Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world-wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Daniel Chang Date

Identifying Genetic Modifiers for Left and Right Sided Orofacial Clefts

By

Daniel Chang
Master of Science in Public Health

Department of Epidemiology

_____

Jennifer Mulle, Ph.D., MHS
Committee Chair

_____

Elizabeth Leslie, Ph.D.
Committee Member

Identifying Genetic Modifiers for Left and Right Sided Orofacial Clefts

By

Daniel Chang

Master of Health Science
Johns Hopkins University
2016

Bachelor of Science
Johns Hopkins University
2014

Thesis Committee Chair: Jennifer Mulle, Ph.D., MHS

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Epidemiology
2020

Abstract

Identifying Genetic Modifiers for Left and Right Sided Orofacial Clefts

By Daniel Chang

Orofacial clefts (OFC) are a common congenital malformation that affect 1 in 700 births globally. Affected individuals are also impacted by other health conditions, increased mortality at all ages, and face additional financial burdens associated with years of treatment. OFCs are classified by sub-type, including cleft lip, cleft lip with palate, and cleft lip with or without palate. Additionally, OFC can present as either bilateral, occurring on both sides of the lip/palate, or unilateral, affecting only a single side. The current paradigm in OFC research and treatment is to group unilateral left and unilateral right clefts together for research studies without accounting for subtypes. However, the non-random distribution of OFC sub-types suggests that these subtypes may be pathogenetically distinct and subtype-specific analyses should be considered. One approach to identify genetic factors contributing to OFC subtypes is to compare cases of OFC sub-types against cases of a different OFC sub-type in a genome-wide association study (GWAS). This approach identifies genetic modifiers that distinguish two subtypes and allows us to develop a better understanding of genetic factors that contribute to differences between OFCs. We analyzed a multi-ethnic dataset comprised of singleton 702 cases, 1293 case-parent trios, and 1626 controls using this modifier approach and identified a candidate 4q28 region with suggestive evidence of association ($p=8.4x10^{-8}$) and opposite direction of effect between right-sided cleft lip (OR=0.64, 0.45-0.90) and left-sided cleft lip (OR=1.84, 1.43-2.37).

Identifying Genetic Modifiers for Left and Right Sided Orofacial Clefts

By

Daniel Chang

Master of Health Science
Johns Hopkins University
2016

Bachelor of Science
Johns Hopkins University
2014

Thesis Committee Chair: Jennifer Mulle, Ph.D., MHS

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Epidemiology
2020

Acknowledgements

I would like to use this opportunity to thank the various individuals who have helped me through my time at Emory University. I would like to thank Dr. Jennifer Mulle for her guidance throughout the thesis process, as well as her advisement through the GME program. I also wish to thank Dr. Elizabeth Leslie for all of the opportunities and mentorship she provided from the very beginning of my time at Emory University. This study was only possible with her support. I also wish to acknowledge all the members of the Leslie lab for their help and advice, especially Kim and Courtney for their continued friendship. I would also like to thank Lili, Yijin, and David for their moral support during our time in the Department of Epidemiology. Finally, I would also like to thank my family for the support they have shown while I pursued the MSPH program, without them I would not have been able to accomplish all that I have during my time at Emory.

Table of Contents

INTRODUCTION

Background on OFC:

Orofacial clefts (OFC) are common congenital craniofacial anomalies that affect 1 in 700 births globally. In the U.S., cleft lip with or without palate affects 10.5 per 10,000 live births while 6.4 per 10,000 live births are affected by cleft palate only (1). In addition to early childhood complications, affected individuals are at a higher risk for dental (2-4) and speech problems (5, 6), ear infections (5, 7), various forms of cancer (8, 9), and mental health concerns (10), and these individuals have an overall higher rate of mortality throughout all ages (11). Beyond its impact on the health of affected individuals, OFCs also incur additional financial burdens, where lifetime costs for cleft related surgeries, hospital stays, orthodontic treatments, and speech therapy can exceed $200,000 (12).

OFCs are categorized based on various characteristics, such as the location and severity of the cleft and the presence or absence of other features. OFCs can be categorized as either syndromic, where clefting is accompanied by other cognitive or physical abnormalities, or nonsyndromic, classified by an absence of additional accompanying conditions (12). OFCs are also classified by the affected structures, which include cleft lip only (CL), cleft lip with palate (CLP), cleft lip with or without palate (CL/P), and cleft palate only (CPO). Additionally, OFCs can be grouped by the side of the lip and/or palate where the clefting occurred. Bilateral cases of CL/P (BCL/P), CLP (BCLP), and CL (BCL) are defined as when the cleft occurs on both sides of the face, whereas unilateral cases are when clefting affects only one side of the face. More specifically, unilateral left CL/P (LCL/P), CLP (LCLP), and CL (LCL) refer to when clefting occurs on the left side of the face, while unilateral right CL/P (RCL/P), CLP (RCLP), and CL (RCL) are classified as such when clefting occurs on the right side.

The prevalence of OFC varies by both subtype and classification. 70% of all CL/P cases are classified as nonsyndromic, while 50% of CPO cases are classified as such (13). Additionally, cases of CLP are twice as common as cases of CL (14). In terms of OFC distribution by laterality, among CL cases, unilateral clefting comprises 75% of all cases. Even among unilateral cases, the distribution of laterality is not balanced, with unilateral left cases having twice the frequency of unilateral right cases (15). The prevalence of OFCs also differs by sex and subtype: there are twice as many CL/P subtypes in males than females, while the reverse has been found in CPO cases, with females having twice the frequency (16). The non-random distribution of OFC between syndromic vs nonsyndromic cases, the various subtypes, and the side by which clefting manifests suggests the presence of underlying genetic factors contribute to the differentiation between OFC cases.

Genetics & OFC

Although the exact causal mechanism of OFC is unknown, OFC pathogenesis is multifactorial. There have been evidence for both environmental factors, such as maternal smoking (17), alcohol consumption during pregnancy (18-20), and folate exposure *in utero* (21), and genetic factors contributing to OFC development. Seasonal factors have also been suggested to affect OFC pathogenesis, with seasons affecting the differing prevalence of different sub-types affecting OFC pathogenesis (22, 23). With regards to the genetic components, early consideration originated from segregation analysis (24) and twin studies (25, 26). Studies on patterns of OFC in families supported the concept of genetic pathways, with evidence of increased risk for CL for individuals with a family history of CL (27), as well as findings that monozygotic twins have higher concordance for OFC than dizygotic twins (26). To further study the complex origins of non-syndromic OFCs, large-scale genomic studies employed methods such as genome-wide association studies (GWAS) (28-

32), linkage analysis (33), and candidate gene approaches (34). These genetic approaches have identified an association between heritability with OFC classification and subtypes. The differences in the genetic pathway between syndromic OFCs and non-syndromic OFCs highlight such distinctions. The genetic pathway of syndromic OFCs can be attributed to a single genetic origin, such as a single genetic locus or chromosomal abnormalities (12). On the other hand, non-syndromic OFCs are genetically complex and involve interaction of multiple genetic and environmental risk factors (12). Given OFC's nature as a complex disease, many studies focus on identifying the genetic pathways of nonsyndromic OFCs.

GWAS approaches have been able to identify approximately 40 candidate risk loci for OFC, such as the association between OFC and the 8q24 loci in a European population(29) or the 1q32 loci in an Asian population in a stratified GWAS (28). Specific genes have also been identified to be associated with non-syndromic OFCs, such as *IRF6* (28, 29), *MAFB* (28, 35), and *ARHGAP29* (28, 30) through GWAS methods. Although GWAS approaches and other study methods have identified potential genetic variants that are associated with OFCs as a whole, OFC sub-types are often are grouped together during analyses. Despite this, the genetic components contributing to each OFC sub-types may differ. Grouping sub-types together could potentially overlook factors that are unique to sub-type pathogenesis. This is especially the case regarding for right- and left-sided OFC, which have not been a focus of genetic studies and are usually grouped together into the umbrella category of unilateral CL, CLP, or CL/P. Given the variations observed in body asymmetry, insight may be gained by studying genetic factors that contribute to lateral asymmetry in OFCs as well.

Non-OFC Laterality

Common non-OFC examples of laterality include left/right handedness, facial asymmetry, and asymmetry in the left/right brain hemispheres. The differences in body asymmetry have also been proposed to have clinical implications. There has been evidence of differential schizophrenia prevalence among those with increased directional asymmetry (36) as well as increased facial asymmetry among children with autism spectrum disorder (37). To better understand left/right asymmetry, both environmental and genetic causal paths have been proposed. Multiple genetic analyses have been used to study handedness, including whole exome sequencing (38) and GWAS approaches (39, 40). Direction of handedness has also been connected with non-laterality outcomes, where studies have also identified specific genes associated with handedness and behavioral disorders (41). Given the potential association between genetic risk factors and body sidedness, there may be underlying genetic causal pathways for the differences between right-sided and left-sided OFCs as well.

OFC laterality

The distribution of clefting between unilateral left, unilateral right, and bilateral clefting is non-random, suggesting that specific factors could be identified that influence laterality differences in OFC. Understanding these factors may yield insights into craniofacial development but also has potential to elucidate the relationship between OFC laterality and associated clinical outcomes. For example, the association between non-OFC laterality with behavior may extend to OFCs as studies have explored academic achievement differences with clefting laterality. One study found that children with left-sided cleft had worse academic performance measured by standardized test scores than those with right-sided cleft (42); however, this result has been inconsistent as a follow-up study with a larger cohort by the same group found that, based on clinical evaluations, RCL/P was associated with a

higher risk of requiring academic support when compared to LCL/P (43). This same study also found evidence of varying health outcomes by laterality, with cases of nonsyndromic CL/P associated with a higher risk of additional malformations, such as heart lesions (43). Some preliminary evidence has been shown for differing health outcomes between BCL and UCL following cleft lip repair as well, with BCL being associated with higher prevalence of comorbidities related to pulmonary, gastrointestinal, neurologic, and cardiac diseases when compared to UCL cases (44). Despite evidence of differences between left- and right-sided clefting, there have been no genome-wide studies to determine possible genetic origins for left, right, and bilateral subtypes of nonsyndromic OFC. Studies thus far have examined genetic risk factors for differences between CL, CLP, and CL/P subtypes, yet modifiers for laterality have not been studied. This thesis project could help develop a better understanding of factors that could modify laterality, which could shed valuable insight into overall OFC etiology. Especially with recent findings of the genetic association of other OFC subtypes, we hope to develop a better understanding of genetic variants of laterality modifiers.

We will use genome-wide association studies to identify possible genetics factors that contribute to laterality subtypes of OFCs. The current paradigm in OFC research and treatment is to group unilateral left and unilateral right clefting together to identify genetic factors that increase risk for OFCs in a general sense, without accounting for subtypes. However, increasing evidence suggests these subtypes should be considered pathogenetically distinct (45). Given that OFCs of different laterality have difference rates of prevalence and affect gender and different populations disproportionately, we could also potentially extend this argument to laterality as well and attempt to identify an underlying factor for these subtypes. For example, OFC asymmetry has been studied for association with other health

outcomes and even academic performance (42, 43). By studying whether there are genetic factors associated with the subtypes, we may be able to further define disease etiology, and hopefully learn about the genetic pathways responsible for the variation in subtypes. Because laterality could be seen as a measure of clefting severity, with bilateral OFC being a more severe case than unilateral OFC, understanding the underlying genetic factors and predicting OFC outcomes could provide better clinical responses. For example, cleft lip/palate surgical procedures differ depending on bilateral or unilateral, and a better understanding of the variants could help identify appropriate treatment, ultimately reducing both medical and cost burdens for patients. Thus far, studies have been done to show the underlying genetic differences between CL & CLP, but none have focused on laterality.

METHODS

Study population

Samples were obtained from a multi-ethnic cohort from the Pittsburgh Orofacial Cleft (POFC) study, with recruitment occurring in 18 sites in North, Central, and South America, Asia, and Europe. These recruitment sites were comprised of OFC treatment centers, with many of these centers being a part of genetic studies by the University of Pittsburgh Center for Craniofacial and Dental Genetics and the University of Iowa (46, 47). Recruitment was approved by the IRB of each recruiting site, as well as the IRB of the University of Pittsburgh and University of Iowa, and informed consent was obtained for each research subject (46, 47). Eligibility was determined on whether or not the individual had OFC. Cases for analyses were selected from unrelated individuals who have cleft lip either with or without palate. Controls were defined as individuals unrelated to cases who have no known history of OFC or other craniofacial anomalies. The final analysis cohort consisted of a total of 702 cases, 1293 trios and 1626 controls.

Genotyping

Samples were genotyped for a combination of 580,000 SNPs using the Illumina HumanCore+Exome platform. An additional 15,980 SNPs were also genotyped to include candidate genes and loci that previous studies have found to be associated with OFCs (46). The dataset analyzed in this dataset underwent quality control using pipelines developed by the University of Washington Genetics Coordinating Center (48). This process involved examining samples for duplicates, batch effects, chromosomal anomalies, familial relatedness, Mendelian errors among relatives, and population structure (46). SNP probe quality was also inspected by examining inter-sample comparisons, missing call rates, separation of clusters during genotype calling, and deviations from the Hardy-Weinberg

equilibrium. After quality control, the final number of genotyped SNPs was 539,473, with 293,633 SNPs having a minor allele frequency 1% or greater (46, 47).

Additionally, unobserved genetic polymorphisms were imputed using the IMPUTE2 software in conjunction with the 1000 Genome Projects (phase 3) as the reference panel, with haplotypes created using the SHAPEIT2 software to "pre-phase" genotyped SNPs that have passed quality control (46, 47). After imputation, masked variant analysis was done to assess imputation accuracy. This resulted in a mean concordance of 0.955 when MAF <0.05 but a mean concordance of 0.960 for SNPs with MAF ≥ 0.05, indicating high-quality imputation (46, 47). Incorporating the final list of imputed SNPs with the genotyped data resulted in 34,985,077 SNPs for analysis.

Modifier GWAS Comparison Groups

SNPs associated with differences between OFC sub-types were identified by conducting GWAS using a modifier approach. In this approach, cases of an OFC sub-type were compared against cases of another sub-type. Initial analysis was conducted on BCL/P cases against unilateral CL/P cases (defined as all cases of either unilateral right or unilateral left cleft lip). Subsequent analyses were conducted separately for BCL against unilateral CL as well as BCLP and unilateral CLP. Similarly, unilateral left and right-sided clefts were compared to assess sidedness differences. Among the CL/P cases, the following groups were analyzed: LCL/P against BCL/P, RCL/P against BCL/P, and LCL/P against RCL/P. The same set of analyses were conducted for CLP, with LCLP analyzed against RCLP, RCLP against BCLP, and LCLP against RCLP. Analysis was done with the CL cases as well, with LCL against BCL, RCL against BCL, and LCL against RCL.

Case-control GWAS Approach

After candidate SNPs were identified in the modifier analyses, additional association tests were conducted using the traditional case-control method to ascertain the direction of effects for each SNP using a common reference group. For the case-control analysis, we examined case groups as previously defined against the 1626 controls.

GWAS Parameters

Given the multi-ethnic nature of the data and to account for the population structure within the dataset, a principal components of ancestry (PCA) analysis was conducted, which resulted in 18 principal components of ancestry (46, 47). Incorporating these PCs as covariates in the model allows us to adjust for the allelic distribution of SNPs resulting from ancestry-specific variations within the population. For each of comparison groups, including the modifier approach and case-control approach, GWAS was conducted using an additive logistic regression and adjusting for both sex and principal components of ancestry using PLINK v1.90b5.3 (49). For an association with a SNP to be considered to be genome-wide significant, the association must meet the p-value threshold of $5x10^{-8}$. Associations were also considered suggestive if p-value < $1x10^{-5}$. SNPs were excluded from analysis if they had a with a minor allele frequency less than 0.05, were not in Hardy–Weinberg equilibrium in controls, or had an imputation information score less than 0.5. Data cleaning, analysis of GWAS results, and generation of figures were done on R v3.6.1 (50-52). Regional association plots were created using LocusZoom (53).

FAT4 Gene Expression Analysis

To determine biological relevance of the 4q28 locus, the Lipinski lab at the University of Wisconsin conducted analyses to characterize the expression of the *FAT4* candidate gene by *in situ* hybridization of mouse embryos using standard protocols.

RESULTS

Population Characteristics:

The distribution of OFC cases in our analysis population was consistent with patterns described in the literature. We observed a greater number of OFC cases in males than in females and a greater number of unilateral cases overall. Among the unilateral OFC cases, there were a greater number of unilateral left cases than unilateral right. This trend was observed for each of the CL/P, CL, and CLP sub-types. Additionally, within our dataset, the majority of OFC cases were male (Table 1). This remained the case after stratifying samples by each sub-type, consistent with the reported male-bias of cleft lip in population-based epidemiological studies of OFCs. We also observed the number of unilateral left cases was approximately 1.5 times the number of unilateral right cases in CL/P, CL, and CLP. When examining distribution by ethnicity, there was uneven distribution of cases by ancestry groups, Among the CL/P and CLP cases, the majority of our cases were of Latin/South American ancestry, followed by the Caucasian, and the Asian ancestry groups having the fewest number of samples. In the CL sub-type, there was an even distribution of cases between the ancestry groups.

Modifier GWAS Results

GWAS was conducted on this multi-ethnic cohort by comparing sub-types of OFC cases against sub-types. A total of 1995 cases were analyzed across all GWAS conducted. The comparison between bilateral clefting against unilateral clefting did not result in any associations at the genome-wide significance level, only SNPs at the suggestive level of significance. The GWAS results for the bilateral against unilateral comparison among CL/P resulted in 57 SNPs across 19 loci that meeting the suggestive threshold for association. In the bilateral against unilateral analysis of the CLP cases, 54 SNPs across 15 loci were found

to be suggestive, while the bilateral against unilateral comparison among CL cases yielded 10 SNPs across 7 loci as suggestive.

We next compared left-sided OFCs against right-side OFCs. In this series of analyses, 29 SNPs across 10 loci were at suggestive significance among CL/P cases, 56 SNPs across 10 loci among the CLP cases, and 31 SNPs across 13 loci among CL cases. Again, we did not identify any associations at genome-wide significance in this analysis (Fig 1, Table 3), however, in the LCL vs. RCL modifier analysis, a locus on chromosome 4q28 approached genome-wide significance, with the lead SNP rs6855309 (p-value = 8.4×10^{-8}; OR = 3.5; 95% CI: 2.21-5.55) (Figure 2C, Table 3).

Overall, the analyses for CL and CLP appeared to differ but comparisons based on p-values alone are confounded by differences in sample size. In order to fairly compare the results of the CL and CLP analyses, we plotted the effects (log odds ratios) of several lead SNPs from the CLP modifier against the effects of the same SNPs in the CL analyses and vice versa (Figure 3) Interestingly, SNPs that were suggestive for significance in the CL analysis had no modifier effect in the CLP analysis and similarly, the most significant SNPs from the CLP had no modifier effect in the CL analysis. Therefore, these loci identified in the CL analysis are truly specific to a subtype and differences in the association results are not likely due to sample size. Given the specificity of the 4q28 locus, we further explored the region by plotting the odds ratio of rs6855309 resulting from the unilateral right against unilateral left analyses for the CL, CLP, and CL/P sub-types. This plot reveals an association is observed only in the CL sub-type (Figure 4A). No associations were found for rs6855309 for the CL/P and CLP analyses (Figure 4B, 4C). Even within the CL sub-type, an association was present only when comparing unilateral left against unilateral right, but not in the

bilateral against unilateral combined analysis nor the bilateral against unilateral right or bilateral against unilateral left analysis (Figure 4D).

Case-Control GWAS Results

To better understand the 4q28 region, we conducted a case-control GWAS with unilateral right cases vs control and unilateral left cases vs control in both CL and CLP sub-types. A total of 546 right-sided OFCs and 854 left-sided OFCs were compared against 1626 controls. GWAS of right-sided CL case-control analysis resulted in an odds ratio of 0.64 (95% CI: 0.45, 0.90) for rs6855309, while the left-side CL analysis resulted in an odds ratio of 1.84 (95% CI: 1.43, 2.37). GWAS of the unilateral left CLP case-control results in an odds ratio of 1.05 (95% CI: 0.89, 1.23) while the unilateral right CLP analysis resulted in an odds ratio of 1.21 (95% CI: 0.99, 1.46). These results are plotted on Figure 5 to highlight the opposite direction of effect. We then examined the LocusZoom regional association plot to further examine the 4q28 loci (Figure 6). The regional plot indicated that presence of the *FAT4* gene proximate to the 4q28 region.

DISCUSSION

GWAS Outcome

OFCs are a heterogeneous group of disorders that share various sub-types. To better understand the underlying genetic factors that may result in OFC sub-types, we conducted a modifier-based approach to GWAS in a multi-ethnic cohort. Our analysis identified a candidate genetic locus with evidence of suggestive association at the 4q28 region, which was specific to the CL sub-type while having opposite direction of effect between RCL and LCL. Additionally, we identified a gene in proximity of the 4q28 locus which may be a candidate gene.

Traditional approaches to GWAS compare cases with a given disease against controls; however, in order to better understand genetic risk factors that underlie the OFC sub-types, we conducted GWAS using a modifier-based approach, which cannot detect loci associated with both subtypes and therefore highlights the modifier loci that distinguish two subtypes. By using this strategy, we hoped to identify potential genetic differences that underlie the laterality and sidedness types of OFCs. To do this, we examined bilateral and unilateral OFC as well as left- and right-side clefting. These analyses were conducted on the CL, CLP, and CL/P sub-types.

Because both CL and CLP have the same 2:1 bias toward unilateral left cleft lip, we initially expected that the CL/P modifier analysis would have the most power to uncover modifying loci. However, no evidence of association was observed at genome-wide significance, given a threshold of p-value < $5x10^{-8}$ (Table 2) in either the laterality or sidedness modifiers. We considered the possibility that potential associations may be lost by combining both CL and CLP groups and conducted analysis separated by these subtypes, revealing the 4q28 locus reached near genome-wide significance in the LCL versus RCL

analysis (Table 2). Additional comparisons of the effect sizes of several lead SNPs from the CL and CLP analyses mostly indicated that the effects of the regions were specific to their respective analysis (Figure 3), with SNPs that were suggestive for significance in the CL analysis (in red in Figure 3) found to have nearly no effect in the CLP analysis and vice versa.

The increased effect size of the 4q28 region was observed only in the CL analysis (Figure 4A), with the lead SNP resulting in a statistically significant odds ratio of 3.51 (95% CI: 2.2, 5.4) (Figure 2, Table 3), reinforcing the possibility that the effect of the 4q28 region is specific to the CL sub-type. To determine the direction of effect of the 4q28 region, a traditional case-control GWAS approach was taken, which revealed an opposite direction of effect for the rs6855309 SNP between left-sided and right-sided CL (Figure 5).

Previous studies focused mainly on studying OFC as a whole, but some previous studies have focused on identifying loci associated with individual sub-types. For example, the CLP sub-type has been associated with the 13q31 (54) and 15q13 loci (55). Additionally, there have been evidence of differential effect modification by a gene on different OFC sub-types, where *IRF6* carries a stronger risk for the CL sub-type compared to the CLP sub-type (56). By identifying a candidate region that is both specific to CL and has an opposite direction of effect between LCL and RCL, we have found suggestive evidence for genetic factors that underlie a mechanism contributing to the differentiation of right-side and left-side CL. By doing so, we provide support for the consideration of analyze unilateral right and unilateral left CL as distinct disorders and warrants a more detailed examination of the 4q28 locus.

FAT4

The associated SNPs in the 4q28 region reside in an intergenic region. The closest protein coding gene was *FAT4,* located approximately 400Kb centromeric to the associated

region (Figure 6). Mutations in *FAT4* have been linked to both Hennekam syndrome (57, 58) and Van Maldergem syndrome(58, 59). Hennekam syndrome is a disorder resulting from malformations in the lymphatic system and can cause facial dysmorphism (60) while Van Maldergerm syndrome is characterized by intellectual disabilities, craniofacial abnormalities, and other skeletal malformations (61). In addition to these syndromes, missense variants in *FAT4* were associated with OFCs in an extended Syrian pedigree (62). We found that *FAT4* is expressed in mouse embryos along the medial nasal process at gestational day 10.75 and gestational day 11.75 (Figure 7). At this developmental stage, the medial nasal process is a precursor to the upper lip, with previous work having found that disrupting the proliferation of mesenchymal cell in the medial nasal process results in cleft lip (63). Therefore, although the two syndromes commonly associated with *FAT4* are not characterized by OFCs, these data show that *FAT4* is expressed at the proper time and place to be involved in the differentiation between left-sided and right-sided cleft lip. Further analysis of the expression and function of *FAT4* in this region could help elucidate possibility of the gene's involvement in the development of left/right differentiation in cleft lip.

Limitations/Future Direction:

Although we identified a promising candidate locus and gene, one limitation of our analysis was the sample size in our analysis cohort, especially among BCL cases, which is the rarest of all OFC subtypes. Because of the diversity of the samples, we had to accounted for population structure by adjusting for a large number of principal components of ancestry, which can reduce statistical power. Larger samples of each ancestry group could provide more statistical power and the ability to identify population-specific modifiers. Similarly, we would like to conduct replication analyses, however, we are limited by the availability of replication data sets, the sample size in the replication set, and the accuracy/completeness of

the phenotype data. Finally, there are computational limitations due to the size of genomic datasets, especially when increasing the sample size; however, newer technologies and informatics methods can alleviate this concern.

Future Direction

The results of the unilateral left versus unilateral right analysis among the CL cases were at the level of suggestive significance. Expanding the sample size and replicating the findings in the 4q28 region would strengthen the findings listed here, however, as mentioned, there are limitations by availability of both genetic and phenotypic data. In addition to replication studies, utilizing candidate gene approaches could also allow us to better characterize the association between *FAT4* and unilateral left and unilateral right CL. Molecular experiments would also help develop a better understanding the potential role of *FAT4*. Functional analysis via overexpression or knock-out mouse models would allow us to determine how the 4q28 SNPs affect *FAT4* expression and would be valuable in developing a better understanding of association between the 4q28 region, the *FAT4* gene, and OFCs.

Conclusion

To develop a better understanding of the differences of the sub-types underlying OFCs, we utilized a modifier-based approach to identify any associations underlying the differences between the comparison groups. This approach allowed us to identify a candidate region, the 4q28 locus, that may be involved in left-side and right-side differentiation in cleft lip. The results of the analysis indicated that there is opposite direction of effect in this region between right-sided and left-sided CL. These results highlight the heterogenous nature of OFCs and how potential genetic mechanisms for OFC pathogenesis may be overlooked by the current paradigm of analyzing OFC sub-types together. Instead, approaching the different subtypes as unique disease entities may be required to fully

character OFC. We also identified the *FAT4* gene in proximity to this region, which may hold further insight into laterality differences of cleft lip. Potential insight into the varying health outcomes associated with the different subtypes could ultimately have clinical implications, such as utilizing genetic markers for screening or genetic counseling, to improve the public health outcome of those affected by OFCs.

REFERENCES

1. Rahimov F, Jugessur A, Murray JC. Genetics of nonsyndromic orofacial clefts. *Cleft Palate Craniofac J* 2012;49(1):73-91.
2. Akcam MO, Evirgen S, Uslu O, et al. Dental anomalies in individuals with cleft lip and/or palate. *The European Journal of Orthodontics* 2010;32(2):207-13.
3. Ribeiro LL, das Neves LT, Costa B, et al. Dental development of permanent lateral incisor in complete unilateral cleft lip and palate. *Cleft Palate Craniofac J* 2002;39(2):193-6.
4. Lourenco Ribeiro L, Teixeira Das Neves L, Costa B, et al. Dental anomalies of the permanent lateral incisors and prevalence of hypodontia outside the cleft area in complete unilateral cleft lip and palate. *Cleft Palate Craniofac J* 2003;40(2):172-5.
5. Jocelyn LJ, Penko MA, Rode HL. Cognition, communication, and hearing in young children with cleft lip and palate and in control children: a longitudinal study. *Pediatrics* 1996;97(4):529-34.
6. Mort K. Seech characteristics of individuals with cleft lip and palate. In: Spriestersbach DC, Sherman D, eds. *Cleft Palate and Communication*. New York: Academic Press, 1968.
7. Nackashi JA, Dedlow RE, Dixon-Wood V. Health care for children with cleft lip and palate: comprehensive services, and infant feeding. In: Wyszynski D, ed. *Cleft Lip and Palate from Origin to Treatment* New York, NY: Oxford University Press, 2002: p.127-58.
8. Dietz A, Pedersen DA, Jacobsen R, et al. Risk of breast cancer in families with cleft lip and palate. *Annals of epidemiology* 2012;22(1):37-42.
9. Menezes R, Marazita ML, McHenry TG, et al. AXIS inhibition protein 2, orofacial clefts and a family history of cancer. *The Journal of the American Dental Association* 2009;140(1):80-4.
10. Christensen K, Mortensen PB. Facial clefting and psychiatric diseases: a follow-up of the Danish 1936–1987 Facial Cleft cohort. *The Cleft palate-craniofacial journal* 2002;39(4):392-6.
11. Christensen K, Juel K, Herskind AM, et al. Long term follow up study of survival associated with cleft lip and palate at birth. *Bmj* 2004;328(7453):1405.
12. Leslie EJ, Marazita ML. Genetics of cleft lip and cleft palate. *Am J Med Genet C Semin Med Genet* 2013;163C(4):246-58.
13. Tolarová MM, Cervenka J. Classification and birth prevalence of orofacial clefts. *American journal of medical genetics* 1998;75(2):126-37.
14. Jensen BL, Kreiborg S, Dahl E, et al. Cleft lip and palate in Denmark, 1976-1981: epidemiology, variability, and early somatic development. *Cleft Palate J* 1988;25(3):258-69.
15. Gundlach KKH, Maus C. Epidemiological studies on the frequency of clefts in Europe and world-wide. *Journal of Cranio-Maxillofacial Surgery* 2006;34:1-2.
16. Mossey PA, Little J, Munger RG, et al. Cleft lip and palate. *The Lancet* 2009;374(9703):1773-85.
17. Shi M, Wehby GL, Murray JC. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. *Birth Defects Research Part C: Embryo Today: Reviews* 2008;84(1):16-29.

18. Murray J. Gene/environment causes of cleft lip and/or palate. *Clinical genetics* 2002;61(4):248-56.
19. Jugessur A, Shi M, Gjessing HK, et al. Genetic determinants of facial clefting: analysis of 357 candidate genes using two national cleft studies from Scandinavia. *PLoS One* 2009;4(4).
20. Boyles AL, DeRoo LA, Lie RT, et al. Maternal alcohol consumption, alcohol metabolism genes, and the risk of oral clefts: a population-based case-control study in Norway, 1996–2001. *American journal of epidemiology* 2010;172(8):924-31.
21. Wehby GL, Cassell CH. The impact of orofacial clefts on quality of life and healthcare use and costs. *Oral Dis* 2010;16(1):3-10.
22. Elliott RF, Jovic G, Beveridge M. Seasonal variation and regional distribution of cleft lip and palate in Zambia. *The Cleft palate-craniofacial journal* 2008;45(5):533-8.
23. Sofianos C, Christofides EA, Phiri SE. Seasonal Variation of Orofacial Clefts. *Journal of Craniofacial Surgery* 2018;29(2):368-71.
24. Marazita ML, Spence MA, Melnick M. Genetic analysis of cleft lip with or without cleft palate in Danish kindreds. *Am J Med Genet* 1984;19(1):9-18.
25. Mitchell LE. Mode of inheritance of oral clefts. In: Wyszyski DF, ed. *Cleft Lip and Palate: From Origin to Treatment*: Oxford University Press, 2002:234-9.
26. Little J, Bryan E. Congenital anomalies in twins. *Semin Perinatol* 1986;10(1):50-64.
27. Sivertsen Å, Wilcox AJ, Skjærven R, et al. Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives. *Bmj* 2008;336(7641):432-4.
28. Beaty TH, Murray JC, Marazita ML, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nature genetics* 2010;42(6):525.
29. Birnbaum S, Ludwig KU, Reutter H, et al. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nature genetics* 2009;41(4):473.
30. Ludwig KU, Mangold E, Herms S, et al. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nature genetics* 2012;44(9):968.
31. Mangold E, Ludwig KU, Birnbaum S, et al. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nature genetics* 2010;42(1):24-6.
32. Grant SF, Wang K, Zhang H, et al. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *The Journal of pediatrics* 2009;155(6):909-13.
33. Marazita ML, Murray JC, Lidral AC, et al. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. *The American Journal of Human Genetics* 2004;75(2):161-73.
34. Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. *Human molecular genetics* 2004;13(suppl_1):R73-R81.
35. Lennon CJ, Birkeland AC, Nuñez JAP, et al. Association of candidate genes with nonsyndromic clefts in Honduran and Colombian populations. *The Laryngoscope* 2012;122(9):2082-7.

36. Hennessy RJ, Lane A, Kinsella A, et al. 3D morphometrics of craniofacial dysmorphology reveals sex-specific asymmetries in schizophrenia. *Schizophr Res* 2004;67(2-3):261-8.
37. Hammond P, Forster-Gibson C, Chudley AE, et al. Face-brain asymmetry in autism spectrum disorders. *Mol Psychiatry* 2008;13(6):614-23.
38. Kavaklioglu T, Ajmal M, Hameed A, et al. Whole exome sequencing for handedness in a large and highly consanguineous family. *Neuropsychologia* 2016;93(Pt B):342-9.
39. McManus IC, Davison A, Armour JA. Multilocus genetic models of handedness closely resemble single-locus models in explaining family data and are compatible with genome-wide association studies. *Ann N Y Acad Sci* 2013;1288:48-58.
40. Eriksson N, Macpherson JM, Tung JY, et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010;6(6):e1000993.
41. Francks C, Maegawa S, Lauren J, et al. LRRTM1 on chromosome 2p12 is a maternally suppressed gene that is associated paternally with handedness and schizophrenia. *Mol Psychiatry* 2007;12(12):1129-39, 057.
42. Gallagher ER, Collett BR, Barron S, et al. Laterality of Oral Clefts and Academic Achievement. *Pediatrics* 2017;139(2).
43. Gallagher ER, Siebold B, Collett BR, et al. Associations between laterality of orofacial clefts and medical and academic outcomes. *Am J Med Genet A* 2018;176(2):267-76.
44. Chouairi F, Mets EJ, Torabi SJ, et al. Cleft lip repair: are outcomes between unilateral and bilateral clefts comparable? *Journal of Plastic Surgery and Hand Surgery* 2020;54(1):29-32.
45. Genisca AE, Frias JL, Broussard CS, et al. Orofacial clefts in the National Birth Defects Prevention Study, 1997-2004. *Am J Med Genet A* 2009;149A(6):1149-58.
46. Leslie EJ, Liu H, Carlson JC, et al. A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic Missense Variant in GRHL3. *Am J Hum Genet* 2016;98(4):744-54.
47. Leslie EJ, Carlson JC, Shaffer JR, et al. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. *Hum Mol Genet* 2016;25(13):2862-72.
48. Laurie CC, Doheny KF, Mirel DB, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic epidemiology* 2010;34(6):591-602.
49. Purcell S NB, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics* 2007;81.
50. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/, 2013.
51. Gordon M, Lumley T. forestplot: Advanced Forest Plot Using 'grid' Graphics. R package version 1.9. https://CRAN.R-project.org/package=forestplot, 2019.
52. Turner SD. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv* 2014:005165.
53. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26(18):2336-7.

54. Jia Z, Leslie EJ, Cooper ME, et al. Replication of 13q31. 1 association in nonsyndromic cleft lip with cleft palate in Europeans. *American journal of medical genetics Part A* 2015;167(5):1054-60.
55. Ludwig KU, Ahmed ST, Böhmer AC, et al. Meta-analysis reveals genome-wide significance at 15q13 for nonsyndromic clefting of both the lip and the palate, and functional analyses implicate GREM1 as a plausible causative gene. *PLoS genetics* 2016;12(3).
56. Rahimov F, Marazita ML, Visel A, et al. Disruption of an AP-2α binding site in an IRF6 enhancer is associated with cleft lip. *Nature genetics* 2008;40(11):1341.
57. Alders M, Al-Gazali L, Cordeiro I, et al. Hennekam syndrome can be caused by FAT4 mutations and be allelic to Van Maldergem syndrome. *Human genetics* 2014;133(9):1161-7.
58. Ivanovski I, Akbaroghli S, Pollazzon M, et al. Van Maldergem syndrome and Hennekam syndrome: Further delineation of allelic phenotypes. *American Journal of Medical Genetics Part A* 2018;176(5):1166-74.
59. Cappello S, Gray MJ, Badouel C, et al. Mutations in genes encoding the cadherin receptor-ligand pair DCHS1 and FAT4 disrupt cerebral cortical development. *Nature genetics* 2013;45(11):1300.
60. Lister Hill National Center for Biomedical Communications. Hennekam syndrome. https://ghr.nlm.nih.gov/condition/hennekam-syndrome; 2020. (Accessed).
61. Mansour S, Swinkels M, Terhal PA, et al. Van Maldergem syndrome: further characterisation and evidence for neuronal migration abnormalities and autosomal recessive inheritance. *European journal of human genetics* 2012;20(10):1024-31.
62. Holzinger ER, Li Q, Parker MM, et al. Analysis of sequence data to identify potential risk variants for oral clefts in multiplex families. *Molecular genetics & genomic medicine* 2017;5(5):570-9.
63. Everson JL, Fink DM, Yoon JW, et al. Sonic hedgehog regulation of Foxf2 promotes cranial neural crest mesenchyme proliferation and is disrupted in cleft lip morphogenesis. *Development* 2017;144(11):2082-91.

TABLES & FIGURES

Table 1: Singleton cases, trios, and controls by recruitment site

Population	Recruitment Site	Control	Singleton Cases	Trio	Total Cases
European	Denmark	0	18	28	46
	Hungary	253	23	82	105
	Spain	0	2	31	33
	Turkey	171	57	115	172
	United States	411	70	150	220
	Total	835	170	406	576
Central/South American	Argentina	30	68	43	111
	Colombia	277	276	405	681
	Guatemala	208	47	55	102
	Puerto Rico	106	33	51	84
	United States	5	25	47	72
	Total	626	449	601	1050
Asian	China	27	32	125	157
	India	38	18	32	51
	Philippines	96	32	127	159
	Total	161	83	284	367
Unspecified Ancestry Group		4	0	2	2
Total		1626	702	1293	1995

Table 2: Non-random distribution of OFC cases by sub-type that differ by gender & ancestry.

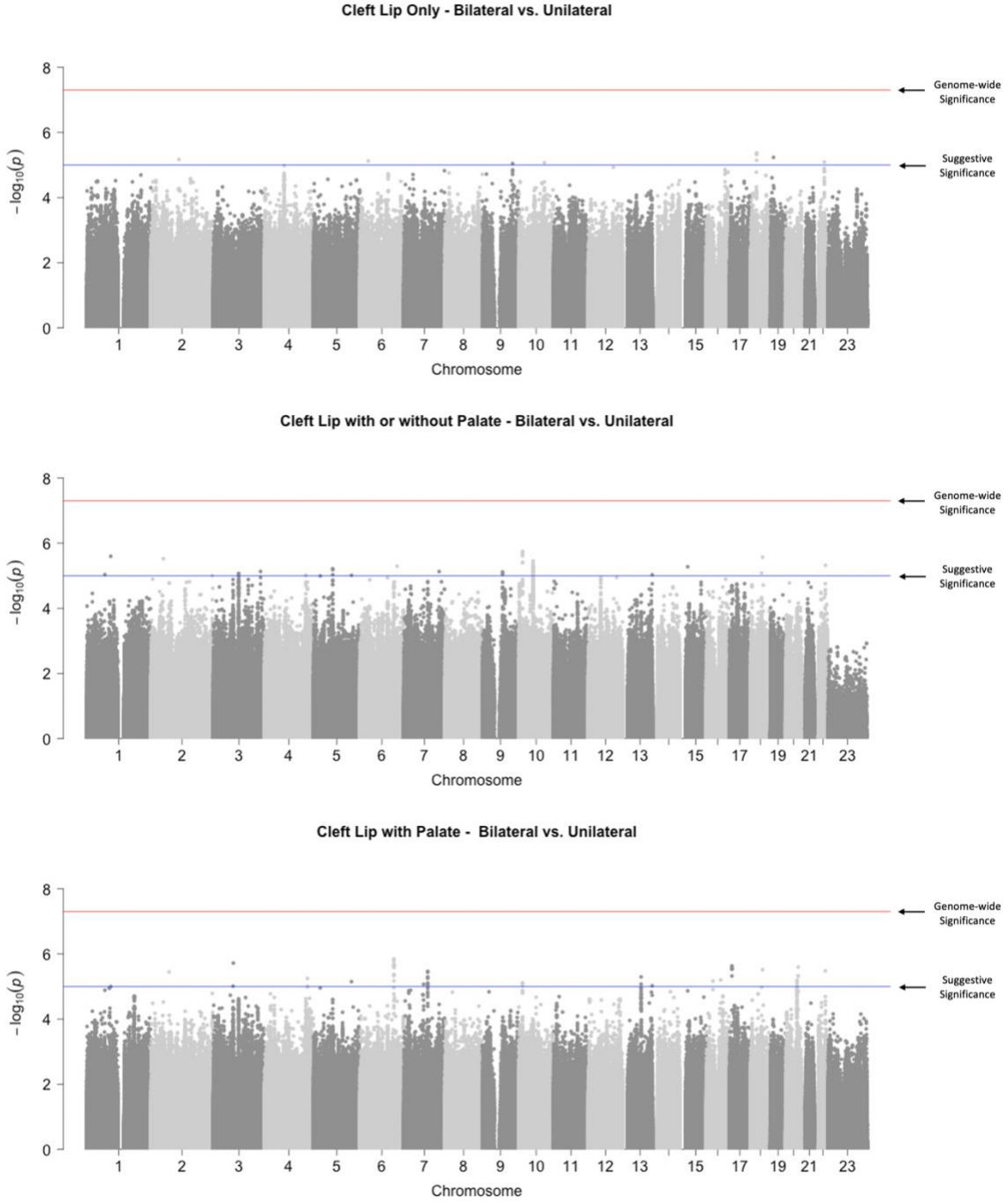
		CL/P: Unilateral vs Bilateral									
		Bilateral (n=522)		Unilateral (n=1400)		Right (n=546)		Left (n=854)		Unknown side (n=4)	
Male	**(%)**	356	(68.20)	822	(58.71)	323	(59.16)	499	(58.43)	1	(25.00)
Female	**(%)**	166	(31.80)	578	(41.29)	223	(40.84)	355	(41.57)	3	(75.00)
Asian	**(%)**	96	(18.39)	268	(19.14)	102	(18.68)	166	(19.44)	0	(00.00)
Caucasian	**(%)**	164	(31.42)	411	(29.36)	134	(24.54)	277	(32.44)	0	(00.00)
LSA	**(%)**	261	(50.00)	720	(51.43)	310	(56.78)	410	(48.01)	4	(100.00)
Unspecified	**(%)**	1	(00.19)	1	(00.07)	0	(00.00)	1	(00.12)	0	(00.00)

		CL: Unilateral vs Bilateral									
		Bilateral (n=28)		Unilateral (n=346)		Right (n=130)		Left (n=216)		Unknown side (n=2)	
Male	**(%)**	17	(60.71)	194	(56.07)	75	(57.69)	119	(55.09)	0	(00.00)
Female	**(%)**	11	(39.29)	152	(43.93)	55	(42.31)	97	(44.91)	2	(100.00)
Asian	**(%)**	10	(35.71)	111	(32.08)	43	(33.08)	68	(31.48)	0	(00.00)
Caucasian	**(%)**	10	(35.71)	118	(34.10)	40	(30.77)	78	(36.11)	0	(00.00)
LSA	**(%)**	8	(28.57)	117	(33.82)	47	(36.15)	70	(32.41)	2	(100.00)
Unspecified	**(%)**	0	(00.00)	0	(00.00)	0	(00.00)	0	(00.00)	0	(00.00)

		CLP: Unilateral vs Bilateral									
		Bilateral (n=494)		Unilateral (n=1054)		Right (n=416)		Left (n=638)		Unknown side (n=2)	
Male	**(%)**	339	(68.62)	628	(59.58)	248	(59.62)	380	(59.56)	1	(50.00)
Female	**(%)**	155	(31.38)	426	(40.42)	168	(40.38)	258	(40.44)	1	(50.00)
Asian	**(%)**	86	(17.41)	157	(14.90)	59	(14.18)	98	(15.36)	0	(00.00)
Caucasian	**(%)**	154	(31.17)	293	(27.80)	94	(22.60)	199	(31.19)	0	(00.00)
LSA	**(%)**	253	(51.21)	563	(53.42)	263	(63.22)	340	(53.29)	2	(100.00)
Unspecified	**(%)**	1	(00.20)	1	(00.09)	0	(00.00)	1	(00.16)	0	(00.00)

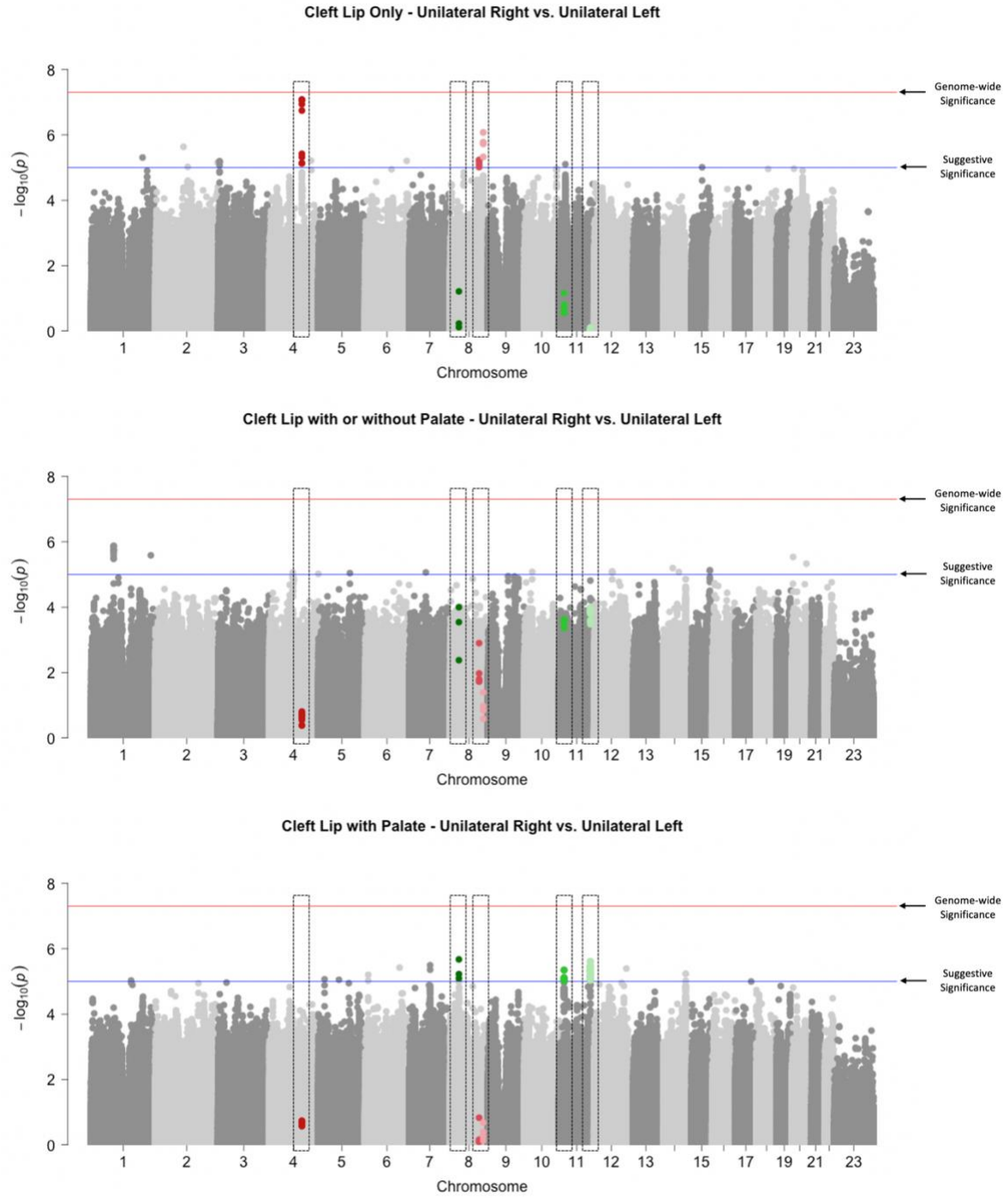
		CL/P: Unilateral Left vs Unilateral Right				**CL: Unilateral Left vs Unilateral Right**				**CLP: Unilateral Left vs Unilateral Right**			
		Unilateral Right (n=546)		Unilateral Left (n=854)		Unilateral Right (n=130)		Unilateral Left (n=216)		Unilateral Right (n=416)		Unilateral Left (n=638)	
Male	**(%)**	323	(59.16)	499	(58.43)	75	(57.69)	119	(55.09)	248	(59.62)	380	(59.56)
Female	**(%)**	223	(40.84)	355	(41.57)	55	(42.31)	97	(44.91)	168	(40.38)	258	(40.44)
Asian	**(%)**	102	(18.68)	166	(19.44)	43	(33.08)	68	(31.48)	59	(14.18)	98	(15.36)
Caucasian	**(%)**	134	(24.54)	277	(32.44)	40	(30.77)	78	(36.11)	94	(22.60)	199	(31.19)
LSA	**(%)**	310	(56.78)	410	(48.01)	47	(36.15)	70	(32.41)	263	(63.22)	340	(53.29)
Unspecified	**(%)**	0	(00.00)	1	(00.12)	0	(0.00)	0	(00.00)	0	(00.00)	1	(00.16)

Figure 1: GWAS Results Bilateral vs Unilateral (1a: CL/P; 1b: CL; 1c: CLP).



We see that none of the bilateral vs unilateral analysis identified any SNPs at genome-wide significance nor any promising candidate SNPs overall.

Figure 2: GWAS Results Unilateral Right v Left (2a: CL/P; 2b: CL; 2c: CLP)



Manhattan plots of the GWAS results for unilateral right against unilateral left analysis for CL, CLP, and CL/P sub-types. Points in red denote three of the most significant regions from the CL analysis. Points in green indicate three of the most significant regions from the CLP analysis.

Table 3A: Most Significant Loci per Analysis – Bilateral vs Unilateral

CL/P: Bilateral vs Unilateral							
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	Candidate Gene	# of SNP
1	1p22.1	rs71650502	2.022	(1.508, 2.711)	2.52E-06	EVI5	1
2	2p16	rs11407371	1.927	(1.463, 2.537)	2.99E-06	FOXN2	1
10	10p13	rs143826861	0.685	(0.587, 0.800)	1.80E-06	FAM107B	7
10	10q21.1	rs17643564	0.647	(0.539, 0.778)	3.50E-06	PCDH15	27
18	18q21.1	rs1787328	0.649	(0.531, 0.771)	2.68E-06	MYO5B	2

CLP: Bilateral vs Unilateral						
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	# of SNP
3	3p12	rs13087163	0.657	(0.553, 0.781)	1.90E-06	2
6	6q23.2	rs7765331	0.405	(0.280, 0.585)	1.43E-06	15
17	17p13	rs71358253	0.485	(0.359, 0.655)	2.32E-06	7
18	18q21.1	rs1787328	0.624	(0.512, 0.761)	3.03E-06	1
20	20q13.1	rs13045716	1.477	(1.255, 1.737)	2.52E-06	8

CL: Bilateral vs Unilateral						
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	# of SNP
2	2q12	rs112499595	0.059	(0.017, 0.203)	6.75E-06	1
6	6p21.3	rs78897911	0.035	(0.008, 0.151)	7.49E-06	1
18	18q11.2	rs72890974	0.079	(0.027, 0.233)	4.24E-06	3
19	19p13.2	rs143383743	0.086	(0.030, 0.249)	5.83E-06	1
22	22q13.1	rs113844586	0.192	(0.093, 0.396)	8.05E-06	1

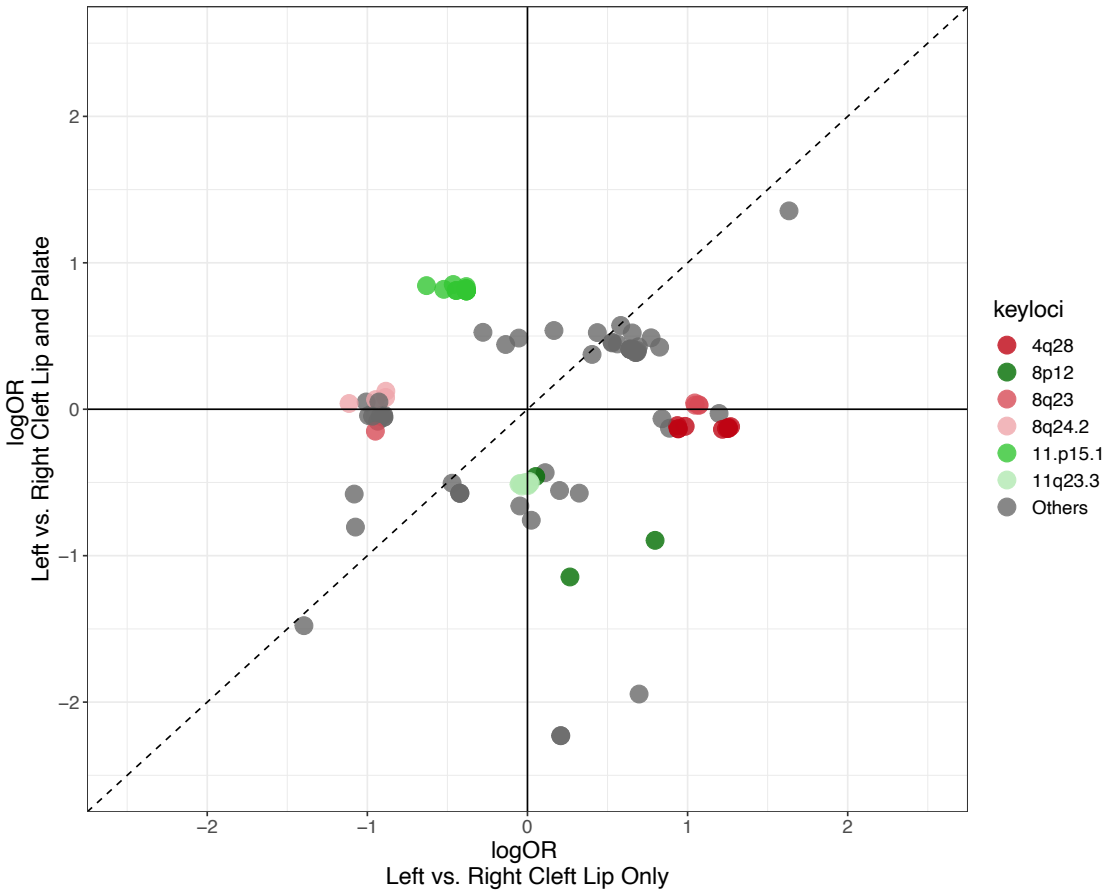
Table 3B: Most Significant Loci per Analysis –Unilateral Right vs Unilateral Left

CL/P: Unilateral Right vs Unilateral Left						
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	# of SNP
1	1p22.3	rs6667102	1.599	(1.322, 1.935)	1.32E-06	14
12	12q15	rs57646976	0.582	(0.459, 0.738)	8.00E-06	4
14	14q22	rs4080563	0.506	(0.377, 0.681)	6.37E-06	1
15	15q26.1	rs2388015	1.611	(1.308, 1.985)	7.41E-06	2
20	20p12.3	rs6116687	1.625	(1.326, 1.991)	2.91E-06	2

CLP: Unilateral Right vs Unilateral Left						
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	# of SNP
6	6q23.2	rs17642884	0.469	(0.340, 0.646)	3.79E-06	1
7	7q21.1	rs2194751	1.556	(1.292, 1.873)	3.20E-06	2
8	8p12	rs56056078	0.318	(0.198, 0.511)	2.12E-06	3
11	11q23.3	rs499804	0.594	(0.478, 0.738)	2.45E-06	27
12	12q24.31	rs34152756	1.691	(1.352, 2.114)	4.05E-06	1

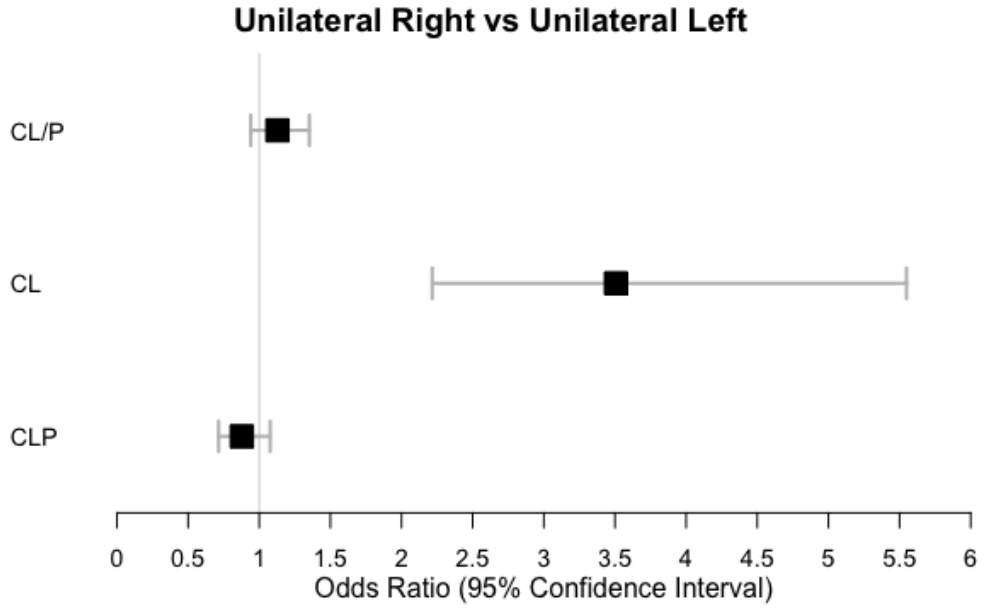
CL: Unilateral Right vs Unilateral Left						
CHR	Region	Lead SNP	Odds Ratio	95% CI	P-value	# of SNP
1	1q32.1	rs12732777	0.365	(0.237, 0.563)	4.93E-06	1
2	2q12	rs139260643	3.309	(2.014, 5.438)	2.32E-06	1
4	4q28	rs3956582	3.507	(2.216, 5.549)	8.40E-08	10
8	8q24.2	rs13267780	0.328	(0.211, 0.511)	8.40E-07	4
8	8q23	rs62520628	2.919	(1.836, 4.64)	5.91E-06	5

Figure 3: Comparison of the Effect Sizes of Lead SNPs



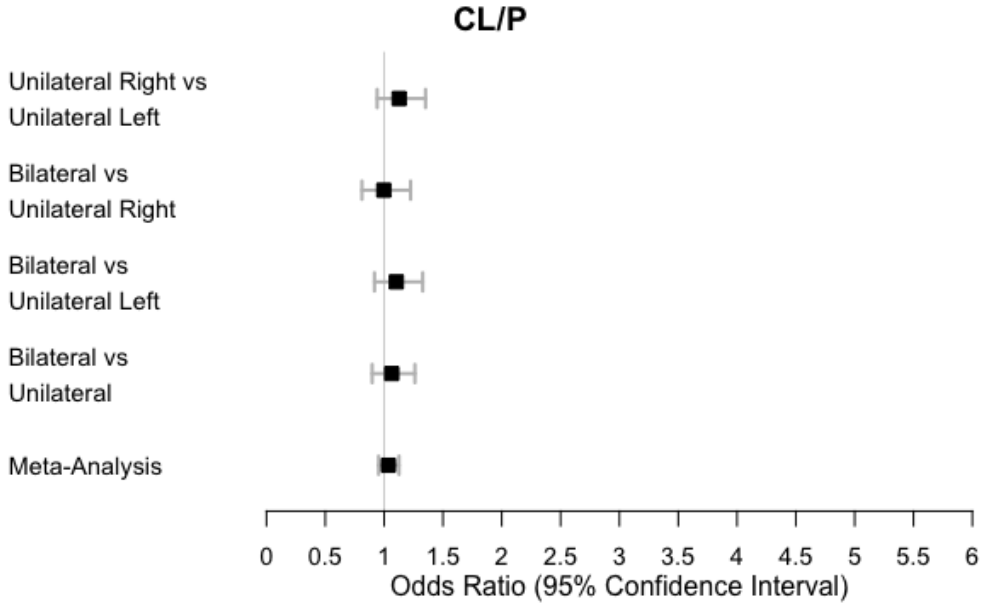
Comparison of log transformed odds ratio of lead SNPs from CL left/right GWAS and log transformed odds ratio of lead SNPs from the CLP left/right GWAS. Points in red are three of the most significant regions from the CL analysis. Points in green are three of the most significant regions from the CLP analysis. The effect size of these SNPs separate by their respective analysis group

Figure 4A: Comparison of the odds ratio of rs6855309 from the unilateral left against unilateral right analyses only in the CL CL/P, CLP, and CL sub-types



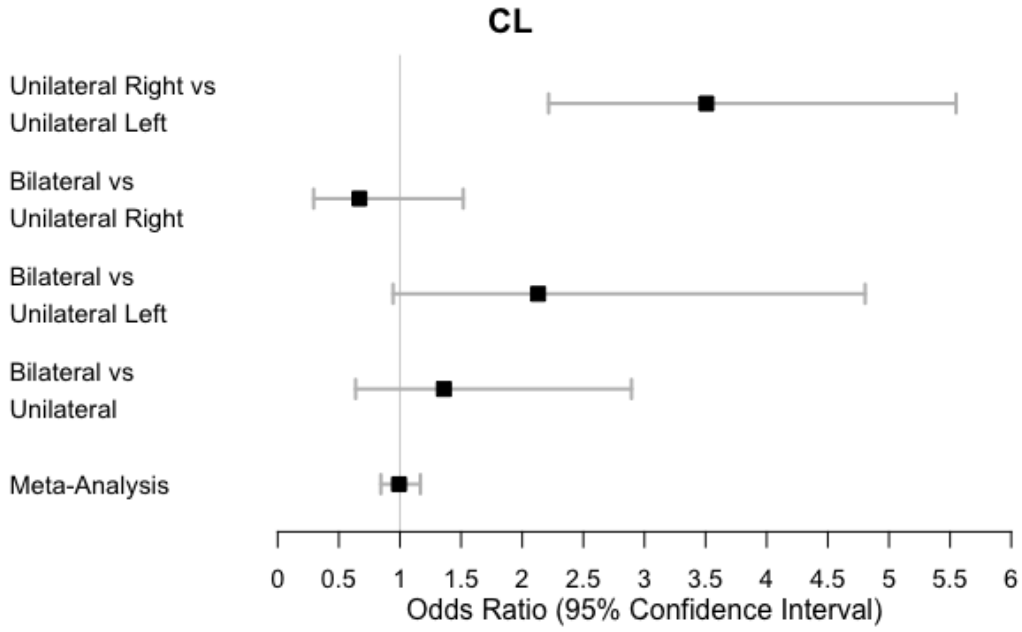
Results from the unilateral right vs unilateral left GWAS reveal a significant increase in the odds ratio only in the CL sub-type.

Figure 4B: Comparison of the odds ratio of rs6855309 resulting from CL/P GWAS



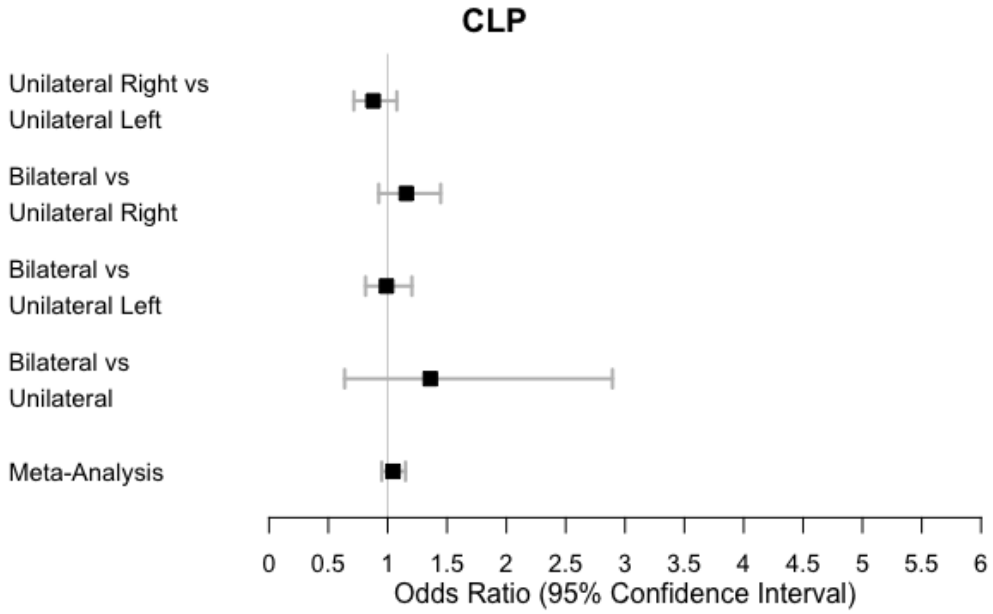
Results from the CL/P subtype-specific GWAS show that there was no significant change in the odds ratio of the rs6855309 SNP

Figure 4C: Comparison of the odds ratio of rs6855309 resulting from CL GWAS



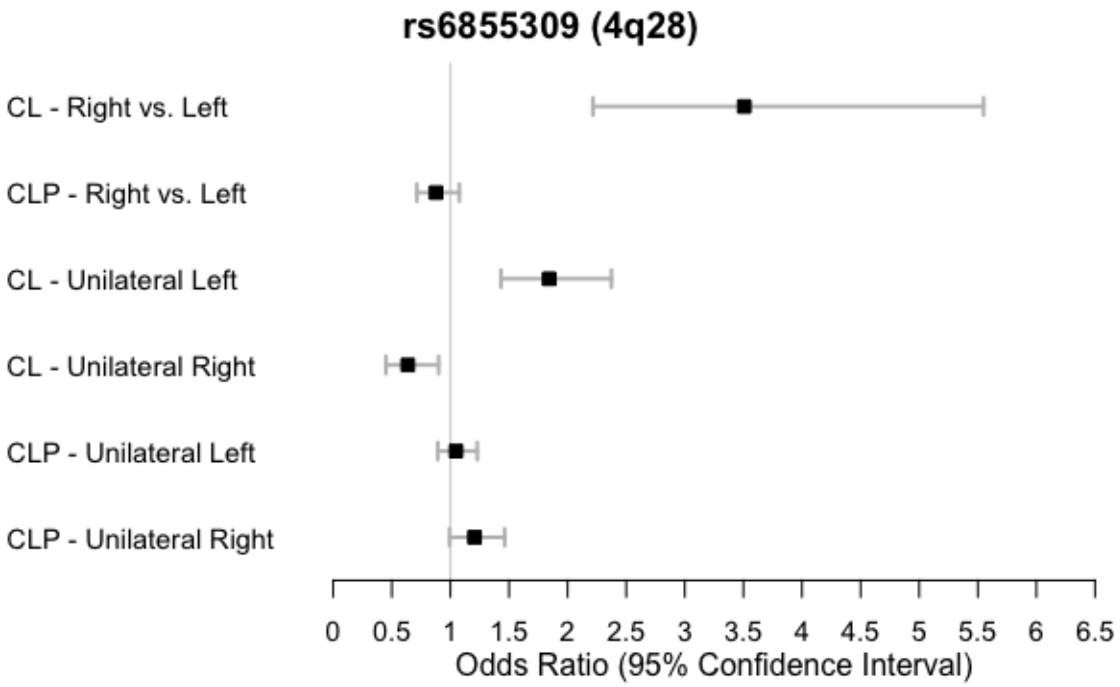
Results from the CL subtype-specific GWAS indicate a significant odds ratio in the unilateral right vs unilateral left modifier GWAS

Figure 4D: Comparison of the odds ratio of rs6855309 resulting from CLP GWAS



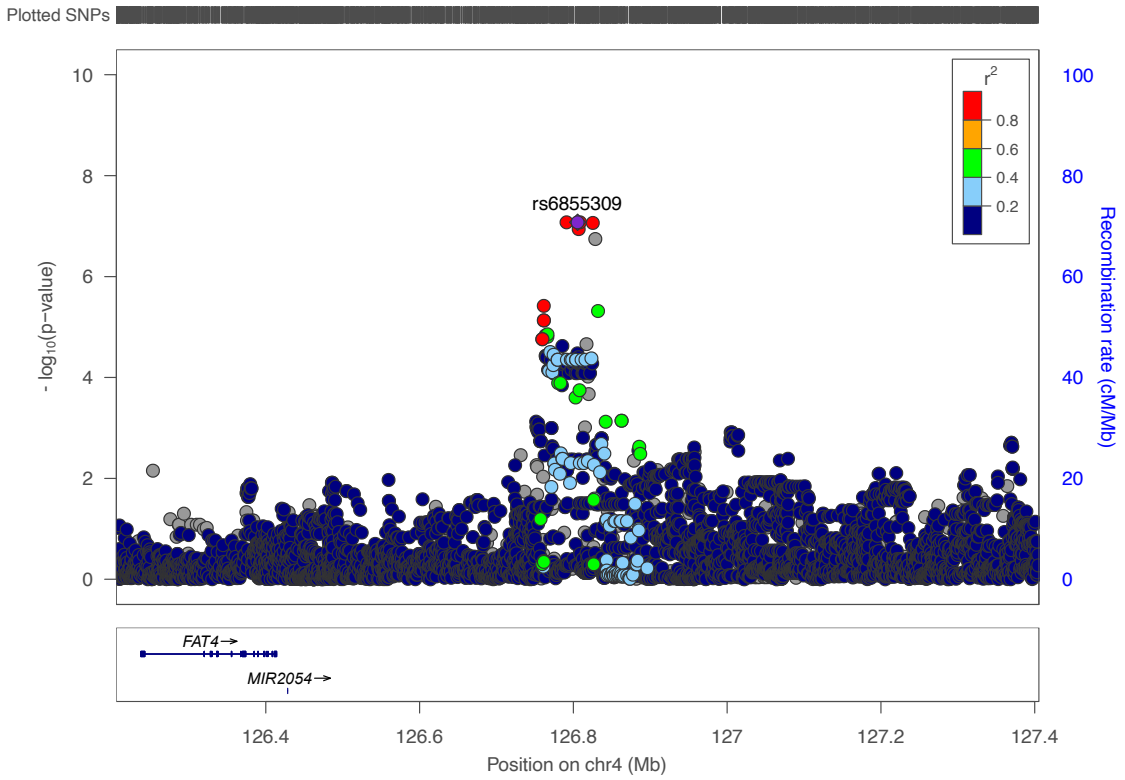
Results from the CLP subtype-specific GWAS show that there was no significant change in the odds ratio of the rs6855309 SNP

Figure 5: Comparison of the s6855309 Odds Ratio by Analysis.



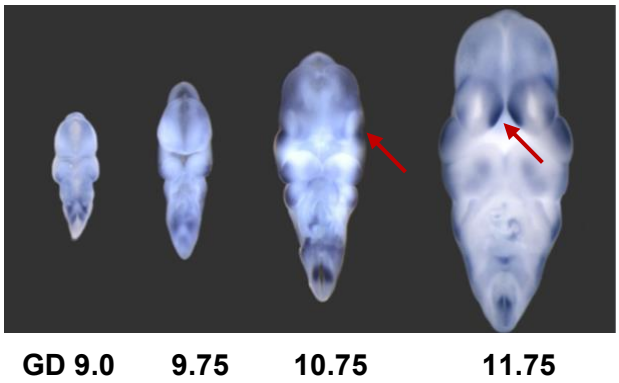
Examining the odds ratios of the rs6855309 SNP from each GWAS indicate that the suggestive association between OFCs and 4q28 by laterality is found only in CL cases. Important to note is that there was an opposite direction of effect between LCL and RCL

Figure 6: Regional Association Plot of 4q28 Reveal Proximity of the *FAT4* Gene.



Focusing in on the 4q28 region revealed that the *FAT4* gene lies slightly over 400Kb upstream of the lead SNP (rs6855309) and is a candidate gene for further analysis

Figure 7: *FAT4* Expression at the Medial Nasal Process at Gestational Timepoints in Embryonic Mice.



Staining for *FAT4* expression at various timepoint of gestational mice demonstrates its expression in the medial nasal process, a precursor for the upper lip