John R Shaffer, Ekaterina Orlova, Myoung Keun Lee, Elizabeth J Leslie, Zachary D Raffensperger, Carrie L Heike, Michael L Cunningham, Jacqueline T Hecht, Chung How Kau, Nichole L Nidey, Lina M Moreno, George L Wehby, Jeffrey C Murray, Cecelia A Laurie, Cathy C Laurie, Joanne Cole, Tracey Ferrara, Stephanie Santorico, Ophir Klein, Washington Mio, Eleanor Feingold, Benedikt Hallgrimsson, Richard A Spritz, Mary L Marazita, and Seth M Weinberg. Genome-Wide Association Study Reveals Multiple Loci Influencing Normal Human Facial Morphology. *PLoS Genetics*, 12(8):e1006149–21, August 2016.

[35] Joanne B Cole, Mange Manyama, Emmanuel Kimwaga, Joshua Mathayo, Jacinda R Larson, Denise K Liberton, Ken Lukowiak, Tracey M Ferrara, Sheri L Riccardi, Mao Li, Washington Mio, Michaela Prochazkova, Trevor Williams, Hong Li, Kenneth L Jones, Ophir D Klein, Stephanie A Santorico, Benedikt Hallgrimsson, and Richard A Spritz. Genomewide Association Study of African Children Identifies Association of SCHIP1 and PDE8A with Facial Size and Shape. *PLoS Genetics*, 12(8):e1006174, August 2016.

[36] Kaustubh Adhikari, Macarena Fuentes-Guajardo, Mirsha Quinto-Sánchez, Javier Mendoza-Revilla, Juan Camilo Chacón-Duque, Victor Acuña-Alonzo, Claudia Jaramillo, William Arias, Rodrigo Barquera Lozano, Gastón Macín Pérez, Jorge Gómez-Valdés, Hugo Villamil-Ramírez, Tábita Hunemeier, Virginia Ramallo, Caio C Silva de Cerqueira, Malena Hurtado, Valeria Villegas, Vanessa Granja, Carla Gallo, Giovanni Poletti, Lavinia Schuler-Faccini, Francisco M Salzano, Maria-Cátira Bortolini, Samuel Canizales-Quinteros, Michael Cheeseman, Javier Rosique, Gabriel

Bedoya, Francisco Rothhammer, Denis Headon, Rolando González-José, David Balding, and Andrés Ruiz-Linares. A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature Communications*, 7(1):11616, May 2016.

[37] Fan Liu, Fedde van der Lijn, Claudia Schurmann, Gu Zhu, M Mallar Chakravarty, Pirro G Hysi, Andreas Wollstein, Oscar Lao, Marleen de Bruijne, M Arfan Ikram, Aad van der Lugt, Fernando Rivadeneira, Andre G Uitterlinden, Albert Hofman, Wiro J Niessen, Georg Homuth, Greig de Zubicaray, Katie L McMahon, Paul M Thompson, Amro Daboul, Ralf Puls, Katrin Hegenscheid, Liisa Bevan, Zdenka Pausova, Sarah E Medland, Grant W Montgomery, Margaret J Wright, Carol Wicking, Stefan Boehringer, Timothy D Spector, Tomáš Paus, Nicholas G Martin, Reiner Biffar, and Manfred Kayser. A Genome-Wide Association Study Identifies Five Loci Influencing Facial Morphology in Europeans. *PLoS Genetics*, 8(9):e1002932–13, September 2012.

[38] Lavinia Paternoster, Alexei I Zhurov, Arshed M Toma, John P Kemp, Beate St Pourcain, Nicholas J Timpson, George McMahon, Wendy McArdle, Susan M Ring, George Davey Smith, Stephen Richmond, and David M Evans. REPOR T Genome-wide Association Study of Three-Dimensional Facial Morphology Identifies a Variant in PAX3 Associated with Nasion Position. *The American Journal of Human Genetics*, 90(3):478–485, March 2012.

[39] Teri A Manolio, Francis S Collins, Nancy J Cox, David B Goldstein, Lucia A Hindorff, David J Hunter, Mark I McCarthy, Erin M Ramos, Lon R Cardon, Aravinda Chakravarti, Judy H Cho, Alan E Guttmacher, Augustine Kong, Leonid Kruglyak, Elaine Mardis, Charles N Rotimi, Montgomery Slatkin, David Valle, Alice S Whittemore, Michael Boehnke, Andrew G Clark, Evan E Eichler, Greg Gibson, Jonathan L Haines, Trudy F C Mackay, Steven A McCarroll, and Peter M Visscher. Finding the missing heritability of complex diseases. *Nature*, 461(7265):747–753, October 2009.

[40] Seongwon Cha, Ji Eun Lim, Ah Yeon Park, Jun-Hyeong Do, Si Woo Lee, Chol Shin, Nam Han Cho, Ji-One Kang, Jeong Min Nam, Jong-Sik Kim, Kwang-Man Woo, Seung-Hwan Lee, Jong Yeol Kim, and Bermseok Oh. Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC GENOMICS*, 19(1):481–17, June 2018.

[41] Alexander I Young. Solving the missing heritability problem. *PLoS Genetics*, 15(6):e1008222, June 2019.

[42] Lorenzo Bomba, Klaudia Walter, and Nicole Soranzo. The impact of rare and low-frequency genetic variants in common disease. *Genome biology*, 18(1):77, April 2017.

[43] Or Zuk, Stephen F Schaffner, Kaitlin Samocha, Ron Do, Eliana Hechter, Sekar Kathiresan, Mark J Daly, Benjamin M Neale, Shamil R Sunyaev, and Eric S Lander. Searching for missing heritability: designing rare variant association studies. *Proceedings of the National Academy of Sciences of the United States of America*, 111(4):E455–64, January 2014.

[44] Evan E Eichler, Jonathan Flint, Greg Gibson, Augustine Kong, Suzanne M Leal, Jason H Moore, and Joseph H Nadeau. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics*, 11(6):446–450, June 2010.

[45] David B Goldstein, Andrew Allen, Jonathan Keebler, Elliott H Margulies, Steven Petrou, Slavé Petrovski, and Shamil Sunyaev. Sequencing studies in human genetics: design and interpretation. *Nature Reviews Genetics*, 14(7):460–470, July 2013.

[46] Gregory V Kryukov, Len A Pennacchio, and Shamil R Sunyaev. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *The American Journal of Human Genetics*, 80(4):727–739, April 2007.

[47] J K Pritchard. Are rare variants responsible for susceptibility to complex diseases? *The American Journal of Human Genetics*, 69(1):124–137, July 2001.

[48] Elizabeth T Cirulli, Simon White, Robert W Read, Gai Elhanan, William J Metcalf, Francisco Tanudjaja, Donna M Fath, Efren Sandoval, Magnus Isaksson, Karen A Schlauch, Joseph J Grzymski, James T Lu, and Nicole L Washington. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. *Nature Communications*, 11(1):542–10, January 2020.

[49] Zhangchen Zhao, Wenjian Bi, Wei Zhou, Peter VandeHaar, Lars G Fritsche, and Seunggeun Lee. UK Biobank Whole-Exome Sequence Binary Phenome Analysis with Robust Region-Based Rare-Variant Test. *American journal of human genetics*, 106(1):3–12, January 2020.

[50] International Multiple Sclerosis Genetics Consortium. Electronic address: chris.cotsapas@yale.edu and International Multiple Sclerosis Genetics Consortium. Low-Frequency and Rare-Coding Variation Contributes to Multiple Sclerosis Risk. *Cell*, 175(6):1679–1687.e7, November 2018.

[51] Dajiang J Liu, Gina M Peloso, Haojie Yu, Adam S Butterworth, Xiao Wang, Anubha Mahajan, Danish Saleheen, Connor Emdin, Dewan Alam, Alexessander Couto Alves, Philippe Amouyel, Emanuele Di Angelantonio, Dominique Arveiler, Themistocles L

Assimes, Paul L Auer, Usman Baber, Christie M Ballantyne, Lia E Bang, Marianne Benn, Joshua C Bis, Michael Boehnke, Eric Boerwinkle, Jette Bork-Jensen, Erwin P Bottinger, Ivan Brandslund, Morris Brown, Fabio Busonero, Mark J Caulfield, John C Chambers, Daniel I Chasman, Y Eugene Chen, Yii-Der Ida Chen, Rajiv Chowdhury, Cramer Christensen, Audrey Y Chu, John M Connell, Francesco Cucca, L Adrienne Cupples, Scott M Damrauer, Gail Davies, Ian J Deary, George Dedoussis, Joshua C Denny, Anna Dominiczak, Marie-Pierre Dubé, Tapani Ebeling, Gudny Eiriksdottir, Tonu Esko, Aliki-Eleni Farmaki, Mary F Feitosa, Marco Ferrario, Jean Ferrières, Ian Ford, Myriam Fornage, Paul W Franks, Timothy M Frayling, Ruth Frikke-Schmidt, Lars G Fritsche, Philippe Frossard, Valentin Fuster, Santhi K Ganesh, Wei Gao, Melissa E Garcia, Christian Gieger, Franco Giulianini, Mark O Goodarzi, Harald Grallert, Niels Grarup, Leif Groop, Megan L Grove, Vilmundur Gudnason, Torben Hansen, Tamara B Harris, Caroline Hayward, Joel N Hirschhorn, Oddgeir L Holmen, Jennifer Huffman, Yong Huo, Kristian Hveem, Sehrish Jabeen, Anne U Jackson, Johanna Jakobsdottir, Marjo-Riitta Jarvelin, Gorm B Jensen, Marit E Jørgensen, J Wouter Jukema, Johanne M Justesen, Pia R Kamstrup, Stavroula Kanoni, Fredrik Karpe, Frank Kee, Amit V Khera, Derek Klarin, Heikki A Koistinen, Jaspal S Kooner, Charles Kooperberg, Kari Kuulasmaa, Johanna Kuusisto, Markku Laakso, Timo Lakka, Claudia Langenberg, Anne Langsted, Lenore J Launer, Torsten Lauritzen, David C M Liewald, Li-An Lin, Allan Linneberg, Ruth J F Loos, Yingchang Lu, Xiangfeng Lu, Reedik Mägi, Anders Malarstig, Ani Manichaikul, Alisa K Manning, Pekka Mäntyselkä, Eirini Marouli, Nicholas G D Masca, Andrea Maschio, James B Meigs, Olle Melander, Andres Metspalu, Andrew P Morris, Alanna C Morrison, Antonella Mulas, Martina Müller-Nurasyid, Patricia B Munroe, Matt J Neville, Jonas B Nielsen, Sune F Nielsen, Børge G Nordestgaard, Jose M Ordovas, Roxana Mehran, Christoper J O'Donnell, Marju Orho-Melander, Cliona M Molony, Pieter Muntendam, Sandosh Padmanabhan, Colin N A Palmer, Dorota Pasko, Aniruddh P Patel, Oluf Pedersen, Markus Perola, Annette Peters, Charlotta Pisinger, Giorgio Pistis, Ozren Polasek, Neil Poulter, Bruce M Psaty, Daniel J Rader, Asif Rasheed, Rainer Rauramaa, Dermot F Reilly, Alex P Reiner, Frida Renstrom, Stephen S Rich, Paul M Ridker, John D Rioux, Neil R Robertson, Dan M Roden, Jerome I Rotter, Igor Rudan, Veikko Salomaa, Nilesh J Samani, Serena Sanna, Naveed Sattar, Ellen M Schmidt, Robert A Scott, Peter Sever, Raquel S Sevilla, Christian M Shaffer, Xueling Sim, Suthesh Sivapalaratnam, Kerrin S Small, Albert V Smith, Blair H Smith, Sangeetha Somayajula, Lorraine Southam, Timothy D Spector, Elizabeth K Speliotes, John M Starr, Kathleen E Stirrups, Nathan Stitziel, Konstantin Strauch, Heather M Stringham, Praveen Surendran, Hayato Tada, Alan R Tall, Hua Tang, Jean-Claude Tardif, Kent D Taylor, Stella Trompet, Philip S Tsao, Jaakko Tuomilehto, Anne Tybjaerg-Hansen, Natalie R van Zuydam, Anette Varbo, Tibor V Varga, Jarmo Virtamo, Melanie Waldenberger, Nan Wang, Nick J Wareham, Helen R Warren, Peter E Weeke, Joshua Weinstock, Jennifer Wessel, James G Wilson, Peter W F Wilson, Ming Xu, Hanieh Yaghootkar, Robin Young, Eleftheria Zeggini, He Zhang, Neil S Zheng, Weihua Zhang, Yan Zhang, Wei Zhou, Yanhua Zhou, Magdalena Zoledziewska, Joanna M M Howson, John Danesh, Mark I McCarthy, Chad A Cowan, Goncalo Abecasis, Panos Deloukas, Kiran Musunuru, Cristen J Willer, and Sekar Kathiresan. Exome-

wide association study of plasma lipids in > 300,000 individuals. *Nature Genetics*, 49(12):1758–1766, October 2017.

[52] Chunyu Liu, Aldi T Kraja, Jennifer A Smith, Jennifer A Brody, Nora Franceschini, Joshua C Bis, Kenneth Rice, Alanna C Morrison, Yingchang Lu, Stefan Weiss, Xiuqing Guo, Walter Palmas, Lisa W Martin, Yii-Der Ida Chen, Praveen Surendran, Fotios Drenos, James P Cook, Paul L Auer, Audrey Y Chu, Ayush Giri, Wei Zhao, Johanna Jakobsdottir, Li-An Lin, Jeanette M Stafford, Najaf Amin, Hao Mei, Jie Yao, Arend Voorman, CHD Exome+ Consortium, ExomeBP Consortium, GoT2DGenes Consortium, T2D-Genes Consortium, Martin G Larson, Megan L Grove, Albert V Smith, Shih-Jen Hwang, Han Chen, Tianxiao Huan, Gulum Kosova, Nathan O Stitziel, Sekar Kathiresan, Nilesh Samani, Heribert Schunkert, Panos Deloukas, Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia, Man Li, Christian Fuchsberger, Cristian Pattaro, Mathias Gorski, CKDGen Consortium, Charles Kooperberg, George J Papanicolaou, Jacques E Rossouw, Jessica D Faul, Sharon L R Kardia, Claude Bouchard, Leslie J Raffel, Andre G Uitterlinden, Oscar H Franco, Ramachandran S Vasan, Christopher J O'Donnell, Kent D Taylor, Kiang Liu, Erwin P Bottinger, Omri Gottesman, E Warwick Daw, Franco Giulianini, Santhi Ganesh, Elias Salfati, Tamara B Harris, Lenore J Launer, Marcus Dörr, Stephan B Felix, Rainer Rettig, Henry Völzke, Eric Kim, Wen-Jane Lee, I-Te Lee, Wayne H-H Sheu, Krystal S Tsosie, Digna R Velez Edwards, Yongmei Liu, Adolfo Correa, David R Weir, Uwe Völker, Paul M Ridker, Eric Boerwinkle, Vilmundur Gudnason, Alexander P Reiner, Cornelia M van Duijn, Ingrid B Borecki, Todd L Edwards, Aravinda Chakravarti, Jerome I Rotter, Bruce M Psaty, Ruth J F Loos, Myriam Fornage, Georg B Ehret, Christopher Newton-Cheh, Daniel Levy, and Daniel I Chasman. Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nature Genetics*, 48(10):1162–1170, October 2016.

[53] Gundula Povysil, Slavé Petrovski, Joseph Hostyk, Vimla Aggarwal, Andrew S Allen, and David B Goldstein. Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nature Reviews Genetics*, 20(12):747–759, December 2019.

[54] Dajiang J Liu, Gina M Peloso, Xiaowei Zhan, Oddgeir L Holmen, Matthew Zawistowski, Shuang Feng, Majid Nikpay, Paul L Auer, Anuj Goel, He Zhang, Ulrike Peters, Martin Farrall, Marju Orho-Melander, Charles Kooperberg, Ruth McPherson, Hugh Watkins, Cristen J Willer, Kristian Hveem, Olle Melander, Sekar Kathiresan, and Goncalo R Abecasis. Meta-analysis of gene-level tests for rare variant association. *Nature Genetics*, 46(2):200–204, February 2014.

[55] Shuang Feng, Dajiang Liu, Xiaowei Zhan, Mary Kate Wing, and Goncalo R Abecasis. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics*, 30(19):2828–2829, October 2014.

[56] Seunggeun Lee, Tanya M Teslovich, Michael Boehnke, and Xihong Lin. General framework for meta-analysis of rare variants in sequencing association studies. *American journal of human genetics*, 93(1):42–53, July 2013.

[57] Paul L Auer and Guillaume Lettre. Rare variant association studies: considerations, challenges and opportunities. *Genome medicine*, 7(1):16, 2015.

[58] Lee, Seunggeung, Abecasis, Goncalo R, Boehnke, Michael, and Lin, Xihong. Rare-Variant Association Analysis: Study Designs and Statistical Tests. *The American Journal of Human Genetics*, 95(1):5–23, July 2014.

[59] Stephan Morgenthaler and William G Thilly. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutation research*, 615(1-2):28–56, February 2007.

[60] Bingshan Li and Suzanne M Leal. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *American journal of human genetics*, 83(3):311–321, September 2008.

[61] Bo Eskerod Madsen and Sharon R Browning. A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genetics*, 5(2):e1000384, February 2009.

[62] Fang Han and Wei Pan. A data-adaptive sum test for disease association with multiple common or rare variants. *Human Heredity*, 70(1):42–54, 2010.

[63] Alkes L Price, Gregory V Kryukov, Paul I W de Bakker, Shaun M Purcell, Jeff Staples, Lee-Jen Wei, and Shamil R Sunyaev. Pooled association tests for rare variants in exon-resequencing studies. *American journal of human genetics*, 86(6):832–838, June 2010.

[64] Michael C Wu, Seunggeun Lee, Tianxi Cai, Yun Li, Michael Boehnke, and Xihong Lin. Rare-variant association testing for sequencing data with the sequence kernel association test. *American journal of human genetics*, 89(1):82–93, July 2011.

[65] Wei Pan. Asymptotic tests of association with multiple SNPs in linkage disequilibrium. *Genetic epidemiology*, 33(6):497–507, September 2009.

[66] Benjamin M Neale, Manuel A Rivas, Benjamin F Voight, David Altshuler, Bernie Devlin, Marju Orho-Melander, Sekar Kathiresan, Shaun M Purcell, Kathryn Roeder,

and Mark J Daly. Testing for an unusual distribution of rare variants. *PLoS Genetics*, 7(3):e1001322, March 2011.

[67] Seunggeun Lee, Michael C Wu, and Xihong Lin. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics*, 13(4):762–775, September 2012.

[68] Jianping Sun, Yingye Zheng, and Li Hsu. A unified mixed-effects model for rare-variant association in sequencing studies. *Genetic epidemiology*, 37(4):334–344, May 2013.

[69] Ingrid E Christophersen, Michiel Rienstra, Carolina Roselli, Xiaoyan Yin, Bastiaan Geelhoed, John Barnard, Honghuang Lin, Dan E Arking, Albert V Smith, Christine M Albert, Mark Chaffin, Nathan R Tucker, Molong Li, Derek Klarin, Nathan A Bihlmeyer, Siew-Kee Low, Peter E Weeke, Martina Müller-Nurasyid, J Gustav Smith, Jennifer A Brody, Maartje N Niemeijer, Marcus Dörr, Stella Trompet, Jennifer Huffman, Stefan Gustafsson, Claudia Schurmann, Marcus E Kleber, Leo-Pekka Lyytikäinen, Ilkka Seppälä, Rainer Malik, Andrea R V R Horimoto, Marco Perez, Juha Sinisalo, Stefanie Aeschbacher, Sébastien Thériault, Jie Yao, Farid Radmanesh, Stefan Weiss, Alexander Teumer, Seung Hoan Choi, Lu-Chen Weng, Sebastian Clauss, Rajat Deo, Daniel J Rader, Svati H Shah, Albert Sun, Jemma C Hopewell, Stephanie Debette, Ganesh Chauhan, Qiong Yang, Bradford B Worrall, Guillaume Paré, Yoichiro Kamatani, Yanick P Hagemeijer, Niek Verweij, Joylene E Siland, Michiaki Kubo, Jonathan D Smith, David R Van Wagoner, Joshua C Bis, Siegfried Perz, Bruce M Psaty, Paul M Ridker, Jared W Magnani, Tamara B Harris, Lenore J Launer, M Benjamin Shoemaker, Sandosh Padmanabhan, Jeffrey Haessler, Traci M Bartz, Melanie Waldenberger, Peter Lichtner, Marina Arendt, Jose E Krieger, Mika Kähönen, Lorenz Risch, Alfredo J Mansur, Annette Peters, Blair H Smith, Lars Lind, Stuart A Scott, Yingchang Lu, Erwin B Bottinger, Jussi Hernesniemi, Cecilia M Lindgren, Jorge A Wong, Jie Huang, Markku Eskola, Andrew P Morris, Ian Ford, Alex P Reiner, Graciela Delgado, Lin Y Chen, Yii-Der Ida Chen, Roopinder K Sandhu, Man Li, Eric Boerwinkle, Lewin Eisele, Lars Lannfelt, Natalia Rost, Christopher D Anderson, Kent D Taylor, Archie Campbell, Patrik K Magnusson, David Porteous, Lynne J Hocking, Efthymia Vlachopoulou, Nancy L Pedersen, Kjell Nikus, Marju Orho-Melander, Anders Hamsten, Jan Heeringa, Joshua C Denny, Jennifer Kriebel, Dawood Darbar, Christopher Newton-Cheh, Christian Shaffer, Peter W Macfarlane, Stefanie Heilmann-Heimbach, Peter Almgren, Paul L Huang, Nona Sotoodehnia, Elsayed Z Soliman, Andre G Uitterlinden, Albert Hofman, Oscar H Franco, Uwe Völker, Karl-Heinz Jöckel, Moritz F Sinner, Henry J Lin, Xiuqing Guo, METASTROKE Consortium of the ISGC, Neurology Working Group of the CHARGE Consortium, Martin Dichgans, Erik Ingelsson, Charles Kooperberg, Olle Melander, Ruth J F Loos, Jari Laurikka, David Conen, Jonathan Rosand, Pim van der Harst, Marja-Liisa Lokki, Sekar Kathiresan, Alexandre Pereira, J Wouter Jukema, Caroline Hayward, Jerome I Rotter, Winfried März, Terho Lehtimäki, Bruno H Stricker, Mina K

Chung, Stephan B Felix, Vilmundur Gudnason, Alvaro Alonso, Dan M Roden, Stefan Kääb, Daniel I Chasman, Susan R Heckbert, Emelia J Benjamin, Toshihiro Tanaka, Kathryn L Lunetta, Steven A Lubitz, Patrick T Ellinor, and AFGen Consortium. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nature Genetics*, 49(6):946–952, June 2017.

[70] Richard D Bagnall, Douglas E Crompton, Slavé Petrovski, Lien Lam, Carina Cutmore, Sarah I Garry, Lynette G Sadleir, Leanne M Dibbens, Anita Cairns, Sara Kivity, Zaid Afawi, Brigid M Regan, Johan Duflou, Samuel F Berkovic, Ingrid E Scheffer, and Christopher Semsarian. Exome-based analysis of cardiac arrhythmia, respiratory control, and epilepsy genes in sudden unexpected death in epilepsy. *Annals of neurology*, 79(4):522–534, April 2016.

[71] Maren E Cannon and Karen L Mohlke. Deciphering the Emerging Complexities of Molecular Mechanisms at GWAS Loci. *American journal of human genetics*, 103(5):637–653, November 2018.

[72] Eddie Cano-Gamez and Gosia Trynka. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Frontiers in genetics*, 11:424, 2020.

[73] GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45(6):580–585, June 2013.

[74] ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414):57–74, September 2012.

[75] Bradley E Bernstein, John A Stamatoyannopoulos, Joseph F Costello, Bing Ren, Aleksandar Milosavljevic, Alexander Meissner, Manolis Kellis, Marco A Marra, Arthur L Beaudet, Joseph R Ecker, Peggy J Farnham, Martin Hirst, Eric S Lander, Tarjei S Mikkelsen, and James A Thomson. The NIH Roadmap Epigenomics Mapping Consortium. *Nature Biotechnology*, 28(10):1045–1048, October 2010.

[76] Michael Wainberg, Nasa Sinnott-Armstrong, Nicholas Mancuso, Alvaro N Barbeira, David A Knowles, David Golan, Raili Ermel, Arno Ruusalepp, Thomas Quertermous, Ke Hao, Johan L M Björkegren, Hae Kyung Im, Bogdan Pasaniuc, Manuel A Rivas, and Anshul Kundaje. Opportunities and challenges for transcriptome-wide association studies. *Nature Genetics*, 51(4):592–599, April 2019.

[77] Xin Qi, Sen Wang, Lu Zhang, Li Liu, Yan Wen, Mei Ma, Shiqiang Cheng, Ping Li, Bolun Cheng, Yanan Du, Xiao Liang, Yan Zhao, Miao Ding, and Feng Zhang. An

integrative analysis of transcriptome-wide association study and mRNA expression profile identified candidate genes for attention-deficit/hyperactivity disorder. *Psychiatry research*, 282:112639, December 2019.

[78] Molly Went, Ben Kinnersley, Amit Sud, David C Johnson, Niels Weinhold, Asta Försti, Mark van Duin, Giulia Orlando, Jonathan S Mitchell, Rowan Kuiper, Brian A Walker, Walter M Gregory, Per Hoffmann, Graham H Jackson, Markus M Nöthen, Miguel Inacio da Silva Filho, Hauke Thomsen, Annemiek Broyl, Faith E Davies, Unnur Thorsteinsdottir, Markus Hansson, Martin Kaiser, Pieter Sonneveld, Hartmut Goldschmidt, Kari Stefansson, Kari Hemminki, Björn Nilsson, Gareth J Morgan, and Richard S Houlston. Transcriptome-wide association study of multiple myeloma identifies candidate susceptibility genes. *Human Genomics*, 13(1):37–8, August 2019.

[79] Loïc Yengo, Julia Sidorenko, Kathryn E Kemper, Zhili Zheng, Andrew R Wood, Michael N Weedon, Timothy M Frayling, Joel Hirschhorn, Jian Yang, Peter M Visscher, and GIANT Consortium. Meta-analysis of genome-wide association studies for height and body mass index in 700000 individuals of European ancestry. *Human Molecular Genetics*, 27(20):3641–3649, October 2018.

[80] Alvaro N Barbeira, Rodrigo Bonazzola, Eric R Gamazon, Yanyu Liang, YoSon Park, Sarah Kim-Hellmuth, Gao Wang, Zhuoxun Jiang, Dan Zhou, Farhad Hormozdiari, Boxiang Liu, Abhiram Rao, Andrew R Hamel, Milton D Pividori, François Aguet, GTEx GWAS Working Group, Lisa Bastarache, Daniel M Jordan, Marie Verbanck, Ron Do, GTEx Consortium, Matthew Stephens, Kristin Ardlie, Mark McCarthy, Stephen B Montgomery, Ayellet V Segrè, Christopher D Brown, Tuuli Lappalainen, Xiaoquan Wen, and Hae Kyung Im. Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *bioRxiv*, 42(D1):814350, May 2020.

[81] Eric R Gamazon, Heather E Wheeler, Kaanan P Shah, Sahar V Mozaffari, Keston Aquino-Michaels, Robert J Carroll, Anne E Eyler, Joshua C Denny, GTEx Consortium, Dan L Nicolae, Nancy J Cox, and Hae Kyung Im. A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics*, 47(9):1091–1098, September 2015.

[82] Alvaro N Barbeira, Scott P Dickinson, Rodrigo Bonazzola, Jiamao Zheng, Heather E Wheeler, Jason M Torres, Eric S Torstenson, Kaanan P Shah, Tzintzuni Garcia, Todd L Edwards, Eli A Stahl, Laura M Huckins, GTEx Consortium, Dan L Nicolae, Nancy J Cox, and Hae Kyung Im. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications*, 9(1):1825, May 2018.

[83] Alexander Gusev, Arthur Ko, Huwenbo Shi, Gaurav Bhatia, Wonil Chung, Brenda W J H Penninx, Rick Jansen, Eco J C de Geus, Dorret I Boomsma, Fred A Wright, Patrick F Sullivan, Elina Nikkola, Marcus Alvarez, Mete Civelek, Aldons J Lusis, Terho Lehtimäki, Emma Raitoharju, Mika Kähönen, Ilkka Seppälä, Olli T Raitakari, Johanna Kuusisto, Markku Laakso, Alkes L Price, Paivi Pajukanta, and Bogdan Pasaniuc. Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics*, 48(3):245–252, February 2016.

[84] Sini Nagpal, Xiaoran Meng, Michael P Epstein, Lam C Tsoi, Matthew Patrick, Greg Gibson, Philip L De Jager, David A Bennett, Aliza P Wingo, Thomas S Wingo, and Jingjing Yang. TIGAR: An Improved Bayesian Tool for Transcriptomic Data Imputation Enhances Gene Mapping of Complex Traits. *American journal of human genetics*, 105(2):258–266, August 2019.

[85] Justin M Luningham, Junyu Chen, Shizhen Tang, Philip L De Jager, David A Bennett, Aron S Buchman, and Jingjing Yang. Bayesian Genome-wide TWAS method to leverage both cis- and trans- eQTL information through summary statistics. *bioRxiv*, 6:e1000888–33, August 2020.

[86] Ek, Weronica E, Rask-Andersen, Mathias, Karlsson, Torgny, Enroth, Stefan, Gyllensten, Ulf, and Johansson, Åsa. Genetic variants influencing phenotypic variance heterogeneity. *Human Molecular Genetics*, 27(5):799–810, March 2018.

[87] Waddington CH. Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811):563–565, November 1942.

[88] Julien F Ayroles, Sean M Buchanan, Chelsea O'Leary, Kyobi Skutt-Kakaria, Jennifer K Grenier, Andrew G Clark, Daniel L Hartl, and Benjamin L de Bivort. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proceedings of the National Academy of Sciences of the United States of America*, 112(21):6706–6711, May 2015.

[89] Xia Shen, Mats Pettersson, Lars Rönnegård, and Örjan Carlborg. Inheritance Beyond Plain Heritability: Variance-Controlling Genes in Arabidopsis thaliana. *PLoS Genetics*, 8(8):e1002839, August 2012.

[90] P Christe, A P Moller, N Saino, and F De Lope. Genetic and environmental components of phenotypic variation in immune response and body size of a colonial bird, Delichon urbica (the house martin). *Heredity*, 85 (Pt 1)(1):75–83, July 2000.

[91] Massimo Pigliucci. Evolution of phenotypic plasticity: where are we going now? *Trends in ecology & evolution*, 20(9):481–486, September 2005.

[92] G Gibson and G Wagner. Canalization in evolutionary genetics: a stabilizing theory? *BioEssays : news and reviews in molecular, cellular and developmental biology*, 22(4):372–380, April 2000.

[93] Christine Queitsch, Todd A Sangster, and Susan Lindquist. Hsp90 as a capacitor of phenotypic variation. *Nature*, 417(6889):618–624, June 2002.

[94] Hiroyuki Kuwahara and Orkun S Soyer. Bistability in feedback circuits as a byproduct of evolution of evolvability. *Molecular Systems Biology*, 8(1):564, January 2012.

[95] Joanna Masel and Mark L Siegal. Robustness: mechanisms and consequences. *Trends in genetics : TIG*, 25(9):395–403, September 2009.

[96] Hiroaki Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, November 2004.

[97] Richard E Lenski, Jeffrey E Barrick, and Charles Ofria. Balancing robustness and evolvability. *PLoS biology*, 4(12):e428, December 2006.

[98] Simon K G Forsberg and Örjan Carlborg. On the relationship between epistasis and genetic variance heterogeneity. *Journal of experimental botany*, 68(20):5431–5438, November 2017.

[99] Maksim V Struchalin, Abbas Dehghan, Jacqueline Cm Witteman, Cornelia van Duijn, and Yurii S Aulchenko. Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. *BMC Genetics*, 11(1):92–10, October 2010.

[100] Or Zuk, Eliana Hechter, Shamil R Sunyaev, and Eric S Lander. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4):1193–1198, January 2012.

[101] Guillaume Paré, Nancy R Cook, Paul M Ridker, and Daniel I Chasman. On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the Women's Genome Health Study. *PLoS Genetics*, 6(6):e1000981, June 2010.

[102] Jian Yang, Ruth J F Loos, Joseph E Powell, Sarah E Medland, Elizabeth K Speliotes, Daniel I Chasman, Lynda M Rose, Gudmar Thorleifsson, Valgerdur Steinthorsdottir, Reedik Mägi, Lindsay Waite, Albert Vernon Smith, Laura M Yerges-Armstrong, Keri L Monda, David Hadley, Anubha Mahajan, Guo Li, Karen Kapur, Veronique Vitart, Jennifer E Huffman, Sophie R Wang, Cameron Palmer, Tonu Esko, Krista Fischer, Jing Hua Zhao, Ayse Demirkan, Aaron Isaacs, Mary F Feitosa, Jian'an Luan, Nancy L Heard-Costa, Charles White, Anne U Jackson, Michael Preuss, Andreas Ziegler, Joel Eriksson, Zoltán Kutalik, Francesca Frau, Ilja M Nolte, Jana V Van Vliet-Ostaptchouk, Jouke Jan Hottenga, Kevin B Jacobs, Niek Verweij, Anuj Goel, Carolina Medina-Gomez, Karol Estrada, Jennifer Lynn Bragg-Gresham, Serena Sanna, Carlo Sidore, Jonathan Tyrer, Alexander Teumer, Inga Prokopenko, Massimo Mangino, Cecilia M Lindgren, Themistocles L Assimes, Alan R Shuldiner, Jennie Hui, John P Beilby, Wendy L McArdle, Per Hall, Talin Haritunians, Lina Zgaga, Ivana Kolcic, Ozren Polasek, Tatijana Zemunik, Ben A Oostra, M Juhani Junttila, Henrik Grönberg, Stefan Schreiber, Annette Peters, Andrew A Hicks, Jonathan Stephens, Nicola S Foad, Jaana Laitinen, Anneli Pouta, Marika Kaakinen, Gonneke Willemsen, Jacqueline M Vink, Sarah H Wild, Gerjan Navis, Folkert W Asselbergs, Georg Homuth, Ulrich John, Carlos Iribarren, Tamara Harris, Lenore Launer, Vilmundur Gudnason, Jeffrey R O'Connell, Eric Boerwinkle, Gemma Cadby, Lyle J Palmer, Alan L James, Arthur W Musk, Erik Ingelsson, Bruce M Psaty, Jacques S Beckmann, Gérard Waeber, Peter Vollenweider, Caroline Hayward, Alan F Wright, Igor Rudan, Leif C Groop, Andres Metspalu, Kay-Tee Khaw, Cornelia M van Duijn, Ingrid B Borecki, Michael A Province, Nicholas J Wareham, Jean-Claude Tardif, Heikki V Huikuri, L Adrienne Cupples, Larry D Atwood, Caroline S Fox, Michael Boehnke, Francis S Collins, Karen L Mohlke, Jeanette Erdmann, Heribert Schunkert, Christian Hengstenberg, Klaus Stark, Mattias Lorentzon, Claes Ohlsson, Daniele Cusi, Jan A Staessen, Melanie M Van der Klauw, Peter P Pramstaller, Sekar Kathiresan, Jennifer D Jolley, Samuli Ripatti, Marjo-Riitta Jarvelin, Eco J C de Geus, Dorret I Boomsma, Brenda Penninx, James F Wilson, Harry Campbell, Stephen J Chanock, Pim van der Harst, Anders Hamsten, Hugh Watkins, Albert Hofman, Jacqueline C Witteman, M Carola Zillikens, Andre G Uitterlinden, Fernando Rivadeneira, Lambertus A Kiemeney, Sita H Vermeulen, Goncalo R Abecasis, David Schlessinger, Sabine Schipf, Michael Stumvoll, Anke Tönjes, Tim D Spector, Kari E North, Guillaume Lettre, Mark I McCarthy, Sonja I Berndt, Andrew C Heath, Pamela A F Madden, Dale R Nyholt, Grant W Montgomery, Nicholas G Martin, Barbara McKnight, David P Strachan, William G Hill, Harold Snieder, Paul M Ridker, Unnur Thorsteinsdottir, Kari Stefansson, Timothy M Frayling, Joel N Hirschhorn, Michael E Goddard, and Peter M Visscher. FTO genotype is associated with phenotypic variability of body mass index. *Nature*, 490(7419):267–272, October 2012.

[103] Lars Rönnegård and William Valdar. Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genetics*, 13(1):63, July 2012.

[104] Abhishek K Sarkar, Po-Yuan Tung, John D Blischak, Jonathan E Burnett, Yang I Li, Matthew Stephens, and Yoav Gilad. Discovery and characterization of variance QTLs in human induced pluripotent stem cells. *PLoS genetics*, 15(4):e1008045, April 2019.

[105] Amanda M Hulse and James J Cai. Genetic variants contribute to gene expression variability in humans. *Genetics*, 193(1):95–108, January 2013.

[106] Jasmien Roosenboom, Karlijne Indencleef, Greet Hens, Hilde Peeters, Kaare Christensen, Mary L Marazita, Peter Claes, Elizabeth J Leslie, and Seth M Weinberg. Testing the face shape hypothesis in twins discordant for nonsyndromic orofacial clefting. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 173(11):2886–2892, September 2017.

[107] Stefan Boehringer, Fedde van der Lijn, Fan Liu, Manuel Günther, Stella Sinigerova, Stefanie Nowak, Kerstin U Ludwig, Ruth Herberz, Stefan Klein, Albert Hofman, Andre G Uitterlinden, Wiro J Niessen, Monique M B Breteler, Aad van der Lugt, Rolf P Würtz, Markus M Nöthen, Bernhard Horsthemke, Dagmar Wieczorek, Elisabeth Mangold, and Manfred Kayser. Genetic determination of human facial morphology: links between cleft-lips and normal variation. *European journal of human genetics : EJHG*, 19(11):1192–1197, November 2011.

[108] S M Weinberg, S D Naidoo, K M Bardi, C A Brandon, K Neiswanger, J M Resick, R A Martin, and M L Marazita. Face shape of unaffected parents with cleft affected offspring: combining three-dimensional surface imaging and geometric morphometrics. *Orthodontics & craniofacial research*, 12(4):271–281, November 2009.

[109] G L Wehby and C H Cassell. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral diseases*, 16(1):3–10, January 2010.

[110] Ronald P Strauss and Cynthia H Cassell. Critical issues in craniofacial care: quality of life, costs of care, and implications of prenatal diagnosis. *Academic pediatrics*, 9(6):427–432, November 2009.

[111] A H Jheon, S Oberoi, R C Solem, and S Kapila. Moving towards precision orthodontics: An evolving paradigm shift in the planning and delivery of customized orthodontic therapy. *Orthodontics & craniofacial research*, 20 Suppl 1:106–113, June 2017.

[112] Manfred Kayser. Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic science international. Genetics*, 18:33–48, September 2015.

[113] Jasmien Roosenboom, Greet Hens, Brooke C Mattern, Mark D Shriver, and Peter Claes. Exploring the Underlying Genetics of Craniofacial Morphology through Various Sources of Knowledge. *BioMed research international*, 2016:3054578, 2016.

[114] Cathy C Laurie, Kimberly F Doheny, Daniel B Mirel, Elizabeth W Pugh, Laura J Bierut, Tushar Bhangale, Frederick Boehm, Neil E Caporaso, Marilyn C Cornelis, Howard J Edenberg, Stacy B Gabriel, Emily L Harris, Frank B Hu, Kevin B Jacobs, Peter Kraft, Maria Teresa Landi, Thomas Lumley, Teri A Manolio, Caitlin McHugh, Ian Painter, Justin Paschall, John P Rice, Kenneth M Rice, Xiuwen Zheng, Bruce S Weir, and GENEVA Investigators. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic epidemiology*, 34(6):591–602, September 2010.

[115] Julie D White, Alejandra Ortega-Castrillón, Harold Matthews, Arslan A Zaidi, Omid Ekrami, Jonatan Snyders, Yi Fan, Tony Penington, Stefan Van Dongen, Mark D Shriver, and Peter Claes. MeshMonk: Open-source large-scale intensive 3D phenotyping. *Scientific reports*, 9(1):6085, April 2019.

[116] Diptavo Dutta, Laura Scott, Michael Boehnke, and Seunggeun Lee. Multi-SKAT: General framework to test for rare-variant association with multiple phenotypes. *Genetic epidemiology*, 43(1):4–23, February 2019.

[117] Diptavo Dutta, Sarah A Gagliano Taliun, Joshua S Weinstock, Matthew Zawistowski, Carlo Sidore, Lars G Fritsche, Francesco Cucca, David Schlessinger, Goncalo R Abecasis, Chad M Brummett, and Seunggeun Lee. Meta-MultiSKAT: Multiple phenotype meta-analysis for region-based association test. *Genetic epidemiology*, 43(7):800–814, October 2019.

[118] Auer, Paul L, Reiner, Alex P, and Leal, Suzanne M. The effect of phenotypic outliers and non-normality on rare-variant association testing. *European journal of human genetics : EJHG*, 24(8):1188–1194, August 2016.

[119] J Li and L Ji. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95(3):221–227, September 2005.

[120] Cory Y McLean, Dave Bristor, Michael Hiller, Shoa L Clarke, Bruce T Schaar, Craig B Lowe, Aaron M Wenger, and Gill Bejerano. GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, 28(5):495–501, May 2010.

[121] Kyoko Watanabe, Erdogan Taskesen, Arjen van Bochoven, and Danielle Posthuma. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8(1):1826–11, November 2017.

[122] Jing Chen, Eric E Bardes, Bruce J Aronow, and Anil G Jegga. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic acids research*, 37(Web Server issue):W305–11, July 2009.

[123] Paul F O'Reilly, Clive J Hoggart, Yotsawat Pomyen, Federico C F Calboli, Paul Elliott, Marjo-Riitta Jarvelin, and Lachlan J M Coin. MultiPhen: Joint Model of Multiple Phenotypes Can Increase Discovery in GWAS. *PLoS ONE*, 7(5):e34861–12, May 2012.

[124] Philipp Rentzsch, Daniela Witten, Gregory M Cooper, Jay Shendure, and Martin Kircher. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic acids research*, 47(D1):D886–D894, January 2019.

[125] Mihir A Kamat, James A Blackshaw, Robin Young, Praveen Surendran, Stephen Burgess, John Danesh, Adam S Butterworth, and James R Staley. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics*, 35(22):4851–4853, June 2019.

[126] Elizabeth J Leslie, Jenna C Carlson, John R Shaffer, Carmen J Buxó, Eduardo E Castilla, Kaare Christensen, Frederic W B Deleyiannis, Leigh L Field, Jacqueline T Hecht, Lina Moreno, Iêda M Orioli, Carmencita Padilla, Alexandre R Vieira, George L Wehby, Eleanor Feingold, Seth M Weinberg, Jeffrey C Murray, and Mary L Marazita. Association studies of low-frequency coding variants in nonsyndromic cleft lip with or without cleft palate. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 173(6):1531–1538, April 2017.

[127] Xiaowei Zhan, Youna Hu, Bingshan Li, Goncalo R Abecasis, and Dajiang J Liu. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics*, 32(9):1423–1426, May 2016.

[128] Lucia Conde, Eran Halperin, Nicholas K Akers, Kevin M Brown, Karin E Smedby, Nathaniel Rothman, Alexandra Nieters, Susan L Slager, Angela Brooks-Wilson, Luz

Agana, Jacques Riby, Jianjun Liu, Hans-Olov Adami, Hatef Darabi, Henrik Hjalgrim, Hui-Qi Low, Keith Humphreys, Mads Melbye, Ellen T Chang, Bengt Glimelius, Wendy Cozen, Scott Davis, Patricia Hartge, Lindsay M Morton, Maryjean Schenk, Sophia S Wang, Bruce Armstrong, Anne Kricker, Sam Milliken, Mark P Purdue, Claire M Vajdic, Peter Boyle, Qing Lan, Shelia H Zahm, Yawei Zhang, Tongzhang Zheng, Nikolaus Becker, Yolanda Benavente, Paolo Boffetta, Paul Brennan, Katja Butterbach, Pierluigi Cocco, Lenka Foretova, Marc Maynadié, Silvia de Sanjosé, Anthony Staines, John J Spinelli, Sara J Achenbach, Timothy G Call, Nicola J Camp, Martha Glenn, Neil E Caporaso, James R Cerhan, Julie M Cunningham, Lynn R Goldin, Curtis A Hanson, Neil E Kay, Mark C Lanasa, Jose F Leis, Gerald E Marti, Kari G Rabe, Laura Z Rassenti, Logan G Spector, Sara S Strom, Celine M Vachon, J Brice Weinberg, Elizabeth A Holly, Stephen Chanock, Martyn T Smith, Paige M Bracci, and Christine F Skibola. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nature Genetics*, 42(8):661–664, August 2010.

[129] Maria Chiara Di Bernardo, Dalemari Crowther-Swanepoel, Peter Broderick, Emily Webb, Gabrielle Sellick, Ruth Wild, Kate Sullivan, Jayaram Vijayakrishnan, Yufei Wang, Alan M Pittman, Nicola J Sunter, Andrew G Hall, Martin J S Dyer, Estella Matutes, Claire Dearden, Tryfonia Mainou-Fowler, Graham H Jackson, Geoffrey Summerfield, Robert J Harris, Andrew R Pettitt, Peter Hillmen, David J Allsup, James R Bailey, Guy Pratt, Chris Pepper, Chris Fegan, James M Allan, Daniel Catovsky, and Richard S Houlston. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nature Genetics*, 40(10):1204–1210, October 2008.

[130] Joseph R Avila, Peter A Jezewski, Alexandre R Vieira, Iêda M Orioli, Eduardo E Castilla, Kaare Christensen, Sandra Daack-Hirsch, Paul A Romitti, and Jeffrey C Murray. PVRL1 variants contribute to non-syndromic cleft lip and palate in multiple populations. *American Journal of Medical Genetics Part A*, 140(23):2562–2570, December 2006.

[131] M A Sozen, K Suzuki, M M Tolarova, T Bustos, J E Fernández Iglesias, and R A Spritz. Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. *Nature Genetics*, 29(2):141–142, October 2001.

[132] K Suzuki, D Hu, T Bustos, J Zlotogora, A Richieri-Costa, J A Helms, and R A Spritz. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. *Nature Genetics*, 25(4):427–430, August 2000.

[133] Zlotogora, J. Syndactyly, Ectodermal Dysplasia, and Cleft Lip/Palate. *Journal of Medical Genetics*, 31(12):957–959, December 1994.

[134] Martyn T Cobourne. The complex genetics of cleft lip and palate. *European journal of orthodontics*, 26(1):7–16, February 2004.

[135] Deniz Aslar Oner and Hakki Tastan. Identification of Novel Variants in the PVRL1 Gene in Patients With Nonsyndromic Cleft Lip With or Without Cleft Palate. *Genetic Testing and Molecular Biomarkers*, 20(5):269–272, May 2016.

[136] S Tongkobpetch, K Suphapeetiporn, P Siriwan, and V Shotelersuk. Study of the poliovirus receptor related-1 gene in Thai patients with non-syndromic cleft lip with or without cleft palate. *International Journal of Oral & Maxillofacial Surgery*, 37(6):550–553, June 2008.

[137] L Scapoli, A Palmieri, M Martinelli, C Vaccari, J Marchesini, F Pezzetti, U Baciliero, E Padula, P Carinci, and F Carinci. Study of the PVRL1 gene in Italian nonsyndromic cleft lip patients with or without cleft palate. *Annals of Human Genetics*, 70(Pt 3):410–413, May 2006.

[138] Michael C Campbell and Sarah A Tishkoff. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review of Genomics and Human Genetics*, 9(1):403–433, 2008.

[139] Julie D White, Karlijne Indencleef, Sahin Naqvi, Ryan J Eller, Jasmien Roosenboom, Myoung Keun Lee, Jiarui Li, Jaaved Mohammed, Stephen Richmond, Ellen E Quillen, Heather L Norton, Eleanor Feingold, Tomek Swigut, Mary L Marazita, Hilde Peeters, Greet Hens, John R Shaffer, Joanna Wysocka, Susan Walsh, Seth M Weinberg, Mark D Shriver, and Peter Claes. Insights into the genetic architecture of the human face. *Biorxiv*, May 2020.

[140] Alexandra C Nica, Stephen B Montgomery, Antigone S Dimas, Barbara E Stranger, Claude Beazley, Inês Barroso, and Emmanouil T Dermitzakis. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genetics*, 6(4):e1000895, April 2010.

[141] Xin He, Chris K Fuller, Yi Song, Qingying Meng, Bin Zhang, Xia Yang, and Hao Li. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *American journal of human genetics*, 92(5):667–680, May 2013.

[142] Claudia Giambartolomei, Damjan Vukcevic, Eric E Schadt, Lude Franke, Aroon D Hingorani, Chris Wallace, and Vincent Plagnol. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genetics*, 10(5):e1004383, May 2014.

[143] Xiaoquan Wen, Roger Pique-Regi, and Francesca Luca. Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLoS Genetics*, 13(3):e1006646, March 2017.

[144] Ignacio González, Sébastien Déjean, Pascal G P Martin, and Alain Baccini. CCA: An R Package to Extend Canonical Correlation Analysis. *Journal of Statistical Software*, 23(12):1–14, January 2008.

[145] Boxiang Liu, Michael J Gloudemans, Abhiram S Rao, Erik Ingelsson, and Stephen B Montgomery. Abundant associations with gene expression complicate GWAS follow-up. *Nature Genetics*, 51(5):768–769, May 2019.

[146] Ge Zhang, Bjarke Feenstra, Jonas Bacelis, Xueping Liu, Lisa M Muglia, Julius Juodakis, Daniel E Miller, Nadia Litterman, Pan-Pan Jiang, Laura Russell, David A Hinds, Youna Hu, Matthew T Weirauch, Xiaoting Chen, Arun R Chavan, Günter P Wagner, Mihaela Pavličev, Mauris C Nnamani, Jamie Maziarz, Minna K Karjalainen, Mika Rämet, Verena Sengpiel, Frank Geller, Heather A Boyd, Aarno Palotie, Allison Momany, Bruce Bedell, Kelli K Ryckman, Johanna M Huusko, Carmy R Forney, Leah C Kottyan, Mikko Hallman, Kari Teramo, Ellen A Nohr, George Davey Smith, Mads Melbye, Bo Jacobsson, and Louis J Muglia. Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. *New England Journal of Medicine*, 377(12):1156–1167, September 2017.

[147] Tulika Gupta, Kanchan Kapoor, Balbir Singh, and Anju Huria. Fetal orbital and ocular biometry at different gestational ages. *Journal of the Anatomical Society of India*, 62(2):157–165, December 2013.

[148] I Goldstein, A Tamir, Z Weiner, and P Jakobi. Dimensions of the fetal facial profile in normal pregnancy. *Ultrasound in Obstetrics and Gynecology*, 35(2):191–194, February 2010.

[149] Vyacheslav M Labunskyy, Dolph L Hatfield, and Vadim N Gladyshev. Selenoproteins: molecular pathways and physiological roles. *Physiological reviews*, 94(3):739–777, July 2014.

[150] Gonzalo A Mardones, Patricia V Burgos, Doug A Brooks, Emma Parkinson-Lawrence, Rafael Mattera, and Juan S Bonifacino. The trans-Golgi network accessory protein p56 promotes long-range movement of GGA/clathrin-containing transport carriers and lysosomal enzyme sorting. *Molecular biology of the cell*, 18(9):3486–3501, September 2007.

[151] Xavier Muñoz, Pilar Amiano, David Celorrio, Miren Dorronsoro, María-José Sánchez, José M Huerta, Aurelio Barricarte, Larraitz Arriola, Carmen Navarro, Esther Molina-Montes, M Dolores Chirlaque, Eva Ardanaz, Laudina Rodriguez, Eric J Duell, Elizabeth Hijona, Marta Herreros-Villanueva, Núria Sala, and Luis Bujanda. Association of alcohol dehydrogenase polymorphisms and life-style factors with excessive alcohol intake within the Spanish population (EPIC-Spain). *Addiction (Abingdon, England)*, 107(12):2117–2127, December 2012.

[152] Po-Hsiu Kuo, Gursharan Kalsi, Carol A Prescott, Colin A Hodgkinson, David Goldman, Edwin J van den Oord, Jeffry Alexander, Cizhong Jiang, Patrick F Sullivan, Diana G Patterson, Dermot Walsh, Kenneth S Kendler, and Brien P Riley. Association of ADH and ALDH genes with alcohol dependence in the Irish Affected Sib Pair Study of alcohol dependence (IASPSAD) sample. *Alcoholism, clinical and experimental research*, 32(5):785–795, May 2008.

[153] Lingjun Zuo, Kesheng Wang, Xiang-Yang Zhang, Xinghua Pan, Guilin Wang, Yunlong Tan, Chunlong Zhong, John H Krystal, Matthew State, Heping Zhang, and Xingguang Luo. Association between common alcohol dehydrogenase gene (ADH) variants and schizophrenia and autism. *Human Genetics*, 132(7):735–743, July 2013.

[154] Douglas W Yao, Luke J O'Connor, Alkes L Price, and Alexander Gusev. Quantifying genetic effects on disease mediated by assayed gene expression levels. *Nature Genetics*, 52(6):626–633, June 2020.

[155] S. M. Bartlett. Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond*, (A 160):113–126, 1937.

[156] H Levene. Robust tests for equality of variances. In *Contributions to probability and statistics Essays in honor of Harold Hotelling*, pages 278–292. Stanford University Press, 1960.

[157] M A Fligner and TJ Killeen. Distribution-free two-sample tests for scale. *Journal of the American Statistical Association*, 71(353):210–213, 1976.

[158] Lars Rönnegård, Majbritt Felleki, Freddy Fikse, Herman A Mulder, and Erling Strandberg. Genetic heterogeneity of residual variance - estimation of variance components using double hierarchical generalized linear models. *Genetics, selection, evolution: GSE*, 42(1):8, March 2010.

[159] Ying Cao, Peng Wei, Matthew Bailey, John S K Kauwe, and Taylor J Maxwell. A Versatile Omnibus Test for Detecting Mean and Variance Heterogeneity. *Genetic epidemiology*, 38(1):51–59, December 2014.

[160] Gorge E.P. Box. A general distribution theory for a class of likelihood criteria. *Biometrika*, 36(3-4):317–346, December 1949.

[161] M Friendly, M Sigal The American Statistician, and 2018. Visualizing tests for equality of covariance matrices. *arXiv*, 2018.

[162] Miao-Xin Li, Juilian M Y Yeung, Stacey S Cherny, and Pak C Sham. Evaluating the effective numbers of independent tests and significant p -value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, 131(5):747–756, May 2012.

[163] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2):175–191, May 2007.

[164] Futao Zhang, Wenhan Chen, Zhihong Zhu, Qian Zhang, Marta F Nabais, Ting Qi, Ian J Deary, Naomi R Wray, Peter M Visscher, Allan F McRae, and Jian Yang. OSCA: a tool for omic-data-based complex trait analysis. *Genome biology*, 20(1):107–13, May 2019.

[165] R Craig Albertson, Tracie L Payne-Ferreira, John Postlethwait, and Pamela C Yelick. Zebrafish acvr2a and acvr2b exhibit distinct roles in craniofacial development. *Developmental Dynamics*, 233(4):1405–1418, August 2005.

[166] Duane L Guernsey, Makoto Matsuoka, Haiyan Jiang, Susan Evans, Christine Macgillivray, Mathew Nightingale, Scott Perry, Meghan Ferguson, Marissa LeBlanc, Jean Paquette, Lysanne Patry, Andrea L Rideout, Aidan Thomas, Andrew Orr, Chris R McMaster, Jacques L Michaud, Cheri Deal, Sylvie Langlois, Duane W Superneau, Sandhya Parkash, Mark Ludman, David L Skidmore, and Mark E Samuels. Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nature Genetics*, 43(4):360–364, February 2011.

[167] Vladimir Camarena, Lei Cao, Clemer Abad, Alexander Abrams, Yaima Toledo, Kimi Araki, Masatake Araki, Katherina Walz, and Juan I Young. Disruption of Mbd5 in mice causes neuronal functional deficits and neurobehavioral abnormalities consistent with 2q23.1 microdeletion syndrome. *EMBO molecular medicine*, 6(8):1003–1015, August 2014.

[168] Ana Belinda Campos-Xavier, Danielle Martinet, John Bateman, Dan Belluoccio, Lynn Rowley, Tiong Yang Tan, Alica Baxová, Karl-Henrik Gustavson, Zvi U Borochowitz, A Micheil Innes, Sheila Unger, Jacques S Beckmann, Lauréane Mittaz, Diana Ballhausen, Andrea Superti-Furga, Ravi Savarirayan, and Luisa Bonafé. Mutations in the heparan-sulfate proteoglycan glypican 6 (GPC6) impair endochondral ossification and cause recessive omodysplasia. *American journal of human genetics*, 84(6):760–770, June 2009.

[169] Adrianna Mostowska, Kamil K Hozyasz, Barbara Biedziak, Piotr Wojcicki, Margarita Lianeri, and Pawel P Jagodzinski. Genotype and haplotype analysis of WNT genes in non-syndromic cleft lip with or without cleft palate. *European Journal of Oral Sciences*, 120(1):1–8, February 2012.

[170] Fenglei He and Yiping Chen. Wnt signaling in lip and palate development. *Frontiers of oral biology*, 16:81–90, 2012.

[171] Tian Yang, Alexander G Bassuk, and Bernd Fritzsch. Prickle1 stunts limb growth through alteration of cell polarity and gene expression. *Developmental dynamics : an official publication of the American Association of Anatomists*, 242(11):1293–1306, November 2013.

[172] Kamil Ahsan, Noor Singh, Manuel Rocha, Christina Huang, and Victoria E Prince. Prickle1 is required for EMT and migration of zebrafish cranial neural crest. *Developmental biology*, 448(1):16–35, April 2019.

[173] Yong Wan, Brandi Lantz, Brian J Cusack, and Heather L Szabo-Rogers. Prickle1 regulates differentiation of frontal bone osteoblasts. *Scientific reports*, 8(1):18021, December 2018.

[174] Tian Yang, Zhonglin Jia, Whitney Bryant-Pike, Anand Chandrasekhar, Jeffrey C Murray, Bernd Fritzsch, and Alexander G Bassuk. Analysis of PRICKLE1 in human cleft palate and mouse development demonstrates rare and common variants involved in human malformations. *Molecular genetics & genomic medicine*, 2(2):138–151, March 2014.

[175] Sun, Xiangqing, Elston, Robert, Morris, Nathan, and Zhu, Xiaofeng. What Is the Significance of Difference in Phenotypic Variability across SNP Genotypes? *The American Journal of Human Genetics*, 93(2):390–397, 2013.