genes

MDPI

Article

Genetic Association of the 8q21.3 Locus with Non-Syndromic Cleft Lip Only in a Vietnamese Population: Case–Control and Family-Based Evidence

Chi Kim Nguyen [1,2], Teruyuki Niimi [1,3], Hisataka Kondo [4], Satoshi Suzuki [1], Hideto Imura [1], Toko Hayakawa [3,5], Hiroo Furukawa [5], Chisato Sakuma [1,3], Ken Kitagawa [1,3], Kumiko Fujiwara [6], Yuki Akiyama [1,3], Ryosuke Miwa [1,3], Nagana Natsume [1,3], Ai Diep Minh Nguyen [2], Loc Nguyen Gia Pham [1,2], Dang Minh Nguyen [2], Tung Minh Duong [2], Van Thi Thao Nguyen [2], Minh Duc Nguyen [2] and Nagato Natsume [1,3,*]

1 Division of Research and Treatment of Oral-Maxillofacial Congenital Anomalies, Aichi Gakuin University, Nagoya 464-8651, Japan; kichinguyen97@gmail.com (C.K.N.); niimi@dpc.agu.ac.jp (T.N.); zsatoshi@hotmail.com (S.S.); h-imura@dpc.agu.ac.jp (H.I.); 9646ken@gmail.com (K.K.); nsfhk427@yahoo.co.jp (Y.A.); miwa.r@outlook.com (R.M.); ag223d15@dpc.agu.ac.jp (N.N.); gialocp@gmail.com (L.N.G.P.)
2 Ho Chi Minh City Odonto-Maxillo-Facial Hospital, Ho Chi Minh City 70000, Vietnam; nguyendiepminhai@gmail.com (A.D.M.N.); drnguyenminhdang1104@gmail.com (D.M.N.); minhtungduong83@gmail.com (T.M.D.); ntthaovan@gmail.com (V.T.T.N.); dr.ducminhrhm@gmail.com (M.D.N.)
3 Cleft Lip and Palate Center, Aichi Gakuin University Dental Hospital, Nagoya 464-8651, Japan; hayakawa@dpc.agu.ac.jp
4 Junior College, Aichi Gakuin University, Nagoya 464-0037, Japan; kondoh@dpc.agu.ac.jp
5 Division of Speech, Hearing, and Language, Aichi Gakuin University Dental Hospital, Nagoya 464-8651, Japan; hfuru@dpc.agu.ac.jp
6 Oral and Maxillofacial Surgery, Osaka Medical and Pharmaceutical University, Takatsuki 569-8686, Japan; kumiko.fujiwara@ompu.ac.jp
* Correspondence: natsume@dpc.agu.ac.jp

Abstract

Background/Objectives: Genetic susceptibility to nonsyndromic orofacial clefts (NSOFCs) varies across populations and phenotypes. This study investigated associations of rs7552 (*CYRIA/FAM49A*) and rs12543318 (8q21.3) with nonsyndromic cleft lip only (NSCLO) and cleft lip with palate (NSCLP) in Vietnamese individuals. Methods: We analyzed 231 case–parent trios (111 NSCLO, 120 NSCLP) and 240 unrelated controls. Both variants were genotyped using TaqMan assays. Case–control associations were evaluated under a prespecified genetic comparison, with complementary family-based transmission disequilibrium tests (TDTs). Secondary sex-stratified and exploratory genetic analyses were also performed. Functional annotations were examined using publicly available databases. Results: Under the prespecified genetic comparison, rs12543318 was associated with NSCLO (CC vs. AC+AA; OR = 1.82, 95% CI: 1.14–2.91; $p = 0.0118$), meeting the Bonferroni threshold ($p < 0.0125$). The effect estimate was larger in males (OR = 2.72, 95% CI: 1.41–5.24; $p = 0.0025$), but the SNP-by-sex interaction was not statistically significant ($p = 0.0846$). The TDT revealed under-transmission of the A allele in NSCLO for rs12543318 (T/U = 47/81; $p = 0.0027$) and rs7552 (T/U = 27/52; $p = 0.0049$). Neither variant showed significant corrected case–control associations with NSCLP. Formal subtype comparisons and parent-of-origin tests were not statistically significant. In silico analyses suggested regulatory links involving *MMP16* and *CYRIA/FAM49A*. Conclusions: The rs12543318 CC genotype was associated with increased odds of NSCLO in Vietnamese individuals, with directionally consistent family-based evidence. For rs7552, transmission distortion was observed despite limited case–control evidence.

Check for updates

Academic Editor: Hui Li

Received: 3 November 2026
Revised: 25 November 2026
Accepted: 27 November 2026
Published: 28 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: *CYRIA*/*FAM49A*; 8q21.3; *MMP16*; cleft lip only; cleft lip/palate

1. Introduction

Orofacial clefts (OFCs) are one of the most common congenital anomalies worldwide [1], with a reported prevalence of approximately 1.4 per 1000 live births in Vietnam [2]. OFCs are classified as syndromic or non-syndromic (NSOFCs), with the latter accounting for over two-thirds of cases [3]. NSOFCs include cleft palate only (NSCPO; MIM #119540), cleft lip only (NSCLO), and cleft lip with palate (NSCLP), with the latter two collectively classified as nonsyndromic cleft lip with or without cleft palate (NSCL/P; MIM #119530) [4]. Their occurrence varies by sex and phenotype, with an approximately 2:1 male-to-female ratio for clefts involving the lip, a reverse 1:2 ratio for cleft palate only, and a 2:1 left-to-right ratio among unilateral cleft lip cases [5].

Normal craniofacial morphology depends on tightly regulated interactions between neural crest-derived facial mesenchyme and the overlying epithelium. Disruption of these developmental processes through genetic and epigenetic interactions contributes to the complex multifactorial etiology of NSOFCs [5,6].

Genome-wide association studies (GWASs) have identified numerous NSOFC susceptibility loci, including *IRF6*, *PAX7*, *NTN1*, 8q24, and the *CYRIA* (formerly *FAM49A*) locus at 2p24.2 [7–9]. However, genetic associations may differ across populations because of variations in allele frequencies, linkage disequilibrium patterns, and genetic backgrounds, highlighting the importance of evaluating established susceptibility variants in diverse populations.

Among these susceptibility regions, 2p24.2 has emerged as an important locus for NSCL/P. The single-nucleotide polymorphism (SNP) rs7552, located in the 3′ untranslated region (3′-UTR) of *CYRIA*/*FAM49A*, was initially identified as a genome-wide significant susceptibility marker in a large multi-ethnic GWAS [8–11]. Subsequent studies replicated its association with OFCs in Chinese and Brazilian populations [12,13]. The association was observed across NSCLO and NSCLP in the dominant model, although the magnitude of the effect appeared to vary between phenotypes [10]. These findings support a potential role of rs7552 in OFC susceptibility while suggesting population- and phenotype-dependent heterogeneity in its genetic effect.

The biological characteristics of *CYRIA*/*FAM49A* provide additional plausibility for its involvement in OFCs. *CYRIA*/*FAM49A* regulates RAC1-dependent actin remodeling, thereby influencing cell polarity and migration processes that may be relevant to cranial neural crest development [14,15]. Furthermore, *CYRIA*/*FAM49A* expression has been demonstrated in the palatal mesenchyme and epithelium of mouse embryos, supporting a potential role in craniofacial development [11]. However, a direct causal relationship between *CYRIA*/*FAM49A* and cleft formation has not been established, and the mechanism through which rs7552 may influence OFC susceptibility remains unclear.

Another susceptibility region of interest is 8q21.3, where rs12543318 was identified as a leading variant in a large multiethnic GWAS of cleft-specific phenotypes [8,16–18]. rs12543318 is an intergenic, noncoding variant located approximately 15 kb downstream of *DCAF4L2* at 8q21.3. Subsequent studies have provided further evidence for its association with OFCs [19,20], although the association was not replicated in a Brazilian population, suggesting potential population-specific effects [21]. Phenotype-specific analyses have also suggested stronger associations of rs12543318 with NSCLO than with NSCLP or NSCPO in Chinese populations, especially in the dominant model [19,20], supporting potential subtype-specific genetic effects within NSCL/P. However, the functional target of rs12543318

remains uncertain, and its annotation near *DCAF4L2* is based largely on genomic proximity. Thus, rs12543318 currently represents a genetic susceptibility marker for which its biological mechanism in cleft development remains to be clarified.

Despite previous evidence implicating rs7552 and rs12543318 in NSOFC susceptibility [12,13,20,22,23], their associations in the Vietnamese population remain insufficiently characterized. Evaluating established susceptibility variants in additional populations is important for assessing the consistency and generalizability of previously reported genetic associations, particularly given population differences in allele frequencies and linkage disequilibrium patterns [24]. Furthermore, although NSCLO and NSCLP share genetic susceptibility factors, their genetic architectures may not be identical [8,10,16,25].

The present study therefore aimed to provide phenotype-resolved estimates of genetic associations and allelic transmission for rs12543318 and rs7552 in a Vietnamese population. The dominant model was prespecified as the primary case–control genetic comparison. Family-based analyses were conducted as a complementary approach to assess allelic transmission and parent-of-origin effects, with reduced susceptibility to population stratification [26]. Sex-stratified analyses were prespecified secondary analyses; pooled NSCL/P analyses were exploratory. By integrating these analytical approaches, this study sought to evaluate the consistency of previously reported associations and characterize genetic association patterns across cleft phenotypes while providing additional evidence for the genetic epidemiology of NSOFCs in Vietnam.

2. Materials and Methods

2.1. Study Design and Participants

This study used an unrelated case–control design as the primary genetic association analysis, complemented by family-based analyses in case–parent trios consisting of an affected child and both biological parents. All participants were screened by maxillofacial surgeons to rule out associated congenital anomalies or recognizable syndromes and were diagnosed with nonsyndromic cleft lip/palate based on clinical examination, treatment history, and medical records at Odonto-Maxillo-Facial Hospital Ho Chi Minh City (HCMC), Vietnam, from 2015 to 2024. Peripheral blood samples were collected at the Odonto-Maxillo-Facial Hospital of Ho Chi Minh City, Vietnam, and applied to dried blood spot cards. Following transfer to Japan, the dried blood spot samples were stored at the World Banking Room, Aichi Gakuin University, Nagoya, until genetic analysis. Written informed consent was obtained from all participants or, for participants younger than 18 years, from their parents or legal guardians. All procedures were conducted in accordance with the ethical principles of the Declaration of Helsinki (World Medical Association, 2013). This study was approved by the Ethics Committee of Aichi Gakuin University (approval No. 78).

This study included 231 Vietnamese patients and their biological parents, comprising 111 patients with nonsyndromic cleft lip only (NSCLO; 56 males and 55 females) and 120 with nonsyndromic cleft lip and palate (NSCLP; 60 males and 60 females), forming 231 case–parent trios. In addition, 240 unrelated healthy controls were recruited from the same geographic region and frequency-matched to the affected participants by sex. All participants self-identified as Vietnamese of Kinh ethnicity, with demographic information including age and sex recorded. Ethnicity was determined by self-report. Genomic ancestry was not independently assessed using ancestry-informative markers or principal component analysis (PCA).

Initial sample size planning for the case–control analysis was performed using G*Power (version 3.1.9.7; Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany) before sample selection from the Aichi Gakuin University World Gene Bank, with reference to previously reported genetic associations of rs12543318 and rs7552 with nonsyndromic orofacial

clefts [20,23,27]. The initial planning considered the overall case–control population rather than individual cleft phenotypes. No separate a prior sample size calculation was performed for the family-based analysis. Assuming a two-sided α of 0.05 and 80% power, a comparison-of-proportions approach yielded minimum sample sizes of approximately 418 participants (for rs12543318) and 308 participants (for rs7552). The final study includes 471 participants for the case–control analysis (231 patients and 240 controls). Sensitivity power analyses were subsequently performed to characterize the detectable effect sizes for each SNP and phenotype. For the case–control analyses, minimum detectable effect sizes (MDEs) were estimated under the prespecified genetic comparison using the observed control genotype frequencies, the corresponding numbers of cases and controls, 80% statistical power, and a two-sided Bonferroni-adjusted significance level of 0.0125. A two-sample comparison-of-proportions framework with pooled variance was used, and the resulting genotype proportions were converted to odds ratios. MDEs were estimated separately for risk-increasing and protective effects.

Regarding sampling criteria, patients selected for the study must have a confirmed clinical diagnosis of NSCL/P, no other congenital malformations or acquired systemic health conditions, and no family history of other genetic disorders. Additionally, both the patient's parents gave informed consent to participate in this study. Participants in the control group were screened to ensure complete absence of congenital conditions or malformations, no personal or family history of orofacial clefts, and willingness to participate in the study.

2.2. Single-Nucleotide Polymorphism Selection and Genotyping

Candidate single nucleotide polymorphisms (SNPs) were prioritized based on previous genome-wide association study (GWAS) evidence for NSOFC. Susceptibility variants identified in the NHGRI-EBI GWAS Catalog were cross-referenced with the NCBI dbSNP database and 1000 Genome Browser to verify genomic coordinates, annotation, allele profiles, and population frequencies [24,28]. Evidence from Asian populations was prioritized because of potentially shared genetic backgrounds with Vietnamese individuals. Budget constraints limited selection to two previously associated variants: rs7552 (*CYRIA*/*FAM49A*) and rs12543318 (8q21.3).

DNA was extracted from dried blood samples using the QIAamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions, with concentration and purity assessed spectrophotometrically before testing. Genotyping for rs7552 and rs12543318 was performed using the TaqMan™ SNP Genotyping Assay on the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA).

Genotypes were determined via StepOnePlus™ software version 2.3 (Thermo Fisher Scientific, Waltham, MA, USA), with unclear clusters or samples with low signal being reanalyzed. Final genotyping call rates were calculated as the proportion of successfully genotyped samples among all samples tested for each SNP. Genotype call confidence scores were summarized by SNP, participant group, and genotype using means, standard deviations, and ranges. Before family-based association testing, all markers were tested for Mendelian genetic inconsistencies. Systematic duplicate genotyping was not performed. Initial genotype call records were not retained; only final call rates are reported.

2.3. Statistical Analysis

Genotype and allele frequencies, including minor allele frequencies (MAFs), were calculated for each SNP, and Hardy–Weinberg equilibrium (HWE) was assessed in controls using a chi-square goodness-of-fit test. Genotype distributions between cases and controls were compared using Pearson's chi-square or Fisher's exact test, as appropriate.

Case–control association analyses were performed separately for NSCLO and NSCLP, which were treated as the two primary phenotype-specific comparisons. The combined NSCL/P phenotype was evaluated as a secondary pooled analysis. Associations of rs7552 and rs12543318 with cleft susceptibility were examined under allelic, genotypic, dominant, recessive, and overdominant models. The primary genetic comparison was prespecified before examining the association results, based on previous studies implicating the CC genotype of rs12543318 and the GG genotype of rs7552 in orofacial cleft susceptibility [20,23]. Accordingly, major-allele homozygotes were compared with carriers of at least one minor allele (CC vs. AC+AA for rs12543318 and GG vs. AG+AA for rs7552). This comparison corresponds to a dominant model for the minor allele, with odds ratios expressed for major-allele homozygotes relative to minor-allele carriers. The dominant model was used to reduce the degrees of freedom and limit the problem of low genotype counts in the rare homozygote group [29].

Additional allelic, genotypic, recessive, and overdominant models, together with the Cochran–Armitage trend test, were evaluated as exploratory analyses to characterize alternative patterns of genetic association and assess the consistency of the observed findings [29]. These analyses were not used to select the primary genetic model.

Sex-stratified analyses were prespecified as secondary analyses based on previously reported sex differences in the prevalence and genetic susceptibility of NSCL/P [30,31]. Potential differences in genetic effects between males and females were formally evaluated under the prespecified genetic comparison using SNP-by-sex interaction terms in logistic regression models, assessed by likelihood-ratio tests. Differences in statistical significance between sex-stratified estimates were not considered sufficient evidence of effect modification by sex.

Family-based transmission disequilibrium tests (TDTs) were performed as complementary analyses to assess preferential allelic transmission from heterozygous parents to affected offspring. Analyses were conducted separately for NSCLO and NSCLP. Transmission odds ratios were calculated as the ratio of transmitted to untransmitted alleles for the designated effect allele. Parent-specific transmission and parent-of-origin (POO) analyses were performed using PLINK 1.9 (−tdt poo). Allelic transmissions from heterozygous fathers and mothers were evaluated separately using chi-square tests. When both parents and the affected offspring were heterozygous, the parental origin of the transmitted allele could not be unambiguously determined. In such cases, PLINK assigned fractional counts of 0.5 to each possible parental transmission, resulting in non-integer transmitted and untransmitted allele counts [32,33]. The formal POO test used an asymptotic Z statistic to compare paternal and maternal transmission odds ratios. Parent-specific transmission and POO analyses were considered exploratory, and nominal p values were reported without further adjustment for multiple testing.

Pooled NSCL/P analyses were performed for comparison with previous studies that evaluated cleft lip with or without cleft palate as a combined phenotype. Parent-of-origin analyses, SNP-by-sex interactions, and pooled phenotype analyses were considered exploratory. Exploratory sex-adjusted case–case logistic regression compared NSCLO with NSCLP as the reference outcome under the prespecified genotype grouping. TDT heterogeneity was assessed by likelihood-ratio tests comparing transmission odds between phenotypes.

Statistical analyses were conducted using R version 4.5.1 within RStudio version 2026.05.0+218 (Posit Software, PBC, Boston, MA, USA). Family-based transmission disequilibrium and parent-of-origin analyses were performed using PLINK version 1.9 (Purcell Lab, Harvard Medical School). All statistical tests were two-sided. The primary case–control hypothesis family was defined before examining the association results and comprised four

tests, corresponding to two candidate SNPs and two primary cleft phenotypes (NSCLO and NSCLP) under the prespecified genetic comparison. Bonferroni correction was therefore applied across these four primary comparisons, yielding a significance threshold of $\alpha = 0.0125$ (0.05/4). The complementary family-based TDT analyses were treated as a separate testing family because they evaluated within-family allelic transmission rather than between-group genotype associations. The same Bonferroni-adjusted significance threshold of $\alpha = 0.0125$ was applied across the four phenotype-specific TDT comparisons. The prespecified secondary sex-stratified analyses comprised eight comparisons (two SNPs × two phenotypes × two sexes) under the primary genetic comparison, with a separate Bonferroni-adjusted significance threshold of $\alpha = 0.00625$ (0.05/8).

Alternative genetic models, genotype trend tests, pooled NSCL/P analyses, SNP-by-sex interactions, and parent-of-origin analyses were considered exploratory. These analyses were not included in the primary Bonferroni correction family because they were intended to characterize additional association patterns and generate hypotheses rather than establish the primary findings. Nominal *p* values were reported without further multiplicity adjustment and were interpreted cautiously. Multiplicity was controlled separately within each defined testing family rather than across the entire set of analyses. Consequently, the complementary and secondary findings were interpreted within their respective analytical frameworks, and exploratory results were considered hypothesis-generating.

Interpretation of the findings emphasized the magnitude, direction, and precision of effect estimates, as reflected by odds ratios and their corresponding 95% confidence intervals, rather than statistical significance alone.

2.4. In Silico Functional Annotation

Functional annotation of rs7552 and rs12543318 was performed using the Ensembl Variant Effect Predictor (VEP), RegulomeDB v2.2, HaploReg v4, and the Genotype-Tissue Expression (GTEx) Portal. Variant coordinates and predicted molecular consequences were annotated using VEP based on the GRCh38 human genome assembly [34]. RegulomeDB was used to evaluate regulatory evidence, including regulatory ranks, chromatin accessibility, transcription factor binding, and quantitative trait locus annotations [35]. HaploReg was used to characterize the regulatory and linkage disequilibrium (LD) context of each index variant [36]. Variants in strong LD ($r^2 \geq 0.80$) were identified using the Asian (ASN) reference population, and their chromatin-state, DNase-hypersensitivity, protein-binding, and predicted transcription factor motif annotations were examined. For rs7552, located within the 3′ untranslated region (3′UTR) of *CYRIA*/*FAM49A*, miRNASNP-v3 was used to predict allele-dependent gains or losses of microRNA-binding sites [37]. The index variants and selected strongly linked variants were subsequently queried in the GTEx Portal (Analysis Release V10) for expression quantitative trait locus (eQTL) and splicing quantitative trait locus (sQTL) associations [38]. Because eQTL and sQTL data from human embryonic craniofacial tissues are unavailable in GTEx, available adult tissues were examined to explore potential regulatory associations [38]. LD information was obtained from the Asian (ASN) reference population in HaploReg and was not independently estimated in the present Vietnamese cohort [36]. All databases were accessed in September–October 2026.

The functional annotation was based exclusively on publicly available database resources and was intended to explore the potential regulatory relevance of the investigated variants. Genomic annotations, computational predictions, and previously reported QTL associations were distinguished according to their respective levels of evidence. No statistical fine-mapping, colocalization analysis, or experimental functional validation was performed. Therefore, the annotations were considered hypothesis-generating rather than evidence establishing causal variants, target genes, or disease-related regulatory mechanisms [39].

3. Results

3.1. Study Population and Genotype Distribution

Of the 933 individuals genotyped, 25 SNP-specific calls required repeat genotyping (13 for rs12543318 and 12 for rs7552). Following repeat genotyping, valid genotype calls were obtained for all individuals for both SNPs (933/933 per SNP), yielding a final genotyping call rate of 100% (1866/1866). Detailed genotyping quality control metrics, including genotype call confidence scores stratified by SNP, participant group, and genotype, are presented in Supplementary Table S1. No Mendelian inconsistencies were detected among the 231 case–parent trios.

This study included 231 patients with NSCL/P (111 NSCLO and 120 NSCLP) and 240 unrelated controls. Among the affected participants, 186 had unilateral clefts, and 45 had bilateral clefts; the distribution was 99/12 in NSCLO and 87/33 in NSCLP, respectively.

Genotype distributions and minor allele frequencies (MAFs) are summarized in Table 1. Both SNPs were in Hardy–Weinberg equilibrium among controls (rs7552, $p = 0.780$; rs12543318, $p = 0.642$). Compared with controls, the minor A allele was less frequent in NSCLO for both rs7552 (0.149 vs. 0.217) and rs12543318 (0.383 vs. 0.471), while smaller differences were observed for NSCLP.

Table 1. Genotype distribution, minor allele frequency, and Hardy–Weinberg equilibrium.

SNP	Gene/ Locus	Group	Genotype *n* (%)	MAF	HWE *p*
rs7552 (G>A)	*CYRIA/FAM49A*	Controls	GG 148 (61.7%)/AG 80 (33.3%)/AA 12 (5.0%)	0.217	0.780
		NSCLO	GG 80 (72.1%)/AG 29 (26.1%)/AA 2 (1.8%)	0.149	0.734
		NSCLP	GG 73 (60.8%)/AG 45 (37.5%)/AA 2 (1.7%)	0.204	0.092
		NSCL/P	GG 153 (66.2%)/AG 74 (32.0%)/AA 4 (1.7%)	0.178	0.140
rs12543318 (C>A)	8q21.3	Controls	CC 69 (28.8%)/AC 116 (48.3%)/AA 55 (22.9%)	0.471	0.642
		NSCLO	CC 47 (42.3%)/AC 43 (38.7%)/AA 21 (18.9%)	0.383	0.058
		NSCLP	CC 37 (30.8%)/AC 64 (53.3%)/AA 19 (15.8%)	0.425	0.318
		NSCL/P	CC 84 (36.4%)/AC 107 (46.3%)/AA 40 (17.3%)	0.405	0.556

Successful genotype calls were obtained for all participants for both rs12543318 and rs7552, with no missing final genotypes.

3.2. Case–Control Association Analysis

Under the prespecified primary genetic comparison, individuals with the rs12543318 CC genotype had 1.82-fold higher odds of NSCLO than individuals carrying at least one A allele (OR = 1.82, 95% CI: 1.14–2.91; $p = 0.0118$), meeting the Bonferroni-adjusted significance threshold. For rs7552, the GG genotype was associated with an elevated point estimate for NSCLO (OR = 1.60, 95% CI: 0.98–2.62; $p = 0.0574$), although the confidence interval included the null value.

For NSCLP, the effect estimates were closer to the null for both rs12543318 (OR = 1.10, 95% CI: 0.69–1.78) and rs7552 (OR = 0.97, 95% CI: 0.62–1.51). However, the confidence intervals did not exclude modest genetic effects.

Exploratory analyses indicated lower odds of NSCLO associated with the A allele of rs12543318 (OR = 0.70, 95% CI: 0.50–0.96) and rs7552 (OR = 0.63, 95% CI: 0.41–0.97). These nominal associations were considered supportive but not confirmatory evidence. In the exploratory pooled NSCL/P analysis, rs12543318 showed a modest allelic association (OR = 0.76, 95% CI: 0.59–0.99), whereas no clear association was observed for rs7552. Detailed results for all genetic models are presented in Table 2.

Table 2. Association of SNPs with NSCLO, NSCLP and NSCL/P compared with controls.

SNP	Genotype/Allele	Case, *n*	Control, *n*	Model/ Comparison	OR (95% CI)	*p* Value	*p*-Trend
NSCLO							
CYRIA/FAM49A rs7552	GG	80	148	Genotypic		0.1231 [a]	0.0355
	AG	29	80				
	AA	2	12				
	G	189	376	Allelic reference	Reference		
	A	33	104	Allelic	0.63 (0.41–0.97)	0.0345 [b]	
	GG vs. AG+AA	80/31	148/92	Dominant	1.60 (0.98–2.62)	0.0574 [b]	
	AA vs. GG+AG	2/109	12/228	Recessive	0.35 (0.08–1.58)	0.2405 [a]	
	AG vs. GG+AA	29/82	80/160	Overdominant	0.71 (0.43–1.17)	0.1748 [b]	
8q21.3 rs12543318	CC	47	69	Genotypic		0.0419 [b]	0.036
	AC	43	116				
	AA	21	55				
	C	137	254	Allelic reference	Reference		
	A	85	226	Allelic	0.70 (0.50–0.96)	0.0292 [b]	
	CC vs. AC+AA	47/64	69/171	Dominant	1.82 (1.14–2.91)	**0.0118** [b]	
	AA vs. CC+AC	21/90	55/185	Recessive	0.78 (0.45–1.38)	0.3978 [b]	
	AC vs. CC+AA	43/68	116/124	Overdominant	0.68 (0.43–1.07)	0.0931 [b]	
NSCLP							
CYRIA/FAM49A rs7552	GG	73	148	Genotypic	—	0.2730 [a]	0.6936
	AG	45	80				
	AA	2	12				
	G	191	376	Allelic reference	Reference		
	A	49	104	Allelic	0.93 (0.63–1.36)	0.6991 [b]	
	GG vs. AG+AA	73/47	148/92	Dominant	0.97 (0.62–1.51)	0.8783 [b]	
	AA vs. GG+AG	2/118	12/228	Recessive	0.32 (0.07–1.46)	0.1548 [a]	
	AG vs. GG+AA	45/75	80/160	Overdominant	1.20 (0.76–1.89)	0.4337 [b]	
8q21.3 rs12543318	CC	37	69	Genotypic	—	0.2907 [b]	0.2425
	AC	64	116				
	AA	19	55				
	C	138	254	Allelic reference	Reference		
	A	102	226	Allelic	0.83 (0.61–1.14)	0.2444 [b]	
	CC vs. AC+AA	37/83	69/171	Dominant	1.10 (0.69–1.78)	0.6827 [b]	
	AA vs. CC+AC	19/101	55/185	Recessive	0.63 (0.36–1.12)	0.1169 [b]	
	AC vs. CC+AA	64/56	116/124	Overdominant	1.22 (0.79–1.90)	0.3711 [b]	
NSCL/P							
CYRIA/FAM49A rs7552	GG	153	148	Genotypic	—	0.1258 [b]	0.1249
	AG	74	80				
	AA	4	12				
	G	380	376	Allelic reference	Reference		
	A	82	104	Allelic	0.78 (0.57–1.08)	0.1310 [b]	
	GG vs. AG+AA	153/78	148/92	Dominant	1.22 (0.84–1.78)	0.3022 [b]	
	AA vs. GG+AG	4/227	12/228	Recessive	0.33 (0.11–1.05)	0.0503 [b]	
	AG vs. GG+AA	74/157	80/160	Overdominant	0.94 (0.64–1.39)	0.7639 [b]	
8q21.3 rs12543318	CC	84	69	Genotypic	—	0.1332 [b]	0.045
	AC	107	116				
	AA	40	55				
	C	275	254	Allelic reference	Reference		
	A	187	226	Allelic	0.76 (0.59–0.99)	0.0410 [b]	
	CC vs. AC+AA	84/147	69/171	Dominant	1.42 (0.96–2.09)	0.0778 [b]	
	AA vs. CC+AC	40/191	55/185	Recessive	0.70 (0.45–1.11)	0.1300 [b]	
	AC vs. CC+AA	107/124	116/124	Overdominant	0.92 (0.64–1.32)	0.6618 [b]	

[a] Fisher's exact test; [b] Pearson chi-square test. The primary genetic comparisons were CC vs. AC+AA for rs12543318 and GG vs. AG+AA for rs7552. Bonferroni correction was applied across four primary tests (two SNPs × two phenotypes), yielding a significance threshold of $p < 0.0125$. Bold values indicate primary associations meeting this threshold. All other genetic models, genotype trend tests, and pooled NSCL/P analyses were exploratory. Nominal *p* values and unadjusted 95% confidence intervals are reported. OR, odds ratio; CI, confidence interval.

3.3. Sex-Stratified Association Analysis

In the prespecified secondary sex-stratified analyses, the rs12543318 CC genotype was associated with elevated odds of NSCLO among males (OR = 2.72, 95% CI: 1.41–5.24; $p = 0.0025$), meeting the Bonferroni-adjusted significance threshold of 0.00625. Among females, the corresponding effect estimate was smaller (OR = 1.18, 95% CI: 0.59–2.35).

Despite the larger effect estimate in males, the formal SNP-by-sex interaction was not statistically significant (interaction OR = 2.30, 95% CI: 0.89–5.95; $p = 0.0846$), indicating insufficient evidence of a difference in genetic effects between males and females (Table A1).

No other prespecified sex-stratified genetic comparisons met the corresponding Bonferroni-adjusted significance threshold. Exploratory analyses under alternative genetic models and pooled NSCL/P analyses are presented in Table 3.

Table 3. Sex-stratified association of SNPs with NSCLO, NSCLP, and NSCL/P.

SNP	Sex	Genotype/ Allele Counts, Case vs. Control	Genotypic *p*-Value	Allelic OR (95% CI) *p*-Value	Dominant OR (95% CI) *p*-Value	Recessive OR (95% CI) *p*-Value	Overdominant OR (95% CI) *p*-Value	*p*-Trend
NSCLO								
rs7552	Female	GG/AG/AA 41/12/2 vs. 70/42/8 G/A 94/16 vs. 182/58	0.1299 [a]	0.53 (0.29–0.98) 0.0407 [b]	2.09 (1.03–4.24) 0.0387 [b]	0.53 (0.11–2.57) 0.7268 [a]	0.52 (0.25–1.09) 0.0797 [b]	0.0484
	Male	GG/AG/AA 39/17/0 vs. 78/38/4 G/A 95/17 vs. 194/46	0.5063 [a]	0.75 (0.41–1.39) 0.3633 [b]	1.24 (0.62–2.44) 0.5434 [b]	0.23 (0.01–4.33) 0.3083 [a]	0.94 (0.47–1.87) 0.8614 [b]	0.3475
rs12543318	Female	CC/AC/AA 18/27/10 vs. 35/54/31 C/A 63/47 vs. 124/116	0.5388 [b]	0.80 (0.51–1.26) 0.3290 [b]	1.18 (0.59–2.35) 0.6342 [b]	0.64 (0.29–1.42) 0.2673 [b]	1.18 (0.62–2.23) 0.6144 [b]	0.3454
	Male	CC/AC/AA 29/16/11 vs. 34/62/24 C/A 74/38 vs. 130/110	0.0053 [b]	0.61 (0.38–0.97) 0.0351 [b]	2.72 (1.41–5.24) **0.0025** [b]	0.98 (0.44–2.17) 0.9559 [b]	0.37 (0.19–0.74) 0.0041 [b]	0.0436
NSCLP								
rs7552	Female	GG/AG/AA 37/21/2 vs. 70/42/8 G/A 95/25 vs. 182/58	0.7234 [a]	0.83 (0.49–1.40) 0.4790 [b]	1.15 (0.61–2.17) 0.6677 [b]	0.48 (0.10–2.35) 0.4996 [a]	1.00 (0.52–1.92) 1.0000 [b]	0.482
	Male	GG/AG/AA 36/24/0 vs. 78/38/4 G/A 96/24 vs. 194/46	0.2940 [a]	1.05 (0.61–1.83) 0.8506 [b]	0.81 (0.43–1.53) 0.5117 [b]	0.21 (0.01–4.04) 0.3028 [a]	1.44 (0.76–2.74) 0.2674 [b]	0.8427

Table 3. *Cont.*

SNP	Sex	Genotype/ Allele Counts, Case vs. Control	Genotypic *p*-Value	Allelic OR (95% CI) *p*-Value	Dominant OR (95% CI) *p*-Value	Recessive OR (95% CI) *p*-Value	Overdominant OR (95% CI) *p*-Value	*p*-Trend
rs12543318	Female	CC/AC/AA	0.0724 [b]	0.74 (0.47–1.15) 0.1782 [b]	1.04 (0.53–2.05) 0.9079 [b]	0.38 (0.16–0.92) 0.0281 [b]	1.71 (0.91–3.20) 0.0917 [b]	0.1791
		18/35/7 vs. 35/54/31						
		C/A						
		71/49 vs. 124/116						
	Male	CC/AC/AA	0.8874 [b]	0.93 (0.60–1.45) 0.7646 [b]	1.17 (0.60–2.30) 0.6437 [b]	1.00 (0.46–2.17) 1.0000 [b]	0.88 (0.47–1.63) 0.6733 [b]	0.7622
		19/29/12 vs. 34/62/24						
		C/A						
		67/53 vs. 130/110						
NSCL/P								
rs7552	Female	GG/AG/AA	0.2540 [b]	0.68 (0.43–1.07) 0.0920 [b]	1.51 (0.88–2.57) 0.1319 [b]	0.50 (0.15–1.72) 0.2670 [b]	0.75 (0.43–1.30)	0.0985
		78/33/4 vs. 70/42/8						
		G/A						
		189/41 vs. 182/58						
	Male	GG/AG/AA	0.1695 [a]	0.91 (0.57–1.44) 0.6755 [b]	0.98 (0.58–1.68) 0.9558 [b]	0.11 (0.01–2.09) 0.1220 [a]	1.18 (0.69–2.03) 0.5494 [b]	0.6567
		75/41/0 vs. 78/38/4						
		G/A						
		191/41 vs. 194/46						
rs12543318	Female	CC/AC/AA	0.1031b	0.77 (0.53–1.10) 0.1510 [b]	1.11 (0.63–1.93) 0.7213 [b]	0.50 (0.26–0.96) 0.0357 [b]	1.43 (0.86–2.39) 0.1719 [b]	0.1516
		36/62/17 vs. 35/54/31						
		C/A						
		134/96 vs. 124/116						
	Male	CC/AC/AA	0.0802 [b]	0.76 (0.53–1.10) 0.1466 [b]	1.79 (1.04–3.07) 0.0354 [b]	0.99 (0.52–1.87) 0.9736 [b]	0.59 (0.35–0.99) 0.0470 [b]	0.161
		48/45/23 vs. 34/62/24						
		C/A						
		141/91 vs. 130/110						

Note: [a] Fisher's exact test; [b] Pearson chi-square test. Sex-stratified analyses under the prespecified primary genetic comparison were considered secondary analyses. Bonferroni correction was applied across eight tests (two SNPs × two phenotypes × two sexes), yielding a significance threshold of $p < 0.00625$. Bold values indicate secondary associations meeting this threshold. Alternative genetic models and pooled NSCL/P analyses were exploratory. Nominal *p* values and unadjusted 95% confidence intervals are reported.

3.4. Family-Based Association Analysis

The complementary family-based TDT analyses provided evidence of preferential allelic transmission in NSCLO. The A allele of rs12543318 was under-transmitted to affected offspring (T/U = 47/81; OR = 0.58, 95% CI: 0.41–0.83; $p = 0.0027$). Similarly, the A allele of rs7552 showed under-transmission (T/U = 27/52; OR = 0.52, 95% CI: 0.33–0.83; $p = 0.0049$). Both findings met the Bonferroni-adjusted significance threshold of 0.0125.

For NSCLP, neither variant met the corrected significance threshold. The transmission odds ratios were 0.68 (95% CI: 0.47–0.99) for rs12543318 and 0.86 (95% CI: 0.55–1.34) for rs7552.

Exploratory pooled NSCL/P analyses also suggested under-transmission of the A allele for both variants. Although some parental transmission tests showed nominal evidence

of transmission distortion, the formal parent-of-origin tests provided insufficient evidence of differential transmission between paternal and maternal alleles (Table 4). Nominally significant transmission from one parent alone does not establish a parent-of-origin effect.

Table 4. Family-based transmission disequilibrium and parent-of-origin analyses.

Gene/Locus	Phenotype	A1/A2	TDT T/U	OR (95% CI)	P_TDT	Paternal T/U	P_Pat	Maternal T/U	P_Mat	P_POO
CYRIA/FAM49A rs7552	NSCLO	A/G	27/52	0.52 (0.33–0.83)	**0.0049**	14/26	0.0578	13/26	0.0374	0.8759
	NSCLP	A/G	37/43	0.86 (0.55–1.34)	0.5023	19/19	1	18/24	0.3545	0.5226
	NSCL/P	A/G	64/95	0.67 (0.49–0.92)	0.014	33/45	0.1742	31/50	0.0348	0.604
8q21.3 rs12543318	NSCLO	A/C	47/81	0.58 (0.41–0.83)	**0.0027**	25/36	0.159	22/45	0.005	0.3403
	NSCLP	A/C	47/69	0.68 (0.47–0.99)	0.0411	24.5/37.5	0.0987	22.5/31.5	0.2207	0.814
	NSCL/P	A/C	94/150	0.63 (0.48–0.81)	0.0003	49.5/73.5	0.0305	44.5/76.5	0.0036	0.578

Note: TDT, transmission disequilibrium test; T, transmitted allele count; U, untransmitted allele count; OR, odds ratio; POO, parent of origin; Pat, paternal; Mat, maternal. ORs are expressed for the A1 allele relative to the A2 allele. The 95% CIs for TDT ORs were calculated from the transmitted and untransmitted allele counts. Bonferroni correction was applied across four phenotype-specific TDT comparisons (two SNPs × two phenotypes), yielding a significance threshold of $p < 0.0125$. Bold values indicate TDT associations meeting this threshold. Pooled NSCL/P and parent-of-origin analyses were exploratory, with nominal p values reported. P_POO represents the formal test of differential transmission between paternal and maternal alleles. Non-integer paternal and maternal T/U counts result from PLINK's allocation of fractional counts (0.5) to ambiguous transmissions when parental origin cannot be determined. P_Pat and P_Mat assess transmission distortion separately for fathers and mothers, respectively, whereas P_POO tests for differences between paternal and maternal transmission odds ratios. All parent-specific and POO p values are nominal and exploratory.

3.5. Statistical Power and Sensitivity Analyses

At 80% statistical power and a Bonferroni-adjusted significance level of 0.0125, the minimum detectable odds ratios ranged from 2.16 to 2.38 in the risk-increasing direction and from 0.37 to 0.47 in the protective direction across the phenotype-specific case–control comparisons (Table A2). For rs12543318 in NSCLO, the observed odds ratio of 1.82 (95% CI: 1.14–2.91) was below the corresponding minimum detectable odds ratio of 2.20 at 80% power (Table A3). The remaining primary case–control effect estimates were also within the range of effects for which the study had less than 80% power under the specified assumptions. Exploratory sex-adjusted case–case comparisons yielded ORs of 1.65 (95% CI: 0.96–2.85; likelihood-ratio $p = 0.0686$) for rs12543318 and 1.66 (95% CI: 0.96–2.89; $p = 0.0698$) for rs7552. TDT heterogeneity tests were also not statistically significant ($p = 0.5427$ and 0.1201, respectively; Table A4). These results do not establish differential effects between NSCLO and NSCLP.

3.6. In Silico Functional Annotation

Database-based functional annotations identified potential regulatory features at both investigated loci (Tables 5 and 6). At 8q21.3, the index SNP rs12543318 was annotated as a noncoding intergenic variant located near *DCAF4L2*. The variant rs11784027 was identified as a strongly linked candidate regulatory variant in the Asian reference population ($r^2 = 0.88$, $D' = 0.98$). Both rs12543318 and rs11784027 showed eQTL associations with *MMP16* expression in adult aortic tissue, identifying *MMP16* as a candidate regulatory target gene rather than an established causal gene at this locus.

Table 5. In silico functional annotation of rs12543318 at 8q21.3.

Annotation	Database	Main Findings	Interpretation
Genomic consequence	Ensembl VEP (GRCh38)	Intergenic variant at chr8:87,856,112 (C>A); predicted impact: MODIFIER	Identifies rs12543318 as a noncoding intergenic variant; genomic location alone does not establish its regulatory target.
Regulatory potential	RegulomeDB v2.2	Rank 1f; probability score 0.12; 28 overlapping peaks.	Indicates potential regulatory relevance based on database annotations, without establishing allele-specific regulatory function.
Chromatin state	RegulomeDB	Predominantly quiescent/low chromatin annotations; no ChIP-seq result displayed in the queried panel.	Provides limited evidence of active regulatory chromatin in the queried datasets; developmental and tissue-specific activity remains unresolved.
Regulatory context	HaploReg v4	Approximately 15 kb downstream (3′) of *DCAF4L2*; DNase annotation and predicted ATF3 motif alteration.	Establishes genomic proximity to *DCAF4L2* and suggests potential regulatory activity. Predicted motif alterations do not establish altered transcription factor binding.
Linkage disequilibrium	HaploReg v4 (ASN)	Strong LD with rs11784027 ($r^2 = 0.88$, $D' = 0.98$).	Identifies rs11784027 as a strongly linked candidate variant in the ASN reference population.
Linked variant	HaploReg v4	rs11784027 located approximately 1.1 kb upstream (5′) of *DCAF4L2*; predicted AP-1 motif alteration.	Identifies an additional candidate regulatory variant within the LD region; its functional contribution remains unvalidated.
eQTL (index variant)	GTEx v10	Associated with *MMP16* expression in artery–aorta ($p = 1.3 \times 10^{-5}$; NES = +0.14).	Identifies *MMP16* as an expression-associated candidate target gene in adult aortic tissue, without establishing a developmental or disease-mediated effect.
eQTL (LD proxy)	GTEx v10	rs11784027 associated with *MMP16* expression in artery–aorta ($p = 3.9 \times 10^{-6}$; NES = +0.15).	Provides concordant expression-association evidence involving *MMP16* but does not distinguish the regulatory contributions of the index SNP and its LD proxy.
sQTL	GTEx v10	No single-tissue sQTL association returned for either variant in the queried results.	No evidence from this query for an effect on RNA splicing; this does not exclude developmental/tissue-specific effects.

Table 6. In silico functional annotation of rs7552 at the *CYRIA*/*FAM49A* locus.

Annotation	Database	Main Findings	Interpretation
Genomic consequence	Ensembl VEP (GRCh38)	3′-UTR variant at chr2:16,552,660; predicted impact: MODIFIER	Identifies rs7552 as a noncoding variant within the *CYRIA*/*FAM49A* 3′-UTR, suggesting potential post-transcriptional regulatory relevance.
Transcript annotation	Ensembl VEP	Additional splice-region and noncoding transcript exon consequences	Indicates additional annotated transcript consequences without demonstrating altered *CYRIA*/*FAM49A* transcript processing.
Regulatory potential	RegulomeDB v2.2	Rank 1f; probability score 0.22271; 75 overlapping peaks	Suggests regulatory relevance based on integrated database annotations; the score does not represent disease-association probability or experimentally established function.

Table 6. *Cont.*

Annotation	Database	Main Findings	Interpretation
Chromatin state	RegulomeDB v2.2	Predominantly quiescent/low chromatin states	Indicates context-dependent chromatin annotations without establishing regulatory activity during craniofacial development.
Chromatin accessibility	RegulomeDB v2.2	Three DNase/ATAC-seq results, including kidney datasets	Identifies accessible chromatin in selected tissues, without direct evidence of accessibility in embryonic craniofacial tissues.
Regulatory motifs	HaploReg v4	Predicted alterations of BRCA1 and TBX5 motifs	Predicts potential allele-dependent changes in transcription factor binding motifs that require experimental validation.
Linkage disequilibrium	HaploReg v4 (ASN)	Multiple linked variants with $r^2 \geq 0.8$, including variants overlapping regulatory features	Identifies multiple linked candidate variants in the ASN reference population, preventing definitive attribution of regulatory effects to rs7552 itself.
eQTL (*CYRIA*/*FAM49A*)	GTEx v10	Brain–frontal cortex (BA9): $p = 3.0 \times 10^{-5}$; NES = −0.076	Identifies an association with *CYRIA*/*FAM49A* expression in adult frontal cortex; its relevance to embryonic craniofacial development remains unknown.
sQTL (*CYRIA*/*FAM49A*)	GTEx v10	Thyroid: $p = 2.0 \times 10^{-6}$; NES = +0.23	Identifies an association with *CYRIA*/*FAM49A* transcript splicing in adult thyroid tissue, without establishing a developmental effect.
Additional eQTLs	GTEx v10	Associations with ENSG00000237633 and ENSG00000279663 in testis and brain tissues.	Indicates additional expression-associated regulatory signals involving other transcripts at the locus in adult testis and brain tissues.
Predicted miRNA-binding gain	miRNA SNP-v3	Target gain with SNP in 3′UTR: hsa-miR-6800-5p, hsa-miR-6802-5p and hsa-miR-4694-5p. For hsa-miR-6800-5p, the predicted duplex free energy was −15.70 kcal/mol, binding free energy −15.88 kcal/mol, and miRmap-reported TargetScan score 23.75.	Predicts allele-dependent changes in potential miRNA-binding sites without demonstrating altered miRNA binding, mRNA stability, or *CYRIA*/*FAM49A* expression.
		Target loss with SNP in 3′UTR: hsa-miR-4781-3p, hsa-miR-576-3p and hsa-miR-10397-3p. For hsa-miR-4781-3p, the predicted duplex free energy was −9.70 kcal/mol, binding free energy was −8.75 kcal/mol, and miRmap-reported TargetScan score was 22.22.	

eQTL, expression quantitative trait locus; sQTL, splicing quantitative trait locus; NES, normalized effect size. The eQTL association in the frontal cortex (BA9) reflects variation in *CYRIA*/*FAM49A* expression, whereas the sQTL association in the thyroid reflects variation in *CYRIA*/*FAM49A* transcript splicing. NES indicates the direction and magnitude of the allelic effect on the corresponding molecular phenotype. NES values for eQTLs and sQTLs are not directly comparable. These associations were identified in adult tissues and do not establish regulatory effects during embryonic craniofacial development.

At the *CYRIA*/*FAM49A* locus, rs7552 was annotated as a 3′-UTR variant within an extensive LD block containing multiple variants with regulatory annotations. GTEx data

identified an eQTL association with *CYRIA*/*FAM49A* expression in adult frontal cortex (BA9; $p = 3.0 \times 10^{-5}$, NES = −0.076) and an sQTL association in adult thyroid tissue ($p = 2.0 \times 10^{-6}$, NES = 0.23). Additionally, miRNASNP-v3 predicted allele-dependent gains and losses of potential miRNA-binding sites within the *CYRIA*/*FAM49A* 3′-UTR.

These findings provide preliminary evidence of potential regulatory relevance at both loci. However, the QTL associations were derived from adult non-craniofacial tissues, and the predicted regulatory motif and miRNA-binding alterations have not been experimentally validated. Neither the causal variants nor the genes mediating NSOFC susceptibility were established by these annotations.

4. Discussion

In this Vietnamese cohort, statistically significant associations were observed in NSCLO, whereas neither variant met the corresponding Bonferroni-adjusted significance thresholds in NSCLP. However, formal comparisons provided insufficient evidence of differential genetic effects between the two cleft phenotypes. The strongest evidence was found for rs12543318 at 8q21.3, which remained significantly associated with NSCLO after multiple testing correction in the case–control analysis, with undertransmission of the A allele observed in the family-based analysis. For rs7552 at the *CYRIA*/*FAM49A* locus, the case–control association did not reach statistical significance after multiple testing correction, although significant undertransmission of the A allele was detected in the NSCLO trios. In the secondary sex-stratified analyses, the effect estimate for rs12543318 in NSCLO was larger in males and met the corresponding Bonferroni-adjusted significance threshold. However, the non-significant SNP-by-sex interaction provided insufficient evidence of differential genetic effects between males and females. Database-based functional annotations identified *MMP16* and *CYRIA*/*FAM49A* as candidate regulatory target genes at the rs12543318 and rs7552 loci, respectively, although the causal variants and disease-relevant regulatory mechanisms remain unresolved.

Although rs12543318 and rs7552 have previously been implicated in orofacial cleft susceptibility, the contribution of the present study extends beyond population-specific replication. By combining phenotype-resolved case–control analyses, family-based transmission tests, and formal comparisons between cleft subtypes, this study provides an integrated assessment of the association patterns of these established susceptibility variants in Vietnamese individuals. The directionally consistent findings for rs12543318 and the differing levels of evidence across analytical designs for rs7552 illustrate the value of complementary approaches in evaluating genetic susceptibility. Furthermore, the absence of statistically significant subtype and sex interactions helps define the limits of the current evidence. Together with database-based functional annotations, these findings provide a population-specific framework for future replication and mechanistic investigations.

The interpretation of the present findings was guided by the prespecified hierarchy of statistical analyses. The primary case–control evidence was confined to the association between rs12543318 and NSCLO under the prespecified genetic comparison, with complementary support from the family-based TDT. Although additional association patterns were observed across alternative genetic models and subgroups, these findings should be regarded as hypothesis-generating rather than confirmatory. Particular emphasis should be placed on the magnitude and precision of the effect estimates, as the relatively wide confidence intervals reflect uncertainty regarding the underlying genetic effects.

4.1. Associations Across Cleft Phenotypes

Although NSCLO and NSCLP are commonly grouped as NSCL/P, their genetic architectures appear to be partially shared rather than identical [16]. This heterogeneity is biologically

plausible given the distinct developmental processes involved in upper lip formation and secondary palatal development [6], although differences in developmental timing do not necessarily imply phenotype-specific mechanisms. Genome-wide studies by Mukhopadhyay et al. (2022), Carlson et al. (2019), Huang et al. (2019), and Dack et al. (2025) have identified both shared susceptibility loci and phenotype-dependent genetic effects, supporting the value of analyzing cleft subtypes separately [8,10,16,25]. Pooling NSCLO and NSCLP may therefore obscure associations concentrated in individual phenotypes, even though combined analyses can improve power to detect shared effects. Our findings support the relevance of phenotype-stratified analyses but do not establish statistically significant heterogeneity between NSCLO and NSCLP. For rs12543318, Pan et al. (2013) identified an association with nonsyndromic orofacial clefts [19]. In contrast, Zhang et al. (2018) reported significant under-transmission of the A allele of rs12543318 (OR = 0.57, $p = 0.0032$) and over-transmission of the CC genotype in Western Han Chinese individuals with NSCLO [20]. Consistent with these findings, our case–control analysis indicated that individuals with the CC genotype had higher odds of NSCLO than carriers of at least one A allele (OR = 1.82, 95% CI: 1.14–2.91). The complementary family-based analysis yielded a similar allelic transmission estimate to that reported by Zhang et al. (OR = 0.58, $p = 0.0027$).

From a biological perspective, these directionally consistent findings raise the hypothesis that the presence of at least one A allele may be associated with reduced susceptibility to NSCLO. However, the observed genotype contrast does not establish a dominant biological mechanism, a direct protective effect of the A allele, or a causal role for rs12543318 [29]. Further functional investigations are required to determine whether this association reflects a regulatory effect of rs12543318 itself or of another functional variant in linkage disequilibrium [40].

4.2. Population Differences

Population genetic background may also contribute to differences across studies. In our Vietnamese controls, the A allele frequencies of rs7552 (0.217) and rs12543318 (0.471) were relatively close to those reported in the Vietnamese Genetic Variation Database (0.241 and 0.369, respectively) and the 1000 Genomes East Asian population (0.268 and 0.418) but differed substantially from European reference frequencies (0.711 and 0.645) [24,41]. These distributions highlight population-level differences in allele frequencies that may influence statistical power and the interpretation of associations across ancestries [42]. Differences in linkage disequilibrium patterns may also affect how well these markers capture underlying functional variants, although this possibility has not been established for the loci examined here [43]. Nevertheless, population background alone cannot explain the inconsistent findings. For rs12543318, de Aquino et al. (2014) did not identify a significant association after multiple-testing correction in a Brazilian cohort [21]. Similarly, although previous studies reported associations involving rs7552 in Chinese, Brazilian, and Kuwaiti populations [12,13,23]-its association with NSCLO in our case–control analysis did not withstand multiple-testing correction, despite significant under-transmission of the A allele in affected trios. Differences in the number of variants and genetic models tested, correction procedures, and sample sizes may therefore contribute to variation in statistical significance across studies [42,44]. Taken together, these findings support the importance of phenotype-specific analyses and population-aware interpretation, while emphasizing the need for independent replication under comparable analytical frameworks [8].

4.3. Case–Control Versus Family-Based Evidence

The case–control and family-based analyses provided complementary but unequal levels of evidence for the two variants. For rs12543318, the increased odds of NSCLO

associated with the CC genotype in the case–control analysis were directionally consistent with the under-transmission of the A allele in the family-based analysis. This agreement provides complementary within-cohort evidence across two distinct analytical approaches.

Because the TDT evaluates allelic transmission within families and is robust to population stratification, it provides additional statistical support for the observed association [26]. Nevertheless, the case–control and family-based analyses included the same affected individuals and therefore do not constitute independent replication [45]. Moreover, the genotype-based case–control comparison and the allelic TDT evaluate different statistical contrasts, and their effect estimates should not be interpreted as directly equivalent [29]. Confirmation of the association in an independent Vietnamese or other Southeast Asian cohort remains necessary.

In contrast, the evidence for rs7552 was stronger in the family-based analysis than in the case–control comparison. Exact binomial sensitivity analyses supported both TDT findings, while conditional power analyses showed lower sensitivity to modest transmission effects for rs7552 because fewer informative transmissions were available. Differences in statistical information and sampling variability may contribute to the contrasting levels of statistical support, but their relative contributions cannot be established from the present analyses [46]. Independent replication is therefore particularly important for clarifying the association involving rs7552.

An additional contribution of the present study is the comparison of association evidence obtained from two complementary analytical designs within the same study population. While rs12543318 showed directionally consistent findings across both approaches, rs7552 demonstrated stronger evidence in the transmission analysis than in the case–control comparison. Reporting these patterns, including the findings that did not reach statistical significance, provides a more comprehensive assessment of the evidence for these established susceptibility variants in the Vietnamese population. However, because the two analyses shared affected individuals, their agreement should not be interpreted as independent replication.

4.4. Functional Interpretation of the 8q21.3 Locus: MMP16 as a Candidate Regulatory Target

The association of rs12543318 with NSCLO identifies a susceptibility signal at 8q21.3 but does not establish the causal variant or the gene mediating this association. Although rs12543318 is located in an intergenic region near *DCAF4L2*, genomic proximity alone does not establish the identity of its regulatory target [40].

Database annotations identified rs11784027 as a strongly linked candidate regulatory variant in the Asian reference population ($r^2 = 0.88$, $D' = 0.98$). Both rs12543318 and rs11784027 showed eQTL associations with *MMP16* expression in adult aortic tissue. These findings prioritize *MMP16* as a candidate regulatory target but do not distinguish the functional contributions of the two variants. The observed associations could reflect long-range cis-regulation or linkage disequilibrium with another functional variant [40].

The biological plausibility of *MMP16* is supported by previous genetic and experimental studies. Lin et al. (2023) identified associations between common variants in the *MMP16* region and NSCL/P in Western Han Chinese individuals, with significant findings in the NSCLO subgroup after multiple-testing correction [22]. Furthermore, *MMP16* encodes a membrane-type matrix metalloproteinase involved in extracellular matrix remodeling, a process relevant to craniofacial morphogenesis. Experimental studies in mice with combined MMP14 and *MMP16* deficiency support the developmental importance of these enzymes, including extracellular matrix permissiveness and mesenchymal proliferation, although the individual contribution of *MMP16* remains incompletely understood [47]. Altered matrix

remodeling could hypothetically influence epithelial–mesenchymal interactions during facial development [6,47].

Collectively, the regional genetic associations, adult-tissue eQTL evidence, and previously reported biological functions provide a rationale for prioritizing *MMP16* for further investigation. However, these findings do not establish that rs12543318 or rs11784027 regulates *MMP16* during embryonic craniofacial development. The LD estimates were derived from an external reference population, and no colocalization or functional validation was performed. Therefore, the proposed regulatory relationship remains hypothetical.

Further fine-mapping, chromatin interaction analyses, such as 3C, 4C, or Hi-C, and functional assays would be needed to determine whether this locus directly regulates *MMP16* expression during craniofacial development.

4.5. Functional Interpretation of rs7552: Potential Regulatory Role of CYRIA/FAM49A

Unlike the intergenic variant rs12543318, rs7552 is located within the 3′ untranslated region (3′UTR) of *CYRIA/FAM49A*, providing a genomic basis for investigating potential post-transcriptional regulatory mechanisms. Nevertheless, its location within *CYRIA/FAM49A* does not establish a direct functional effect on this gene.

Computational predictions from miRNASNP-v3 indicated allele-dependent gains and losses of potential miRNA-binding sites, including those involving hsa-miR-6800-5p and hsa-miR-4781-3p. These predictions raise the possibility that rs7552 may influence miRNA-mediated regulation of *CYRIA/FAM49A* transcripts, although altered miRNA binding, mRNA stability, and gene expression have not been experimentally demonstrated.

In addition to these computational predictions, previously reported GTEx data identified an eQTL association with *CYRIA/FAM49A* expression in adult frontal cortex and an sQTL association in adult thyroid tissue. These empirical QTL associations provide additional evidence for the potential regulatory relevance of the locus but do not establish a causal relationship between rs7552 and *CYRIA/FAM49A* expression or splicing [39].

Previous experimental studies have demonstrated that CYRI proteins participate in RAC1-dependent regulation of actin dynamics and cell migration [14,48]. These cellular processes are relevant to craniofacial morphogenesis, particularly the migration and organization of cranial neural crest-derived cells [49]. However, whether rs7552 influences these developmental processes through altered *CYRIA/FAM49A* regulation remains unknown. Changes in cell migration and organization could hypothetically disrupt epithelial–mesenchymal interactions during facial morphogenesis [6,49].

Furthermore, the available regulatory data were derived from adult non-craniofacial tissues, and the extensive LD region surrounding rs7552 prevents definitive attribution of the observed regulatory associations to the index SNP itself. Therefore, *CYRIA/FAM49A* should be considered a biologically plausible candidate regulatory target rather than an established causal gene for NSCLO. Functional validation in developmentally relevant models is needed to clarify these relationships.

4.6. Sex-Stratified Genetic Associations and Potential Effect Modification by Sex

Sex differences in the occurrence of NSCL/P provided a rationale for the prespecified secondary sex-stratified analyses. Epidemiological studies have reported an approximately 2:1 male-to-female ratio [3,50], while Pham et al. (2023) reported a ratio of 1.41:1 among 1413 Vietnamese patients with CL/P [51]. Previous genome-wide interaction studies have also identified loci at which genetic associations with NSCL/P differed between males and females, suggesting a potential contribution of sex-dependent genetic regulation to orofacial cleft susceptibility [30,31].

In the present study, the effect estimate for the rs12543318 CC genotype in NSCLO was larger in males (OR = 2.72, 95% CI: 1.41–5.24) than in females (OR = 1.18, 95% CI: 0.59–2.35). Although the association in males met the Bonferroni-adjusted significance threshold, the formal SNP-by-sex interaction was not statistically significant (likelihood-ratio $p = 0.0846$). Statistical significance in one subgroup but not another does not establish a difference in genetic effects between subgroups [52].

From a biological perspective, sex-dependent gene expression, hormonal regulation, epigenetic modifications, and interactions with sex chromosomes have been proposed as potential contributors to sex differences in orofacial cleft susceptibility [31]. However, whether any of these mechanisms influence the association between rs12543318 and NSCLO remains unknown. Therefore, the observed sex-stratified pattern should be considered hypothesis-generating rather than evidence of a confirmed sex-dependent genetic effect. Larger independent cohorts with sufficient statistical power to evaluate SNP-by-sex interactions are required.

4.7. Study Limitations and Future Directions

Several limitations should be acknowledged. First, the relatively modest sample size limited statistical power, particularly for phenotype-specific and sex-stratified analyses. Sensitivity power analyses indicated that the available sample was primarily powered to detect relatively large genetic effects. Notably, the observed odds ratio for rs12543318 in NSCLO (OR = 1.82) was below the corresponding minimum detectable odds ratio of 2.20 at 80% power. Therefore, non-significant findings, particularly in NSCLP, do not exclude modest genetic effects [53]. The limited subgroup sample sizes may also have reduced statistical power to detect SNP-by-sex interactions. Although the effect estimate for rs12543318 in NSCLO was larger in males, the non-significant interaction and relatively wide confidence intervals preclude firm conclusions regarding effect modification by sex [31,52]. All available samples meeting the eligibility criteria were included in the present study. Additional eligible samples were not available within the current study cohort. Therefore, increasing the sample size would require further recruitment or collaboration with independent cohorts.

Second, although Bonferroni correction was applied within the defined primary, complementary family-based, and secondary sex-stratified testing families, additional exploratory analyses were not adjusted for multiplicity [54]. Consequently, the overall risk of false-positive findings across the full set of analyses cannot be excluded, and exploratory associations require independent confirmation.

Third, although the case–control and family-based analyses provided directionally consistent evidence for rs12543318 in NSCLO, both approaches included the same affected individuals. Their concordance therefore represents complementary within-cohort evidence rather than independent replication. Replication in a separate cohort is necessary to establish the reproducibility and generalizability of the observed association [45].

Fourth, although all participants self-identified as Kinh Vietnamese and were recruited from the same geographical region, genomic ancestry was not assessed using ancestry-informative markers or PCA. Previous population-genetic studies have reported broadly shared ancestry patterns and population-level genetic clustering among Kinh Vietnamese individuals [41,55]. However, shared self-reported ethnicity and geographical origin do not exclude subtle genetic substructure. Residual population stratification may therefore have influenced the case–control association estimates. The complementary family-based TDT, which evaluates allelic transmission within families, provides additional evidence that is less susceptible to confounding by population stratification [26].

Fifth, the investigation was restricted to two candidate SNPs, limiting assessment of broader genetic variation. Finally, functional interpretations were based on publicly

available database annotations, computational predictions, and regulatory associations derived primarily from adult non-craniofacial tissues. The tissue- and developmental stage-specific nature of gene regulation limits the extrapolation of these findings to embryonic craniofacial development. Furthermore, LD estimates were obtained from an external Asian reference population and may not fully reflect the genetic architecture of the Vietnamese cohort. In the absence of statistical fine-mapping, colocalization analysis, and experimental validation, neither the causal variants nor the genes mediating NSOFC susceptibility could be definitively established.

Future studies should prioritize replication in larger independent Vietnamese and other Southeast Asian cohorts incorporating genomic ancestry assessment, particularly to evaluate potential phenotype- and sex-related differences in genetic effects, followed by fine-mapping and functional validation in developmentally relevant craniofacial models to clarify the potential roles of *CYRIA*/*FAM49A* and *MMP16*.

5. Conclusions

This study extends the existing evidence on NSOFC susceptibility by providing phenotype-resolved genetic association and allelic transmission estimates for rs12543318 and rs7552 in a Vietnamese population. The CC genotype of rs12543318 was associated with increased odds of NSCLO compared with carriers of at least one A allele. The concordant under-transmission of the A allele provided complementary within-cohort evidence, consistent with previous findings in Western Han Chinese individuals, but did not constitute independent replication.

For rs7552, significant transmission distortion was observed despite limited case–control evidence. Differential genetic effects between cleft phenotypes or sexes remain unconfirmed. Although database-based annotations prioritize *MMP16* and *CYRIA*/*FAM49A* as candidate regulatory target genes, the causal variants and disease-relevant biological mechanisms remain to be established. These findings contribute to the genetic epidemiology of NSOFCs in Vietnam and provide a basis for larger independent replication studies in Vietnam and other Southeast Asian populations and functional investigations.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/genes17121536/s1. Table S1: Final genotyping call rates and genotype call confidence scores.

Author Contributions: Conceptualization, C.K.N., T.N., H.K. and S.S.; methodology, C.K.N. and H.I.; validation, L.N.G.P., V.T.T.N., D.M.N. and T.M.D.; formal analysis, C.K.N.; investigation, H.K., S.S., N.N. (Nagaga Natsume), Y.A., and M.D.N.; resources, H.I., M.D.N. and N.N. (Nagaga Natsume); data curation, C.K.N. and A.D.M.N.; writing—original draft preparation, C.K.N.; writing—review and editing, C.K.N., R.M., A.D.M.N., T.H., C.S., K.K. and N.N. (Nagato Natsume); visualization, C.K.N.; supervision, N.N. (Nagato Natsume), M.D.N. and H.K.; project administration, N.N. (Nagato Natsume), M.D.N., T.N., H.I., H.F., K.F. and T.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted under the research protocol entitled "Genetic Analysis Research by the Joint Research Organization for Genes Related to Oral Diseases," approved by Aichi Gakuin University on 21 August 2018 (Approval No. 78). All procedures have been carried out in accordance with the ethical guidelines of the Declaration of Helsinki. Peripheral blood samples were collected at the Odonto-Maxillo-Facial Hospital of Ho Chi Minh City and were transferred to Aichi Gakuin University under a Memorandum of Understanding (MOU) between the two institutions, which was in effect from 2003 to 2025. A subsequent amendment extending the study period through 26 August 2028 was approved on 7 July 2025 (Approval No. 750).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patients to publish this paper.

Data Availability Statement: The datasets used and analyzed during this study are available from the corresponding author upon reasonable request.

Acknowledgments: We sincerely thank the doctors and technicians at the Laboratory and Maxillofacial Surgery Departments, as well as the General Planning Department of Ho Chi Minh City Odonto-Maxillo-Facial Hospital, for their help in conducting this research. Grammarly was used for grammatical editing. The authors reviewed and verified all AI-assisted content and take full responsibility for the final manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NSOFC	Non-syndromic orofacial cleft
NSCLO	Non-syndromic cleft lip only
NSCLP	Non-syndromic cleft lip and palate
NSCL/P	Non-syndromic cleft lip with or without cleft palate
NSCPO	Non-syndromic cleft palate only
SNP	Single-nucleotide polymorphism
TDT	Transmission disequilibrium test
POO	Parent of origin
MDE	Minimum detectable effect
MAFs	Minor allele frequencies
HWE	Hardy–Weinberg equilibrium

Appendix A

Table A1. Sex-stratified associations and genotype-by-sex interactions.

SNP	Genetic Comparison	Phenotype	Female OR (95% CI)	*p*-Value	Male OR (95% CI)	*p*-Value	Interaction OR (95% CI)	*p*-Value Interaction
rs12543318	CC vs. AC, AA	NSCLO	1.18 (0.59–2.35)	0.6342	2.72 (1.41–5.24)	0.0025	2.30 (0.89–5.95)	0.0846
		NSCLP	1.04 (0.53–2.05)	0.9079	1.17 (0.60–2.30)	0.6437	1.13 (0.43–2.93)	0.8074
rs7552	GG vs. AG, AA	NSCLO	2.09 (1.03–4.24)	0.0407	1.24 (0.62–2.44)	0.5437	0.59 (0.22–1.58)	0.2924
		NSCLP	1.15 (0.61–2.17)	0.6677	0.81 (0.43–1.53)	0.512	0.70 (0.29–1.73)	0.4425

NSCLO, nonsyndromic cleft lip only; NSCLP, nonsyndromic cleft lip with palate; SNP, single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; LRT, likelihood ratio test. Sex-stratified associations are presented as ORs with 95% CIs. Genotype-by-sex interactions were evaluated using logistic regression models, including a genotype-by-sex interaction term. Interaction *p*-values were obtained using likelihood ratio tests comparing models with and without the interaction term. All tests were two-sided.

Table A2. Minimum detectable effect sizes at 80% statistical power for case–control and family-based analyses.

Analysis	SNP	Phenotype	Sample Size/Informative Transmissions	MDE OR, Risk/Over-Transmission	MDE OR, Protective/Under-Transmission	Observed OR (95% CI)
Case–control	rs7552	NSCLO	111 cases/240 controls	2.38	0.46	1.60 (0.98–2.62)
		NSCLP	120 cases/240 controls	2.32	0.47	0.97 (0.62–1.51)
	rs12543318	NSCLO	111 cases/240 controls	2.2	0.37	1.82 (1.14–2.91)
		NSCLP	120 cases/240 controls	2.16	0.38	1.10 (0.69–1.78)

Table A2. *Cont.*

Analysis	SNP	Phenotype	Sample Size/Informative Transmissions	MDE OR, Risk/Over-Transmission	MDE OR, Protective/Under-Transmission	Observed OR (95% CI)
Family-based TDT	rs7552	NSCLO	79 transmissions	2.28	0.44	0.52 (0.33–0.83)
	rs12543318	NSCLO	128 transmissions	1.85	0.54	0.58 (0.41–0.83)

Note: MDE, minimum detectable effect; OR, odds ratio; NSCLO, nonsyndromic cleft lip only; NSCLP, nonsyndromic cleft lip and palate; TDT, transmission disequilibrium test. Minimum detectable effect sizes (MDEs) were estimated at 80% statistical power using a two-sided Bonferroni-adjusted significance threshold of $\alpha = 0.0125$. For case–control analyses, MDEs were calculated separately for each SNP and phenotype under the dominant genetic comparison, using the observed control genotype frequencies and available case–control sample sizes. A two-sample comparison-of-proportions approach with pooled variance was used to identify detectable genotype-frequency differences, which were subsequently expressed as odds ratios for risk-increasing and protective effects. Family-based MDEs were estimated using an exact two-sided binomial framework conditional on the observed number of informative transmissions, assuming independent informative transmissions. Case–control odds ratios and transmission odds ratios represent different estimands and should not be directly compared.

Table A3. Sensitivity analysis of transmission disequilibrium using exact binomial tests in NSCLO trios.

SNP	Transmitted A (T)	Untransmitted A (U)	Informative Transmissions (N)	Transmission Rate (%)	TDT OR	Exact *p*-Value
rs12543318	47	81	128	36.72	0.58	0.0034
rs7552	27	52	79	34.18	0.519	0.0066

TDT, transmission disequilibrium test; NSCLO, nonsyndromic cleft lip only; OR, odds ratio. Exact two-sided binomial tests were performed under the null hypothesis of equal transmission probabilities ($p = 0.5$). Transmission rates were calculated as T/(T + U), and transmission odds ratios as T/U. Both variants remained statistically significant at the Bonferroni-adjusted threshold of $\alpha = 0.0125$.

Table A4. Exploratory heterogeneity between NSCLO and NSCLP.

Analysis	SNP	Contrast	OR (95% CI)	*p* (LRT)
Sex-adjusted case–case	rs12543318	CC vs. AC+AA	1.65 (0.96–2.85)	0.0686
Sex-adjusted case–case	rs7552	GG vs. AG+AA	1.66 (0.96–2.89)	0.0698
TDT heterogeneity	rs12543318	A; CLO vs. CLP	0.85 (0.51–1.43)	0.5427
TDT heterogeneity	rs7552	A; CLO vs. CLP	0.60 (0.32–1.14)	0.1201

Case–case ORs compare NSCLO with NSCLP as the reference outcome, adjusted for sex. TDT ORs are ratios of transmission odds (NSCLO/NSCLP). *p* values are nominal; LRT, likelihood-ratio test.

References

1. Salari, N.; Darvishi, N.; Heydari, M.; Bokaee, S.; Darvishi, F.; Mohammadi, M. Global prevalence of cleft palate, cleft lip and cleft palate and lip: A comprehensive systematic review and meta-analysis. *J. Stomatol. Oral Maxillofac. Surg.* **2022**, *123*, 110–120. [CrossRef] [PubMed]
2. Natsume, N.; Kawai, T.; Le, H. Letter to the Editor. *Cleft Palate Craniofac. J.* **1998**, *35*, 183. [CrossRef] [PubMed]
3. IPDTOC Working Group. Prevalence at birth of cleft lip with or without cleft palate: Data from the International Perinatal Database of Typical Oral Clefts (IPDTOC). *Cleft Palate Craniofac. J.* **2011**, *48*, 66–81. [CrossRef] [PubMed]
4. Jugessur, A.; Farlie, P.G.; Kilpatrick, N. The genetics of isolated orofacial clefts: From genotypes to subphenotypes. *Oral Dis.* **2009**, *15*, 437–453. [CrossRef] [PubMed]
5. Dixon, M.J.; Marazita, M.L.; Beaty, T.H.; Murray, J.C. Cleft lip and palate: Understanding genetic and environmental influences. *Nat. Rev. Genet.* **2011**, *12*, 167–178. [CrossRef] [PubMed]
6. Hammond, N.L.; Dixon, M.J. Revisiting the embryogenesis of lip and palate development. *Oral Dis.* **2022**, *28*, 1306–1326. [CrossRef] [PubMed]
7. Mahamad Irfanulla Khan, A.N.; Prashanth, C.S.; Srinath, N. Genetic etiology of cleft lip and cleft palate. *AIMS Mol. Sci.* **2020**, *7*, 328–348. [CrossRef]

8. Mukhopadhyay, N.; Feingold, E.; Moreno-Uribe, L.; Wehby, G.; Valencia-Ramirez, L.C.; Restrepo Muneton, C.P.; Padilla, C.; Deleyiannis, F.; Christensen, K.; Poletta, F.A.; et al. Genome-wide association study of multiethnic nonsyndromic orofacial cleft families identifies novel loci specific to family and phenotypic subtypes. *Genet. Epidemiol.* **2022**, *46*, 182–198. [CrossRef] [PubMed]
9. Yang, Y.; Suzuki, A.; Iwata, J.; Jun, G. Secondary Genome-Wide Association Study Using Novel Analytical Strategies Disentangle Genetic Components of Cleft Lip and/or Cleft Palate in 1q32.2. *Genes* **2020**, *11*, 1280. [CrossRef] [PubMed]
10. Huang, L.; Jia, Z.; Shi, Y.; Du, Q.; Shi, J.; Wang, Z.; Mou, Y.; Wang, Q.; Zhang, B.; Wang, Q.; et al. Genetic factors define CPO and CLO subtypes of nonsyndromicorofacial cleft. *PLoS Genet.* **2019**, *15*, e1008357. [CrossRef] [PubMed]
11. Yu, Y.; Zuo, X.; He, M.; Gao, J.; Fu, Y.; Qin, C.; Meng, L.; Wang, W.; Song, Y.; Cheng, Y.; et al. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. *Nat. Commun.* **2017**, *8*, 14364. [CrossRef] [PubMed]
12. Chen, C.; Guo, Q.; Shi, J.; Jiao, X.; Lv, K.; Liu, X.; Jiang, Y.; Hui, X.; Song, T. Genetic variants of MGMT, RHPN2, and FAM49A contributed to susceptibility of nonsyndromic orofacial clefts in a Chinese population. *J. Oral Pathol. Med.* **2018**, *47*, 796–801. [CrossRef] [PubMed]
13. Machado, R.A.; Nogueira, E.N.; Martelli-Junior, H.; Reis, S.R.; Persuhn, D.C.; Coletta, R.D. 2p24.2 (rs7552) is a susceptibility locus for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Clin. Genet.* **2018**, *93*, 1199–1204. [CrossRef] [PubMed]
14. Fort, L.; Batista, J.M.; Thomason, P.A.; Spence, H.J.; Whitelaw, J.A.; Tweedy, L.; Greaves, J.; Martin, K.J.; Anderson, K.I.; Brown, P.; et al. Fam49/CYRI interacts with Rac1 and locally suppresses protrusions. *Nat. Cell Biol.* **2018**, *20*, 1159–1171. [CrossRef] [PubMed]
15. Whitelaw, J.A.; Lilla, S.; Paul, N.R.; Fort, L.; Zanivan, S.; Machesky, L.M. CYRI/Fam49 Proteins Represent a New Class of Rac1 Interactors. *Commun. Integr. Biol.* **2019**, *12*, 112–118. [CrossRef] [PubMed]
16. Dack, K.; Ludwig, K.U.; Stergiakouli, E.; Sandy, J.; Aryee, S.; Davey Smith, G.; Davies, A.; Wren, Y.; Sharp, G.C.; Humphries, K.; et al. Genetic heterogeneity and homogeneity among orofacial cleft subtypes: Genome-wide association studies in the cleft collective. *Hum. Mol. Genet.* **2025**, *34*, 1934–1950. [CrossRef] [PubMed]
17. Garg, P.; Ludwig, K.U.; Bohmer, A.C.; Rubini, M.; Steegers-Theunissen, R.; Mossey, P.A.; Mangold, E.; Sharp, A.J. Genome-wide analysis of parent-of-origin effects in non-syndromic orofacial clefts. *Eur. J. Hum. Genet.* **2014**, *22*, 822–830. [CrossRef] [PubMed]
18. Ludwig, K.U.; Mangold, E.; Herms, S.; Nowak, S.; Reutter, H.; Paul, A.; Becker, J.; Herberz, R.; Alchawa, T.; Nasser, E.; et al. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nat. Genet.* **2012**, *44*, 968–971. [CrossRef] [PubMed]
19. Pan, Y.; Han, Y.; Zhang, H.; Zhou, L.; Li, D.; Cai, Q.; Ma, J.; Zhang, W.; Wang, L. Association and cumulative effects of GWAS-identified genetic variants for nonsyndromic orofacial clefts in a Chinese population. *Environ. Mol. Mutagen.* **2013**, *54*, 261–267. [CrossRef] [PubMed]
20. Zhang, B.H.; Huang, N.; Shi, J.Y.; Shi, B.; Jia, Z.L. Homozygote C/C at rs12543318 was risk factor for non-syndromic cleft lip only from Western Han Chinese population. *J. Oral Pathol. Med.* **2018**, *47*, 620–626. [CrossRef] [PubMed]
21. De Aquino, S.N.; Messetti, A.C.; Hoshi, R.; Borges, A.; Viena, C.S.; Reis, S.R.; Oliveira Swerts, M.S.; Graner, E.; Martelli-Junior, H.; Coletta, R.D. Analysis of susceptibility polymorphisms for nonsyndromic cleft lip with or without cleft palate in the Brazilian population. *Birth Defects Res. A Clin. Mol. Teratol.* **2014**, *100*, 36–42. [CrossRef] [PubMed]
22. Lin, Y.; Shi, J.; Shi, B.; Jia, Z. MMP16 as NSCL +/- P Susceptible Gene in Western Han Chinese. *Cleft Palate Craniofac. J.* **2023**, *60*, 1625–1631. [CrossRef] [PubMed]
23. Abdelhafez, N.; Aladsani, A.; Alkharafi, L.; Al-Bustan, S. Association of selected gene variants with nonsyndromic orofacial clefts in Kuwait. *Gene* **2025**, *934*, 149028. [CrossRef] [PubMed]
24. Genomes Project, C.; Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; Mccarthy, S.; Mcvean, G.A.; et al. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74. [CrossRef] [PubMed]
25. Carlson, J.C.; Anand, D.; Butali, A.; Buxo, C.J.; Christensen, K.; Deleyiannis, F.; Hecht, J.T.; Moreno, L.M.; Orioli, I.M.; Padilla, C.; et al. A systematic genetic analysis and visualization of phenotypic heterogeneity among orofacial cleft GWAS signals. *Genet. Epidemiol.* **2019**, *43*, 704–716. [CrossRef] [PubMed]
26. Laird, N.M.; Lange, C. Family-based designs in the age of large-scale gene-association studies. *Nat. Rev. Genet.* **2006**, *7*, 385–394. [CrossRef] [PubMed]
27. Faul, F.; Erdfelder, E.; Lang, A.G.; Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **2007**, *39*, 175–191. [CrossRef] [PubMed]
28. Cerezo, M.; Sollis, E.; Ji, Y.; Lewis, E.; Abid, A.; Bircan, K.O.; Hall, P.; Hayhurst, J.; John, S.; Mosaku, A.; et al. The NHGRI-EBI GWAS Catalog: Standards for reusability, sustainability and diversity. *Nucleic Acids Res.* **2025**, *53*, D998–D1005. [CrossRef] [PubMed]
29. Balding, D.J. A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* **2006**, *7*, 781–791. [CrossRef] [PubMed]
30. Awotoye, W.; Comnick, C.; Pendleton, C.; Zeng, E.; Alade, A.; Mossey, P.A.; Gowans, L.J.J.; Eshete, M.A.; Adeyemo, W.L.; Naicker, T.; et al. Genome-wide Gene-by-Sex Interaction Studies Identify Novel Nonsyndromic Orofacial Clefts Risk Locus. *J. Dent. Res.* **2022**, *101*, 465–472. [CrossRef] [PubMed]

31. Carlson, J.C.; Nidey, N.L.; Butali, A.; Buxo, C.J.; Christensen, K.; Deleyiannis, F.W.; Hecht, J.T.; Field, L.L.; Moreno-Uribe, L.M.; Orioli, I.M.; et al. Genome-wide interaction studies identify sex-specific risk alleles for nonsyndromic orofacial clefts. *Genet. Epidemiol.* **2018**, *42*, 664–672. [CrossRef] [PubMed]
32. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [CrossRef] [PubMed]
33. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]
34. Mclaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [CrossRef] [PubMed]
35. Dong, S.; Zhao, N.; Spragins, E.; Kagda, M.S.; Li, M.; Assis, P.R.; Jolanki, O.; Luo, Y.; Cherry, J.M.; Boyle, A.P.; et al. Annotating and prioritizing human non-coding variants with RegulomeDB v.2. *Nat. Genet.* **2023**, *55*, 724–726. [CrossRef] [PubMed]
36. Ward, L.D.; Kellis, M. HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **2016**, *44*, D877–D881. [CrossRef] [PubMed]
37. Liu, C.J.; Fu, X.; Xia, M.; Zhang, Q.; Gu, Z.; Guo, A.Y. miRNASNP-v3: A comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. *Nucleic Acids Res.* **2021**, *49*, D1276–D1281. [CrossRef] [PubMed]
38. Consortium, G.T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **2020**, *369*, 1318–1330. [CrossRef] [PubMed]
39. Cano-Gamez, E.; Trynka, G. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front. Genet.* **2020**, *11*, 424. [CrossRef] [PubMed]
40. Schierding, W.; Cutfield, W.S.; O'sullivan, J.M. The missing story behind Genome Wide Association Studies: Single nucleotide polymorphisms in gene deserts have a story to tell. *Front. Genet.* **2014**, *5*, 39. [CrossRef] [PubMed]
41. Le, V.S.; Tran, K.T.; Bui, H.T.P.; Le, H.T.T.; Nguyen, C.D.; Do, D.H.; Ly, H.T.T.; Pham, L.T.D.; Dao, L.T.M.; Nguyen, L.T. A Vietnamese human genetic variation database. *Hum. Mutat.* **2019**, *40*, 1664–1675. [CrossRef] [PubMed]
42. Sham, P.C.; Purcell, S.M. Statistical power and significance testing in large-scale genetic studies. *Nat. Rev. Genet.* **2014**, *15*, 335–346. [CrossRef] [PubMed]
43. Rosenberg, N.A.; Huang, L.; Jewett, E.M.; Szpiech, Z.A.; Jankovic, I.; Boehnke, M. Genome-wide association studies in diverse populations. *Nat. Rev. Genet.* **2010**, *11*, 356–366. [CrossRef] [PubMed]
44. Conneely, K.N.; Boehnke, M. Meta-analysis of genetic association studies and adjustment for multiple testing of correlated SNPs and traits. *Genet Epidemiol.* **2010**, *34*, 739–746. [CrossRef] [PubMed]
45. NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. *Nature* **2007**, *447*, 655–660. [CrossRef] [PubMed]
46. Haldar, T.; Ghosh, S. Power comparison between population-based case-control studies and family-based transmission-disequilibrium tests: An empirical study. *Indian J. Hum. Genet.* **2011**, *17*, S27–S31. [CrossRef] [PubMed]
47. Shi, J.; Son, M.Y.; Yamada, S.; Szabova, L.; Kahan, S.; Chrysovergis, K.; Wolf, L.; Surmak, A.; Holmbeck, K. Membrane-type MMPs enable extracellular matrix permissiveness and mesenchymal cell proliferation during embryogenesis. *Dev. Biol.* **2008**, *313*, 196–209. [CrossRef] [PubMed]
48. Thomas, P.S.; Kim, J.; Nunez, S.; Glogauer, M.; Kaartinen, V. Neural crest cell-specific deletion of Rac1 results in defective cell-matrix interactions and severe craniofacial and cardiovascular malformations. *Dev. Biol.* **2010**, *340*, 613–625. [CrossRef] [PubMed]
49. Cordero, D.R.; Brugmann, S.; Chu, Y.; Bajpai, R.; Jame, M.; Helms, J.A. Cranial neural crest cells on the move: Their roles in craniofacial development. *Am. J. Med. Genet. A* **2011**, *155A*, 270–279. [CrossRef] [PubMed]
50. Marazita, M.L. The evolution of human genetic studies of cleft lip and cleft palate. *Annu. Rev. Genom. Hum. Genet.* **2012**, *13*, 263–283. [CrossRef] [PubMed]
51. Pham, L.N.G.; Niimi, T.; Suzuki, S.; Imura, H.; Osakabe, R.; Hayami, K.; Natsume, N.; Nguyen, M.D.; Le, A.K.; Natsume, N. Descriptive Study of Orofacial Clefts in Vietnam on the occasion of the 50th Anniversary of the Establishment of Diplomatic Relations between Japan and Vietnam. *J. JAICOH* **2023**, *1*, 2–13.
52. Altman, D.G.; Bland, J.M. Interaction revisited: The difference between two estimates. *BMJ* **2003**, *326*, 219. [CrossRef] [PubMed]
53. Altman, D.G.; Bland, J.M. Absence of evidence is not evidence of absence. *BMJ* **1995**, *311*, 485. [CrossRef] [PubMed]
54. Bender, R.; Lange, S. Adjusting for multiple testing—When and how? *J. Clin. Epidemiol.* **2001**, *54*, 343–349. [CrossRef] [PubMed]
55. Macholdt, E.; Arias, L.; Duong, N.T.; Ton, N.D.; Van Phong, N.; Schroder, R.; Pakendorf, B.; Van Hai, N.; Stoneking, M. The paternal and maternal genetic history of Vietnamese populations. *Eur. J. Hum. Genet.* **2020**, *28*, 636–645. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.