Supplementary Figures and Tables

Table S1: Participant demographics based on genetically defined ancestry groups

	Cleft lip only		Cleft lip and palate		Controls
	Bilateral	Unilateral	Bilateral	Unilateral	
Number of participants	44 (8.1%)	502	572 (31.9%)	1221	1626
Number of families	44	434	530	1123	1626
Number with Asian ancestry	15 (34.1%)	184 (36.7%)	121 (21.4%)	205 (16.9%)	161 (9.9%)
Number with European ancestry	13 (29.5%)	151 (30.1%)	159 (28.1%)	315 (26.0%)	835 (51.4%)
Number with Latin/South American ancestry	15 (34.1%)	167 (33.3%)	285 (50.4%)	692 (57.0%)	626 (38.5%)
Number with Unspecified ancestry	1 (2.3%)	0 (0%)	1 (0.2%)	1 (0.1%)	4 (0.2%)

Table S6: Comparison of subtype-specific analyses

Comparison	Percent of SNPs with same effect size that are known	Percent of SNPs with different effect sizes that are known	Odds ratio	95% CI	P-value
BCL-UCL	1.8%	26.8%	0.05	0.02-0.11	2.41×10^{-20}
BCLP-UCLP	8.6%	44.0%	0.12	0.04-0.28	1.04×10^{-7}
BCL-BCLP	7.7%	31.4%	0.18	0.12-0.26	6.89×10^{-20}
UCL-UCLP	43.7%	54.3%	0.65	0.47-0.88	0.005

Table S9: Loci with suggestive significance in the modifier analysis of cleft lip with or without cleft palate

Locus	Top variant	Reference allele	Alternate allele	AF[1]	OR[2]	95%CI	P
20p11.23	rs12480480	G	A	0.17	1.49	1.37-1.62	1.72E-06
4p15.33	rs11373830	C	CA	0.07	1.99	1.73-2.3	2.47E-06
4p16.1	rs149050104	C	T	0.06	1.82	1.61-2.05	3.94E-06

[1]Allele frequency of alternate allele

[2]Risk of BCL over UCL

Table S10: Replication results for 20p11 in cleft lip

SNP	CHR	BP	Reference allele	Alternate allele	AF[1]	OR[2]	95% CI	P-value
rs6113495	20	22061487	G	A	0.46	1.05	0.57-1.92	0.8735
rs6113496	20	22061632	T	G	0.40	0.74	0.40-1.35	0.33
rs6047869	20	22062683	A	C	0.40	0.74	0.40-1.35	0.33
20:22062945	20	22062945	G	A	0.15	1.14	0.34-3.81	0.8314
rs115747027	20	22063265	T	C	0.40	0.74	0.40-1.35	0.33
20:22063347	20	22063347	T	C	0.13	3.83	1.64-8.95	0.001897
rs113999308	20	22063421	C	T	0.28	1.17	0.40-3.43	0.763
rs117160568	20	22063660	A	G	0.46	1.05	0.57-1.93	0.8575

[1]Allele frequency of alternate allele

[2]Risk of BCL over UCL

Table S11: Replication results for 20p11 in cleft lip and palate

SNP	CHR	BP	Reference allele	Alternate allele	AF[1]	OR[2]	95% CI	P-value
rs6113495	20	22061487	G	A	0.44	1.03	0.84-1.26	0.7304
rs6113496	20	22061632	T	G	0.39	0.90	0.73-1.11	0.3611
rs6047869	20	22062683	A	C	0.39	0.91	0.74-1.12	0.3846
20:22062934	20	22062934	T	G	0.05	0.94	0.55-1.60	0.8225
20:22062945	20	22062945	G	A	0.16	1.34	0.92-1.94	0.1232
rs115747027	20	22063265	T	C	0.39	0.90	0.73-1.11	0.3611
20:22063347	20	22063347	T	C	0.18	1.12	0.80-1.57	0.476
rs113999308	20	22063421	C	T	0.28	1.30	0.96-1.75	0.08084
rs117160568	20	22063660	A	G	0.44	1.03	0.84-1.26	0.7304

[1]Allele frequency of alternate allele

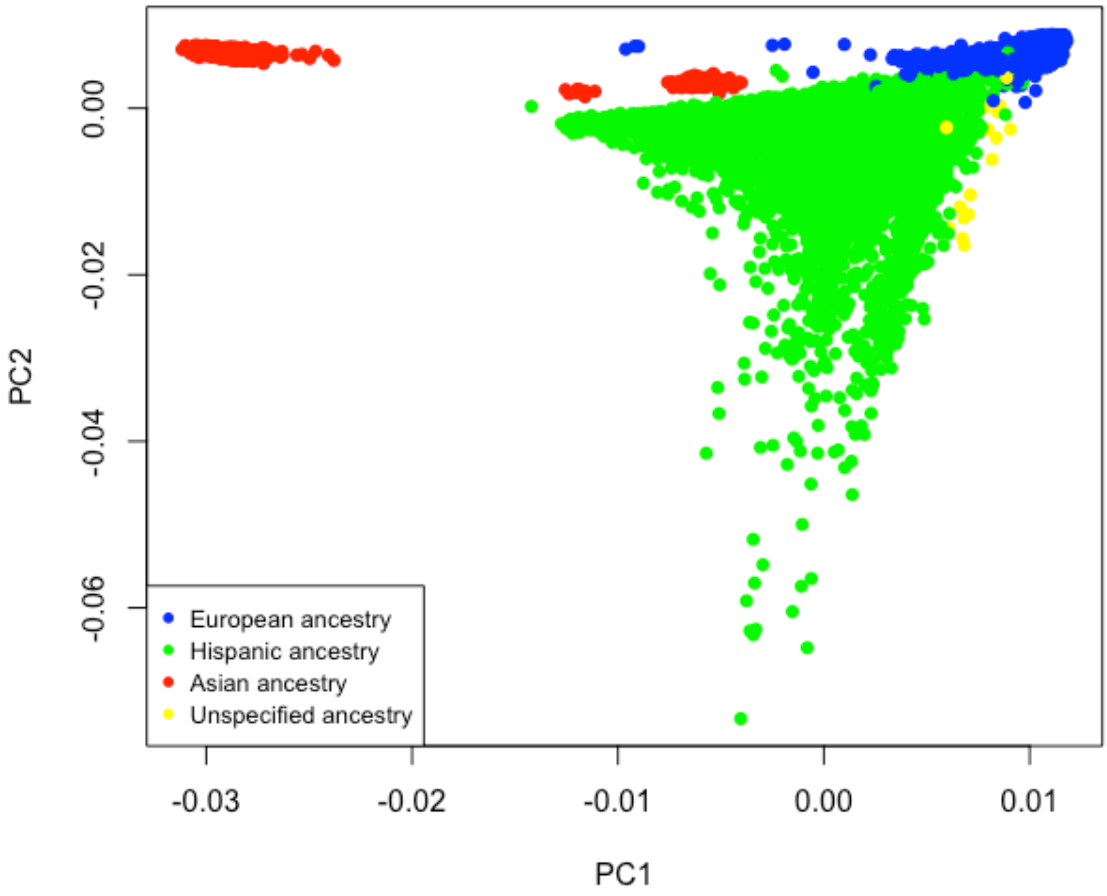
[2]Risk of BCLP over UCLP

Table S12: Comparison of known loci with modifier analyses

Comparison	Known SNPs	Odds ratio	95% CI	P-value
BCL v UCL	rs200656853; rs35640367	0.26	0.04-0.84	0.03
BCLP v UCLP	rs71776173; rs134124	0.53	0.08-1.71	0.32

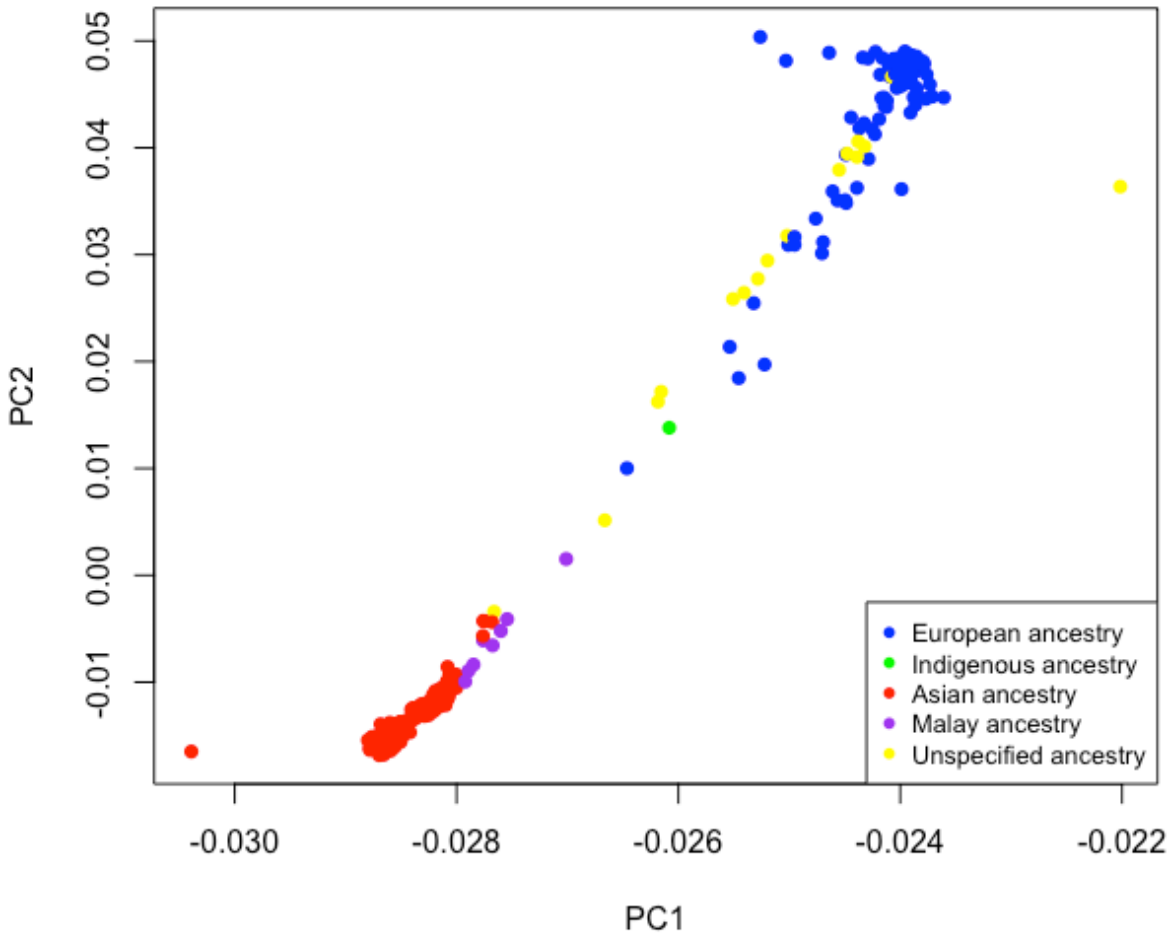
Supplementary figures

Figure S1: PC plot for the POFC cohort



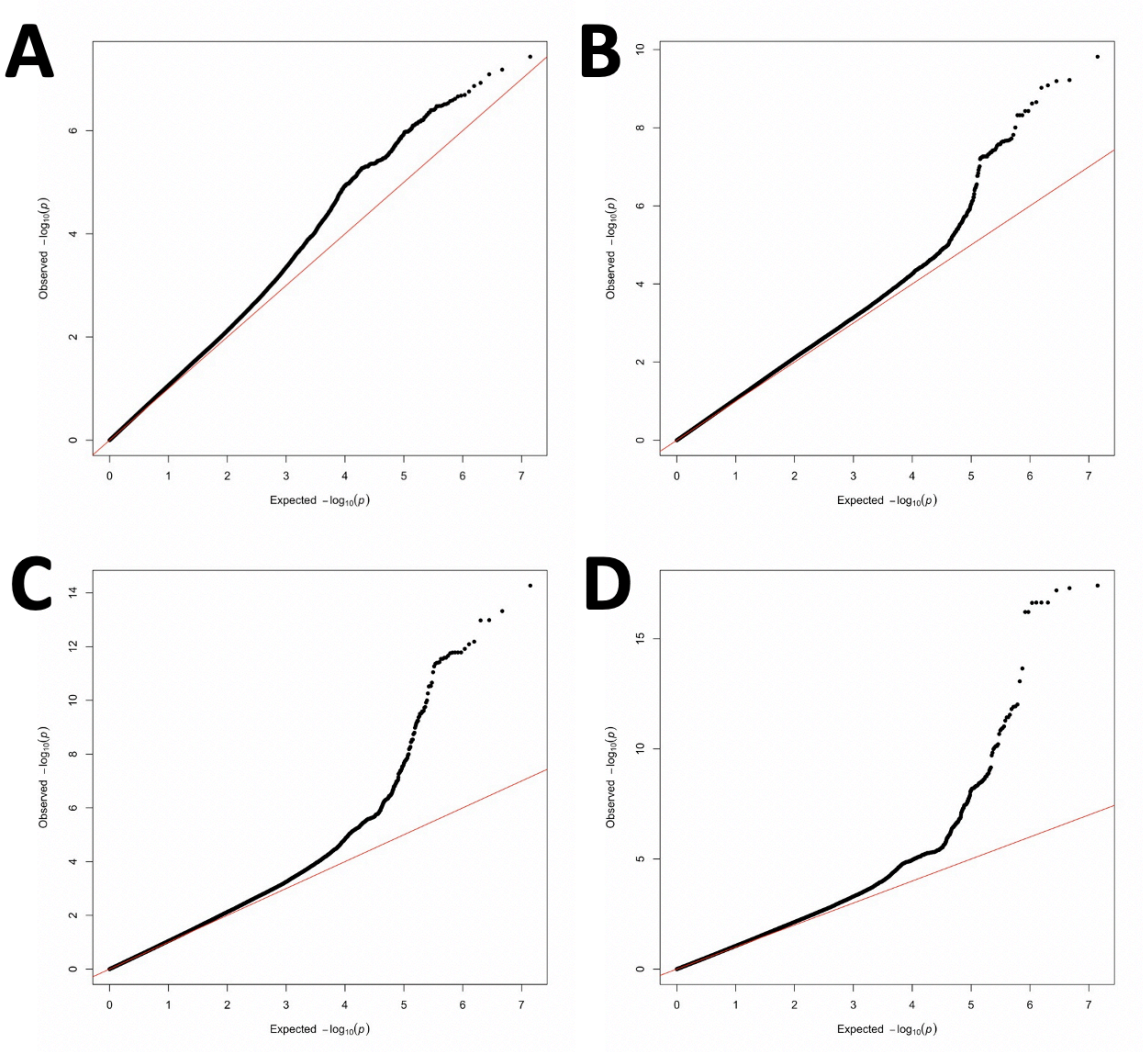
The first two principal components (PCs) for the POFC dataset used in this analysis, showing that a majority of the cohort is of European, Asian, or Hispanic ancestry.

Figure S2: PC plot for the GENEVA cohort



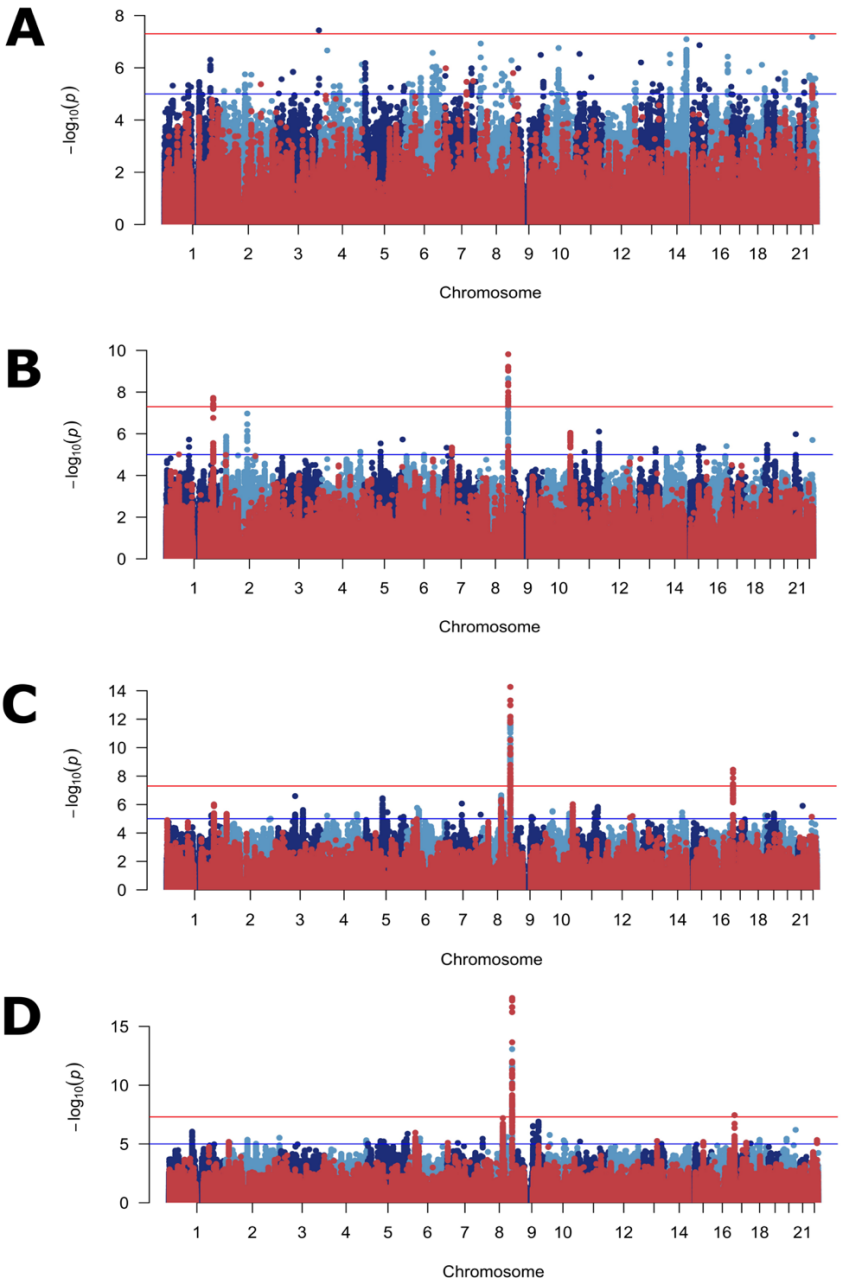
The first two principal components (PCs) for the GENEVA dataset used for replication in this analysis, showing that a majority of the cohort is of European and Asian ancestry.

Figure S3: QQ plots for the subtype-specific analyses



Q-Q plots of the −log10(p-values) of SNPs in the bilateral CL (A), unilateral CL (B), bilateral CLP (C), and unilateral CLP (D) subtype-specific case-control analysis. The genomic inflation factors are 1.10, 1.07, 1.05, and 1.06, respectively.

Figure S4: Manhattan plots for the subtype-specific analyses.



Manhattan plots of $-\log_{10}$(p-values) from the case-control analyses of bilateral cleft lip (A), unilateral cleft lip (B), bilateral cleft lip and palate (C), and unilateral cleft lip and palate (D). Lines indicate suggestive (blue) and genome-wide (red) thresholds for statistical significance. The points in red indicate SNPs within 50kb of a locus previously implicated in nonsyndromic clefting. The genomic inflation factors are 1.10, 1.07, 1.05, and 1.06 respectively.

Figure S5: Regional association plots showing $-\log_{10}$(p-values) for the novel genome-wide suggestive peaks at 14q32 in the modifier analysis in bilateral CL (A), unilateral CL (B), bilateral CLP (C), and unilateral CLP (D). Plots were generated using LocusZoom. The recombination overlay (blue line, right y-axis) indicates the boundaries of the LD block. Points are color coded according to pairwise LD (r^2) with the index SNP.

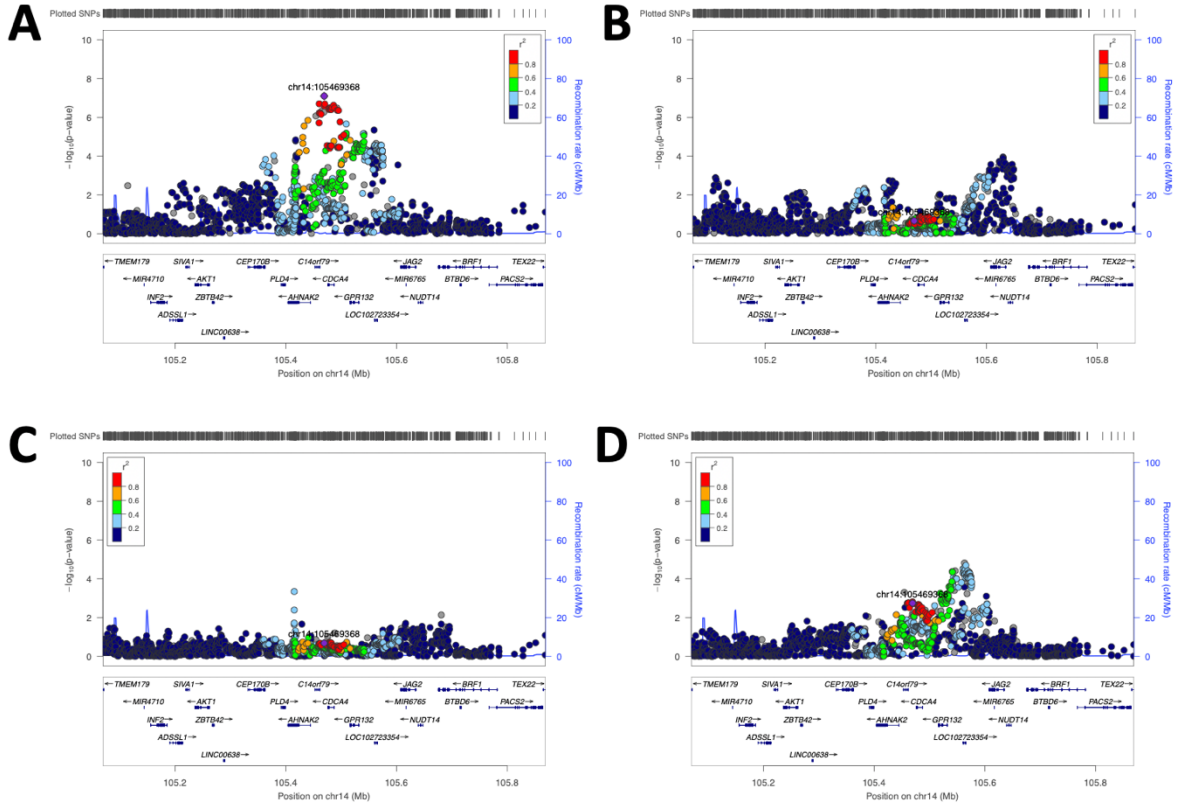
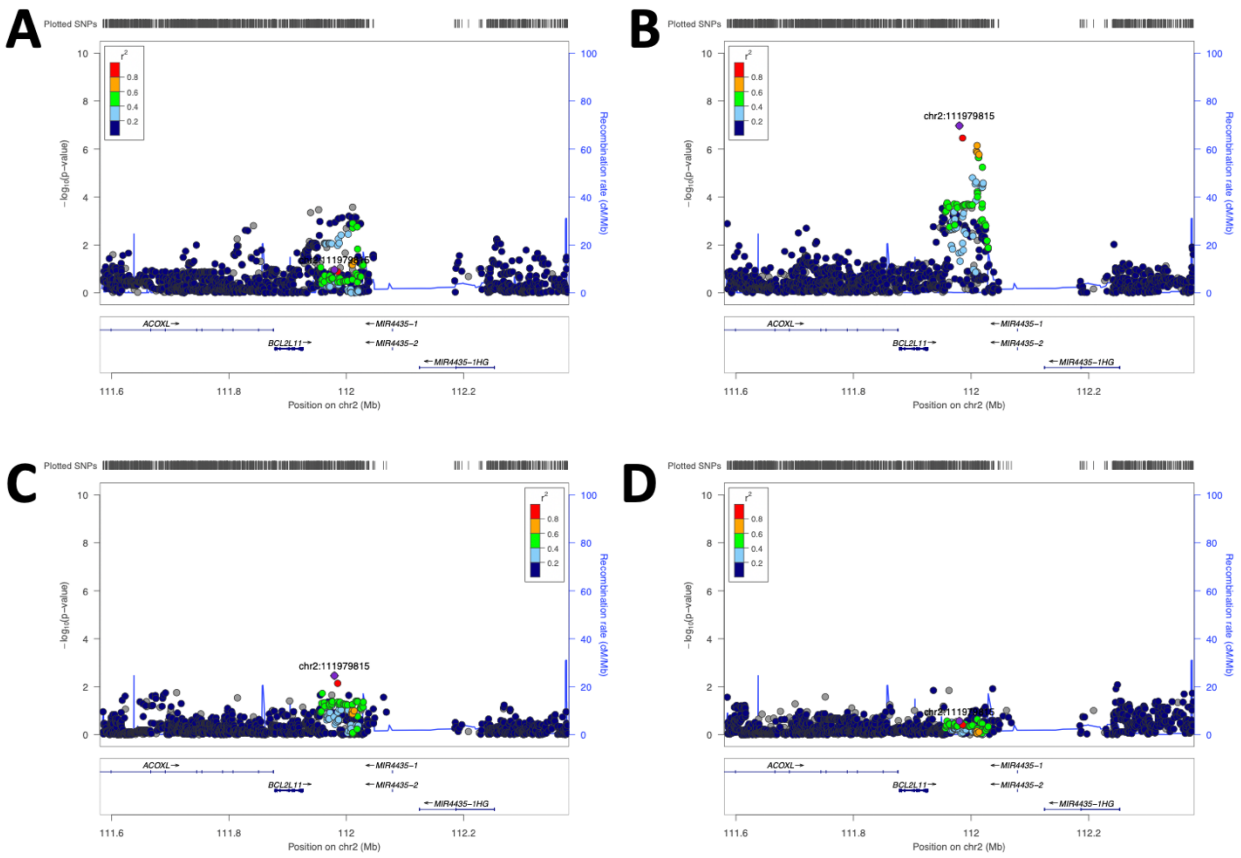
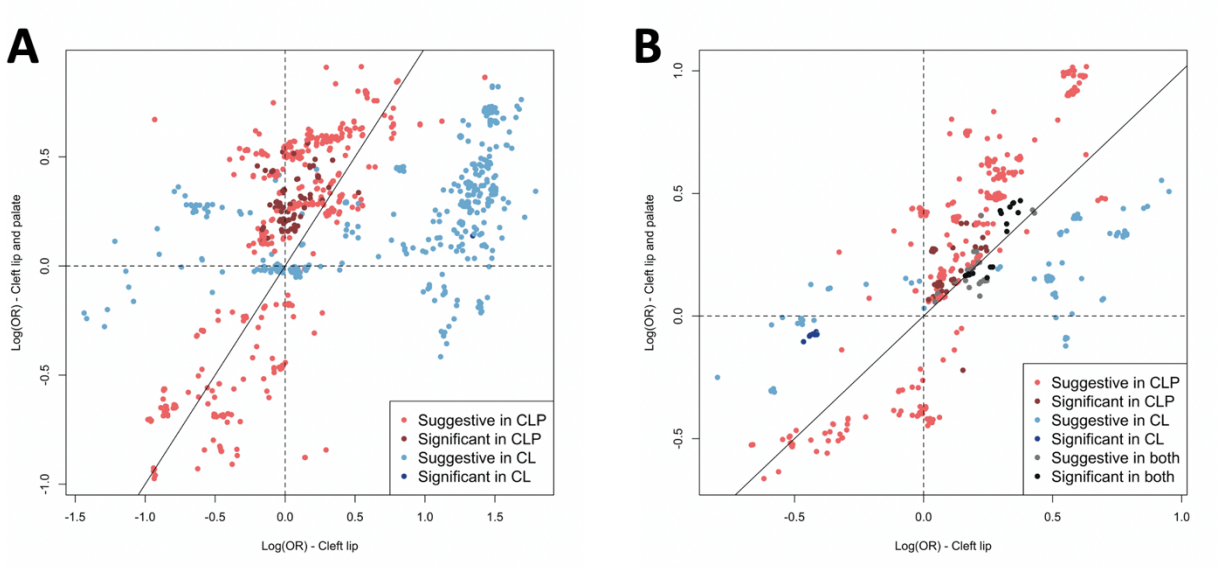


Figure S6: Regional association plots for subtype-specific analyses.



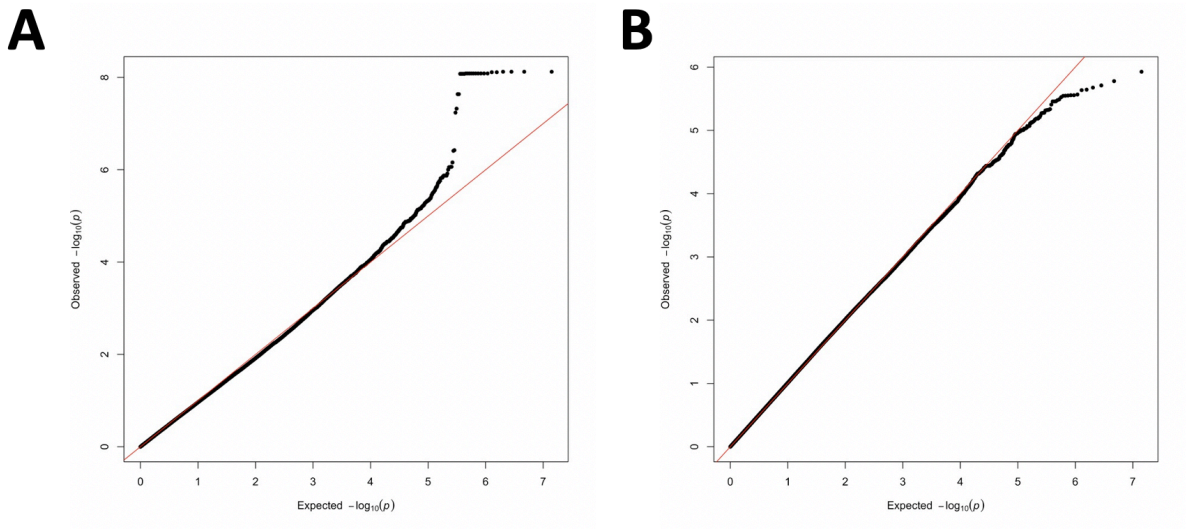
Regional association plots showing $-\log_{10}$(p-values) for the novel genome-wide suggestive peaks at 2q13 in the modifier analysis in bilateral CL (A), unilateral CL (B), bilateral CLP (C), and unilateral CLP (D). Plots were generated using LocusZoom. The recombination overlay (blue line, right y-axis) indicates the boundaries of the LD block. Points are color coded according to pairwise LD (r^2) with the index SNP.

Figure S7: Comparison of subtypes across cleft type



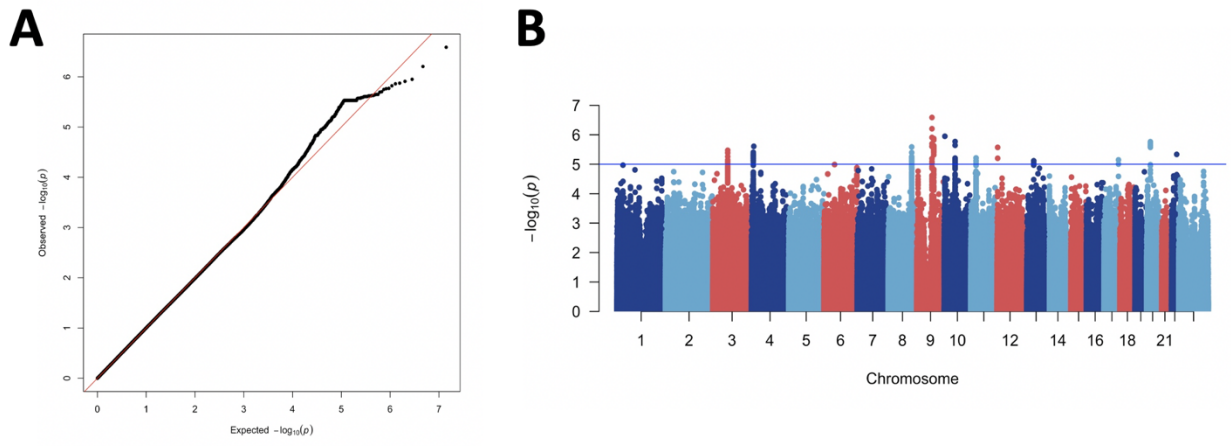
The log odds ratio for SNPs that were suggestive ($p < 1 \times 10^{-5}$) or significant ($p < 5 \times 10^{-8}$) in the subtype-specific case-control analyses in were compared between BCL and BCLP (A), UCL and UCLP (B)

Figure S8: Q-Q plot of modifier analyses



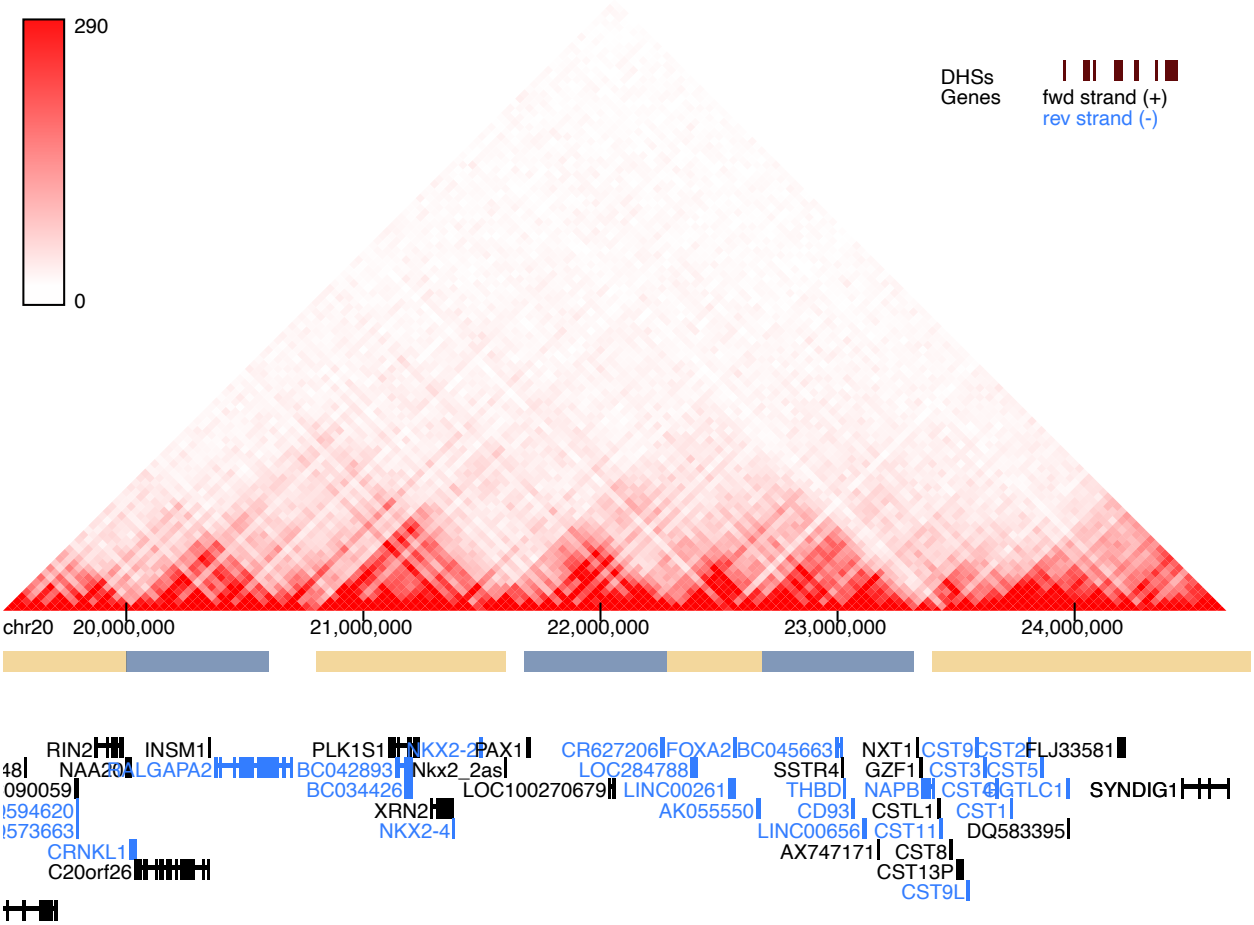
Q-Q plots of the −log10(p-values) of SNPs in the CL (A) and CLP (B) modifier analysis. For the analysis in CL, the genomic inflation factor was 0.96; in CLP, the genomic inflation factor was 1.01.

Figure S9: Analysis of CL/P



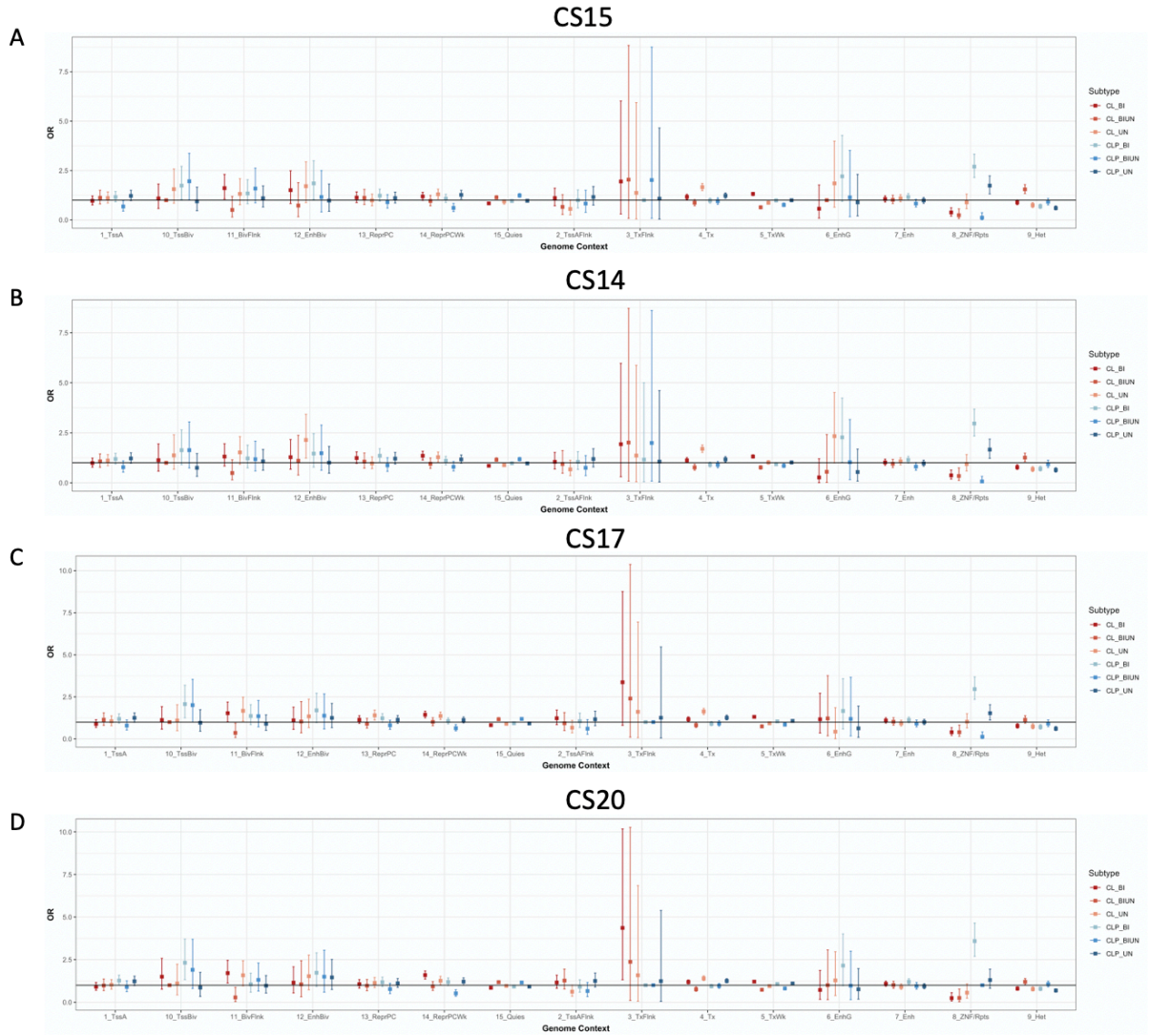
Genome-wide results for the bilateral vs unilateral modifier analysis with CL and CLP combined. Q-Q plots of the −log10(p-values) of SNPs in the CL/P analysis (A). The genomic inflation factor is 1.00. Manhattan plots of $-\log_{10}$(p-values) from the modifier analysis in participants with cleft lip with or without cleft palate (B). Lines indicate suggestive (blue) and genome-wide (red) thresholds for statistical significance.

Figure S10: Topologically-associated domain (TAD) of the 20p11 locus.



Topologically-associated domain (TAD) surrounding the 20p11 locus that was significantly associated with bilateral CL. The associated SNPs were in the same TAD as *Pax1*.

Figure S11: Functional enrichment



Enrichment of the top SNPs associated in either the CL modifier analysis, CLP modifier analysis, and each subtype analysis ($p < 1 \times 10^{-3}$) were tested in each genomic region defined during craniofacial development.