

Genome-wide study of gene-by-sex interactions identifies risks for cleft palate

Kelsey Robinson¹, Randy Parrish^{1,2}, Wasiu Lanre Adeyemo³, Terri H. Beaty⁴, Azeez Butali⁵, Carmen J. Buxó⁶, Lord JJ Gowans⁷, Jacqueline T. Hecht⁸, Lina Moreno⁹, Jeffrey C. Murray¹⁰, Gary M. Shaw¹¹, Seth M. Weinberg¹², Harrison Brand¹³, Mary L. Marazita¹², David J. Cutler¹, Michael P. Epstein¹, Jingjing Yang¹, Elizabeth J. Leslie^{1*}

1-Department of Human Genetics, Emory University, Atlanta, GA, 30322, USA

2- Department of Biostatistics and Bioinformatics, Emory University School of Public Health, Emory University, Atlanta, GA, 30322, USA

3-Department of Oral and Maxillofacial Surgery, College of Medicine, University of Lagos, Lagos, 101017, Nigeria

4-Department of Epidemiology, Johns Hopkins University, Baltimore, MD, 21218, USA

5-Department of Oral Biology, Radiology, and Medicine, University of Iowa, Iowa City, IA, 52242, USA

6-School of Dental Medicine, University of Puerto Rico, San Juan, Puerto Rico 00925, USA

7-Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

8-Department of Pediatrics, McGovern Medical School University of Texas Health at Houston, Houston, TX, 77030, USA

9-Department of Orthodontics & The Iowa Institute for Oral Health Research, University of Iowa, Iowa City, IA, 52242, USA

10-Department of Pediatrics, University of Iowa, Iowa City, IA, 52242, USA

11-Department of Pediatrics, Stanford University, Stanford, CA, 94305, USA

12-Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, School of Dental Medicine, and Department of Human Genetics, School of Public Health, University of Pittsburgh, Pittsburgh, PA, 15260, USA

13-Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, 02114, USA

Abstract

Structural birth defects affect 3-4% of all live births and, depending on the type, tend to manifest in a sex-biased manner. Orofacial clefts (OFCs) are the most common craniofacial structural birth defects and are often divided into cleft lip with or without cleft palate (CL/P) and cleft palate only (CP). Previous studies have found sex-specific risks for CL/P, but these risks have yet to be evaluated in CP. CL/P is more common in males and CP is more frequently observed in females, so we hypothesized there would also be sex-specific differences for CP. Using a trio-based cohort, we performed sex-stratified genome-wide association studies (GWAS) based on proband sex followed by a genome-wide gene-by-sex (GxS) interaction testing. There were 13 loci significant for GxS interactions, with the top finding in *LTBP1* (RR=3.37 [2.04 - 5.56], $p=1.93 \times 10^{-6}$). *LTBP1* plays a role in regulating TGF- β bioavailability, and knockdown in both mice and zebrafish lead to craniofacial anomalies. Further, there is evidence for differential expression of *LTBP1* between males and females in both mice and humans. Therefore, we tested the association between the imputed genetically regulated gene expression of genes with significant GxS interactions and the CP phenotype. We found significant association for *LTBP1* in cell cultured fibroblasts in female probands ($p=0.0013$) but not in males. Taken altogether, we show there are sex-specific risks for CP that are otherwise undetectable in a combined sex cohort, and *LTBP1* is a candidate risk gene, particularly in females.

Introduction

Structural birth defects occur in 3-4% of all live births (Lary and Paulozzi 2001; Sokal, Tata, and Fleming 2014) and account for 20% of infant mortality (Copeland and Kirby 2007; Ely and Driscoll 2019). There are several factors associated with increased risk for structural birth defects, including maternal and paternal ages, environmental exposures, and genetic variation. In general, more males are born with structural birth defects than females (Tennant et al. 2011; Cui et al. 2005). Although there are varying degrees of both phenotypic and genetic heterogeneity among different structural birth defects, this sex discrepancy suggests inherent sex-specific risks. One such example is that of orofacial clefts (OFCs). For clefts involving the upper lip (cleft lip with or without a cleft palate, CL/P), males are affected at a rate about twice that of females, while for cleft palate (CP), females are affected more frequently (Nasreddine, El Hajj, and Ghassibe-Sabbagh 2021; Mai et al. 2019).

OFCs are a common congenital anomaly, occurring in approximately 1 in 1000 live births (Mai et al. 2019). The need for surgical intervention at an early age, advanced orthodontic care, and speech and hearing therapies contribute to a substantial individual and public health burden. As such, better understanding of their etiology can lead to prevention and/or improved health outcomes. OFCs are etiologically heterogeneous, with environmental exposures such as maternal multivitamin use (Yoshida et al. 2020), and maternal alcohol consumption (DeRoo et al. 2016) or smoking (Kummet et al. 2016; Little, Cardy, and Munger 2004) contributing to occurrence risk, as well as genetic factors. CP is highly heritable with estimates around 80-90% (Sivertsen et al. 2008; Grosen et al. 2010), but CP is still understudied compared to CL/P (Dixon et al. 2011; Leslie and Marazita 2013; Carlson et al. 2018; Awotoye et al. 2022).

Several models exist to explain how sex biases manifest in structural birth defects and in CP. An X-linked CP with ankyloglossia phenotype has been reported with variants in *TBX22*, primarily affecting males but appearing to follow a semi-dominant X-linked inheritance pattern (Marçano et al. 2004). Although seemingly counterintuitive, aberrant and incomplete X inactivation secondary to p53 loss has been suggested as a mechanism for increased incidence of neural tube defects in females (Delbridge et al. 2019). Further, there are other X-linked disorders in which female heterozygotes exhibit a wide spectrum of phenotypes as compared to more severely affected males (Pinto et al. 2010; Savige et al. 2016). Another likely contributing factor is the timing of palatal closure in embryogenesis: female embryos undergo this process approximately one week later than males (Burdi and Silvey 1969). This delay in closure allows more time for disruption of the growth and fusion of palatal shelves compared to male embryos and may ultimately contribute to a lower genetic liability in females (*i.e.*, higher susceptibility to CP due to genetic variation). Several genome-wide association studies (GWAS) for CP are published; however, all are sex-combined cohorts (Leslie et al. 2016; Beaty et al. 2011; Butali et al. 2019; Huang et al. 2019; He et al. 2020; Robinson et al. 2023). Stratification by sex in CL/P has led to discovery of novel associated sex-specific factors (Carlson et al. 2018; Awotoye et al. 2022), but no such study has been conducted for CP. Although the observed sex bias is opposite for CP, it is plausible there are similar sex-specific risk factors. One genome-wide interaction study identified SNPs associated with environmental exposures such as maternal smoking, alcohol intake, or vitamin use (Beaty et al. 2011), but none have specifically examined gene-by-sex interactions.

Taken altogether, we hypothesized sex may influence the occurrence of CP. To test this hypothesis, we first performed sex-stratified GWAS, followed by gene-by-sex (GxS) interaction analysis on 435 cleft palate case-parent trios. After identification of 13 loci with significant interactions, we also tested the association of the predicted genetically regulated gene expression with OFCs in the full cohort and by sex for the nearest genes for these significant loci.

Methods

Study population and phenotyping

The study population was recruited from domestic and international sites in North America (United States, Puerto Rico), Europe (Spain, Turkey, Hungary), South America (Argentina), Asia (China, Singapore, Taiwan, the Philippines), and Africa (Nigeria, Ghana). In total, 435 case-parent trios were ascertained through probands affected with CP. We did not exclude subjects based on parent affection status or the presence of additional clinical features. There were 19 trios (4.4%) with an affected parent and 46 trios (10.6%) with minor or major clinical features in addition to CP. Proband sex was confirmed genetically for all cases during quality control processes.

Sample preparation and whole genome sequencing

Whole genome sequencing was performed at the Center for Inherited Disease Research (CIDR) at Johns Hopkins University (Baltimore, MD). Prior to sequencing, a Fragment Analyzer system was used to test samples for adequate quantity and quality of genomic DNA, which were then processed with an Illumina InfiniumQCArray-24v1-0 array to confirm sex, relatedness, and known duplicates. For each sample, 500-750ng of genomic DNA was sheared to 400-600bp fragments, then processed with the Kapa Hyper Prep kit for End-Repair, A-Tailing, and Ligation of IDT (Integrated DNA Technologies) unique dual-indexed adapters according to the Kapa protocol to create a final PCR-free library.

Sequencing was performed on 150bp paired-end reads on a NovaSeq 6000 platform with base calling through the Illumina Real Time Analysis software (version 3.4.4). Files were demultiplexed from binary format (BCL) to individual fastq files with Illumina Isas bcl2fastq (version 1.37.1) and aligned to the human hg38 reference sequence (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39). Alignment, variant calling, and quality control was done with the DRAGEN Germline v3.7.5 pipeline on the Illumina BaseSpace Sequence Hub platform, which produced single sample VCF files. Contamination for any cross-human sample was checked with the DRAGEN contamination detection tool. We performed genotype concordance with existing array-based genotypes using CIDRSeqSuite (version 7.5.0), and genotype concordance checks amongst replicate samples was performed in Picard GenotypeConcordance (Picard 2019). After data quality steps, samples with at least 80% of the genome at 20X coverage or autosomal coverage at 30X underwent joint calling to generate a multi-sample VCF.

Quality control

Variants underwent quality control in the multi-sample VCF. Only variants aligning to the standard chromosomes (1-22, X, Y) were included. Variants with any flag other than "PASS" and a minor allele count (MAC) of <2 were removed. Genotypes were set to missing if the quality score was <20 or had a read depth of <10, and then variants with missingness >10% were removed. Sample-level metrics were then evaluated for heterozygous/homozygous ratio, transition/transversion (Ts/Tv) ratio, and silent/replacement rate. Outlier samples with values outside of 3 standard deviations from the cohort mean were discarded. Samples with high missing data (>5% missing) were removed.

Statistical analysis

We performed GWAS using transmission disequilibrium tests (TDT) on the full cohort and stratified by proband sex. The TDT uses the rate of transmission of the minor allele to determine association. This is beneficial as it is not confounded by population stratification, however, it is only informative at sites for which a parent is heterozygous. TDTs were performed using the R package *trio* (Schwender H 2023) for trios in which all individuals had a genotype missing rate of <5% and did not exceed a Mendelian error rate of 2%. Variants were removed based on the following filters: minor allele frequency (MAF) <3%, Mendelian error rate >0.1%, Hardy-Weinberg

disequilibrium as determined by exact test p-value of $<1 \times 10^{-6}$, and missingness rate of $>5\%$. Relative risk is reported in respect to the alternate allele. We considered $p < 1 \times 10^{-5}$ to be the threshold of suggestive significance and $p < 5 \times 10^{-8}$ of genome-wide significance.

We then performed a TDT-based gene-by-sex interaction analysis in *trio*, using the *colGxE()* function, which uses a two-step procedure as described by Gauderman et. al (Gauderman et al. 2010). We first evaluated all variants using a likelihood ratio test with two degrees of freedom (LRT 2df), which simultaneously tests the SNP main effect and environmental interaction. For this function, a χ^2 -distribution is used as the asymptotic null distribution for p-values determination.

SNPs reaching suggestive levels of association ($p < 1 \times 10^{-5}$) were then evaluated specifically for gene-by-sex interactions. The GxS relative risk (RR) is reported as the RR in females divided by the RR in males. Because these tests are considered independent of one another, we used the number of suggestive SNPs in the LRT 2df test to calculate the significance threshold for the GxS step based on a Bonferroni correction ($0.05/71$ SNPs; $p < 7.04 \times 10^{-4}$). We performed a *post hoc* power calculation in Quanto (version 1.2.4) to evaluate our power to detect interactions in the same direction with different magnitude or antagonistic interactions.

Genetically regulated gene expression imputation and association testing

For each test gene, we first trained sex-specific gene expression imputation models with cis-SNPs as predictors using GTEx V8 reference data (31) for all 49 tissues in female samples and 48 tissues in male samples. According to the GTEx eQTL analysis protocols, covariates including sequencing protocol (PCR-free), sequencing platform (NovaSeq 6000), top five genotype principal components, and top probabilistic estimation of expression residual (PEER) factors (32) were regressed out from the gene expression quantitative traits (log₂ transformed transcripts per million). For each tissue, we used the TIGAR-V2 tool (33) to train both nonparametric Dirichlet process regression (DPR) model and penalized the linear regression model with an Elastic-net penalty. Second, the trained sex-specific gene expression imputation models, i.e., the effect sizes of cis-SNP predictors (referred to as eQTL weights), were used to impute genetically regulated gene expression using the corresponding sex-specific GWAS data. For each gene, the sex-specific association between the genetically regulated gene expression (imputed by the DPR or by the Elastic-Net model) and CP were tested by the S-PrediXcan Z-score statistic (34) as implemented in the TIGAR-V2 tool. Third, for each gene, sex-specific test p-values based on the DPR and the Elastic-Net model were then combined by using the aggregated Cauchy association test (ACAT) (35). Previous studies have shown the ACAT method can improve power by leveraging multiple tests based on different statistical models of the same null hypothesis. As a result, we obtained sex-specific ACAT p-values for testing the association between the genetically regulated gene expression of test genes and the CP phenotype. A significant ACAT p-value means the genetic effect of the test gene (aggregated over multiple test cis-SNPs with non-zero eQTL weights) on CP is potentially mediated through the corresponding gene expression.

Rare Variants

A bed file for hg38 coordinates for *LTBP1* was generated using the UCSC Genome Browser. All variants within this region were extracted from a multisample VCF with vcfTools (version 0.1.13) followed by annotation with Annovar (version 201910) (Wang, Li, and Hakonarson 2010). The commands “dbnsfp42a” and “revel” were included to obtain SIFT, CADD, and REVEL scores for all variants. Additionally, prediction values from AlphaMissense (Cheng et al. 2023) were accessed from the publicly available data. Predicted damaging thresholds were as follows: SIFT ≤ 0.05 , CADD ≥ 20 , REVEL ≥ 0.75 , and AlphaMissense ≥ 0.564 .

Results

Using a dataset of 435 trios, we performed sex-stratified TDTs of 262 female and 173 male probands based on genetic sex (Fig 1, Online Resource 1, 2). For all evaluations, we tested 6,774,204 SNPs. There were 16 SNPs in 9 loci of suggestive significance in the female analysis, and 3 SNPs in 3 loci in males (Table 1). The significant loci from each test were non-overlapping between sexes—that is, they were only significant in either male or female probands. Additionally, each of these loci were more significant with higher odds ratios (ORs) within the sex-stratified TDTs than compared to the full cohort (Online Resource 3). To explore these sex-specific differences more formally, we next performed a GxS analysis to find sex-specific interactions.

We first tested SNPs with the LRT 2df model of association considering both gene and GxS effects (Online Resource 4). There were no SNPs at genome-wide significance in this step, though there were 71 SNPs in 32 regions with $p < 1 \times 10^{-5}$ (suggestive significance threshold) (Table 2). Of those, 26 SNPs in 13 regions were considered significant for gene-by-sex interactions (Bonferroni correction for 71 SNPs, $p < 7.04 \times 10^{-4}$) (Online Resource 5). We noted that for all GxS significant SNPs, the direction of effect was opposite in males versus females (Fig 2a, b). A *post-hoc* power analysis shows given our sample size, we have good power to detect large differences (including opposite effects) when SNPs have a MAF of at least ~15%. However, we have very low power to detect modest interactions, or even large differences at lower MAFs (Online Resource 6). We did not detect significant GxS interactions in the remaining 45 SNPs in 19 regions, suggesting that these associations are primarily driven by the effect of the genotype rather than sex-specific interactions (Fig 2c).

For the 13 loci from the sex-stratified TDTs (Table 1), two of the three loci from the male-specific analysis, and six of nine loci from the female-specific analysis, we found significant for GxS effects. Using the results from all four tests, we compared the relative risks for the ten most significant SNPs in the GxS to those from the sex-stratified and sex-combined TDTs, shown in Fig 3.

The most significant finding for GxS interaction was at 2p22.3, within intron 28/33 of latent TGF-beta binding protein 1 (*LTBP1*). We expanded our analysis of *LTBP1* to look at rare variants (MAF <0.05%) in the full cohort and identified 21 unique protein-altering variants in 25 probands (Online Resource 7). 14 variants in 15 probands had at least one damaging prediction as determined by CADD, REVEL, or AlphaMissense scores. We performed a Fisher's exact test on the difference in the number of predicted damaging rare variants in male versus female probands and found these were not statistically significant ($p=0.22$).

Lastly, there is evidence for differential expression of *LTBP1* between sexes in both the pituitary of mice (Recouvreux et al. 2013) and whole blood of humans (Recouvreux et al. 2013). Therefore, we evaluated the association between the sex-specifically imputed genetically regulated gene expression and the CP phenotype for 12 genes present at the GxS significant loci. We also evaluated the association between the imputed genetically regulated gene expression and the CP phenotype in all samples. There were 11 genes with nominal significance ($p < 0.05$) in at least one of the tested tissues (Table 3). After correction for 12 loci, there were only two loci remaining significant: *GTF2A1L* for females in the substantia nigra of the brain ($p=0.0010$), and *LTBP1* for females in cultured fibroblast cells ($p=0.0013$). The latter association was also observed to a lesser extent for the full cohort ($p=0.0125$).

Discussion

CP is more prevalent in females than males, yet little is known about what contributes to these sex-specific risks. Here, we initially found 12 loci with suggestive evidence of association with CP when stratifying our cohort by proband sex, followed by identification of 13 loci with suggestive evidence of gene-by-sex interactions in CP probands. In all instances, the identified SNPs increased risk of CP for one sex while appearing to reduce the risk for the other. Given our sample size, this was expected as we have more power to detect these large differences in effects; much

larger sample sizes would be needed to detect effects of different magnitude in the same direction in males and females. We also found that most loci were associated with a higher relative risk in females compared to males, which supports the epidemiological data where females are more commonly affected.

In the LRT 2df analysis, we identified 13 loci associated with CP but not significant for GxS interactions. This suggests that these associations are primarily gene-driven, where the genotype influences the risk for CP regardless of patient sex. We reported the effects of these gene-driven SNPs in our previously published GWAS results for the full cohort (Robinson et al. 2023), and our results were consistent across both studies, as expected. However, none of the loci with significant GxS interactions were significant in that study, as the risks in males and females likely negated each other. We also compared these loci to sex-specific risks previously reported for OFCs (Carlson et al. 2018; Awotoye et al. 2022), but none were replicated in our cohort. Altogether, these findings illustrate heterogeneity in genetic architecture both between CP and CL/P as well as between males and females, and highlight the importance of investigating sex-specific differences and/or GxS interactions in CP. While our study contributes to the knowledge of the genetic basis of CP and sex-specific risks, it is not without limitation. Our sample size of 435 trios is underpowered to identify gene-by-sex interactions in the same direction but with differences in the magnitude of risk. Further, due to the use of TDTs for case-parent trios, we are restricted to evaluation of sites for which parents are heterozygous.

Our top finding was an intronic SNP in *LTBP1* at 2p22.3. There are several lines of evidence suggesting that *LTBP1* could play a role in palatogenesis. *LTBP1* is an extracellular matrix protein and a component of the large latent complex (Peeters et al. 2022). This complex binds to a second complex of latency associated proteins and mature TGF- β , working to regulate its biological availability in the extracellular matrix. Disruption of the TGF- β signaling pathway is associated with CP (Iwata, Parada, and Chai 2011; Taya, O'Kane, and Ferguson 1999), so dysregulation caused by variants in *LTBP1* could hypothetically lead to disrupted palatogenesis. We further identified 21 unique, rare variants (MAF <0.5%) in 12 males and 13 females. Of the 21 total, 15 were predicted to be damaging by at least one *in silico* tool. These were found in 9 males and 6 females, although this difference was not significant with a Fisher's exact test.

From a clinical standpoint, bi-allelic truncating variants in *LTBP1* have been reported as causal for autosomal recessive cutis laxa syndrome type IIE (OMIM: 619451). In addition to connective tissue abnormalities, affected individuals present with craniofacial dysmorphisms including high arched palates in 62.5% of cases (Pottie et al. 2021). Further, *LTBP1* directly competes with GARP (glycoprotein A repetitions predominant, encoded by *LRRC32*), a type I membrane protein that plays a role in maturation, processing, and tethering of TGF- β to the cell surface (Tran et al. 2009). Both *LTBP1* and GARP compete for binding via a disulfide link to the same cysteine (Cys-4) on proTGF β 1, and they direct TGF- β to either the cell surface or deposited within the ECM (Wang et al. 2012). Recently, homozygous variants in *LRRC32* have been reported to cause a cleft palate syndrome (Hexner-Erlichman et al. 2022), suggesting imbalance in this process can disrupt palatogenesis.

In support of clinical data, morpholino knockdown of *ltbp1* in zebrafish causes severe jaw malformations in which the mandible has a reduced length and width. Similarly, knockout of *Ltbp-1* in mice resulted in a modified facial profile including a shortened maxilla and mandible, as well as reduced cellular activation and differentiation of myofibroblasts, supporting reduced availability of *Tgfb* when *Ltbp-1* is disrupted (Drews et al. 2008). In humans, shortened mandibles are associated with CP in Pierre Robin sequence (Resnick et al. 2018). Our cohort included 18 trios reported to have Pierre Robin sequence and an additional 31 with clinical features consistent with syndromic CP. While this was too few to perform a separate analysis, we carried out sensitivity analyses in which we repeated the GxS for *LTBP1* without PRS cases, as well as without PRS cases or any suspected to be syndromic. Our results did not meaningfully change, suggesting that—in this study—Pierre Robin or syndromic cases are not driving the signal in this gene.

Although the mechanism by which variants in or around *LTBP1* mediate their effect on a sex-specific basis is unknown, there is evidence for sex-based differences in expression of this gene. A study in mice found that *Ltbp1* expression in the pituitary was lower in females than males, and that both sexes showed reduced expression in response to estradiol, suggesting a role for sex hormones in this differentiation (Recouvreux et al. 2013). A similar finding has been shown in human peripheral blood, with *LTBP1* expression belonging to the top 1% of sex-biased genes (based on log fold change) where males had higher levels of *LTBP1* mRNA than females (Jansen et al. 2014). To explore this based on our GWAS data, we performed an association analysis for CP and the imputed genetically regulated gene expression and found a female-specific significant association for *LTBP1* in cultured fibroblast cells. Despite a lack of access to a more relevant tissue (i.e., palatal tissue), the association of *LTBP1* expression with CP in female probands is of interest as fibroblasts have many of the same traits as mesenchymal stem cells, including an ability to differentiate into chondrocyte and osteoblasts (Denu et al. 2016).

In summary, we have demonstrated that there are gene-by-sex interactions for CP which cannot be detected in combined-sex cohorts as opposite allelic risks can negate each other. We also identified *LTBP1* as a gene of interest for CP risk, particularly for females, and show that it may mediate its effect on the phenotype via gene expression. In conclusion, sex-specific effects contribute to the genetic heterogeneity of CP and should be further explored as sample sizes continue to grow.

Acknowledgements

We would like to thank participants and their families, without whom this research would not be possible. We would also like to thank Elena Serna and Rosa Martinez for collecting samples and maintaining the UHealth database. Sequencing services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268201700006I. Patient recruitment, assembly of phenotypic information, sequencing services, and data analysis were supported by National Institutes of Health (NIH) grants: X01-HG010835 (EL), R01-DE016148 (MM, SW), R01-DE030342 (EL), R01-DE011931 (JH), R01-DE028300 (AB), R01-DE014581 (TB), R37-DE008559 (JM), R00-DE024571 (CB), S21-MD001830 (CB) U54-GM133807 (CB), T32-GM008490 (KR), F31-DE032588 (KR). This work was in part supported through cooperative agreements under PA 96043 from the Centers for Disease Control and Prevention to the Centers for Birth Defects Research and Prevention participating in the National Birth Defect Prevention Study. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the California Department of Public Health.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

Sequence and phenotype data is available from the Database of Genotypes and Phenotypes (dbGaP) under study accession phs002220.v1.p1.

Ethics Approval

This study was performed in line with the principles of the Declaration of Helsinki. Institutional review board approval was obtained for each local recruitment site and coordinating center: University of Iowa, University of Pittsburgh, and Emory University. Informed consent was obtained from either the parents or all individual participants included in the study.

Fig 1 Miami plot highlighting differences in suggestive SNPs in females (top) and males (bottom). The dotted line represents suggestive significance at $p=1 \times 10^{-5}$, and the solid line represents $p=5 \times 10^{-8}$

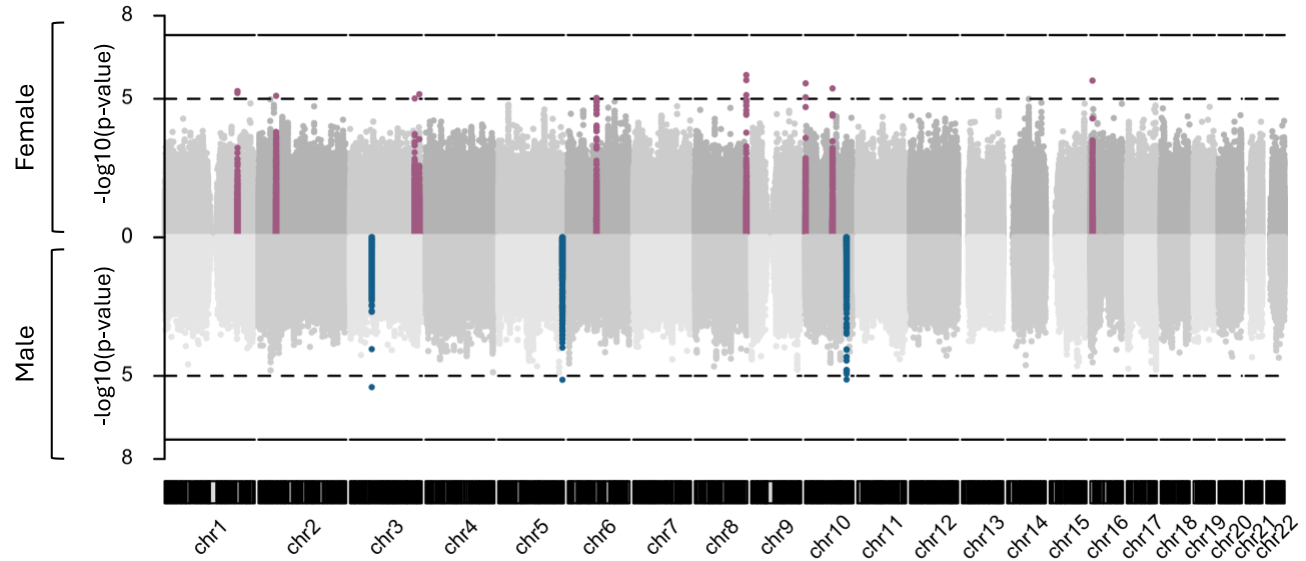


Fig 2 Relative risks for males and females demonstrating GxS effect SNPs at (A) 2p22.3 and (B) 15q21.1, and (C) gene-driven SNP at 112q21.1 with 95% confidence intervals. The dotted line at 1 represents null risk. Light blue and dark blue represent heterozygous and homozygous males, respectively while light and dark pink represent heterozygous and homozygous females, respectively

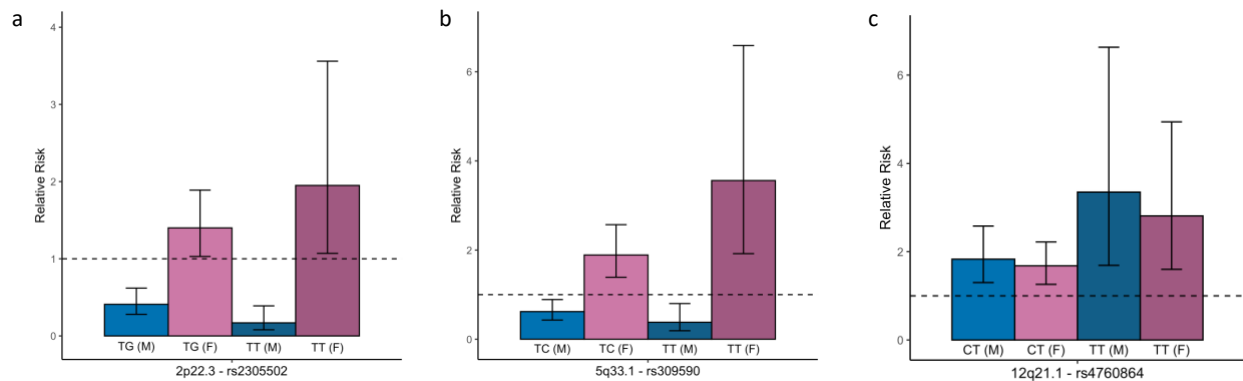


Fig 3 Comparing relative risks with 95% confidence intervals across the top 10 SNPs for GxS effects (top) and two gene-driven SNPs (bottom) to relative risk from TDTs in the full cohort and stratified by sex. Risks are shown for GxS (black), female only TDT (pink), male only TDT (blue), and full cohort TDT (green). An asterisk next to the locus name denotes relative risk by heterozygosity shown in Fig 2

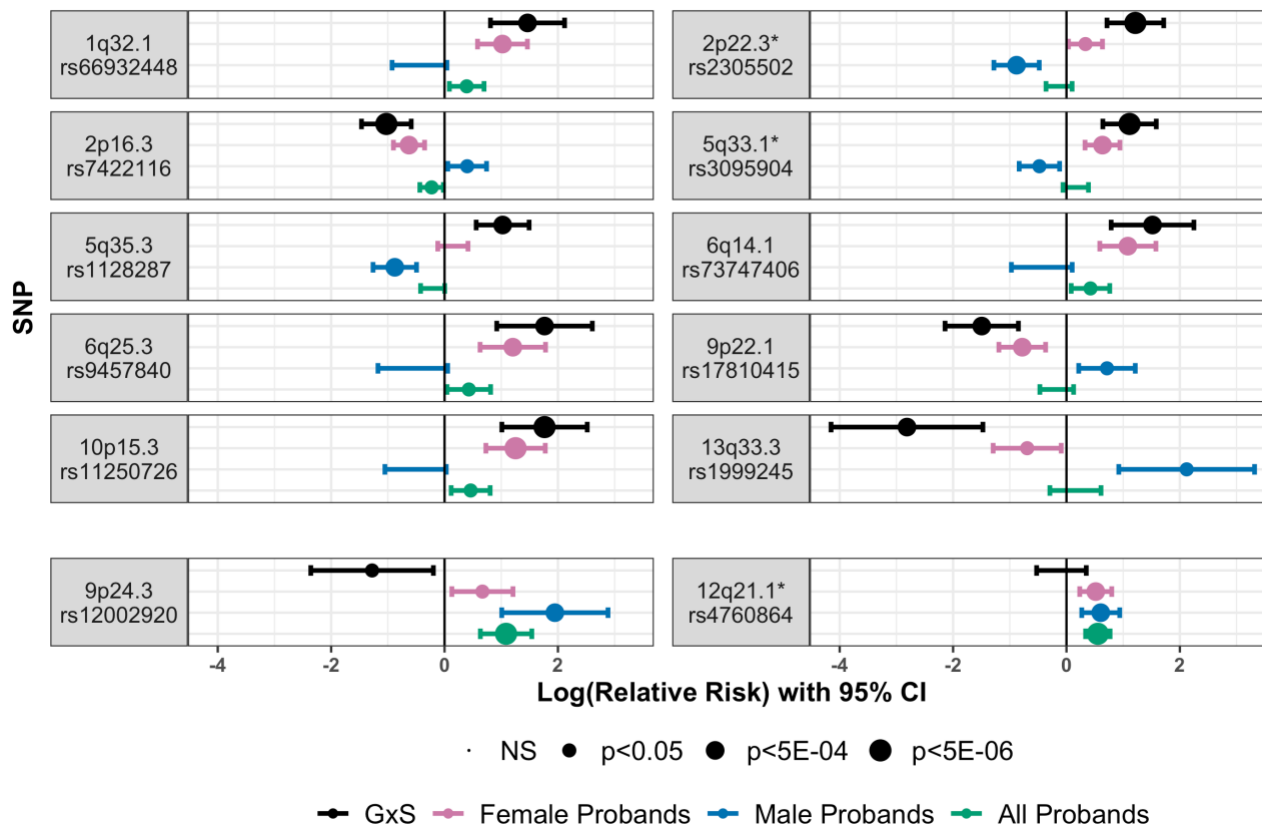


Table 1: Loci of suggestive significance from sex stratified TDTs									
Locus (nearest gene)	CHR:BP (rsID)	Ref/Alt	Female RR	Pvalue	Male RR	Pvalue	All RR	Pvalue	Alt AF [^]
Male									
3p14.2* (FHIT)	3:60050040 (rs212048)	A/G	1.04 [0.77 - 1.4]	0.82	0.37 [0.24 - 0.56]	3.92E-06	0.71 [0.56 - 0.9]	5.11E-03	0.33
5q35.3* (RAB24)	5:177301457 (rs1128287)	G/T	1.16 [0.89 - 1.51]	0.28	0.42 [0.28 - 0.61]	7.22E-06	0.81 [0.66 - 1]	0.05	0.8
10q25.3 (ABLIM1, FAM160B1)	10:114803389 (rs7075095)	T/C	1.12 [0.8 - 1.57]	0.5	2.66 [1.73 - 4.07]	7.40E-06	1.6 [1.23 - 2.07]	3.81E-04	0.43
Female									
1q32.1* (CACNA1S)	1:201074717 (rs66932448)	ACT/A	2.78 [1.79 - 4.31]	5.31E-06	0.64 [0.4 - 1.04]	0.07	1.48 [1.09 - 2.01]	0.01	0.72
2p16.3* (GTF2A1L)	2:48656339 (rs7422116)	T/A	0.53 [0.41 - 0.7]	7.83E-06	1.49 [1.06 - 2.1]	0.02	0.8 [0.65 - 0.98]	0.03	0.43
3q26.32 (KCNMB2)	3:178712151 (rs9290663)	A/T	0.52 [0.39 - 0.69]	9.72E-06	0.98 [0.68 - 1.41]	0.93	0.66 [0.53 - 0.83]	3.02E-04	0.21
3q28 (LINCR-0002)	3:191499526 (rs9790048)	T/A	0.48 [0.35 - 0.66]	6.82E-06	0.85 [0.57 - 1.26]	0.42	0.6 [0.47 - 0.76]	4.26E-05	0.89
6q14.1* (SH3BGRL2)	6:79611392 (rs73476169)	T/C	3.05 [1.86 - 4.99]	9.37E-06	0.81 [0.48 - 1.36]	0.43	1.7 [1.21 - 2.38]	2.23E-03	0.21
8q24.3* (MIR4472-1)	8:142171910 (rs55671251)	G/A	2.29 [1.63 - 3.2]	1.39E-06	0.82 [0.56 - 1.19]	0.29	1.48 [1.16 - 1.88]	1.66E-03	0.19
10p15.3* (ADARB2)	10:1683542 (rs11250726)	A/G	3.5 [2.07 - 5.91]	2.77E-06	0.6 [0.35 - 1.03]	0.06	1.58 [1.12 - 2.24]	8.66E-03	0.17
10q22.3* (LRMDA)	10:76007479 (rs11001574)	T/C	1.93 [1.46 - 2.56]	4.22E-06	1.25 [0.89 - 1.75]	0.2	1.62 [1.31 - 2.01]	9.81E-06	0.19
16p13.3 (RBFOX1)	16:6057173 (rs17139255)	T/G	0.46 [0.34 - 0.64]	2.21E-06	0.94 [0.63 - 1.4]	0.76	0.61 [0.47 - 0.77]	6.00E-05	0.18
*Locus also identified in GxS analysis ^gnomAD v 3.1.2									

Locus (nearest gene)	CHR:BP* (rsID)	Ref/Alt Allele	GxS RR	GxS p-value	LRT 2df p-value	Alt AF [^]
1q32.1 (CACNA1S)	1:201074717 (rs66932448)	ACT/A	4.32 [2.25 - 8.31]	<u>1.14E-05</u>	1.52E-06	0.72
1q41 (LYPAL1)	1:219268643 (rs59611530)	G/GAAT	1.01 [0.65 - 1.56]	0.96	2.39E-06	0.47
2p22.3 (BIRC6)	2:32447229 (rs555243726)	G/A	0.21 [0.08 - 0.54]	1.16E-03	2.02E-06	0.11
2p22.3 (LTBP1)	2:33361709 (rs2305502)	T/G	3.37 [2.04 - 5.56]	<u>1.93E-06</u>	3.21E-06	0.27
2p21 (MTA3)	2:42509602 (rs57081889)	C/T	1.22 [0.79 - 1.88]	0.37	9.30E-06	0.39
2p16.3 (GTF2A1L)	2:48656339 (rs7422116)	T/A	0.36 [0.23 - 0.56]	<u>4.49E-06</u>	1.92E-06	0.41
3p14.2 (FHIT)	3:60050040 (rs212048)	A/G	2.82 [1.68 - 4.75]	<u>9.41E-05</u>	5.82E-06	0.34
4p14	4:39803327 (rs10000967)	T/C	0.48 [0.21 - 1.09]	0.08	6.24E-07	0.10
5q14.1 (HOMER1)	5:79403345 (rs79156100)	T/C	0.38 [0.14 - 1.07]	0.07	1.43E-06	0.01
5q33.1 (NDST1)	5:150536578 (rs309590)	T/C	3.04 [1.9 - 4.87]	<u>3.65E-06</u>	5.36E-06	0.25
5q35.3 (RAB24)	5:177301457 (rs1128287)	G/T	2.79 [1.75 - 4.44]	<u>1.68E-05</u>	8.76E-06	0.78
6q14.1 (SH3BGRL2)	6:79609515 (rs73747406)	T/A	4.56 [2.2 - 9.47]	<u>4.56E-05</u>	6.92E-06	0.20
6q25.3 (SLC22A1)	6:160120316 (rs9457840)	T/C	5.83 [2.51 - 13.55]	<u>4.10E-05</u>	9.18E-06	0.05
8p23.1 (RP1L1)	8:10669884 (rs55782357)	A/T	0.01 [0 - 0.13]	1.48E-03	5.81E-08	0.09
8q24.3 (MIR4472-1)	8:142171910 (rs55671251)	G/A	2.8 [1.69 - 4.64]	<u>6.50E-05</u>	1.82E-06	0.19
9p24.3 (LOC105375951)	9:1705036 (rs12002920)	G/T	0.28 [0.09 - 0.82]	0.02	1.42E-07	0.09
9p22.1 (SAXO1)	9:19032909 (rs17810415)	A/G	0.22 [0.12 - 0.43]	<u>5.93E-06</u>	9.00E-06	0.19
9q21.12 (TRPM3)	9:70937619 (rs67149721)	A/G	0.18 [0.05 - 0.72]	0.02	1.20E-06	0.10
10p15.3 (ADARB2)	10:1683542 (rs11250726)	A/G	5.83 [2.75 - 12.39]	<u>4.43E-06</u>	3.04E-07	0.17
10q21.2 (RTKN2)	10:62241751 (rs112160719)	T/A	0.06 [0.01 - 0.56]	0.01	5.75E-06	0.08
10q22.3 (LRMDA)	10:76007479 (rs11001574)	T/C	1.55 [1 - 2.4]	0.05	6.23E-06	0.19
11q14.3 (FAT3)	11:92264783 (rs76502406)	C/T	0.74 [0.3 - 1.8]	0.5	2.46E-06	0.01
11q24.2	11:124488761 (rs375612889)	C/T	1.25 [0.44 - 3.53]	0.68	8.43E-06	0.07
12q21.1 (LOC105369838)	12:72976347 (rs4760864)	C/T	0.92 [0.59 - 1.42]	0.69	1.88E-06	0.25
13q33.3	13:106844528 (rs78276410)	C/T	0.06 [0.02 - 0.23]	<u>3.85E-05</u>	3.40E-06	0.03
14q22.1 (FRMD6)	14:51469594 (rs10431684)	C/A	1.53 [0.97 - 2.41]	0.07	3.68E-06	0.57
15q26.2	15:96732159 (rs36062094)	T/C	1.37 [0.88 - 2.14]	0.17	7.63E-06	0.23
16p13.3 (RBFox1)	16:6057173 (rs17139255)	T/G	0.49 [0.3 - 0.82]	6.02E-03	5.58E-06	0.19

16q24.2 (<i>C16orf95</i>)	16:87082043 (rs145642800)	CAGA/C	0.28 [0.12 - 0.63]	2.09E-03	8.87E-06	0.06
17p13.2 (<i>ANKFY1</i>)	17:4224286 (rs58695167)	G/A	1.57 [0.98 - 2.5]	0.06	4.65E-06	0.50
17q24.2 (<i>ARSG</i>)	17:68398218 (rs7222535)	A/G	1.85 [1 - 3.43]	0.05	3.43E-06	0.11
17q25.3 (<i>RPTOR</i>)	17:80629981 (rs62068342)	A/G	0.11 [0.03 - 0.43]	1.49E-03	8.77E-06	0.06
Locus is significant with Bonferroni correction for 71 tested SNPs ($p < 7.04 \times 10^{-4}$) ^gnomAD v 3.1.2						

Table 3: Tissues and genes with significant association between imputed gene expression and phenotype				
Gene (Locus)	Tissue	All	Male	Female
<i>CACNA1S</i> (1q32.1)	Minor Salivary Gland	NS	NS	0.0385
<i>LTBP1</i> (2p22.3)	Artery - Tibial	NS	NS	0.0193
	Cells Cultured - Fibroblasts	0.0125	NS	0.0013
<i>GTF2A1L</i> (2p16.3)	Artery - Aorta	NS	NS	0.0385
	Brain - Cerebellum	NS	NS	0.0273
	Brain - Substantia nigra	NS	NS	0.0010
	Colon - Sigmoid	NS	NS	0.0051
	Pancreas	NS	NS	0.0311
	<i>STON1-GTF2A1L</i> (2p16.3)	Brain - Spinal cord cervical C1	NS	NS
	Colon - Sigmoid	NS	NS	0.0237
<i>FHIT</i> (3p14.2)	Esophagus - Mucosa	NS	NS	0.0248
<i>NDST1</i> (5q33.1)	Brain - Substantia nigra	NS	NS	0.0181
	Cells Cultured - Fibroblasts	NS	NS	0.0306
	Minor Salivary Gland	NS	NS	0.0436
<i>RAB24</i> (5q35.3)	Brain - Hypothalamus	NS	NS	0.0343
	Cells - EBV-transformed lymphocytes	NS	0.0052	NS
	Esophagus - Muscularis	NS	0.0405	NS
	Minor Salivary Gland	NS	0.0265	NS
	Whole Blood	NS	0.0151	NS
<i>SH3BGRL2</i> (6q14.1)	Whole Blood	NS	NS	0.0055
<i>SLC22A1</i> (6q25.3)	Brain - Cerebellar Hemisphere	NS	NS	0.0112
	Brain - Hypothalamus	NS	NS	0.0380
	Cells - EBV-transformed lymphocytes	NS	0.0106	NS
<i>SAXO1</i> (9p22.1)	Prostate	NS	0.0195	NA
<i>ADARB2</i> (10p15.3)	Brain - Anterior cingulate cortex BA24	NS	NS	0.0325
	Brain - Amygdala	0.0353	NS	NS
	Brain - Caudate basal ganglia	0.0142	NS	NS
	Brain - Cerebellar Hemisphere	0.0354	NS	NS

NS=not significant, $p > 0.05$; NA=not available

- Awotoye, W., C. Comnick, C. Pendleton, E. Zeng, A. Alade, P. A. Mossey, L. J. J. Gowans, M. A. Eshete, W. L. Adeyemo, T. Naicker, C. Adeleke, T. Busch, M. Li, A. Petrin, J. Olotu, M. Hassan, J. Pape, S. E. Miller, P. Donkor, D. Anand, S. A. Lachke, M. L. Marazita, A. A. Adeyemo, J. C. Murray, D. Albokhari, N. Sobreira, and A. Butali. 2022. 'Genome-wide Gene-by-Sex Interaction Studies Identify Novel Nonsyndromic Orofacial Clefts Risk Locus', *Journal of Dental Research*, 101: 465-72.
- Beaty, T. H., I. Ruczinski, J. C. Murray, M. L. Marazita, R. G. Munger, J. B. Hetmanski, T. Murray, R. J. Redett, M. D. Fallin, K. Y. Liang, T. Wu, P. J. Patel, S. C. Jin, T. X. Zhang, H. Schwender, Y. H. Wu-Chou, P. K. Chen, S. S. Chong, F. Cheah, V. Yeow, X. Ye, H. Wang, S. Huang, E. W. Jabs, B. Shi, A. J. Wilcox, R. T. Lie, S. H. Jee, K. Christensen, K. F. Doheny, E. W. Pugh, H. Ling, and A. F. Scott. 2011. 'Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate', *Genet Epidemiol*, 35: 469-78.
- Burdi, A. R., and R. G. Silvey. 1969. 'Sexual differences in closure of the human palatal shelves', *Cleft Palate J*, 6: 1-7.
- Butali, A., P. A. Mossey, W. L. Adeyemo, M. A. Eshete, L. J. J. Gowans, T. D. Busch, D. Jain, W. Yu, L. Huan, C. A. Laurie, C. C. Laurie, S. Nelson, M. Li, P. A. Sanchez-Lara, W. P. Magee, K. S. Magee, A. Auslander, F. Brindopke, D. M. Kay, M. Caggana, P. A. Romitti, J. L. Mills, R. Audu, C. Onwuamah, G. O. Oseni, A. Owais, O. James, P. B. Olaitan, B. S. Aregbesola, R. O. Braimah, F. O. Oginni, A. O. Oladele, S. A. Bello, J. Rhodes, R. Shiang, P. Donkor, S. Obiri-Yeboah, F. K. N. Arthur, P. Twumasi, P. Agbenorku, G. Plange-Rhule, A. A. Oti, O. M. Ogunlewe, A. A. Oladega, A. A. Adekunle, A. O. Erinoso, O. O. Adamson, A. A. Elufowoju, O. I. Ayelomi, T. Hailu, A. Hailu, Y. Demissie, M. Derebew, S. Eliason, M. Romero-Bustillos, C. Lo, J. Park, S. Desai, M. Mohammed, F. Abate, L. O. Abdur-Rahman, D. Anand, I. Saadi, A. V. Oladugba, S. A. Lachke, B. A. Amendt, C. N. Rotimi, M. L. Marazita, R. A. Cornell, J. C. Murray, and A. A. Adeyemo. 2019. 'Genomic analyses in African populations identify novel risk loci for cleft palate', *Hum Mol Genet*, 28: 1038-51.
- Carlson, J. C., N. L. Nidey, A. Butali, C. J. Buxo, K. Christensen, F. W. Deleyiannis, J. T. Hecht, L. L. Field, L. M. Moreno-Urbe, I. M. Orioli, F. A. Poletta, C. Padilla, A. R. Vieira, S. M. Weinberg, G. L. Wehby, E. Feingold, J. C. Murray, M. L. Marazita, and E. J. Leslie. 2018. 'Genome-wide interaction studies identify sex-specific risk alleles for nonsyndromic orofacial clefts', *Genet Epidemiol*, 42: 664-72.
- Cheng, Jun, Guido Novati, Joshua Pan, Clare Bycroft, Akvilė Žemgulytė, Taylor Applebaum, Alexander Pritzel, Lai Hong Wong, Michal Zielinski, Tobias Sargeant, Rosalia G. Schneider, Andrew W. Senior, John Jumper, Demis Hassabis, Pushmeet Kohli, and Žiga Avsec. 2023. 'Accurate proteome-wide missense variant effect prediction with AlphaMissense', *Science*, 381: eadg7492.
- Copeland, G. E., and R. S. Kirby. 2007. 'Using birth defects registry data to evaluate infant and childhood mortality associated with birth defects: an alternative to traditional mortality assessment using underlying cause of death statistics', *Birth Defects Res A Clin Mol Teratol*, 79: 792-7.
- Cui, Wei, Chang-Xing Ma, Yiwei Tang, Vivian Chang, P.V. Rao, Mario Ariet, Michael B. Resnick, and Jeffrey Roth. 2005. 'Sex differences in birth defects: A study of opposite-sex twins', *Birth Defects Research Part A: Clinical and Molecular Teratology*, 73: 876-80.
- Delbridge, Alex R. D., Andrew J. Kueh, Francine Ke, Natasha M. Zamudio, Farrah El-Saafin, Natasha Jansz, Gao-Yuan Wang, Megan Iminoff, Tamara Beck, Sue Haupt, Yifang Hu, Rose E. May, Lachlan Whitehead, Lin Tai, William Chiang, Marco J. Herold, Ygal Haupt, Gordon K. Smyth, Tim Thomas, Marnie E. Blewitt, Andreas Strasser, and Anne K. Voss.

2019. 'Loss of p53 Causes Stochastic Aberrant X-Chromosome Inactivation and Female-Specific Neural Tube Defects', *Cell Reports*, 27: 442-54.e5.
- Denu, R. A., S. Nemcek, D. D. Bloom, A. D. Goodrich, J. Kim, D. F. Mosher, and P. Hematti. 2016. 'Fibroblasts and Mesenchymal Stromal/Stem Cells Are Phenotypically Indistinguishable', *Acta Haematol*, 136: 85-97.
- DeRoo, L. A., A. J. Wilcox, R. T. Lie, P. A. Romitti, D. A. Pedersen, R. G. Munger, L. M. Moreno Uribe, and G. L. Wehby. 2016. 'Maternal alcohol binge-drinking in the first trimester and the risk of orofacial clefts in offspring: a large population-based pooling study', *Eur J Epidemiol*, 31: 1021-34.
- Dixon, M. J., M. L. Marazita, T. H. Beaty, and J. C. Murray. 2011. 'Cleft lip and palate: understanding genetic and environmental influences', *Nature reviews. Genetics*, 12: 167-78.
- Drews, Falko, Sebastian Knöbel, Markus Moser, Kai G. Muhlack, Simone Mohren, Christian Stoll, Andreas Bosio, Axel M. Gressner, and Ralf Weiskirchen. 2008. 'Disruption of the latent transforming growth factor- β binding protein-1 gene causes alteration in facial structure and influences TGF- β bioavailability', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1783: 34-48.
- Ely, D. M., and A. K. Driscoll. 2019. 'Infant Mortality in the United States, 2017: Data From the Period Linked Birth/Infant Death File', *Natl Vital Stat Rep*, 68: 1-20.
- Gauderman, W. J., D. C. Thomas, C. E. Murcray, D. Conti, D. Li, and J. P. Lewinger. 2010. 'Efficient genome-wide association testing of gene-environment interaction in case-parent trios', *American Journal of Epidemiology*, 172: 116-22.
- Grosen, Dorthe, Cécile Chevrier, Axel Skytthe, Camilla Bille, Kirsten Mølsted, Ase Sivertsen, Jeffrey C. Murray, and Kaare Christensen. 2010. 'A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: support for the multifactorial threshold model of inheritance', *Journal of Medical Genetics*, 47: 162-68.
- He, M., X. Zuo, H. Liu, W. Wang, Y. Zhang, Y. Fu, Q. Zhen, Y. Yu, Y. Pan, C. Qin, B. Li, R. Yang, J. Wu, Z. Huang, H. Ge, H. Wu, Q. Xu, Y. Zuo, W. Chen, Y. Qin, Z. Liu, S. Chen, H. Zhang, F. Zhou, H. Yan, Y. Yu, L. Yong, G. Chen, B. Liang, R. A. Cornell, L. Zong, L. Wang, D. Zou, L. Sun, and Z. Bian. 2020. 'Genome-wide Analyses Identify a Novel Risk Locus for Nonsyndromic Cleft Palate', *Journal of Dental Research*, 99: 1461-68.
- Hexner-Erichman, Z., B. Fichtman, Y. Zehavi, M. Khayat, H. Jabaly-Habib, L. S. Izhaki-Tavor, M. Dessau, O. Elpeleg, and R. Spiegel. 2022. 'A Novel Homozygous Missense Variant in the LRRC32 Gene Is Associated With a New Syndrome of Cleft Palate, Progressive Vitreoretinopathy, Growth Retardation, and Developmental Delay', *Front Pediatr*, 10: 859034.
- Huang, L., Z. Jia, Y. Shi, Q. Du, J. Shi, Z. Wang, Y. Mou, Q. Wang, B. Zhang, Q. Wang, S. Ma, H. Lin, S. Duan, B. Yin, Y. Lin, Y. Wang, D. Jiang, F. Hao, L. Zhang, H. Wang, S. Jiang, H. Xu, C. Yang, C. Li, J. Li, B. Shi, and Z. Yang. 2019. 'Genetic factors define CPO and CLO subtypes of nonsyndromic orofacial cleft', *PLoS Genet*, 15: e1008357.
- Iwata, J., C. Parada, and Y. Chai. 2011. 'The mechanism of TGF- β signaling during palate development', *Oral Dis*, 17: 733-44.
- Jansen, R., S. Batista, A. I. Brooks, J. A. Tischfield, G. Willemsen, G. van Grootheest, J. J. Hottenga, Y. Milaneschi, H. Mbarek, V. Madar, W. Peyrot, J. M. Vink, C. L. Verweij, E. J. de Geus, J. H. Smit, F. A. Wright, P. F. Sullivan, D. I. Boomsma, and B. W. Penninx. 2014. 'Sex differences in the human peripheral blood transcriptome', *BMC Genomics*, 15: 33.
- Kummet, C. M., L. M. Moreno, A. J. Wilcox, P. A. Romitti, L. A. DeRoo, R. G. Munger, R. T. Lie, and G. L. Wehby. 2016. 'Passive Smoke Exposure as a Risk Factor for Oral Clefts-A Large International Population-Based Study', *American Journal of Epidemiology*, 183: 834-41.

- Lary, Joseph M., and Leonard J. Paulozzi. 2001. 'Sex differences in the prevalence of human birth defects: A population-based study', *Teratology*, 64: 237-51.
- Leslie, E. J., H. Liu, J. C. Carlson, J. R. Shaffer, E. Feingold, G. Wehby, C. A. Laurie, D. Jain, C. C. Laurie, K. F. Doheny, T. McHenry, J. Resick, C. Sanchez, J. Jacobs, B. Emanuele, A. R. Vieira, K. Neiswanger, J. Standley, A. E. Czeizel, F. Deleyiannis, K. Christensen, R. G. Munger, R. T. Lie, A. Wilcox, P. A. Romitti, L. L. Field, C. D. Padilla, E. M. Cutiongco-de la Paz, A. C. Lidral, L. C. Valencia-Ramirez, A. M. Lopez-Palacio, D. R. Valencia, M. Arcos-Burgos, E. E. Castilla, J. C. Mereb, F. A. Poletta, I. M. Orioli, F. M. Carvalho, J. T. Hecht, S. H. Blanton, C. J. Buxo, A. Butali, P. A. Mossey, W. L. Adeyemo, O. James, R. O. Braimah, B. S. Aregbesola, M. A. Eshete, M. Deribew, M. Koruyucu, F. Seymen, L. Ma, J. E. de Salamanca, S. M. Weinberg, L. Moreno, R. A. Cornell, J. C. Murray, and M. L. Marazita. 2016. 'A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic Missense Variant in GRHL3', *Am J Hum Genet*, 98: 744-54.
- Leslie, E. J., and M. L. Marazita. 2013. 'Genetics of cleft lip and cleft palate', *Am J Med Genet C Semin Med Genet*, 163C: 246-58.
- Little, J., A. Cardy, and R. G. Munger. 2004. 'Tobacco smoking and oral clefts: a meta-analysis', *Bull World Health Organ*, 82: 213-8.
- Mai, C. T., J. L. Isenburg, M. A. Canfield, R. E. Meyer, A. Correa, C. J. Alverson, P. J. Lupo, T. Riehle-Colarusso, S. J. Cho, D. Aggarwal, R. S. Kirby, and Network National Birth Defects Prevention. 2019. 'National population-based estimates for major birth defects, 2010-2014', *Birth Defects Res*, 111: 1420-35.
- Marçano, A. C., K. Doudney, C. Braybrook, R. Squires, M. A. Patton, M. M. Lees, A. Richieri-Costa, A. C. Lidral, J. C. Murray, G. E. Moore, and P. Stanier. 2004. 'TBX22 mutations are a frequent cause of cleft palate', *J Med Genet*, 41: 68-74.
- Nasreddine, Ghenwa, Joelle El Hajj, and Michella Ghassibe-Sabbagh. 2021. 'Orofacial clefts embryology, classification, epidemiology, and genetics', *Mutation Research/Reviews in Mutation Research*, 787: 108373.
- Peeters, S., P. De Kinderen, J. A. N. Meester, A. Verstraeten, and B. L. Loeys. 2022. 'The fibrillinopathies: New insights with focus on the paradigm of opposing phenotypes for both FBN1 and FBN2', *Hum Mutat*, 43: 815-31.
- Pinto, Louise L. C., Taiane A. Vieira, Roberto Giugliani, and Ida V. D. Schwartz. 2010. 'Expression of the disease on female carriers of X-linked lysosomal disorders: a brief review', *Orphanet Journal of Rare Diseases*, 5: 14.
- Pottie, Lore, Christin S. Adamo, Aude Beyens, Steffen Lütke, Piyanoot Tapaneeyaphan, Adelbert De Clercq, Phil L. Salmon, Riet De Rycke, Alper Gezdirici, Elif Yilmaz Gulec, Naz Khan, Jill E. Urquhart, William G. Newman, Kay Metcalfe, Stephanie Efthymiou, Reza Maroofian, Najwa Anwar, Shazia Maqbool, Fatima Rahman, Ikhlass Altweijri, Monerah Alsaleh, Sawsan Mohamed Abdullah, Mohammad Al-Owain, Mais Hashem, Henry Houlden, Fowzan S. Alkuraya, Patrick Sips, Gerhard Sengle, and Bert Callewaert. 2021. 'Bi-allelic premature truncating variants in *LTBP1* cause cutis laxa syndrome', *The American Journal of Human Genetics*, 108: 1095-114.
- Recouvreux, M. V., L. Lapyckyj, M. A. Camilletti, M. C. Guida, A. Ornstein, D. B. Rifkin, D. Becu-Villalobos, and G. Díaz-Torga. 2013. 'Sex differences in the pituitary transforming growth factor- β 1 system: studies in a model of resistant prolactinomas', *Endocrinology*, 154: 4192-205.
- Resnick, Cory M., Judy A. Estroff, Tessa D. Kooiman, Carly E. Calabrese, Maarten J. Koudstaal, and Bonnie L. Padwa. 2018. 'Pathogenesis of Cleft Palate in Robin Sequence: Observations From Prenatal Magnetic Resonance Imaging', *Journal of Oral and Maxillofacial Surgery*, 76: 1058-64.
- Robinson, Kelsey, Trenell J. Mosley, Kenneth S. Rivera-González, Christopher R. Jabbarpour, Sarah W. Curtis, Wasiu Lanre Adeyemo, Terri H. Beaty, Azeez Butali, Carmen J. Buxó,

- David J. Cutler, Michael P. Epstein, Lord J. J. Gowans, Jacqueline T. Hecht, Jeffrey C. Murray, Gary M. Shaw, Lina Moreno Uribe, Seth M. Weinberg, Harrison Brand, Mary L. Marazita, Robert J. Lipinski, and Elizabeth J. Leslie. 2023. 'Trio-based GWAS identifies novel associations and subtype-specific risk factors for cleft palate', *Human Genetics and Genomics Advances*, 4.
- Savidge, J., D. Colville, M. Rheault, S. Gear, R. Lennon, S. Lagas, M. Finlay, and F. Flinter. 2016. 'Alport Syndrome in Women and Girls', *Clin J Am Soc Nephrol*, 11: 1713-20.
- Schwender H, Li Q, Berger P, Neumann C, Taub M, Ruczinski I. 2023. 'trio: Testing of SNPs and SNP Interactions in Case-Parent Trio Studies'.
- Sivertsen, Ase, Allen J. Wilcox, Rolv Skjaerven, Hallvard Andreas Vindenes, Frank Abyholm, Emily Harville, and Rolv Terje Lie. 2008. 'Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives', *BMJ (Clinical research ed.)*, 336: 432-34.
- Sokal, R., L. J. Tata, and K. M. Fleming. 2014. 'Sex prevalence of major congenital anomalies in the United Kingdom: a national population-based study and international comparison meta-analysis', *Birth Defects Res A Clin Mol Teratol*, 100: 79-91.
- Taya, Y., S. O'Kane, and M. W. Ferguson. 1999. 'Pathogenesis of cleft palate in TGF-beta3 knockout mice', *Development*, 126: 3869-79.
- Tennant, Peter W.G., S. Dilhani Samarasekera, Tanja Pless-Mulloli, and Judith Rankin. 2011. 'Sex differences in the prevalence of congenital anomalies: A population-based study', *Birth Defects Research Part A: Clinical and Molecular Teratology*, 91: 894-901.
- Tran, D. Q., J. Andersson, R. Wang, H. Ramsey, D. Unutmaz, and E. M. Shevach. 2009. 'GARP (LRRC32) is essential for the surface expression of latent TGF-beta on platelets and activated FOXP3+ regulatory T cells', *Proc Natl Acad Sci U S A*, 106: 13445-50.
- Wang, K., M. Li, and H. Hakonarson. 2010. 'ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data', *Nucleic Acids Res*, 38: e164.
- Wang, R., J. Zhu, X. Dong, M. Shi, C. Lu, and T. A. Springer. 2012. 'GARP regulates the bioavailability and activation of TGFβ', *Molecular biology of the cell*, 23: 1129-39.
- Yoshida, S., M. Takeuchi, C. Kawakami, K. Kawakami, and S. Ito. 2020. 'Maternal multivitamin intake and orofacial clefts in offspring: Japan Environment and Children's Study (JECS) cohort study', *BMJ Open*, 10: e035817.