

SCIENTIFIC REPORTS



OPEN

Two-stage comprehensive evaluation of genetic susceptibility of common variants in *FBXO38*, *AP3B2* and *WHAMM* to severe chronic periodontitis

Received: 30 May 2015
Accepted: 09 November 2015
Published: 08 December 2015

Dong Shang^{1,2,*}, Li Dong^{3,*}, Lingfang Zeng^{1,4}, Rui Yang⁵, Jing Xu⁶, Yue Wu⁷, Ran Xu¹, Hong Tao¹ & Nan Zhang¹

Chronic periodontitis is an oral disorder characterized with gingival inflammation and bone destruction. As the sixth-most prevalent condition affecting more than 743 million people around the world, it is classified as one of the seven destructive oral disorders. Early genetic epidemiological evidence indicated a major role for genetics in periodontal disease development. In this study, we conducted a two-stage comprehensive evaluation of the genetic susceptibility of *FBXO38*, *AP3B2* and *WHAMM* with the diagnosis of severe chronic periodontitis. A total of 5,065 study subjects from the Han Chinese population consisting of 1,264 cases and 3,801 healthy controls were recruited, and 65 single nucleotide markers related to the three candidate genes were genotyped to investigate the susceptibility of patients with these polymorphisms to severe chronic periodontitis. To increase the coverage of genetic markers, we implemented imputation techniques to extend the number of tested markers to 416. Single marker and haplotype-based analyses were performed, and significant results were obtained for *FBXO38* (rs10043775, $P = 0.0009$) and *AP3B2* (rs11631963-rs11637433, CA, $P = 9.98 \times 10^{-5}$; rs1864699-rs2099259-rs2278355, ATC, $P = 3.84 \times 10^{-8}$). Our findings provide direct evidence for the association of *FBXO38* and *AP3B2* with severe chronic periodontitis in the Han Chinese population.

Chronic periodontitis (CP) is a multi-factor complex oral disorder characterized by slowly progressing alveolar bone destruction and attachment loss. A recent systematic review showed that severe CP, as the sixth-most prevalent condition, affected more than 743 million people (10.8%) around the world and that its incidence rate was 701 cases per 100,000 person years¹. Although CP is usually not painful, it is one of the important risk factors for early tooth loss and imposes a significant health burden to the patients². In addition, numerous studies have shown significant links between periodontal disease and systemic conditions including cardiovascular disease³⁻⁵, type 2 diabetes mellitus⁶⁻⁸, adverse pregnancy outcomes⁹, and osteoporosis^{10,11}. Therefore, understanding the aetiology and pathology of CP can help to reveal the pathological mechanisms of other major complex disorders and in turn shine light on the development of novel drugs and treatment strategies for these disorders.

Early genetic epidemiological evidence indicated that genetics played a major role in the onset and development of periodontal disease. A twin study of periodontal disease conducted by Corey *et al.* identified significant differences in proband-wise concordance rates between monozygotic (0.38) and dizygotic twins (0.16)¹². Another twin study has estimated the heritability of adult periodontitis to be approximately 50%¹³. In the past decade, with the

¹Department of Stomatology, the First Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China.

²Department of Respiratory, the First Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China.

³Department of Computer Science, Shandong Medical College, Jinan, China. ⁴Department of Pediatric Stomatology, Jinan Stomatological Hospital, Jinan, China. ⁵Tian-rui Institute of Stomatology, Xi'an, China. ⁶Department of Emergency, the First Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China. ⁷Department of Cardiology, the First Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.T. (email: taohong0228@126.com) or N.Z. (email: nanzhangnz@126.com)

development of high throughput genotyping technology, genome-wide association study (GWAS) has been shown to be a powerful tool in identifying susceptibility genes for complex disorder¹⁴. GWAS of chronic periodontitis have identified several susceptibility genes including *C5AR1*¹⁵, *DLG2*¹⁵, *FAM180A*¹⁶, *MFS16*¹⁶, *NCR2*¹⁷, *NPY*¹⁷, *EMR1*¹⁷, *VAV1*¹⁷, *GLT6D1*¹⁸, *WHAMM*¹⁹ and *AP3B2*¹⁹. However, most of these findings were obtained from samples of European ancestry. Thus, validation of these results in other populations is still necessary.

In our study, we aimed to thoroughly evaluate the genetic risk of three candidate gene polymorphisms (i.e., *FBXO38*, *AP3B2* and *WHAMM*) on severe CP in the Han Chinese population. All three genes have been shown to be significant (or suggestively significant) susceptibility genes to CP or CP-related symptoms in recent published GWAS papers^{19,20} using samples with European ancestry. To investigate the potential susceptibility of these candidate genes, we enrolled 5,065 study subjects of Han Chinese ancestry and genotyped 65 (this number increased by 416 after imputation) pre-selected single nucleotide polymorphism (SNP) markers related to these three genes. A two-stage strategy (discovery and validation stage) was applied in our study, and the potential association signals were tested at both the single SNP and haplotype levels. Additionally, to increase the genomic coverage of SNPs, we also imputed SNPs within two 5-Mb genomic regions that covered the three candidate genes on chromosome 5 and chromosome 15.

Materials and Methods

Subjects The subjects involved in this study included 5,065 individuals ranging from 28–75 years of age. Among them, there were 1,264 cases and 3,801 healthy control subjects. All of the study subjects were enrolled from October 2010 to August 2014. Severe chronic periodontitis patients were recruited from the inpatient and outpatient clinical services at the First Affiliated Hospital of Xi'an Jiaotong University and the Stomatology Hospital of the Fourth Military Medical University. Severe chronic periodontitis was defined as a patient who presented at least two interproximal sites with clinical attachment loss (CA loss) \geq 6 mm, not in the same tooth, and probing depth (PD) \geq 5 mm in one or more interproximal sites, excluding third molars. Three examiners in our study were trained and calibrated by the author N. Z. in accordance with aforementioned diagnosis criteria. All four examiners are dentists. They performed clinical examinations by using a manual probe (UNC-15, Hu-Friedy Manufacturing Company, Inc., Chicago, IL, USA), and all were calibrated on 40 chronic periodontitis patients probed twice before conducting periodontal examinations in our study. In order to facilitate the interpretation of measurement reliability, we reported kappa statistics and inter/intra-class correlation coefficients for PD. The Kappa index and intra-class correlation coefficient for PD ranged from 0.77 to 0.92 and 0.73 to 0.87 for intra-examiner calibration. Inter-examiner Kappa index and intra-class correlation coefficient of agreement in measurements varied from 0.81 to 0.91 and 0.77 to 0.86 for PD (Supplemental Table S1). These results evidenced a high degree of consistency in the measurements and ensured the validity of subsequent periodontal examinations.

The healthy control subjects were enrolled when there were no signs of any periodontal disease at the time of sample collection and no history of periodontal disease. All individuals enrolled in this study were non-smokers. All of our subjects were unrelated Han Chinese residing in Shaanxi Province from similar socio-economic levels. Participants who reported the following characteristics were excluded from the study: use of orthodontic appliances, chronic anti-inflammatory drugs or immunosuppressive chemotherapy, antibiotics within the previous 3 months, chronic inflammatory diseases, a history of diabetes mellitus, hepatitis, HIV infection, nephritis, bleeding or autoimmune disorders, diseases with severe commitment of immune function, current pregnancy or breastfeeding, or received periodontal therapy in the preceding 6 months. The peripheral blood sample was drawn from a vein into a sterile tube containing ethylenediaminetetraacetic acid (EDTA). The plasma samples were stored at -80°C .

The research protocol was approved by the Ethical Committee of Xi'an Jiaotong University. We obtained written informed consent from all participants after the objectives and procedures of this study were fully explained. All of the experiments on humans were performed in accordance with the ethical guidelines of the Declaration of Helsinki (version 2002).

SNP selection and genotyping. In this project, we implemented a two-stage experimental cost saving strategy including 1) a discovery stage in which we genotyped a relatively larger set of markers in a smaller subset of subjects (1,783 subjects consisting of 471 cases and 1,312 controls) and 2) a validation stage in which we only genotyped a smaller set of markers that passed the screening P value threshold (0.05) in a relatively larger set of subjects (3,282 subjects with 793 cases and 2,489 controls). Three candidate genes from two genomic regions including *FBXO38*, *AP3B2* and *WHAMM* were selected for this study. All of the SNP markers located within ± 10 kb of the gene regions and with minor allele frequency (MAF) greater or equal to 0.05 in the Hapmap Han Chinese data²¹ were selected. In addition, another set of SNPs related to these genes that were reported to be significant in recent GWAS papers was also selected. The total number of SNPs genotyped in the discovery stage was 65.

Genomic DNA was isolated from the peripheral blood leukocytes according to the manufacturer's protocol (Genomic DNA kit, Axygen Scientific Inc., CA, USA). The samples of patients and controls were mixed on the same plates, and a double-blind procedure was performed. The DNA was stored at -80°C for genotyping. The genotyping was conducted for all of the SNPs using the Sequenom MassARRAY matrix-assisted laser desorption ionization-time of the flight mass spectrometry platform (Sequenom, San Diego, CA, USA) on the genomic DNA isolated from the peripheral leukocytes. Final data were released using Typer Analyzer software (Sequenom).

Statistical analysis. We conducted a comprehensive power analysis with Genetic Power Calculator²². In the power analysis, we calculated the statistical power by setting different relative risk (1.1–1.5) and MAF (0.05, 0.1, 0.2, 0.3, 0.4) of the underlying risk allele. We set the SNP marker allele to be 0.287 (average MAF of the 65 SNPs genotyped in the screening stage). Results of the statistical power analysis were summarized in Supplemental Table S2. As shown in the Supplemental Table S2, around 80% statistical power could be achieved if the underlying risk allele has a relative risk of 1.2–1.3. We utilized Polyphen2²³ to predict the potential functional effects of significant

	Stage 1 (discovery)	Stage 2 (validation)	Total
Gender			
Male (%)	896 (50)	1711(45)	2607 (51)
Female (%)	887 (50)	1571(55)	2458 (49)
Status			
Case (%)	471 (26)	793 (21)	1264 (25)
Control (%)	1312 (74)	2489 (79)	3801 (75)
Age			
Mean (SD.)	48.83 (13.12)	48.76 (13.05)	48.78 (13.08)
Range (years)	28–74	28–75	28–75

Table 1. Characteristic information of study subjects. SD means standard deviation.

non-synonymous SNPs. The statistical computing software R²⁴ was utilized for analysis of general purpose. We implemented a logistic model using genetic analysis software Plink²⁵ to investigate the association between genetic polymorphisms and disorder status in both the discovery and validation stage. Bonferroni correction was applied in the validation stage to address the multiple comparison problem. Gender and age were included in the logistic model as a covariate. Plink was also utilized to test for Hardy-Weinberg equilibrium (HWE) in control samples and to calculate the minor allele frequency (MAF) for each marker in both stages. In addition, the haplotype frequency estimation and haplotype-based association analysis were also conducted by Plink. We constructed the LD plot using Haploview²⁶ based on the data in the discovery stage. Imputation based on our discovery stage data was implemented using IMPUTE2 software²⁷ with HapMap phase III CHB+JPT+CHD data as a reference panel. Association tests based on the imputed dosage data were conducted with the software SNPTEST v2²⁸. We utilized the parameter of “average certainty” calculated in IMPUTE2 as the main indicator of imputation quality. The threshold of this indicator was chosen by exploring the patterns of Q-Q plots based on the P-values of results of association analysis based on multiple marker sets obtained using different certainty thresholds.

Results

The characteristics of these study subjects were summarized in Table 1. Seven SNPs passed the discovery stage screening threshold ($P < 0.05$) and were sent for genotyping in the validation stage. The results of single SNP-based association analysis of these 7 SNPs in both stages were summarized in Table 2. The full association results of discovery stage could be found in Supplemental Table S3. As presented in Table 2, only 1 SNP, rs10043775 ($P = 0.0009$), passed the Bonferroni corrected P-value threshold ($0.05/7 \approx 0.007$) in the validation stage. Ten linkage disequilibrium (LD) blocks were identified on the two candidate genomic regions. Haplotype association analysis was performed based on the data from the discovery stage. The LD blocks and significant findings of haplotype analysis based on the discovery stage genotype data were shown in Fig. 1. As suggested in Fig. 1, three significant haplotypes were successfully identified. One significant haplotype covered part of the *FBX32* gene region (rs10072051-rs10044061-rs10041283, GGC, $P = 2.78 \times 10^{-6}$), and the other two (rs11631963-rs11637433, CA, $P = 9.98 \times 10^{-5}$; rs1864699-rs2099259-rs2278355, ATC, $P = 3.84 \times 10^{-8}$) covered part of the *AP3B2* gene region. The full results of haplotype analysis were summarized in Supplemental Table S4.

The imputation was implemented for two 5-Mb genomic regions including a 1.46×10^8 – 1.51×10^8 region on chromosome 5 (including *FBXO38*) and a 8.0×10^7 – 8.5×10^7 region on chromosome 15 (including *AP3B2* and *WHAMM*). We chose 0.8 as the average certainty threshold to exclude those potential imputed SNPs with low accuracy. This threshold was chosen by exploring the patterns of Q-Q plots based on the P-values of association analysis based on multiple marker sets obtained by using different certainty thresholds (Supplemental Figure S1). In addition, we only focused on the common SNPs, so we also applied $MAF \geq 0.01$ as another filter criterion. After applying these filters, 416 SNPs were successfully imputed and tested for association based on the data of 65 genotyped SNPs in the discovery stage. We summarized the imputed SNPs with nominal significance ($P < 0.05$) from association analysis in Supplemental Table S5, and the visualization of the results of association tests based on the two imputed genomic regions were shown in Fig. 2 (using the most significant genotyped SNP as a reference SNP). The most significant SNPs within these regions were two imputed SNPs: rs4362936 (chr5, $P = 0.0248$) and rs11636500 (chr15, $P = 1.95 \times 10^{-9}$). In addition, we also explored the regional association pattern of these two imputed regions using the most significant imputed SNPs as reference SNPs, and the results were shown in Supplemental Figs S2 and S3.

Discussion

In a large-scale candidate gene-based genetic association study of severe CP from a Han Chinese population, we obtained several significant association signals. The SNP rs10043775, which is in the exonic region of *FBXO38*, passed the discovery stage screening ($P = 0.0121$) and was further confirmed in the validation stage ($P = 0.0009$). The C allele carriers of this SNP increased the odds of suffering from severe CP by 24.1% compared to the reference allele carriers. Rs10043775 is a missense variant that causes a Serine to Proline change; however, according to the analysis of Polyphen2, this change may not cause any damage to the encoded protein. In a recently published GWAS study based on samples of European ancestry, rs10043775 was found to be suggestively associated with periodontal pathogen colonization²⁰. The phenotype utilized in this study was pathogen colonization but not CP diagnosis. In this sense, our finding successfully established a direct link between rs10043775 (and the

SNP	CHR	POS	Gene	Stage 1 (discovery)						Stage 2 (validation)					
				HWE	MAF_A	MAF_U	A1	P	OR	HWE	MAF_A	MAF_U	A1	P	OR
rs10043775	5	148425557	FBXO38	0.8736	0.2866	0.2435	G	0.0121	1.24	0.9107	0.2844	0.2431	G	0.0009	1.24
rs10477376	5	148405651	FBXO38	0.5725	0.2484	0.2130	G	0.0248	1.23	0.4591	0.2339	0.2053	G	0.0164	1.18
rs17108251	5	148415019	FBXO38	0.6213	0.2378	0.2024	G	0.0266	1.22	0.4610	0.2377	0.2188	G	0.1138	1.11
rs10044061	5	148396934	FBXO38	0.6698	0.2548	0.2195	T	0.0270	1.22	0.8603	0.2610	0.2280	T	0.0073	1.19
rs10068216	5	148393755	FBXO38	0.7643	0.2569	0.2233	C	0.0379	1.20	1.0000	0.2509	0.2328	C	0.1395	1.10
rs9325097	5	148440782	FBXO38	0.3198	0.2420	0.2111	A	0.0496	1.19	0.4959	0.2503	0.2292	A	0.0847	1.12
rs2890313	15	82748701	AP3B2/FSD2	0.9108	0.1741	0.1441	C	0.0303	1.24	0.8163	0.1627	0.1529	C	0.3451	1.08

Table 2. Results of association analysis based on single SNP that passed the p value threshold of discovery stage. The Significant association signal was indicated in bold. The P value thresholds used in discovery stage and validation stage are 0.05 and 0.007 (0.05/7), respectively. MAF_A and MAF_U stand for the MAF of cases and controls, respectively.

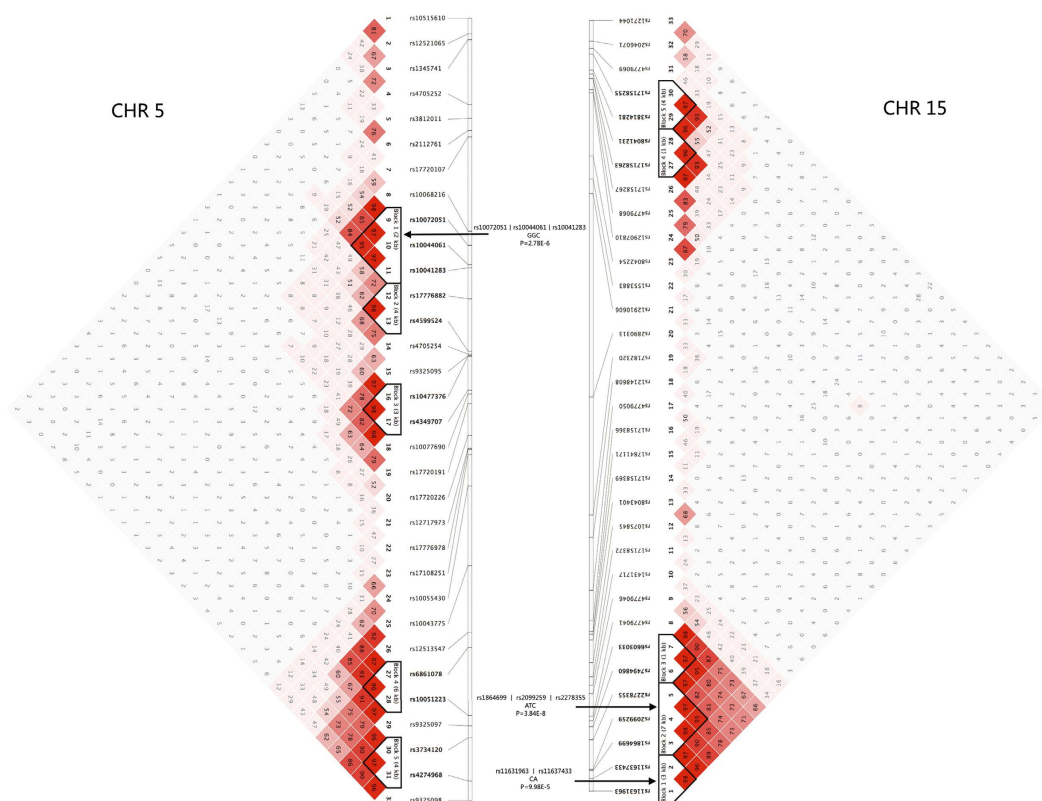


Figure 1. LD structure and significantly associated haplotypes based on the discovery stage data. Ten LD blocks was indicated by the shaded matrices, and three significant haplotypes were identified in the relevant LD blocks indicated by arrows.

gene *FBXO38*) and severe CP with Han Chinese population samples. The statistical power of this GWAS study was more or less limited by its relatively small sample size ($n = 1,020$) and huge number of tests, and therefore, only “suggestive” results were obtained. In our study, we performed a candidate gene-based study and recruited far more samples ($n = 5,065$) to guarantee the sufficient statistical power. In addition to the single marker-based analysis, haplotype analysis produced more significant findings. Among the three significant findings in haplotype analysis, the significant haplotype rs10072051-rs10044061-rs10041283 (GGC, $P = 2.78 \times 10^{-6}$), which covered the *FBXO38* gene region, was a replication of the results from single marker-based analysis. The two other significant haplotypes, rs11631963-rs11637433 (CA, $P = 9.98 \times 10^{-5}$) and rs1864699-rs2099259-rs2278355 (ATC, $P = 3.84 \times 10^{-8}$), covered part of the *AP3B2* gene region. This gene was previously identified as a significant susceptibility gene for CP using samples of European ancestry¹⁹. In addition to these significant findings, we failed to detect any significant signals for WHAMM. Neither single marker nor haplotype-based association analysis resulted in any positive findings for this gene despite it being reported to be a significant candidate gene for CP in a previous GWAS study¹⁹. There may be several reasons for this inconsistent result with the previous study. One is

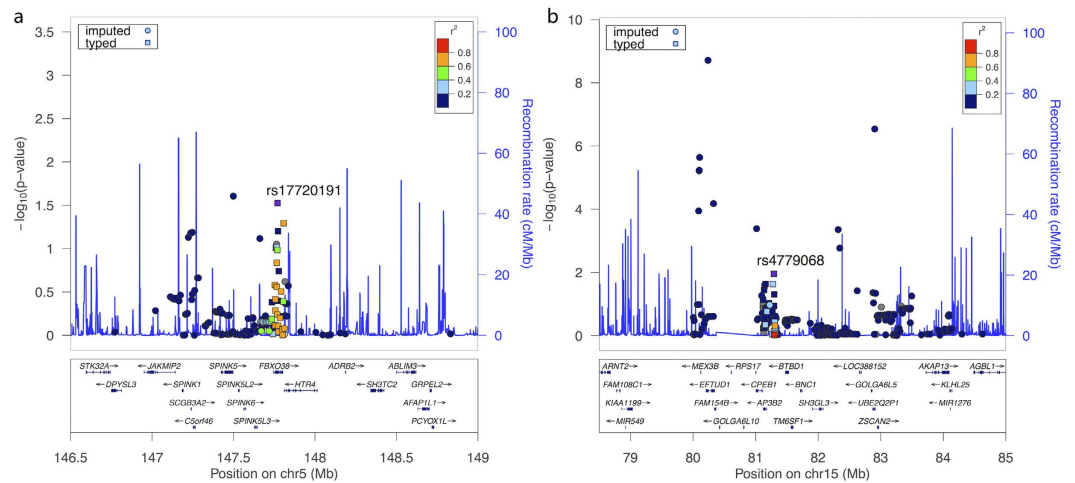


Figure 2. Regional association plots based on imputed regions on chromosome 5 (a) and chromosome 15 (b). Imputed SNPs were indicated as circle and genotyped SNPs were indicated as square. The most significant genotyped SNPs were chosen as reference SNPs in both plots (rs17720191 and rs4779068).

that CP, as a complex disorder, can be associated with multiple genes due to genetic heterogeneity. The difference in ethnic ancestry of the study samples (Han Chinese vs. European) may also affect the results of the association test. Moreover, the statistical method used for the association test in the GWAS study was a gene-centric method, which is different from single marker-based association tests. The significant signal for *WHAMM* identified with this method is due to the contributions of multiple common SNPs within the genetic region, and it is difficult to replicate this result by conducting single marker-based association studies with only a few genotyped SNPs.

Genotype imputation is a widely used statistical technique to predict or impute genotypes that are not directly assayed in a sample of individuals²⁸. Simulations have shown that imputation can lead to a boost in the power of GWAS²⁹. For candidate gene-based association studies, imputed SNPs may sometimes show larger associations and imputation can be used as a powerful tool for fine-mapping those untyped variants when funding is limited³⁰. Through this technique, we successfully imputed and tested 416 SNPs, which increased our marker set by approximately 8 times (416 vs. 65). The strict filtering criteria (certainty > 0.8 and MAF > 0.01) we implemented ensured that we could exclude most if not all of the inaccurately imputed markers. The information from these imputed SNPs offered us a new way to scrutinize our genetic data. The most significant SNPs for both regions (chr5 and chr15) were imputed but not genotyped SNPs. However, when carefully checking the regional association plots made using these top significant SNPs as reference SNPs (Supplemental Figs S2 and S3), we failed to identify a clustering pattern, which might indicate that these top hits might be spurious.

Without further evidence based on functional studies in animal models, it is too early to determine the pathological significance of these susceptibility genes and severe CP. However, it is still meaningful to have a general discussion of the biological functions of these genes. F-box only protein 38, a protein encoded by the gene *FBXO38* (5q32), is a transcriptional co-activator of Krueppel-like factor 7 (*Klf7*), and *Klf7* has long been related to neuronal development³¹. Previous studies have linked *FBXO38* to distal spinal muscular atrophy³¹ and cancer development and progression³². *WHAMM* (15q25.2) encodes a protein that regulates the membrane dynamics and functions at the interface of microtubule and actin cytoskeletons³³. It is related to the process of actin nucleation, mediating Golgi membrane association and microtubule binding³⁴. Adaptor protein-3, beta-2 subunit is a protein encoded by *AP3B2* (15q25.2). As a subunit of Adaptor protein-3 (AP3), it helps to constitute this heterotetrameric vesicle-coat protein complex, which is believed to play an important role in neuron-specific functions³⁵. Early mouse studies have shown that overexpression of the neuronal AP3 subunit, including Ap3b2, can produce a large number of small volume vesicles that release smaller amounts of neurotransmitters compared to mice with AP3 deletions³⁴. AP3, as a protein complex related to neuron functions, has been found to be related to some psychiatric disorders such as schizophrenia³⁵. Despite all of these data, the underlying mechanisms that link these neuronal functions and membrane dynamics to severe CP are still unknown and more research would be necessary.

As a common SNP-based candidate gene association study, our study suffers from several limitations. First, by targeting a relatively small set of common SNPs, we dropped all of the rare variants within these three gene regions. Despite the limitation of statistical analysis methods for rare variant-based association analysis³⁶, recent genetic epidemiology studies have indicated that the rare variants have played an important role in the onset and development of complex disorders^{37–41}. Second, candidate gene-based association study designs only focus on a few potential susceptibility genes but ignore the fact that the biological mechanisms behind most of the complex traits or disorders always involve multiple genes and that these genes interact with each other in a systematic manner^{42–46}. Single gene-based analysis without incorporating knowledge of biological pathways may only offer us limited knowledge about CP^{47,48}. In addition, we obtained the smoking status of study subjects solely based on self-report and its validity might be questionable.

Conclusion

In summary, our study successfully identified two candidate genes, *FBXO38* and *AP3B2*, to be significantly associated with the diagnosis of severe CP in Han Chinese population samples. In the future, further studies are still needed to validate our findings in other cohorts and populations. Additionally, a targeted or exome sequencing study design and pathway-based analysis may provide more evidence to reveal the genetic predisposition of severe CP.

References

- Eke, P. I., Dye, B., Wei, L., Thornton-Evans, G. & Genco, R. J. Prevalence of periodontitis in adults in the United States: 2009 and 2010. *J. Dent. Res.* **91**, 914–920 (2012).
- Ravald, N. & Johansson, C. S. Tooth loss in periodontally treated patients. A long-term study of periodontal disease and root caries. *J. Clin. Periodontol.* **39**, 73–79 (2012).
- Beck, J., Garcia, R., Heiss, G., Vokonas, P. S. & Offenbacher, S. Periodontal disease and cardiovascular disease. *J. Periodontol.* **67**, 1123–1137 (1996).
- Scannapieco, F. A., Bush, R. B. & Paju, S. Associations between periodontal disease and risk for atherosclerosis, cardiovascular disease, and stroke. A systematic review. *Ann. Periodontol.* **8**, 38–53 (2003).
- Glickman, L. T., Glickman, N. W., Moore, G. E., Goldstein, G. S. & Lewis, H. B. Evaluation of the risk of endocarditis and other cardiovascular events on the basis of the severity of periodontal disease in dogs. *J. Am. Vet. Med. Assoc.* **234**, 486–494 (2009).
- Saremi, A. *et al.* Periodontal disease and mortality in type 2 diabetes. *Diabetes Care* **28**, 27–32 (2005).
- Demmer, R. T., Jacobs, D. R. & Desvarieux, M. Periodontal disease and incident type 2 diabetes: Results from the first national health and nutrition examination survey and its epidemiologic follow-up study. *Diabetes Care* **31**, 1373–1379 (2008).
- Novak, M. J., Potter, R. M., Blodgett, J. & Ebersole, J. L. Periodontal disease in Hispanic Americans with type 2 diabetes. *J. Periodontol.* **79**, 629–636 (2008).
- Xiong, X., Buekens, P., Fraser, W. D., Beck, J. & Offenbacher, S. Periodontal disease and adverse pregnancy outcomes: A systematic review. *BJOG An Int. J. Obstet. Gynaecol.* **113**, 135–143 (2006).
- Megson, E., Kapellas, K. & Bartold, P. M. Relationship between periodontal disease and osteoporosis. *Int. J. Evid. Based. Healthc.* **8**, 129–139 (2010).
- Guiglia, R. *et al.* Osteoporosis, jawbones and periodontal disease. *Med. Oral Patol. Oral Cir. Bucal* **18**(1), e93–e99 (2013).
- Corey, L. A., Nance, W. E., Hofstede, P. & Schenkein, H. A. Self-reported periodontal disease in a Virginia twin population. *J. Periodontol.* **64**, 1205–1208 (1993).
- Michalowicz, B. S. *et al.* Evidence of a substantial genetic basis for risk of adult periodontitis. *J. Periodontol.* **71**, 1699–1707 (2000).
- Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am. J. Hum. Genet.* **90**, 7–24 (2012).
- Feng, P. *et al.* Genome wide association scan for chronic periodontitis implicates novel locus. *BMC Oral Health* **14**, 84 (2014).
- Teumer, A. *et al.* Genome-wide association study of chronic periodontitis in a general German population. *J. Clin. Periodontol.* **40**, 977–985 (2013).
- Divaris, K. *et al.* Exploring the genetic basis of chronic periodontitis: A genome-wide association study. *Hum. Mol. Genet.* **22**, 2312–2324 (2013).
- Schaefer, A. S. *et al.* A genome-wide association study identifies *GLT6D1* as a susceptibility locus for periodontitis. *Hum. Mol. Genet.* **19**, 553–562 (2010).
- Rhodin, K. *et al.* Chronic periodontitis genome-wide association studies: gene-centric and gene set enrichment analyses. *J. Dent. Res.* **93**, 882–890 (2014).
- Divaris, K. *et al.* Genome-wide Association Study of Periodontal Pathogen Colonization. *J. Dent. Res.* **91**, S21–S28 (2012).
- Frazer, K. A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
- Purcell, S., Cherny, S. S. & Sham, P. C. Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics*, **19**(1), 149–150 (2003).
- Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* Chapter 7, Unit 7.20; doi: 10.1002/0471142905.hg0720s76 (2013).
- Ripley, B. D. The R project in statistical computing. *MSOR Connect.* **1**, 23–25 (2001).
- Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
- Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
- Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
- Spencer, C. C. A., Su, Z., Donnelly, P. & Marchini, J. Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet.* **5**(5), e1000477 (2009).
- Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
- Sumner, C. J. *et al.* A dominant mutation in *FBXO38* causes distal spinal muscular atrophy with calf predominance. *Am. J. Hum. Genet.* **93**, 976–983 (2013).
- Wan, D. *et al.* Large-scale cDNA transfection screening for genes related to cancer development and progression. *Proc. Natl. Acad. Sci. USA* **101**, 15724–15729 (2004).
- Campellone, K. G., Webb, N. J., Znameroski, E. A. & Welch, M. D. WHAMM Is an Arp2/3 Complex Activator That Binds Microtubules and Functions in ER to Golgi Transport. *Cell* **134**, 148–161 (2008).
- Grabner, C. P., Price, S. D., Lysakowski, A., Cahill, A. L. & Fox, A. P. Regulation of large dense-core vesicle volume and neurotransmitter content mediated by adaptor protein 3. *Proc. Natl. Acad. Sci. USA* **103**, 10035–10040 (2006).
- Hashimoto, R. *et al.* Association analysis between schizophrenia and the AP-3 complex genes. *Neurosci. Res.* **65**, 113–115 (2009).
- Zhang, T. X., Xie, Y. R. & Rice, J. P. Application of noncollapsing methods to the gene-based association test: a comparison study using Genetic Analysis Workshop 18 data. *BMC Proc.* **8**, S53 (2014).
- Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat. Rev. Genet.* **11**, 415–425 (2010).
- Guan, F. *et al.* Association of *PDE4B* polymorphisms and schizophrenia in Northwestern Han Chinese. *Hum. Genet.* **131**, 1047–1056 (2012).
- Guan, F. *et al.* *MIR137* gene and target gene *CACNA1C* of *miR-137* contribute to schizophrenia susceptibility in Han Chinese. *Schizophr. Res.* **152**, 97–104 (2014).
- Guan, F. *et al.* A population-based association study of 2q32.3 and 8q21.3 loci with schizophrenia in Han Chinese. *Journal of psychiatric research.* **47**, 712–717 (2013).
- Guan, F. *et al.* Association study of a new schizophrenia susceptibility locus of 10q24.32–33 in a Han Chinese population. *Schizophr. Res.* **138**, 63–68 (2012).

42. Robinson, R. Common disease, multiple rare (and distant) variants. *PLoS Biol.* **8**, e1000293 (2010).
43. Chen, G., Guan, F., Lin, H., Li, L. & Fu, D. Genetic analysis of common variants in the HDAC2 gene with schizophrenia susceptibility in Han Chinese. *Journal of Human Genetics.* **60**(9), 479–484 (2015).
44. Goh, K., Cusick, M. E., Valle, D., Childs, B. & Vidal, M. The human disease network. *Proc. Natl. Acad. Sci. USA* **104**(21), 8685–8690 (2007).
45. Guan, F. *et al.* Evaluation of genetic susceptibility of common variants in CACNA1D with schizophrenia in Han Chinese. *Sci Rep.* **5**, 12935 (2015).
46. Zhang, B. *et al.* Common variants in *SLC1A2* and schizophrenia: Association and cognitive function in patients with schizophrenia and healthy individuals. *Schizophr Res.* **S0920-9964**(15), 30012–30018; doi: 10.1016/j.schres.2015.10.012 (2015).
47. Sun, J. *et al.* Schizophrenia gene networks and pathways and their applications for novel candidate gene selection. *PLoS One* **5**(6), e11351 (2010).
48. Zhang, T. X., Beaty, T. H. & Ruczinski, I. Candidate pathway based analysis for cleft lip with or without cleft palate. *Stat. Appl. Genet. Mol. Biol.* **11**, 1–19 (2012).

Acknowledgements

We thank the participation of volunteers, patients and their families. This research was supported by and National Natural Science Foundation of China (No. 81401563), China Postdoctoral Science Foundation Funded Project (No. T70927 and M532029), Ph.D. Programs Foundation of Ministry of Education of China (No. 2013021120078) and Fundamental Research Funds for the Central Universities (No. 08142024 and 08143003). The funding sources had no role in the design of this study, the collection, analysis and interpretation of data, the writing of the report, or the decision to submit the paper for publication.

Author Contributions

D.S., H.T. and L.Z. have done sample collection, genotyping, data analysis and written the manuscript which was critically revised by N.Z.; L.D. has helped in data analysis; L.Z. has helped in DNA sample preparation and genotyping; J.X. and Y.W. conducted healthy subject screening; R.Y. and R.X. have clinically characterized the patients whose biological samples have been included in the study; H.T. and N.Z. have conceptualized and led this project including arranging for the required funds. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Shang, D. *et al.* Two-stage comprehensive evaluation of genetic susceptibility of common variants in *FBXO38*, *AP3B2* and *WHAMM* to severe chronic periodontitis. *Sci. Rep.* **5**, 17882; doi: 10.1038/srep17882 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>