Journal section: Oral Medicine and Pathology
Publication Types: Research

doi:10.4317/medoral.21439
http://dx.doi.org/doi:10.4317/medoral.21439

New insights from GWAS for the cleft palate among han Chinese population

Shi-Jun Duan [1#], Ning Huang [2#], Jia-Yu Shi [4#], Bi-He Zhang [3], Sha He [2], Jian Ma [2], Qiong-Qiong Yu [3], Bing Shi [2], Zhong-Lin Jia [1]

[1] PhD, State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China
[2] MD, PhD, State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China
[3] Msc, State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China
[4] PhD, Division of Growth and Development and Section of Orthodontics, School of Dentistry, University of California, Los Angeles, USA

Correspondence:
State Key Laboratory of Oral Diseases,
West China Hospital of Stomatology,
Sichuan University,
No.14, 3rd Section, Renmin Nan Road,
Chengdu, China, 610041,
zhonglinjia@sina.com

Please cite this article in press as: Duan SJ, Huang N, Shi JY, Zhang BH, He S, Ma J, Yu QQ, Shi B, Jia ZL. New insights from GWAS for the cleft palate among han Chinese population. Med Oral Patol Oral Cir Bucal. (2017), doi:10.4317/medoral.21439

Received: 27/05/2016
Accepted: 02/12/2016

Abstract

Background: Genome wide association studies (GWAS) already have identified tens of susceptible loci for non-syndromic cleft lip with or without cleft palate (NSCL/P). However, whether these loci associated with nonsyndromic cleft palate only (NSCPO) remains unknown.

Material and Methods: In this study, we replicated 38 SNPs (Single nucleotide polymorphisms) which has the most significant *p* values in published GWASs, genotyping by using SNPscan among 144 NSCPO trios from Western Han Chinese. We performed the transmission disequilibrium test (TDT) on individual SNPs and gene-gene (GxG) interaction analyses on the family data; Parent-of-Origin effects were assessed by separately considering transmissions from heterozygous fathers versus heterozygous mothers to affected offspring.

Results: Allelic TDT results showed that T allele at rs742071 (PAX7) (p=0.025, $OR_{transmission}$=3.00, 95%CI: 1.09-8.25) and G allele at rs2485893 (10kb 3' of SYT14) were associated with NSCPO (p=0.0036, $OR_{transmission}$= 0.60, 95%CI: 0.42-0.85). Genotypic TDT based on 3 pseudo controls further confirmed that rs742071 (p-value=0.03, $OR_{transmission}$=3.00, 95%CI: 1.09-8.25) and rs2485893 were associated with NSCPO under additive model (p-value= 0.02, $OR_{transmission}$= 0.66, 95%CI: 0.47-0.92). Genotypic TDT for epistatic interactions showed that rs4844913 (37kb 3' of DIEXF) interacted with rs11119388 (SYT14) (p-value=1.80E-08) and rs6072081 (53kb 3' of MAFB) interacted with rs6102085 (33kb 3' of MAFB) (p-value=3.60E-04) for NSCPO, suggesting they may act in the same pathway in the etiology of NSCPO.

Conclusions: In this study, we found that rs742071 and rs2485893 were associated NSCPO from Han Chinese population; also, interactions of rs4844913:rs11119388 and rs6072081:rs6102085 for NSCPO were identified, gene-gene interactions have been proposed as a potential source of the remaining heritability, these findings provided new insights of the previous GWAS.

Key words: *GWAS, NSCPO, TDT, parent-of-origin effects, epistatic interactions.*

Introduction

Cleft palate (CP) is a common birth defect, which has a lower birth prevalence compared to cleft lip with/ without cleft palate (CL/P): 1/2500 live births vs. 1/700; but CP shows less variability in birth prevalence across populations compared to CL/P (1).

In nonsyndromic cleft palate only (NSCPO), affected individuals have no other physical or developmental anomalies. Most studies suggest that about 50% of CP is nonsyndromic (2). Both population studies and family studies suggested that genetic factors played a critical role in the etiology of NSCPO (3,4). Among first degree relatives, the relative risks of recurrence risks were 56 for cleft palate only and 32 for any cleft lip when compared to the general population in Norway (3); several genes have been identified for syndromic forms of CP, few have been identified as influencing risk to NSCPO. The etiology of this complex trait has been widely studied in order to search for the risk factors and to design strategies for prevention.

Genome wide association studies entail the matching of a given human genome sequence with an annotated, high-resolution map of common genetic variation; They are contributing a lot to our understanding of diseases to which there is a genetic predisposition (5). Genome wide association studies (GWAS) already have identified tens of susceptible loci for cleft lip with or without cleft palate (CL/P).

However, whether these loci associated with nonsyndromic cleft palate only for Han Chinese remains unknown. In this study, we replicated 38 SNPs from 19 genes/ regions of 11 chromosomes from previous GWAS (6-10) and other studies (11-13) with prior compelling evidence contributing to NSCL/P to investigate their roles in Han Chinese population.

Material and Methods

- Samples

Our samples consisted of 144 complete case-parent trios with nonsydromic cleft palate only (NSCPO), 59 males, 82 females and 3 unknown gender of the probands. All subjects were self-identified as Western Han Chinese, they were recruited between 2008 and 2013 from the Cleft Surgery Department of West China Hospital of Stomatology, Sichuan University. Written informed consent was obtained from parents on behalf of the children and all affected individuals old enough to give their own consent in this study. The consent procedure and this study approved by the Hospital Ethics Committee (HEC) of West China Hospital of Stomatology, Sichuan University.

- Genotyping

Venous blood samples were drawn from participants and DNA was extracted by phenol chloroform extraction protocol. The SNP genotyping work was performed using a custom-by-design 2x48-Plex SNPscan™ Kit (Cat#:G0104, Genesky Biotechnologies Inc., Shanghai, China).

- Statistical analysis

The unaffected parents were underwent Hardy-Weinberg Equilibrium (HWE) analysis and minor allele frequency (MAF) determination. HWE, MAF, allelic TDT and parent-of-origin effects, were calculated using PLINK (14). Pairwise LD as both their D' and r^2 were computed for all the SNPs using the haploview program (http://www.broad.mit.edu/haploview/haploview). Genotypic TDT and Likelihood ratio test for epistatic interactions based on genotypic TDTs were determined with R Package trio (v1.4.23) (15), all two-way interactions comprised a matrix in genotype format were tested using the function colGxG, without specifying the genes, and the interactions between all 38 (38-1)/2 pairs of the SNPs in a matrix were tested. We used a Bonferonni correction for 38 tests to determine a threshold for formal significance of p=0.0013.

Results

- Transmission Disequilibrium Test

All SNPs were passed the HWE test (p >0.05) (Table 1). Allelic TDT results showed that T allele at rs742071 (PAX7) (p=0.025, $OR_{transmission}$=3.0, 95%CI: 1.09-8.25) and G allele at rs2485893 (10kb 3' of SYT14) were associated with NSCPO (p= 0.0036, $OR_{transmission}$= 0.60, 95%CI: 0.42-0.85) (Table 2). Genotypic TDT based on 3 pseudo controls further confirmed that rs742071 (p-value=0.03, $OR_{transmission}$=3.0, 95%CI: 1.09-8.25) and rs2485893 are associated with NSCPO under additive model (p-value=0.02, $OR_{transmission}$=0.66, 95%CI: 0.47-0.92) (Table 3).

- Parent-of-Origin Effects

There was no significant difference of minor allele transmissions between the maternal and paternal for all SNPs (data not shown). However, we found an excess of maternal transmission of the allele G at rs2485893 (p=0.026), allele A at rs8001641 (p=0.018) and allele G at rs17563 (p=0.045) compared with the paternal (Table 4), which might warrant future investigations.

- Gene by Gene Interactions

The Genotypic TDT for epistatic interactions showed that rs4844913 (43kb 3' of DIEXF) interacted with rs11119388 (SYT14) (p=1.80E-08) and rs6072081 (53kb 3' of MAFB) interacted with rs6102085 (33kb 3' of MAFB) (p= 3.60E-04) for NSCPO (Fig. 1).

- Pair-wise Linkage Disequilibrium and Haplotype Analysis

We calculated the pair-wise linkage disequilibrium (LD) of SNPs on chromosome 1 based the association results above. There were strong LD between six pairs of SNPs (rs4920552-rs766325, rs126280-rs642961, rs2064163-12063989, rs4844913-rs9429830, rs4844913-rs227178

Table 1. Minor allele frequency and Hardy-Weinberg Equilibrium Test of the SNPs for NSCPO.

CHR	Gene	SNP	Position(Hg19)	Location	Minor Allele	Major Allele	MAF	HWpval
1	*PAX7*	rs4920522	18940380	Intergenic	T	C	0.22	0.38
1	*PAX7*	rs766325	18956458	Upstream	A	G	0.19	0.85
1	*PAX7*	rs6695765	18979320	Intronic	C	T	0.34	0.43
1	*PAX7*	rs742071	18979874	Intronic	T	G	0.04	0.34
1	*ABCA4*	rs560426	94553438	Intronic	C	T	0.32	0.69
1	*IRF6*	rs2235371	209964080	Exon	T	C	0.45	0.55
1	*IRF6*	rs642961	209989270	Upstream	A	G	0.19	0.34
1	*DIEXF*	rs126280	210019824	Intronic	A	G	0.19	1.00
1	*DIEXF*	rs2064163	210048819	Intergenic	G	T	0.50	0.81
1	*DIEXF*	rs12063989	210049893	Intergenic	C	T	0.48	1.00
1	*DIEXF*	rs4844913	210068117	Intergenic	G	A	0.38	0.90
1	*SYT14*	rs9429830	210110537	Upstream	C	T	0.41	0.44
1	*SYT14*	rs11119388	210174417	Intronic	A	G	0.49	0.48
1	*SYT14*	rs227178	210216946	Intronic	C	T	0.35	0.20
1	*SYT14*	rs2485893	210348155	Intergenic	G	A	0.38	0.45
1	*SLC25A24*	rs6677101	108699730	Intronic	T	G	0.42	0.47
2	*THADA*	rs7590268	43540125	Intronic	G	T	0.04	0.34
3	*EPHA3*	rs7632427	89534377	Downstream	C	T	0.17	0.68
4	*GRID2*	rs12506428	93830884	Intronic	T	C	0.48	0.64
8	*DCAF4L2*	rs12543318	88868340	Intergenic	A	C	0.35	0.90
8	*LOC728724*	rs987525	129946154	Intergenic	A	C	0.09	0.71
8	*EPHX2*	rs6558002	27389542	Intronic	C	T	0.14	0.23
10	*VAX1*	rs7078160	118827560	Intronic	A	G	0.44	0.12
10	*VAX1*	rs4752028	118834991	Intronic	C	T	0.33	0.43
13	*SPRY2*	rs9574565	80668874	Intergenic	T	C	0.12	1.00
13	*SPRY2*	rs8001641	80692811	Intergenic	A	G	0.16	0.83
14	*BMP4*	rs17563	54417522	Exon	G	A	0.29	0.25
15	*FMN1*	rs1258763	33050423	Downstream	T	C	0.08	0.67
15	*TPM1*	rs7179658	63312695	Downstream	C	T	0.17	1.00
17	*NTN1*	rs9788972	8919630	Upstream	A	G	0.20	0.27
17	*NTN1*	rs9915089	8952894	Intronic	T	C	0.16	0.83
17	*NTN1*	rs8081823	8965551	Intronic	A	G	0.41	0.47
20	*MAFB*	rs6072081	39261054	Intergenic	G	A	0.42	0.72
20	*MAFB*	rs6065259	39261979	Intergenic	A	G	0.40	0.14
20	*MAFB*	rs17820943	39268516	Intergenic	T	C	0.43	0.55
20	*MAFB*	rs13041247	39269074	Intergenic	C	T	0.43	0.63
20	*MAFB*	rs11698025	39274083	Intergenic	A	G	0.35	0.60
20	*MAFB*	rs6102085	39281629	Intergenic	A	G	0.45	0.34

Note: CHR, chromosome; MAF, minor allele frequency; HWpval, *p*-values for Hardy-Weinberg Equilibrium test.

and rs9429830-227178) with D'>0.93 and r^2>0.80, which distributed on two haplotypes (Fig. 2). Based on this, we tried to test if these pairs of SNPs segregate together among NSCPO by carrying out the haplotype analysis. The results did not show any significance (data not shown).

Discussion

Rs742071 is located in the intron of PAX7 which is involved in neural crest induction and is expressed in cranial neural crest cells, and mice lacking Pax7 have malformations of the nasal and maxillary structures (16). The PAX7 was a second tier GWAS hit (8), later it was

Table 2. Allelic TDT results of the SNPs for NSCPO.

Gene	SNP	Position(Hg19)	A1	A2	T/U	OR(95%CI)	CHISQ	*P*
PAX7	rs4920522	18940380	T	C	42/45	0.93(0.61-1.42)	0.10	0.75
PAX7	rs766325	18956458	A	G	44/41	1.07(0.70-1.64)	0.11	0.74
PAX7	rs6695765	18979320	C	T	68/59	1.15(0.81-1.63)	0.64	0.42
PAX7	**rs742071**	**18979874**	**T**	**G**	**15/5**	**3.00(1.09-8.25)**	**5.00**	**0.025**
ABCA4	rs560426	94553438	C	T	59/63	0.94(0.66-1.34)	0.13	0.72
IRF6	rs2235371	209964080	T	C	71/64	1.11(0.79-1.56)	0.36	0.55
IRF6	rs642961	209989270	A	G	41/39	1.05(0.68-1.63)	0.05	0.82
DIEXF	rs126280	210019824	A	G	41/40	1.03(0.66-1.59)	0.01	0.91
DIEXF	rs2064163	210048819	G	T	69/74	0.93(0.67-1.29)	0.17	0.68
DIEXF	rs12063989	210049893	C	T	75/65	1.15(0.83-1.61)	0.71	0.40
DIEXF	rs4844913	210068117	G	A	58/74	0.78(0.56-1.11)	1.94	0.16
SYT14	rs9429830	210110537	C	T	47/59	0.80(0.54-1.17)	1.36	0.24
SYT14	rs11119388	210174417	A	G	68/80	0.85(0.62-1.17)	0.97	0.32
SYT14	rs227178	210216946	C	T	59/78	0.76(0.54-1.06)	2.64	0.10
SYT14	**rs2485893**	**210348155**	**G**	**A**	**51/85**	**0.60(0.42-0.85)**	**8.50**	**0.0036**
SLC25A24	rs6677101	108699730	T	G	65/64	1.02(0.72-1.43)	0.01	0.93
THADA	rs7590268	43540125	G	T	6/13	0.46(0.18-1.21)	2.58	0.11
EPHA3	rs7632427	89534377	C	T	33/44	0.75(0.48-1.18)	1.57	0.21
GRID2	rs12506428	93830884	T	C	65/79	0.82(0.59-1.14)	1.36	0.24
DCAF4L2	rs12543318	88868340	A	C	58/71	0.82(0.58-1.16)	1.31	0.25
LOC728724	rs987525	129946154	A	C	23/22	1.05(0.58-1.88)	0.02	0.88
EPHX2	rs6558002	27389542	C	T	44/31	1.42(0.90-2.25)	2.25	0.13
VAX1	rs7078160	118827560	A	G	66/58	1.14(0.80-1.62)	0.52	0.47
VAX1	rs4752028	118834991	C	T	58/62	0.94(0.65-1.34)	0.13	0.72
SPRY2	rs9574565	80668874	T	C	24/36	0.67(0.40-1.12)	2.40	0.12
SPRY2	rs8001641	80692811	A	G	44/30	1.47(0.92-2.33)	2.65	0.10
BMP4	rs17563	54417522	G	A	69/54	1.28(0.90-1.82)	1.83	0.18
FMN1	rs1258763	33050423	T	C	19/21	0.90(0.49-1.68)	0.10	0.75
TPM1	rs7179658	63312695	C	T	34/42	0.81(0.52-1.27)	0.84	0.36
NTN1	rs9788972	8919630	A	G	40/46	0.87(0.57-1.33)	0.42	0.52
NTN1	rs9915089	8952894	T	C	33/41	0.80(0.51-1.23)	0.86	0.35
NTN1	rs8081823	8965551	A	G	60/67	0.90(0.63-1.27)	0.39	0.53
MAFB	rs6072081	39261054	G	A	68/64	1.06(0.76-1.50)	0.12	0.73
MAFB	rs6065259	39261979	A	G	64/55	1.16(0.81-1.67)	0.68	0.41
MAFB	rs17820943	39268516	T	C	71/60	1.18(0.84-1.67)	0.92	0.34
MAFB	rs13041247	39269074	C	T	71/61	1.16(0.83-1.64)	0.76	0.38
MAFB	rs11698025	39274083	A	G	63/55	1.15(0.80-1.65)	0.54	0.46
MAFB	rs6102085	39281629	A	G	68/62	1.10(0.78-1.55)	0.28	0.60

Note: A1, Minor allele; A2, Major allele; T, minor allele transmitted, U, minor allele un-transmitted; OR, odds ratios for the transmissions; 95%CI, 95% confidence interval. CHISQ, chi-square; *P*, *p*-values; Bold characters show the items with *p* values less than 0.01.

Table 3. Genotypic TDT Based on 3 Pseudo Controls for NSCPO under different genetic models

Gene	SNPs	Additive				Dominant				Recessive			
		Coef	OR(95%CI)	*P*	Trios	Coef	OR(95%CI)	*P*	Trios	Coef	OR(95%CI)	*P*	Trios
PAX7	rs4920522	-0.07	0.93(0.62-1.42)	0.75	76	-0.16	0.85(0.54-1.37)	0.51	73	0.37	1.45(0.51-4.07)	0.49	16
PAX7	rs766325	0.11	1.12(0.74-1.70)	0.6	77	0.05	1.06(0.67-1.67)	0.81	76	0.54	1.72(0.55-5.31)	0.35	13
PAX7	rs6695765	0.19	1.20(0.85-1.70)	0.29	92	0.24	1.27(0.80-2.02)	0.3	86	0.17	1.18(0.62-2.26)	0.61	44
PAX7	**rs742071**	1.10	**3.00(1.09-8.25)**	**0.03**	19	1.28	3.6(1.19-10.87)	0.02	19	--	--	--	1
ABCA4	rs560426	-0.10	0.91(0.64-1.29)	0.59	101	-0.14	0.87(0.57-1.35)	0.54	87	-0.03	0.97(0.49-1.93)	0.93	39
IRF6	rs2235371	0.06	1.06(0.76-1.48)	0.73	106	-0.09	0.91(0.57-1.47)	0.71	75	0.25	1.28(0.76-2.15)	0.35	63
IRF6	rs642961	0.07	1.08(0.7-1.65)	0.74	77	-0.03	0.97(0.61-1.54)	0.91	74	0.98	2.66(0.69-10.32)	0.16	9
DIEXF	rs126280	0.05	1.05(0.68-1.61)	0.83	76	0.00	1.00(0.63-1.59)	1	74	0.46	1.59(0.44-5.82)	0.48	10
DIEXF	rs2064163	-0.07	0.93(0.67-1.29)	0.68	107	0.00	1.00(0.61-1.64)	1	73	-0.16	0.85(0.51-1.41)	0.53	72
DIEXF	rs12063989	0.13	1.13(0.82-1.58)	0.45	109	0.15	1.16(0.72-1.89)	0.54	76	0.13	1.14(0.69-1.90)	0.60	67
DIEXF	rs4844913	-0.17	0.84(0.60-1.17)	0.31	102	-0.21	0.81(0.52-1.27)	0.36	84	-0.18	0.83(0.46-1.52)	0.55	54
SYT14	rs9429830	-0.11	0.90(0.62-1.30)	0.57	81	-0.22	0.80(0.48-1.33)	0.39	65	0.03	1.03(0.55-1.91)	0.94	47
SYT14	rs11119388	-0.13	0.88(0.64-1.21)	0.42	112	0.03	1.03(0.64-1.67)	0.9	77	-0.36	0.70(0.42-1.17)	0.17	75
SYT14	rs227178	-0.23	0.80(0.57-1.11)	0.18	106	-0.15	0.86(0.56-1.33)	0.51	90	-0.51	0.60(0.31-1.16)	0.13	52
SYT14	**rs2485893**	-0.42	**0.66(0.47-0.92)**	**0.02**	103	-0.39	0.68(0.44-1.04)	0.07	89	-0.71	0.49(0.24-1.00)	0.05	52
SLC25A24	rs6677101	-0.02	0.98(0.70-1.39)	0.93	99	0.00	1.00(0.62-1.61)	1	76	-0.04	0.96(0.54-1.70)	0.88	55
THADA	rs7590268	-0.77	0.46(0.18-1.21)	0.12	19	-0.77	0.46(0.18-1.21)	0.12	19	--	--	--	0
EPHA3	rs7632427	-0.29	0.75(0.48-1.18)	0.21	65	-0.36	0.70(0.41-1.18)	0.18	58	-0.13	0.88(0.32-2.39)	0.80	19
GRID2	rs12506428	-0.15	0.86(0.62-1.19)	0.36	108	-0.32	0.73(0.46-1.15)	0.17	80	0.02	1.02(0.61-1.71)	0.95	67
DCAF4L2	rs12543318	-0.15	0.86(0.61-1.21)	0.38	103	-0.09	0.92(0.59-1.42)	0.7	87	-0.37	0.69(0.35-1.37)	0.29	45
LOC728724	rs987525	0.00	1.00(0.56-1.78)	1	42	-0.10	0.91(0.49-1.67)	0.75	42	1.10	3.00(0.42-21.3)	0.27	4
EPHX2	rs6558002	0.32	1.38(0.87-2.17)	0.17	64	0.45	1.56(0.92-2.65)	0.1	63	-0.21	0.81(0.22-2.99)	0.75	13
VAX1	rs7078160	0.13	1.14(0.8-1.61)	0.48	98	0.30	1.35(0.81-2.26)	0.25	69	-0.04	0.96(0.55-1.68)	0.89	57
VAX1	rs4752028	-0.11	0.89(0.63-1.27)	0.53	93	-0.07	0.93(0.60-1.43)	0.74	88	-0.31	0.73(0.33-1.64)	0.45	35
SPRY2	rs9574565	-0.43	0.65(0.39-1.08)	0.1	53	-0.51	0.60(0.34-1.05)	0.07	53	0.00	1.00(0.20-4.96)	1.00	8
SPRY2	rs8001641	0.37	1.45(0.92-2.29)	0.11	64	0.25	1.29(0.76-2.20)	0.35	59	0.91	2.48(0.93-6.62)	0.07	17
BMP4	rs17563	0.22	1.25(0.88-1.78)	0.21	94	0.10	1.11(0.71-1.72)	0.65	88	0.60	1.83(0.95-3.53)	0.07	38
FMN1	rs1258763	-0.15	0.86(0.47-1.60)	0.64	36	-0.14	0.87(0.45-1.68)	0.67	36	-0.29	0.75(0.08-6.71)	0.80	5
TPM1	rs7179658	-0.25	0.78(0.50-1.21)	0.26	69	-0.43	0.65(0.40-1.07)	0.09	66	0.57	1.77(0.60-5.24)	0.30	14
NTN1	rs9788972	-0.12	0.89(0.59-1.36)	0.59	69	-0.10	0.91(0.55-1.49)	0.7	66	-0.26	0.77(0.28-2.15)	0.62	21
NTN1	rs9915089	-0.18	0.83(0.53-1.31)	0.43	65	-0.20	0.82(0.49-1.36)	0.44	62	-0.17	0.84(0.26-2.72)	0.77	15
NTN1	rs8081823	-0.12	0.88(0.63-1.25)	0.48	95	-0.17	0.85(0.52-1.37)	0.5	74	-0.11	0.90(0.50-1.60)	0.72	56
MAFB	rs6072081	0.06	1.06(0.76-1.48)	0.73	104	0.39	1.48(0.92-2.39)	0.11	82	-0.47	0.63(0.33-1.17)	0.14	54
MAFB	rs6065259	0.15	1.16(0.81-1.66)	0.41	96	0.44	1.55(0.94-2.55)	0.08	75	-0.29	0.75(0.39-1.43)	0.38	46
MAFB	rs17820943	0.17	1.18(0.84-1.66)	0.34	107	0.23	1.26(0.78-2.03)	0.34	77	0.12	1.13(0.65-1.96)	0.67	56
MAFB	rs13041247	0.15	1.16(0.83-1.63)	0.39	108	0.20	1.22(0.76-1.96)	0.41	78	0.12	1.13(0.65-1.96)	0.67	56
MAFB	rs11698025	0.13	1.14(0.80-1.64)	0.47	95	0.07	1.07(0.68-1.70)	0.77	79	0.31	1.36(0.71-2.58)	0.35	41
MAFB	rs6102085	0.06	1.06(0.76-1.49)	0.73	104	0.07	1.07(0.64-1.77)	0.8	68	0.07	1.07(0.64-1.80)	0.79	64

Note: Coef, Coefficient; OR, odds ratios for the transmissions; 95%CI, 95% confidence interval. *P*, *p*-value; Bold characters show the items with *p* values less than 0.01.

Table 4. Parent-of-origin effect of effect of the SNPs.

Gene	SNP	Minor Allele	Paternal			Maternal			Z	*P*
			T/U	CHISQ	*P*	T/U	CHISQ	*P*		
PAX7	rs4920522	T	26.5/24.5	0.08	0.78	15.5/20.5	0.69	0.40	0.82	0.41
PAX7	rs766325	A	25.5/20.5	0.54	0.46	18.5/20.5	0.10	0.75	0.73	0.46
PAX7	rs6695765	C	31.5/28.5	0.15	0.70	36.5/30.5	0.54	0.46	-0.22	0.82
PAX7	rs742071	T	8.5/3.5	2.08	0.15	6.5/1.5	3.13	0.077	-0.52	0.60
ABCA4	rs560426	C	31.5/25.5	0.63	0.43	27.5/37.5	1.54	0.21	1.43	0.15
IRF6	rs2235371	T	30/35	0.38	0.54	41/29	2.06	0.15	-1.44	0.15
IRF6	rs642961	A	22.5/22.5	0.00	1.00	18.5/16.5	0.11	0.74	-0.25	0.80
DIEXF	rs126280	A	22/23	0.02	0.88	19/17	0.11	0.74	-0.35	0.73
DIEXF	rs2064163	G	34/33	0.01	0.9	35/41	0.47	0.49	0.56	0.58
DIEXF	rs12063989	C	32/33	0.02	0.9	43/32	1.61	0.20	-0.96	0.34
DIEXF	rs4844913	G	32/37	0.36	0.55	26/37	1.92	0.17	0.59	0.56
SYT14	rs9429830	C	23.5/30.5	0.91	0.34	23.5/28.5	0.48	0.49	-0.17	0.86
SYT14	rs11119388	A	35.5/35.5	0.00	1.00	32.5/44.5	1.87	0.17	0.95	0.34
SYT14	rs227178	C	31.5/37.5	0.52	0.47	27.5/40.5	2.49	0.11	0.62	0.54
SYT14	rs2485893	G	27.5/43.5	3.61	**0.058**	23.5/41.5	4.99	**0.026**	0.31	0.76
SLC25A24	rs6677101	T	35.5/30.5	0.38	0.54	29.5/33.5	0.25	0.61	0.79	0.43
THADA	rs7590268	G	3/5	0.50	0.48	3/8	2.27	0.13	0.47	0.64
EPHA3	rs7632427	C	16.5/15.5	0.03	0.86	16.5/28.5	3.20	0.074	1.30	0.19
GRID2	rs12506428	T	28/43	3.17	**0.075**	37/36	0.01	0.91	-1.35	0.18
DCAF4L2	rs12543318	A	36/39	0.12	0.73	22/32	1.85	0.17	0.82	0.41
LOC728724	rs987525	A	9/12	0.43	0.51	14/10	0.67	0.41	-1.03	0.30
EPHX2	rs6558002	C	18.5/15.5	0.26	0.61	25.5/15.5	2.44	0.12	-0.68	0.50
VAX1	rs7078160	A	34.5/23.5	2.09	0.15	31.5/34.5	0.14	0.71	1.31	0.19
VAX1	rs4752028	C	31/23	1.19	0.28	27/39	2.18	0.14	1.79	**0.073**
SPRY2	rs9574565	T	13/18	0.81	0.37	11/18	1.69	0.19	0.32	0.75
SPRY2	rs8001641	A	22.5/21.5	0.02	0.88	21.5/8.5	5.63	**0.018**	-1.75	**0.081**
BMP4	rs17563	G	33.5/33.5	0.00	1.00	35.5/20.5	4.02	0.045	-1.49	0.14

Table 4 continue. Parent-of-origin effect of effect of the SNPs.

FMN1	rs1258763	T	11.5/12.5	0.04	0.84	7.5/8.5	0.06	0.8	0.06	0.95
TPM1	rs7179658	C	21/21	0.00	1.00	13/21	1.88	0.17	1.02	0.31
NTN1	rs9788972	A	17.5/26.5	1.84	0.17	22.5/19.5	0.21	0.64	-1.28	0.20
NTN1	rs9915089	T	17/25	1.52	0.22	16/16	0.00	1.00	-0.82	0.42
NTN1	rs8081823	A	32/32	0.00	1.00	28/35	0.78	0.38	0.63	0.53
MAFB	rs6072081	G	30/31	0.02	0.9	38/33	0.35	0.55	-0.50	0.62
MAFB	rs6065259	A	28.5/25.5	0.17	0.68	35.5/29.5	0.55	0.46	-0.20	0.84
MAFB	rs17820943	T	32/27	0.42	0.52	39/33	0.50	0.48	0.01	0.99
MAFB	rs13041247	C	32/28	0.27	0.61	39/33	0.50	0.48	-0.10	0.92
MAFB	rs11698025	A	26/25	0.02	0.89	37/30	0.73	0.39	-0.46	0.65
MAFB	rs6102085	A	34.5/29.5	0.39	0.53	33.5/32.5	0.02	0.90	0.36	0.72

Note: T: minor allele transmitted, U, minor allele untransmitted; CHISQ, chi-square; *P*, *p*-values; Z: vector of the large sample Z statistic; Bold characters show the items with *p* values less than 0.01.

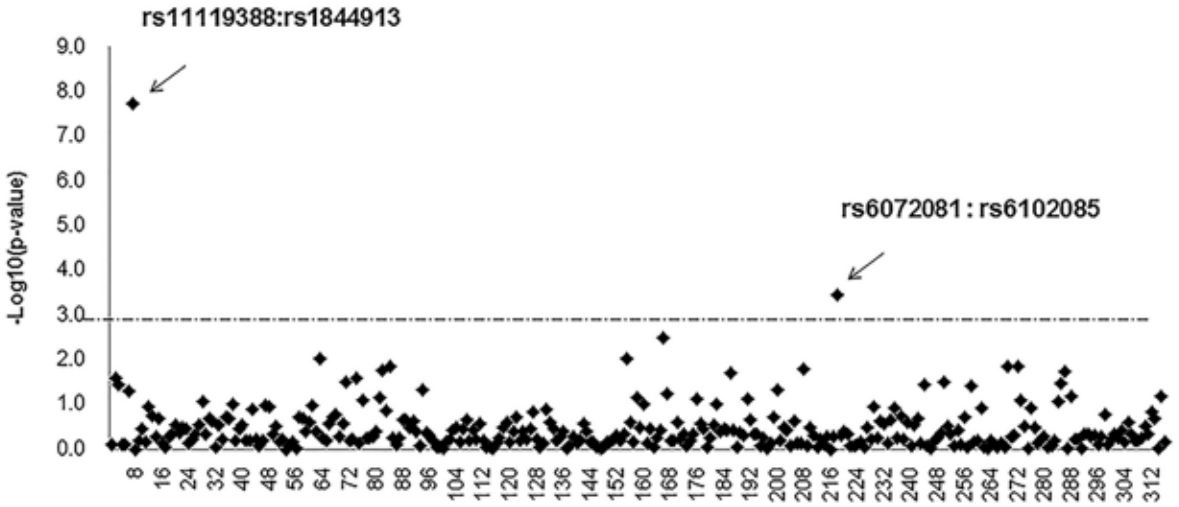

Fig. 1. Genotypic TDT for Epistatic Interactions of Pair-wise SNPs.

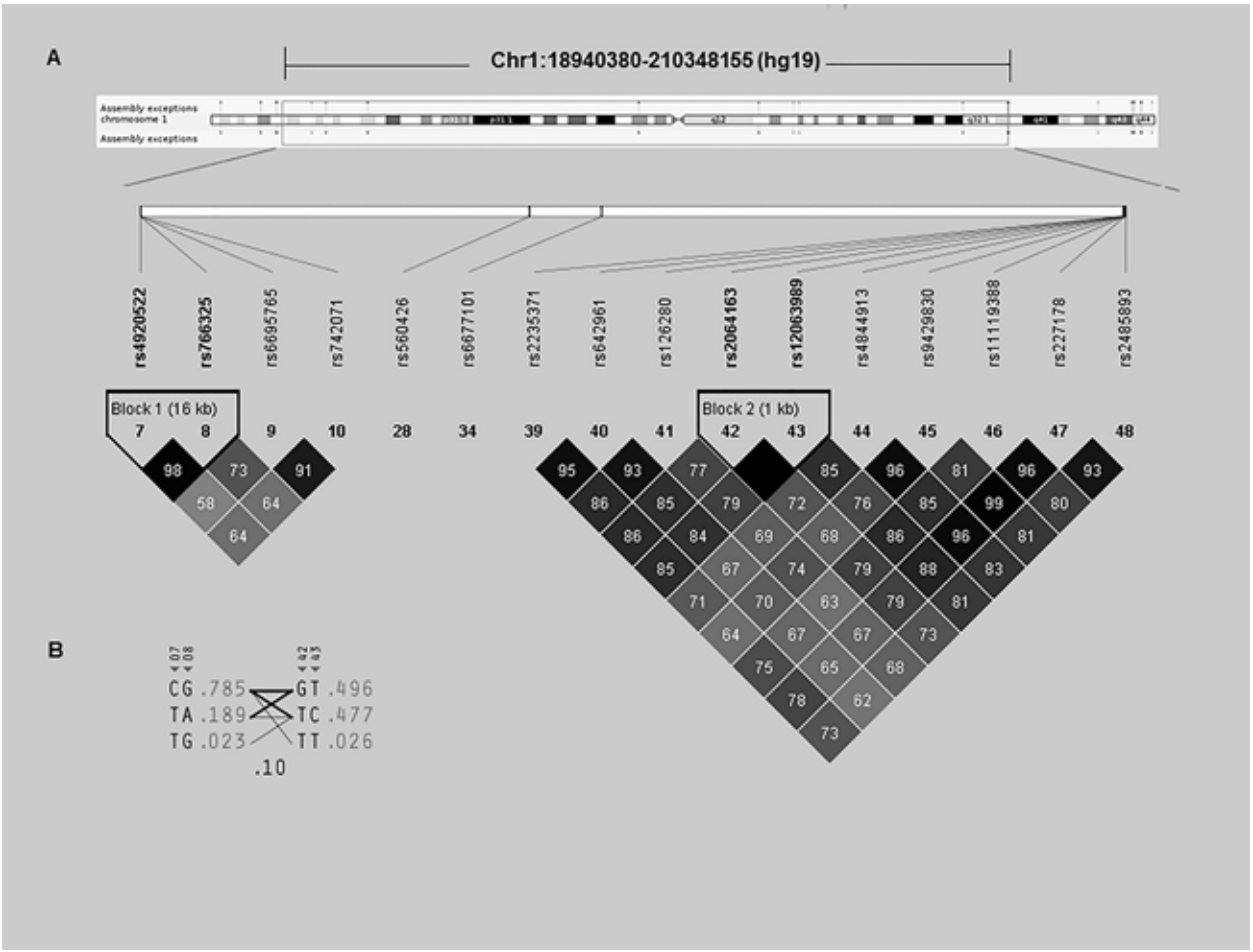

Fig. 2. Pair-wise Linkage Disequilibrium (A) and Haplotype (B) of SNPs at Chromosome 1.

confirmed by replication among European and Southeast Asian (17) and GWAS meta-analysis (10). Recently, Leslie *et al.* 2015 performed targeted sequencing of 13 regions from GWASs and other studies in 1,409 Asian and European trios, and carried out a series of statistical and functional analyses, the results indicated that de novo mutation p.Ala259Val disrupted PAX7 function and might contribute to CLP pathogenesis in this individual (18). Although intronic SNPs do not typically alter protein structure, associations with intronic variants have been reported for a number of complex diseases. In this study, we found rs742071 (PAX7) was associ-

ated with NSCPO and had a larger genetic effects compared with the associations with NSCL/P from previous GWASs (8,10); Motif analyses by HaploReg indicated that the T allele of rs742071 could greatly alter the affinity of Sin3Ak-20_disc4 (score: 2.1-14.1).

Rs742071 had the minor allele frequency as 0.04 among Han Chinese population, indicating it may need larger sample size to validate its significance; and with the limited sample size, rs742071 did not pass the threshold of the Bonferonni correction p value in the current study. We will add more samples to study it and other variants at PAX7 gene among NSCPO among Han Chinese population.

Rs2485893 was associated with CL/P among Asian ancestry with p value 7.86E-07 by GWAS (8). In this study, we found rs2485893 (10kb 3' of SYT14) was also found to be associated with NSCPO. Marked parent of origin effects were seen for rs2485893 alleles, over-transmission was seen preferentially from mothers compared with fathers (Table 4). Motif analyses by HaploReg indicated that the G allele of rs2485893 could greatly alter the affinity of AFP1 (score: 2.7-12.3).

Numerous studies have shown that highly conserved non-coding elements act as developmental enhancers in vivo (19-21). Non-coding conserved elements around rs742071 and rs2485893 therefore might represent putative regulatory elements for PAX7 and SYT14, we will perform functional studies to elucidate their roles in human craniofacial development.

Gene-gene interactions have been proposed as a potential source of the remaining heritability. Genotypic TDT for epistatic interactions showed that rs4844913 interacts with rs11119388 (SYT14) and rs6072081 interacts with rs6102085 for NSCPO, which provided new insights for the previous GWASs.

In summary, we replicated 38 SNPs contributing to NSCL/P to investigate their roles in NSCPO among Han Chinese population. In this study, we found that rs742071 and rs2485893 were associated NSCPO from Han Chinese population; also, interactions of rs4844913:rs11119388 and rs6072081:rs6102085 for NSCPO were identified, which may provide new insights for the previous GWASs.

References

1. Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, et al. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. Genet Epidemiol. 2011. 35:469-478.

2. Shi M, Wehby GL, Murray JC. Review on genetic variants and maternal smoking in the etiology of oral clefts and other birth defects. Birth Defects Res C Embryo Today. 2008. 84:16-29.

3. Sivertsen A, Wilcox AJ, Skjaerven R, Vindenes HA, Abyholm F, Harvilles E, et al. Familial risk of oral clefts by morphological type and severity: population based cohort study of first degree relatives. BMJ. 2008; 336: 432-4.

4. Grosen D, Chevrier C, Skytthe A, Bille C, Molsted K, Sivertsen A, et al. A cohort study of recurrence patterns among more than 54,000 relatives of oral cleft cases in Denmark: Support for the multifactorial threshold model of inheritance. J Med Genet. 2009.47:162–168.

5. Christensen K, Murray JC. What genome-wide association studies can do for medicine. N Engl J Med. 2007. 356:1094-1097.

6 Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, et al. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009. 41:473-477.

7. Grant SF, Wang K, Zhang H, Glaberson W, Annaiah K, Kim CE, et al. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr. 2009. 155:909-913.

8. Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010. 42:525-529.

9. Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, et al. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet . 2010. 42:24-26.

10. Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, 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.

11. Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, et al. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet. 2002. 32:285-9.

12, Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, et al. Disruption of an AP-2alpha binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet. 2008. 40:1341-1347.

13, Suzuki S, Marazita ML, Cooper ME, Miwa N, Hing A, Jugessur A, et al. Mutations in BMP4 are associated with subepithelial, microform, and overt cleft lip. Am J Hum Genet. 2009. 84:406-411.

14, Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007.81:559-575.

15, Schwender H, Taub MA, Beaty TH, Marazita ML, Ruczinski I. Rapid testing of SNPs and gene-environment interactions in case-parent trio data based on exact analytic parameter estimation. Biometrics. 2012. 68:766-773.

16, Mansouri A, Stoykova A, Torres M, Gruss P. Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. Development. 1996. 122:831–838.

17, Beaty TH, Taub MA, Scott AF, Murray JC, Marazita ML, Schwender H, et al. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio study. Hum Genet. 2013. 132:771-781.

18, Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, et al. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. Am J Hum Genet. 2015. 96:397-411.

19. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, et al. In vivo enhancer analysis of human conserved non-coding sequences. Nature. 2006. 444:499-502.

20. Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic acids research. 2007. 35(Database issue):D88-92.

21. Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. Nature. 2009. 461:199-205.#, these two authors contributed equally to this work.

Acknowledgements
We are particularly thankful to the patients and their families who participated in this study. This project was supported by the National Natural Science Foundation of China (No. 81600849 and 30530730) and Scientific Research Funds for Young Teachers of Sichuan University (No. 2015SCU11999).

#, these two authors contributed equally to this work.

Conflict of Interest
The authors have declared that no conflict of interest exist.