Identification of Novel Genome-Wide Associations for Oral Inflammatory Traits

Yanjiao Jin
Capital Medical University School of Stomatology: Beijing Stomatological Hospital
Jie Yang
Tianjin Medical University
Shuyue Zhang
Tianjin Medical University
Jin Li (jli01@tmu.edu.cn)
Tianjin Medical University https://orcid.org/0000-0001-7024-3591
Songlin Wang
Capital Medical University School of Stomatology: Beijing Stomatological Hospital

Research article

Keywords: oral diseases, genome-wide association studies, inflammation, immunity, pleiotropic loci, meta-analysis

Posted Date: November 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-104530/v1

Abstract

Background: Oral diseases impact the majority of the world's population. The following traits are common in oral inflammatory diseases: mouth ulcers, painful gums, bleeding gums, loose teeth, and toothache. Despite the prevalence of genome-wide association studies, the associations between these traits and common genomic variants, and whether pleiotropic loci are shared by some of these traits remain poorly understood.

Methods: In this work, we conducted multi-trait joint analyses based on the summary statistics of genome-wide association studies of these five oral inflammatory traits from the UK Biobank, each of which is comprised of over 10,000 cases and over 300,000 controls. We estimated the genetic correlations between the five traits. We conducted fine-mapping and functional annotation based on multi-omics data to better understand the biological functions of the potential causal variants at each locus. To identify the pathways in which the candidate genes were mainly involved, we applied gene-set enrichment analysis, and further performed protein-protein interaction (PPI) analyses.

Results: We identified 39 association signals that surpassed genome-wide significance, including three that were shared between two or more oral inflammatory traits, consistent with a strong correlation. Among these genome-wide significant loci, two were novel for both painful gums and toothache. We performed fine-mapping and identified causal variants at each novel locus. Further functional annotation based on multi-omics data suggested *IL10* and *IL12A/ TRIM59* as potential candidate genes at the novel pleiotropic loci, respectively. Subsequent analyses of pathway enrichment and protein-protein interaction networks suggested the involvement of candidate genes at genome-wide significant loci in immune regulation.

Conclusions: Our results highlighted the importance of immune regulation in the pathogenesis of oral inflammatory diseases. Some common immune-related pleiotropic loci or genetic variants are shared by multiple oral inflammatory traits. These findings will be beneficial for risk prediction, prevention, and therapy of oral inflammatory diseases.

Background

Oral inflammatory diseases, including gingivitis, periodontitis, endodontitis, and mouth ulcers, affect most of the world's population. Periodontal diseases affect approximately 90% of adults in the US [1]. The genetic factors in the occurrence of oral inflammatory diseases have been well established in twin, family, candidate gene, and genome-wide association studies (GWAS) [2–8]. A comprehensive understanding of the genetic and genomic factors of oral inflammatory diseases is necessary to enable the development of risk assessment tools and therapeutic modalities and for the realization of precision oral health care.

Growing evidence in recent years has established the contributions of genetic factors to oral inflammatory diseases. For example, the *IL37* locus is associated with high IL-1β levels in gingival

crevicular fluid in periodontal inflammation [7]. Two *IL37* variants also result in decreased expression of IL-37, leading to the upregulation of IL-1β and IL-6, constituting a hyper-inflammatory trait [7]. The *IL-37V1* locus is associated with severe chronic periodontal disease in the dental ARIC dataset population [7]. Another study found 97 variants that alter the odds of developing nonspecific mouth ulcers and replicated these in an independent cohort [8]. Further *in silico* functional analyses provided evidence for the role of T cell regulation in the etiology of mouth ulcers [8].

Although the exact mechanisms for the pathogenesis of different kinds of oral inflammatory diseases still need to be elucidated, an important takeaway from these studies is that genetic factors that regulate immune and inflammatory conditions may play an important role in the occurrence and susceptibility of these diseases [9]. However, the loci contributing to common traits or diseases may be disorder-specific or shared between disorders. An example of the latter would be the loci with pleiotropic effects that were identified through a meta-analysis study across eight psychiatric disorders [10]. They found that the candidate genes at the pleiotropic loci show heightened expression in the brain throughout the lifespan, beginning prenatally in the second trimester, and play prominent roles in neurodevelopmental processes, thus having a broad effect on the occurrence and susceptibility of multiple psychiatric disorders [10]. The findings of pleiotropic loci in human diseases challenge the classical and categorical classification of illness. Given that immune-related genetic factors may have a profound effect on the occurrence of oral inflammatory diseases, and the regulation of the immune system is often broad and systematic, we hypothesized that pleiotropic loci may also affect multiple oral inflammatory traits. If that is true, multiple phenotypes (or clinical traits) may share common gene variants. Alternatively, despite some potential similarities in pathogenesis among oral inflammatory diseases, few genetic variants to date have explicitly been associated with two or more clinically defined oral inflammatory disorders [11]. We believe that this may be because the clinical diagnosis of a disease as a phenotype definition is an aggregate of multiple clinical manifestations, including multiple symptoms and signs. More importantly, individuals diagnosed with the same disease are often not highly consistent in the expression patterns of all the manifestations, which may weaken the association with the genetic variants. However, historically, disease phenotypes were defined according to the clinical diagnosis and classification. Contrasting this, a recent investigation of complex traits used principal component analysis combined with subgingival microbial composition data and biomarkers of the tissue inflammatory response [6]. A properly defined phenotype is important for determining the underlying variant(s).

Here, we examined pleiotropic effects in a dataset that included five clinical traits: mouth ulcers (MU), painful gums (PG), bleeding gums (BG), loose teeth (LT), and toothache (TA), each of which is comprised of over 10,000 cases and over 300,000 controls. These clinical manifestations are often seen in major oral inflammatory diseases, including mouth ulcers, gingivitis, periodontitis, and endodontitis. We address the following main questions regarding the genetic basis of oral inflammatory diseases: (1) Are there pleiotropic loci associated with risks for multiple clinical traits of oral inflammatory diseases? (2) Can we identify the functional features of pleiotropic loci that could explain their broad effects on the immune system and inflammation in oral diseases?

Methods

Cohort description and quality control filtering

We included five UKBB cohorts of oral inflammatory traits, MU, PG, BG, LT, and TA, in our study, all of which were downloaded from the website: https://www.ukbiobank.ac.uk. The number of cases and controls in each cohort is listed in Additional file 1: Table S1.

We first conducted quality control (QC) filtering on the summary statistical data of each cohort. We excluded variants with low confidence and those with minor allele frequency (MAF) < 0.01. The number of SNPs before and after filtering is shown in Additional file 1: Table S1.

Joint analyses of correlated disease traits

After QC, we performed a correlated-trait joint analysis of the five oral inflammatory traits based on the summary statistics of each trait using the MTAG (Multi-Trait Analyses of GWAS) [12] with the default parameter setting. MTAG allows for sample overlap using bivariate LD score regression [12].

Define independent loci

We performed clumping with PLINK [13] (V1.9) to define independent significant loci with the setting of –clump-r2=0.4, –clump-kb=500, –clump-p1=5e-08, –clump-p2=5e-02. Then, we merged any adjacent loci with an in-between distance of < 400kb and chose the SNP with the smallest P-value as the index SNP for the region. The region of major histocompatibility complex (MHC) at chr6:25-35Mb was considered a single locus.

Next, we conducted stepwise conditional-joint analyses with the software GCTA-COJO [14] to identify possible independent signals within each locus using the 1000G EUR phase3 panel [15] as the reference panel.

Fine-mapping

We first performed LD-score calculation based on the 1000G EUR phase3 panel using PLINK, and then, we applied the shotgun stochastic search algorithm of the FINEMAP software [16] to identify causal variants (with 95% probability). For each independent significant locus, variants with $r^2 > 0.6$ with the index SNP and the association P-value in MTAG analyses < 1×10^{-4} were selected as causal variants.

Heritability analyses

Pre-computed LD scores from the 1000G EUR phase3 panel were downloaded from the website http://data.broadinstitute.org/alkesgroup/LDSCORE/ and applied to calculate the SNP heritability based on GWAS summary statistics and MTAG analysis results of the five oral inflammatory traits.

Genetic correlation analyses

We conducted analyses of genetic correlation between the five diseases via LD score regression (LDSC) analyses [17] based on MTAG results of variants in the HapMap3 dataset as recommended [18]. LDSC allows for sample overlap between cohorts by applying a self-estimated intercept in the analyses.

Functional annotation

We performed functional annotation of all the variants by FUMA GWAS [19]. The expression quantitative trait loci (eQTLs) were annotated based on the PsychENCODE [20]: scRNA eQTLs [21] , DICE[22] , eQTLGen [23], Blood eQTLs [24], and Salivary Gland of GTEx v8 [25], and GTEx v7 [26]. The false discovery rate was controlled at 0.05 for significant SNP-gene pairs. For chromatin interaction annotation, the Hi-C dataset of immune tissues and cells, Spleen and GM12878(GSE87112) [27], were included for analyses.

Histone modification site analyses

Haploreg [28] was applied to explore the epigenomic annotations of all the index SNPs and variants in the credible sets based on the epigenomics databases of ROADMAP [29]. Epigenomic data from immune cell types were applied.

Gene-set-enrichment analyses

We used MAGMA [30] implemented in FUMA GWAS to perform gene mapping, gene expression profiling, and gene-set-enrichment analyses, with KEGG as the pathway resource database. Bonferroni correction was adopted to adjust for multiple-testing correction. The gene windows were set as the transcript region of each gene, and the GTEx v8 dataset was selected as a reference. We next applied protein-protein interaction network analyses with STRING [31] to all 39 candidate genes at the GWS loci. The minimum required interaction score was set to 0.4.

Results

Multi-trait joint analyses and identification of genome-wide significant (GWS) novel loci

To systematically examine the shared genetic basis of MU, PG, BG, LT, and TA, we conducted a multi-trait joint analysis based on the summary statistics of 7194361 single nucleotide polymorphisms (SNPs) common to these five correlated traits (Additional file 1: Table S1). Quantile-quantile (QQ) plots and genomic inflation factor (Additional file 2: Fig. S1) suggested no evidence of inflation due to confounding factors. In our joint-analyses, 36 independent loci reached the GWS threshold ($P < 5\times10^{-8}$) after clumping (MU:28, PG: 4, BG: 1, LT: 1, TA: 4; Additional file 3: Data file S1). Compared with the original UK Biobank (UKBB) summary statistics and the loci summarized by the GWAS catalog [32] of these five traits, there were three GWS novel loci, including the *IL10* and *IL12A-AS1* loci for PG and TA, and the *BTN3A1* locus for TA. Interestingly, although BG and LT traits were also considered as important manifestations of periodontitis, these three loci did not reach GWS association with BG or LT. Furthermore, step-wise

conditional analyses yielded independent GWS signals at three loci for trait MU. Therefore, we found 39 independent GWS signals, including three novel loci for PG and TA traits (Table 1, Fig. 1, Additional file 4: Fig. S2 and Additional file 5: Fig. S3).

Table 1 The novel significant loci from MTAG analysis

Trait	rsID	chr	pos	A1/A2	Closest gene	MTAG beta	MTAG P-value
PG	rs1518110	1	206944861	C/A	*IL10*	-0.0109	1.76E-14
	rs17753641	3	159647674	G/A	*IL12A-AS1*	0.0259	5.4E-49
TA	rs1518110	1	206944861	C/A	*IL10*	-0.0129	1.77E-11
	rs17753641	3	159647674	G/A	*IL12A-AS1*	0.0313	4.05E-40
	rs3799378	6	26404374	G/A	*BTN3A1*	-0.0103	9.66E-09

PG: painful gums; TA: toothache; chr: chromosome; pos: position (hg19); A1: alternative allele; A2: reference allele

Among the 39 GWS associations, three were pleiotropic, associated with two or more of the oral inflammatory traits (Additional file 6: Data file S2). The index SNPs at the three novel loci (rs1518110, rs17753641, and rs3799378) showed consistent directions of effect across the five traits, although the associations were not genome-wide significant for BG or LT traits (Fig. 2). Among the three novel loci, two were pleiotropic. The locus at 1q32.1, peaking at SNP rs1518110, was associated with PG and TA. SNP rs1518110 was located in the intron of gene interleukin 10 (*IL10*), the product of which is a cytokine mainly produced by monocytes and plays an important role in immune regulation and inflammation. Another pleiotropic novel GWS locus associated with both PG and TA was at 3q25.33 with rs17753641 as the lead SNP located within the IL12A antisense RNA 1 (*IL12A-AS1*), an RNA gene belonging to the lncRNA class. Another gene at this locus*, IL12A* which also plays an important role in the regulation of inflammation and immunity, was also a potential candidate gene. We also found a novel association of the human leukocyte antigen (HLA) locus with TA, the index SNP of which was rs3799378, located in the intron of gene *BTN3A1*.

Heritability and genetic correlation of the five oral disease traits

We computed heritability based on the multi-trait analysis of GWAS (MTAG) results to estimate the proportion of phenotypic variance explained by common variants. The results ranged from 0.018 to 0.029 for the five oral disease traits (MU h^2 = 0.029, PG h^2 = 0.026, BG h^2 = 0.024, LT h^2 = 0.018, TA h^2 = 0.023), each of which was increased compared with the heritability estimated from the original UKBB GWAS summary statistics (Additional file 7: Fig. S4 and Additional file 8: Data file S3).

Next, we estimated the genetic correlations between the five traits. Traits PG and TA were most closely related (r_g = 0.995), whereas the correlation between MU and LT was lowest (r_g = 0.343) (Fig. 3 and

Additional file 8: Data file S3). PG, BG, and LT reflect the different aspects or stages of periodontitis and showed an average genetic correlation of >0.7.

Fine-mapping of plausible variants set at novel loci

In GWAS, the index SNP might be a tag SNP; thus, we conducted fine-mapping to identify plausible causal variants at the novel GWS loci (Additional file 9: Data file S4). The lead SNPs rs1518110 and rs17753641 were in the credible sets, together with other potential causal variants in tight linkage disequilibrium (LD).

Functional annotation of causal SNPs at each novel locus

To better understand the biological functions of the potential causal variants at each locus, we conducted functional annotation. We examined functional genomics datasets, including histone modification, expression quantitative trait loci (eQTL), and high-throughput chromosome conformation capture (Hi-C) data in immune cell types and periodontal tissues. All plausible variants overlapped with histone modification marks in immune cells and tissues (Additional file 10: Data file S5). We observed the eQTL relationship between causal variants rs1518110 at locus 1q32.1 and gene *IL10* from databases of DICE [22], eQTLGen [23], and BIOSQTL [33]. Two other eQTL genes at this locus were *IL24* and *FAIM3*, which encodes an Fc receptor for IgM. The gene showing an eQTL relationship with the causal variant rs17753641 at the locus of 3q25.33 was *TRIM59* functioning in the regulation of innate immunity (Additional file 11: Fig. S5, Additional file 12: Data file S6, Additional file 13: Data file S7). Another nearby gene, *IL12A*, was also a likely candidate gene based on its eQTL relationship ($P = 1.8 \times 10^{-5}$ in pituitary) and biological function. Interestingly, SNP rs17753641 was in a strong enhancer/promoter region in diverse cell types with transcription factor ELF1 bound to it in B-lymphocyte cell type GM12878, demonstrated by ChIP-Seq experiments. ELF1 is a key player in the antiviral immune response and its associations with PG and TA were indexed by SNPs rs113780118 ($P = 4.06 \times 10^{-02}$) and rs76226306 ($P = 4.89 \times 10^{-02}$). The third locus for TA was in the HLA region with the index SNP rs3799378 in the intron of gene *BTN3A1*. This SNP had an eQTL relationship for 12 genes in diverse cell types. Collectively, these findings from epigenomics and transcriptome databases strongly suggested the role of causal variants at the three novel loci in the transcriptional regulation of nearby genes and genes at a distance, in line with their potential role in the pathogenesis of periodontitis and other oral inflammatory diseases.

Gene-enrichment analyses

To identify the pathways in which the candidate genes at all the GWS loci were mainly involved, we applied gene-set enrichment analysis using Gene Ontologue and Kyoto Encyclopedia of Genes and Genomes (KEGG) datasets. We observed a predominant enrichment of immune signaling pathways, such as the process of cytokine and cytokine receptor interaction, chemokine signaling pathway, and JAK STAT signaling pathway for all three traits. We also found the enrichment of autoimmune disease pathways, such as systemic lupus erythematosus (SLE), consistent with the fact that SLE is a systemic,

chronic inflammatory condition with different clinical manifestations, including oral inflammation. (Additional file 14: Fig. S6, Additional file 15: Data file S8).

We further performed protein-protein interaction (PPI) analyses (Additional file 16: Fig. S7) based on 39 candidate genes located at the GWS loci of any oral inflammatory trait. The results of the PPI network analyses showed an enrichment P-value of 4.34×10^{-11}, suggesting close coordination between these proteins in cellular activity. We saw a large module centered on *STAT4*, with *IL10*, *IRF8*, *and NOD2* as other hub proteins in this module, reflecting the autoinflammatory nature of the five traits.

Discussion

In this GWAS meta-analysis of multiple oral inflammatory traits, each of which was comprised of over 10,000 cases and over 300,000 controls, we identified a total of 39 associations that surpassed genome-wide significance. Of the 39, three were shared between two or more oral inflammatory traits, two were novel loci for both PG and TA, and one was novel locus for TA. We also identified potential causal variants at each novel locus. Further functional annotation based on the multi-omics data suggested that *IL10* and *IL12A/TRIM59* may be potential candidate genes at the novel pleiotropic loci, respectively. Subsequent analyses of pathway enrichment and protein-protein interaction networks suggested the involvement of candidate genes at genome-wide significant loci in immune regulation. Our results provide several major insights into the genome-wide associations for oral inflammatory traits and their shared genetic structure.

First, the genetic contribution to oral diseases and their traits has been well established [34]. Similar to other cardiovascular, cerebrovascular, and neuropsychiatric diseases and diabetes, the susceptibility to oral inflammatory diseases can also be explained by the “common disease common variant” hypothesis that proposes if a heritable disease is common in the population, then the associated genetic variants will also be common in the population. GWAS are a powerful approach with agnostic nature and coverage of a substantial proportion of the common variants in the human genome. We identified 39 associations that surpassed genome-wide significance, most of which had allele frequencies higher than 0.1. This roughly reflects the consistency between the incidence of the phenotypes of oral disease and genomic variants.

Second, novel genome-wide significant loci were identified for oral inflammatory traits, including *IL10*, *IL12A/TRIM59*, and *BTN3A1*. Although previous studies have suggested the possible role of IL-10 and interleukin 12 (IL-12) in the development and progression of oral inflammatory diseases, especially periodontitis, no study has clearly shown their association with oral inflammatory diseases at the genome-wide level. IL-10 is mainly produced by monocytes, macrophages, and $CD4^+$ Th2 cells, and is strongly involved in the downregulation of the inflammatory response, thereby preventing damage to the host. It stimulates the production of anti-inflammatory mediators such as interleukin 1 (IL-1) receptor antagonist and downregulates the release of pro-inflammatory mediators such as tumor necrosis factor alpha (TNF-α), IL-1b, interleukin 6 (IL-6), and interleukin 8 (IL-8) [35]. It plays a major role in controlling

inflammation and preventing excessive tissue damage caused by bacterial and viral infections as well as proinflammatory responses. Low IL-10 levels lead to enhanced release of pro-inflammatory cytokines [36]. Mutations within IL10 are associated with increased susceptibility to rheumatoid arthritis [37] and human immunodeficiency virus type 1 (HIV-1) infection [38]. Interestingly, recent studies have shown associations between periodontitis and HIV-1 infection [39] and rheumatoid arthritis [40]. Multiple studies have also investigated the association between *IL10* gene polymorphisms and the risk of periodontitis [41]. However, the reported results were inconsistent and remained inconclusive [41]. Three SNPs in the promoter region of the *IL10* gene, including rs180087, rs1800872, and rs1800896, have been extensively investigated for their association with periodontitis [42]. A recent meta-analysis showed that rs1800871 and rs1800872 were associated with increased susceptibility to periodontitis in Latin American populations but not in Asian and Caucasian populations. There was also no significant association between rs1800896 and periodontitis risk [42]. We confirmed at the genome-wide level the role of gene *IL10* with causal variant rs1518110 at locus 1q32.1 for the risk of oral inflammatory traits of PG and TA. However, we observed no associations between the locus of *IL10* and the traits BG and LT, which are also important manifestations of periodontitis. It is likely that a subgroup of oral inflammatory diseases with *IL10* gene defects may exist. Our study also suggested that the inconsistent results in previous publications may be due to the complex clinical manifestations in periodontitis.

Both the potential candidate genes *IL12A* and *TRIM59* at the novel locus of 3q25.33 play important roles in the regulation of inflammation and immunity. IL-12 is a heterodimeric inflammatory cytokine, composed of a 35 kDa chain (p35) and a 40 kDa chain (p40), which are encoded by genes *IL12A* (p35) and *IL12B* (p40), respectively. In periodontal disease, the level of IL-12 in the gingival crevicular fluid is positively correlated with the severity of inflammation [8]. During the development of the periodontal disease, leukocytes in the periodontium are dominated by T-lymphocytic cells (T cells), which play a role in eliminating pathogenic microbes and maintaining tissue homeostasis. IL-12 is an important regulatory cytokine in the activation of this cell-mediated immunity [43]. *IL12A-AS1* has been reported to be associated with Behceh's disease [44]. A recent GWAS identified multiple associated loci located in or near genes or enriched in pathways related to T cell immunity, including a common variant near *IL12A*, with large effects on the risk of mouth ulcers [8]. *IL12B* genetic variants are associated with susceptibility to chronic periodontitis in a Taiwanese population [45]. Our findings suggested that *IL12A* at the genome-wide level may be a candidate gene for oral inflammatory diseases. However, we do not know whether the *IL12* polymorphisms have any ethnic differences in their association with the risk of oral inflammatory disease, as the data we analyzed were from the UKBB. The candidate gene *TRIM59* at the locus of 3q25.33 functions in the regulation of innate immunity [46]. Previous studies have shown its role in the progression of different types of cancers [47]. Our work suggested its possible role in some traits of oral inflammatory disease. Notably, SNP rs17753641 at locus 3q25.33 was in a strong enhancer/promoter region in diverse cell types with transcription factor ELF1 bound to it in B-lymphocyte cell type GM12878. ELF1 itself is a key player in the antiviral immune response [48]. Butyrophilins-3A (BTN3A) as a cell-surface receptor is expressed on multiple cell types, including immune cells and some malignant cells [49]. BTN3A plays an important role in T cell proliferation and cytokine production [49]. Despite its key

role in inflammation and immunity, few studies have investigated the associations with oral inflammatory diseases. In short, our analysis suggested the role of causal variants at the three novel loci in the transcriptional regulation of nearby genes and genes at a distance, in line with their potential role in the pathogenesis of oral inflammatory diseases.

Third, our study showed that three loci were shared between two or more oral inflammatory traits and confirmed the hypothesis that multiple oral inflammation-related phenotypes may share common pleiotropic loci or genetic variants. As described above, the genes *IL10*, *IL12A*, *TRIM59*, and *CCR3* all play important roles in inflammation and immunity. Many studies support the view that cytokines play a pivotal role at all stages of the immune response in oral inflammatory diseases. However, few cytokine-related loci or genetic variants have been associated with two or more oral inflammatory diseases or their traits. This could be due to the limitations of how the phenotypes are defined. GWAS of oral inflammatory diseases defined by clinical diagnosis criteria with an aggregate of multiple clinical manifestations have had modest success to date. Another reason for the difficulty of obtaining conclusive findings may be the low power to detect small genetic effects, which is common to all GWAS and the relatively modest sample sizes. To identify associations with small effects confidently, very large numbers of cases and controls will need to be studied. In this GWAS meta-analysis, we applied clinical presentations; however, we did not simply use the definition of disease diagnosis to describe the phenotypes. We also applied a large sample size from UKBB with over 10,000 cases and over 300,000 controls for each of the five traits. Thus, we successfully obtained three pleiotropic loci and the results indicating multiple pairwise genetic correlations between oral inflammatory traits.

Finally, our analyses suggested the involvement of candidate genes at genome-wide significant loci in immune regulation and the importance of immune regulation in the pathogenesis of oral inflammatory diseases. On one hand, confirmation of common variants of common oral traits can ultimately be beneficial for precise screening, prevention, and therapy of oral inflammatory diseases. For example, a genomic assessment could help determine which patients should visit the dentist more often and would benefit from additional preventive visits. On the other hand, future studies should focus on the molecular mechanisms involved in the identification of variants that affect gene activities and influence susceptibility to oral inflammatory diseases. Upon replication and mechanistic studies, the findings will enhance our understanding of the pathogenesis of oral inflammatory diseases and aid in the development of novel drugs targeted at impaired immune regulation.

Our study had several limitations. First, individual-level data were not available in our analysis; therefore, the related behavioral, lifestyle, and systemic risk factors, such as age, diet, smoking, diabetes, obesity, and vitamin D deficiency are not known [5]. We do not know whether there are significant differences in the distributions of these non-genetic risk factors among the five traits, and to what extent, this imbalance of distributions may have weakened our power to detect associations and pleiotropic loci. Second, we also do not know the extent of the overlap of the individuals in the five traits, which may overestimate the multiple pairwise genetic correlations between oral inflammatory traits. Third, our designation of pleiotropic loci versus non-pleiotropic loci refers only to the observed effects on the five common traits. If

more traits were included in the analyses, some of the “non-pleiotropic” loci may have been categorized as pleiotropic loci. Lastly, the results should be interpreted cautiously when applied in non-Caucasian populations. Future studies should look to include data from more diverse ethnic populations.

Conclusions

In our cross-trait genome-wide meta-analysis, we identified 39 genomic loci. Of which three loci (*IL10*, *CCR3*, and *IL12A/TRIM59*) had pleiotropic effects, two were novel loci (*IL10*, *IL12A/TRIM59*) for both PG and TA, and one was novel locus (*BTN3A1*) for TA. We also identified causal variants at each novel locus. Further functional annotation based on multi-omics data suggested *IL10* and *IL12A/ TRIM59* as potential candidate genes at the novel pleiotropic loci, respectively. Subsequent analyses of pathway enrichment and protein-protein interaction networks suggested the involvement of candidate genes at genome-wide significant loci in immune regulation. Our results highlighted the importance of immune regulation in the pathogenesis of oral inflammatory diseases.

Abbreviations

GWAS: Genome-wide association study; MU: Mouth ulcers; PG: Painful gums; BG: Bleeding gums, LT: Loose teeth; TA: Toothache; UKBB: UK Biobank; GWS: Genome-wide significant; SNPs: Single nucleotide polymorphisms; IL-10: Interleukin 10; MTAG: Multi-trait analysis of GWAS; LD: Linkage disequilibrium; eQTL: Expression quantitative trait loci; Hi-C: High-throughput chromosome conformation capture; SLE: Systemic lupus erythematosus; PPI: Protein-protein interaction; IL-12: Interleukin 12; TNF-α: Tumor necrosis factor alpha; IL-1: Interleukin 1; IL-6: Interleukin 6; IL-8: Interleukin 8; HIV-1: Human immunodeficiency virus type 1; BTN3A: Butyrophilins-3A; QC: Quality control; MAF: Minor allele frequency; MHC: Major histocompatibility complex; LDSC: LD score regression; KEGG: Kyoto Encyclopedia of Genes and Genomes.

Declarations

-Ethics approval and consent to participate

Not applicable.

- Consent for publishing

Not applicable.

-Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files (Additional files 1-16).

-Competing interests

The authors declare that they have no competing interests.

-Funding

This research was supported by the Scientific and Technological Research Grant of Tianjin Health Bureau (2013KZ118 to YJ). The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, approval, and decision to submit the manuscript for publication.

-Authors' contributions

YJ, JL, and SW conceived the study; YJ and JL designed the study. JL and SW supervised the study and oversaw data collection and analysis. YJ and JY collected the data; JL guided the data analysis; YJ, JY, and SZ conducted the analysis. YJ and JY wrote the first draft of the paper. JL and SW reviewed and edited the paper. All authors approved the final version of the manuscript for publication.

-Acknowledgments

We thank Dr. Xuefeng Shi, MD, PhD (Tianjin Eye Hospital, Clinical College of Ophthalmology, Tianjin Medical University, Tianjin, China) for participating in insightful discussions about the design and results of the study and providing helpful suggestions while editing the manuscript.

References

1. Eke PI, Dye BA, Wei L, Slade GD, Thornton-Evans GO, Borgnakke WS, Taylor GW, Page RC, Beck JD, Genco RJ. Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012. J Periodontol. 2015;86(5):611–22.
2. Schaefer AS, Richter GM, Nothnagel M, Manke T, Dommisch H, Jacobs G, Arlt A, Rosenstiel P, Noack B, Groessner-Schreiber B, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet. 2010;19(3):553–62.
3. Divaris K, Monda KL, North KE, Olshan AF, Reynolds LM, Hsueh WC, Lange EM, Moss K, Barros SP, Weyant RJ, et al. Exploring the genetic basis of chronic periodontitis: a genome-wide association study. Hum Mol Genet. 2013;22(11):2312–24.
4. Teumer A, Holtfreter B, Volker U, Petersmann A, Nauck M, Biffar R, Volzke H, Kroemer HK, Meisel P, Homuth G, et al. Genome-wide association study of chronic periodontitis in a general German population. J Clin Periodontol. 2013;40(11):977–85.
5. Rhodin K, Divaris K, North KE, Barros SP, Moss K, Beck JD, Offenbacher S. Chronic periodontitis genome-wide association studies: gene-centric and gene set enrichment analyses. J Dent Res. 2014;93(9):882–90.
6. Offenbacher S, Divaris K, Barros SP, Moss KL, Marchesan JT, Morelli T, Zhang S, Kim S, Sun L, Beck JD, et al. Genome-wide association study of biologically informed periodontal complex traits offers

novel insights into the genetic basis of periodontal disease. Hum Mol Genet. 2016;25(10):2113–29.

7. Offenbacher S, Jiao Y, Kim SJ, Marchesan J, Moss KL, Jing L, Divaris K, Bencharit S, Agler CS, Morelli T, et al. GWAS for Interleukin-1 beta levels in gingival crevicular fluid identifies IL37 variants in periodontal inflammation. Nat Commun. 2018;9(1):3686.
8. Dudding T, Haworth S, Lind PA, Sathirapongsasuti JF, Tung JY, Mitchell R, Colodro-Conde L, Medland SE, Gordon S, Elsworth B, et al. Genome wide analysis for mouth ulcers identifies associations at immune regulatory loci. Nat Commun. 2019;10(1):1052.
9. Morelli T, Agler CS, Divaris K. Genomics of periodontal disease and tooth morbidity. Periodontol 2000. 2020;82(1):143–56.
10. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell. 2019;179(7):1469–82.
11. Nibali L, Di Iorio A, Tu YK, Vieira AR. Host genetics role in the pathogenesis of periodontal disease and caries. J Clin Periodontol. 2017;44(Suppl 18):52–78.
12. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, Nguyen-Viet TA, Wedow R, Zacher M, Furlotte NA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet. 2018;50(2):229–37.
13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
14. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet. 2012;44(4):369–75, S1-S3.
15. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
16. Benner C, Spencer CC, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics. 2016;32(10):1493–501.
17. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, Duncan L, Perry JR, Patterson N, Robinson EB, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236–41.
18. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291–5.
19. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
20. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro F, Clarke D, Gu M, Emani P, Yang YT, et al. Comprehensive functional genomic resource and integrative model for the human brain. Science. 2018; 362(6420).

21. van der Wijst M, Brugge H, de Vries DH, Deelen P, Swertz MA, Franke L. Single-cell RNA sequencing identifies celltype-specific cis-eQTLs and co-expression QTLs. Nat Genet. 2018;50(4):493–7.
22. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G, et al. Impact of genetic polymorphisms on human immune cell gene expression. Cell. 2018;175(6):1701–15.
23. Võsa U, Claringbould A, Westra H, Bonder M, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R, Kasela S. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalyses. bioRxiv. 2018; 447367.
24. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet. 2013;45(10):1238–43.
25. Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204–13.
26. Aguet F, Barbeira A, Bonazzola R, Brown A, Castel S, Jo B, Kasela S, Kim-Hellmuth S, Liang Y, Oliva M. The GTEx Consortium atlas of genetic regulatory effects across human tissues. bioRxiv. 2019; 787903.
27. Watanabe K, Umicevic MM, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. Nat Commun. 2019;10(1):3222.
28. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40(Database issue):D930–4.
29. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–30.
30. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11(4):e1004219.
31. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607–13.
32. Buniello A, MacArthur J, Cerezo M, Harris LW, Hayhurst J, Malangone C, McMahon A, Morales J, Mountjoy E, Sollis E, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005–12.
33. Zhernakova DV, Deelen P, Vermaat M, van Iterson M, van Galen M, Arindrarto W, van T HP, Mei H, van Dijk F, Westra HJ, et al. Identification of context-dependent expression quantitative trait loci in whole blood. Nat Genet. 2017;49(1):139–45.
34. Agler CS, Shungin D, Ferreira ZA, Schmadeke P, Basta PV, Luo J, Cantrell J, Pahel TJ, Meyer BD, Shaffer JR, et al. Protocols, methods, and tools for genome-wide association studies (GWAS) of

dental traits. Methods Mol Biol. 2019; 1922:493–509.

35. Armingohar Z, Jorgensen JJ, Kristoffersen AK, Schenck K, Dembic Z. Polymorphisms in the interleukin-10 gene and chronic periodontitis in patients with atherosclerotic and aortic aneurysmal vascular diseases. J Oral Microbiol. 2015;7:26051.
36. Smith AJ, Humphries SE. Cytokine and cytokine receptor gene polymorphisms and their functionality. Cytokine Growth Factor Rev. 2009;20(1):43–59.
37. Liu Q, Yang J, He H, Yu Y, Lyu J. Associations between interleukin-10 polymorphisms and susceptibility to rheumatoid arthritis: a meta-analysis and meta-regression. Clin Rheumatol. 2018;37(12):3229–37.
38. Tsiara CG, Nikolopoulos GK, Dimou NL, Pantavou KG, Bagos PG, Mensah B, Talias M, Braliou GG, Paraskeva D, Bonovas S, et al. Interleukin gene polymorphisms and susceptibility to HIV-1 infection: a meta-analysis. J Genet. 2018;97(1):235–51.
39. Polvora T, Nobre A, Tirapelli C, Taba MJ, Macedo LD, Santana RC, Pozzetto B, Lourenco AG, Motta A. Relationship between human immunodeficiency virus (HIV-1) infection and chronic periodontitis. Expert Rev Clin Immunol. 2018;14(4):315–27.
40. Ceccarelli F, Saccucci M, Di Carlo G, Lucchetti R, Pilloni A, Pranno N, Luzzi V, Valesini G, Polimeni A. Periodontitis and rheumatoid arthritis: the same inflammatory mediators? Mediators Inflamm. 2019; 2019:6034546.
41. Albuquerque CM, Cortinhas AJ, Morinha FJ, Leitao JC, Viegas CA, Bastos EM. Association of the IL-10 polymorphisms and periodontitis: a meta-analysis. Mol Biol Rep. 2012;39(10):9319–29.
42. Yang SL, Huang SJ. Interleukin-10 polymorphisms (rs1800871, rs1800872 and rs1800896) and periodontitis risk: A meta-analysis. Arch Oral Biol. 2019;97:59–66.
43. Vignali DA, Kuchroo VK. IL-12 family cytokines: immunological playmakers. Nat Immunol. 2012;13(8):722–8.
44. Sousa I, Shahram F, Francisco D, Davatchi F, Abdollahi BS, Ghaderibarmi F, Nadji A, Mojarad SN, Xavier JM, Oliveira SA. Brief report: association of CCR1, KLRC4, IL12A-AS1, STAT4, and ERAP1 With Behcet's disease in Iranians. Arthritis Rheumatol. 2015;67(10):2742–8.
45. Hu KF, Huang KC, Ho YP, Lin YC, Ho KY, Wu YM, Yang YH, Tsai CC. Interleukin-10 (-592 C/A) and interleukin-12B (+ 16974 A/C) gene polymorphisms and the interleukin-10 ATA haplotype are associated with periodontitis in a Taiwanese population. J Periodontal Res. 2009;44(3):378–85.
46. Ozato K, Shin DM, Chang TH, Morse HR. TRIM family proteins and their emerging roles in innate immunity. Nat Rev Immunol. 2008;8(11):849–60.
47. Wang M, Chao C, Luo G, Wang B, Zhan X, Di D, Qian Y, Zhang X. Prognostic significance of TRIM59 for cancer patient survival: A systematic review and meta-analysis. Medicine. 2019;98(48):e18024.
48. Seifert LL, Si C, Saha D, Sadic M, de Vries M, Ballentine S, Briley A, Wang G, Valero-Jimenez AM, Mohamed A, et al. The ETS transcription factor ELF1 regulates a broadly antiviral program distinct from the type I interferon response. PLoS Pathog. 2019;15(11):e1007634.

49. de Groot N, Groen R, Orie V, Bruijnesteijn J, de Groot NG, Doxiadis G, Bontrop RE. Analysis of macaque BTN3A genes and transcripts in the extended MHC: conserved orthologs of human gammadelta T cell modulators. Immunogenetics. 2019;71(8–9):545–59.

Figures

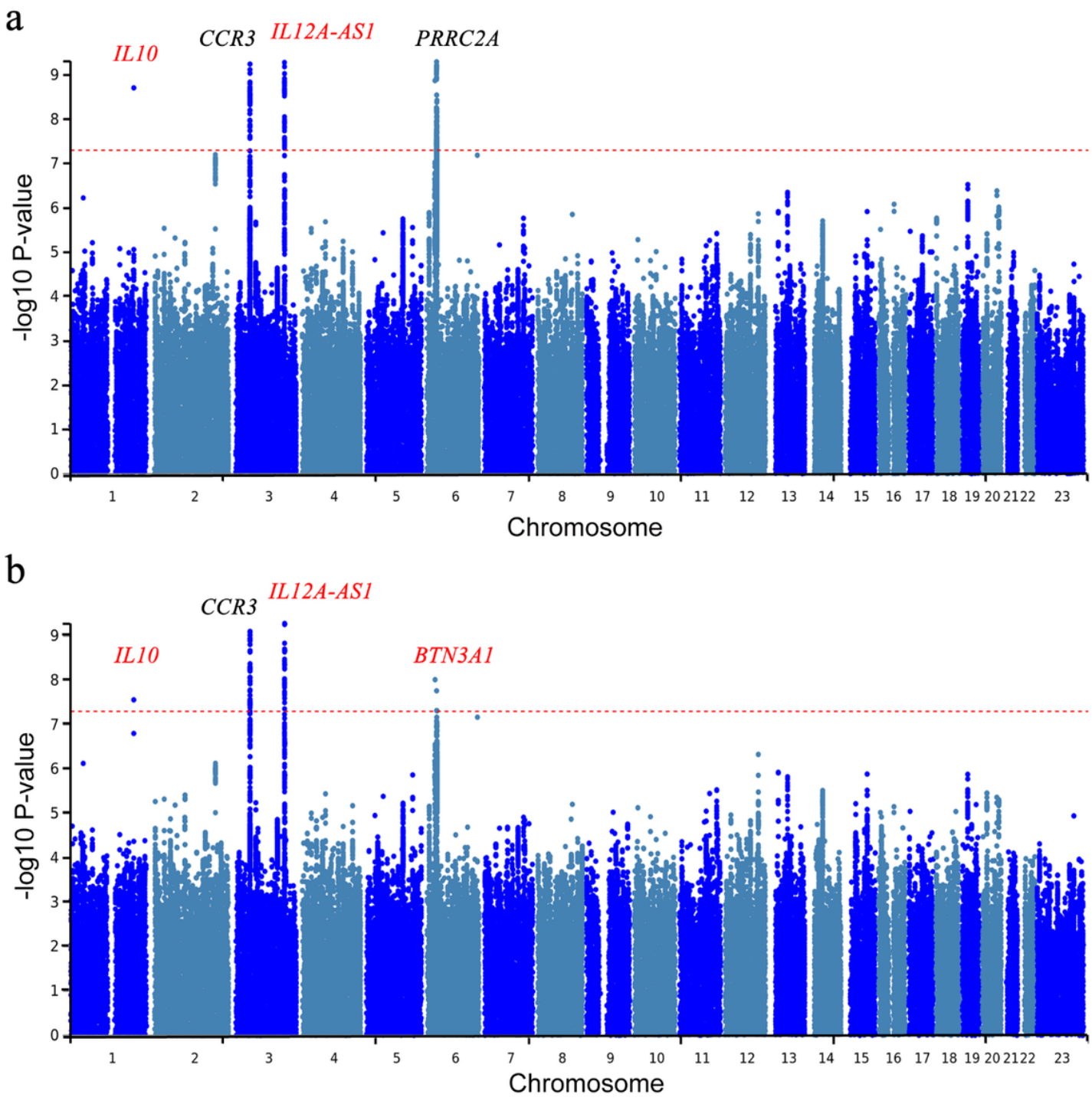

Figure 1

Manhattan plots for MTAG-analysis results of the two oral inflammatory traits. (a) Painful gums (PG) and (b) Toothache (TA). The x-axis shows the genomic position and the y-axis indicates the statistical significance as -log10 (P); the red horizontal line indicates the genome-wide significance (GWS) threshold for P <5×10-8. The y-axis is truncated at 5×10-10 for visualization purposes and the candidate genes for each GWS locus is annotated on the figure.

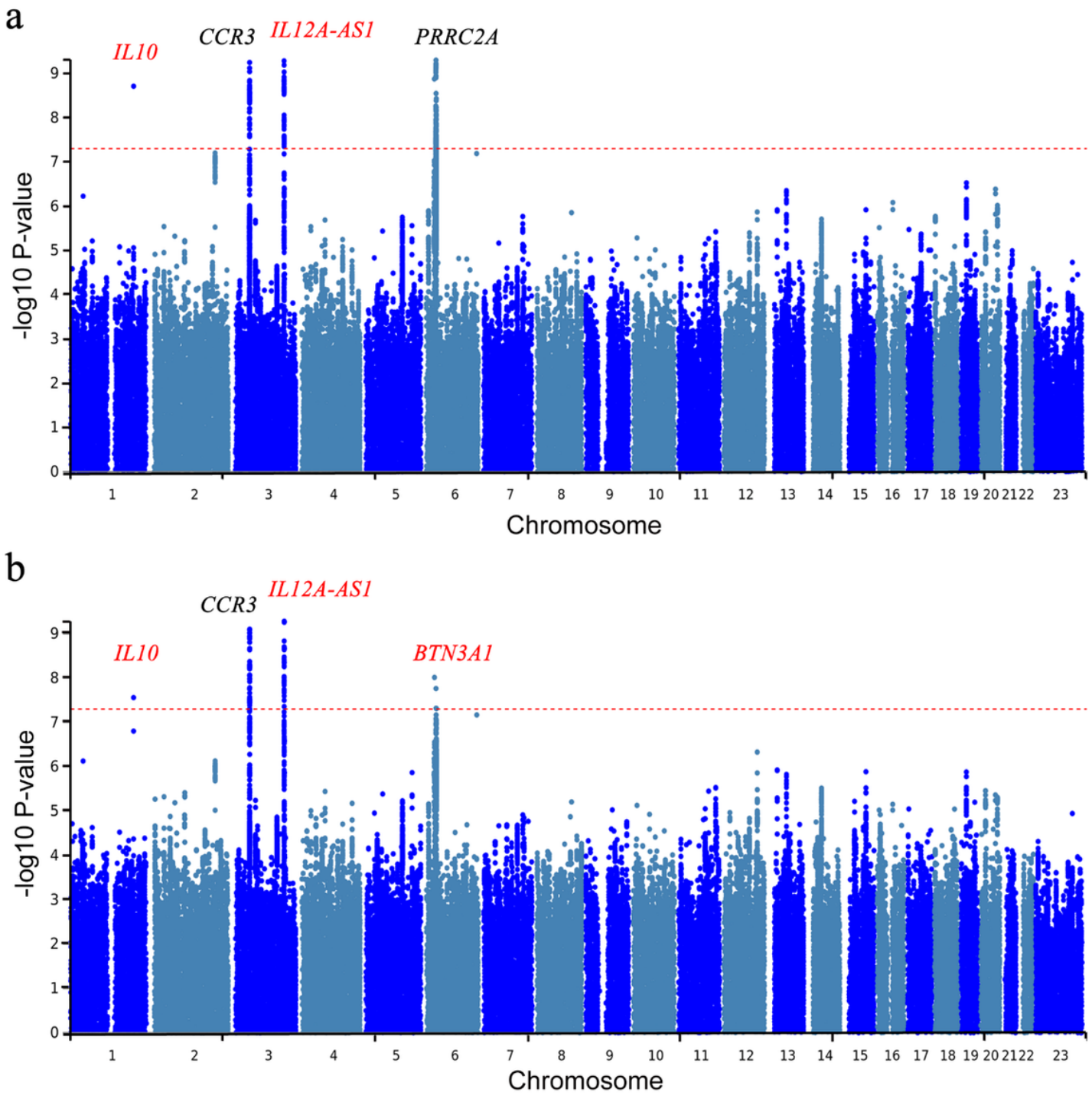

Figure 1

Manhattan plots for MTAG-analysis results of the two oral inflammatory traits. (a) Painful gums (PG) and (b) Toothache (TA). The x-axis shows the genomic position and the y-axis indicates the statistical

significance as -log10 (P); the red horizontal line indicates the genome-wide significance (GWS) threshold for P <5×10-8. The y-axis is truncated at 5×10-10 for visualization purposes and the candidate genes for each GWS locus is annotated on the figure.

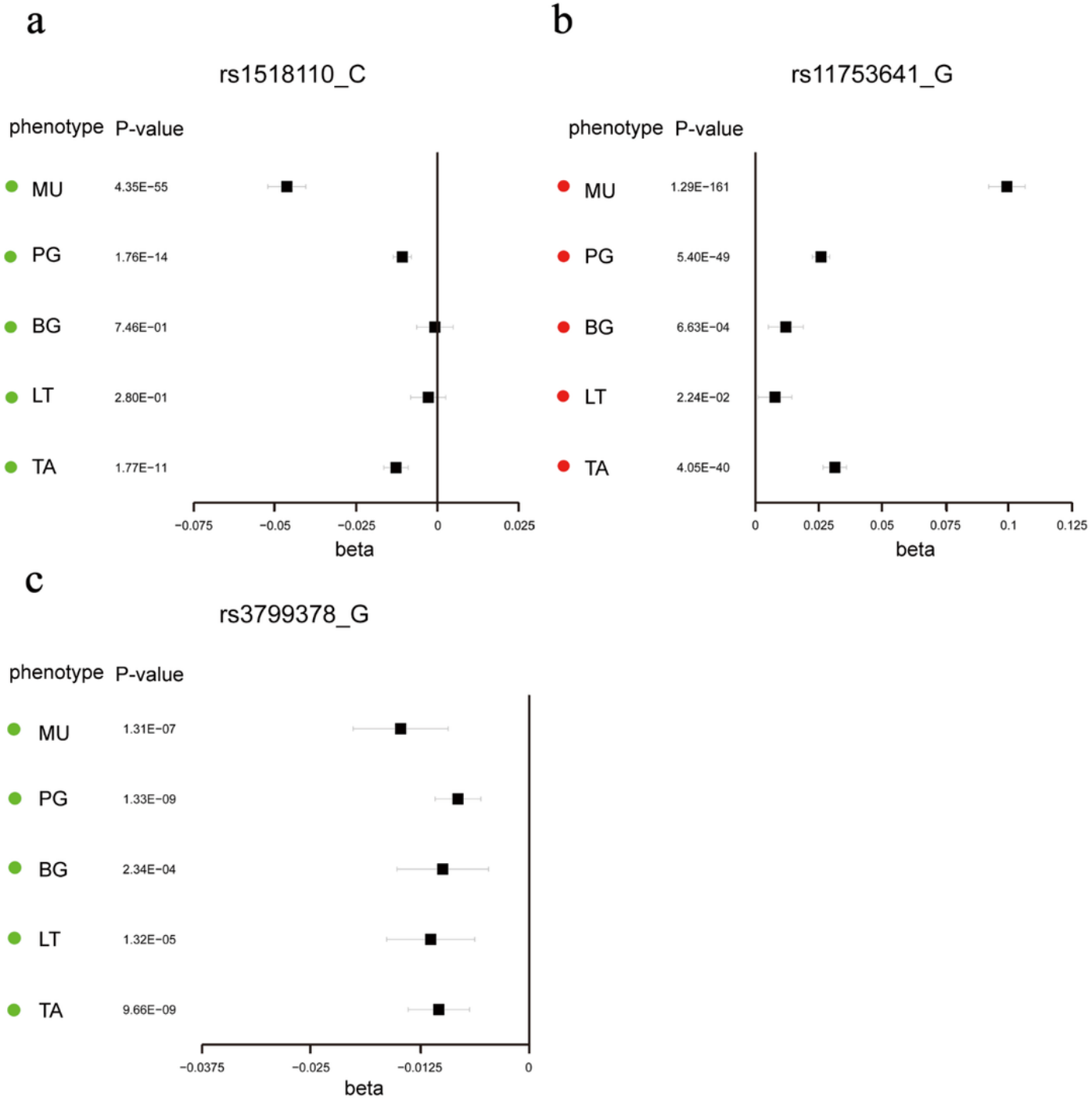

Figure 2

The forest plots of the index SNPs at the novel loci. The effect size, confidence interval, and effect direction of each index SNP at the novel loci are shown. The risk effect is shown in red, and the protective

effect is shown in green. (a) SNP rs1518110; (b) SNP rs17753641; and (c) SNP rs3799378.

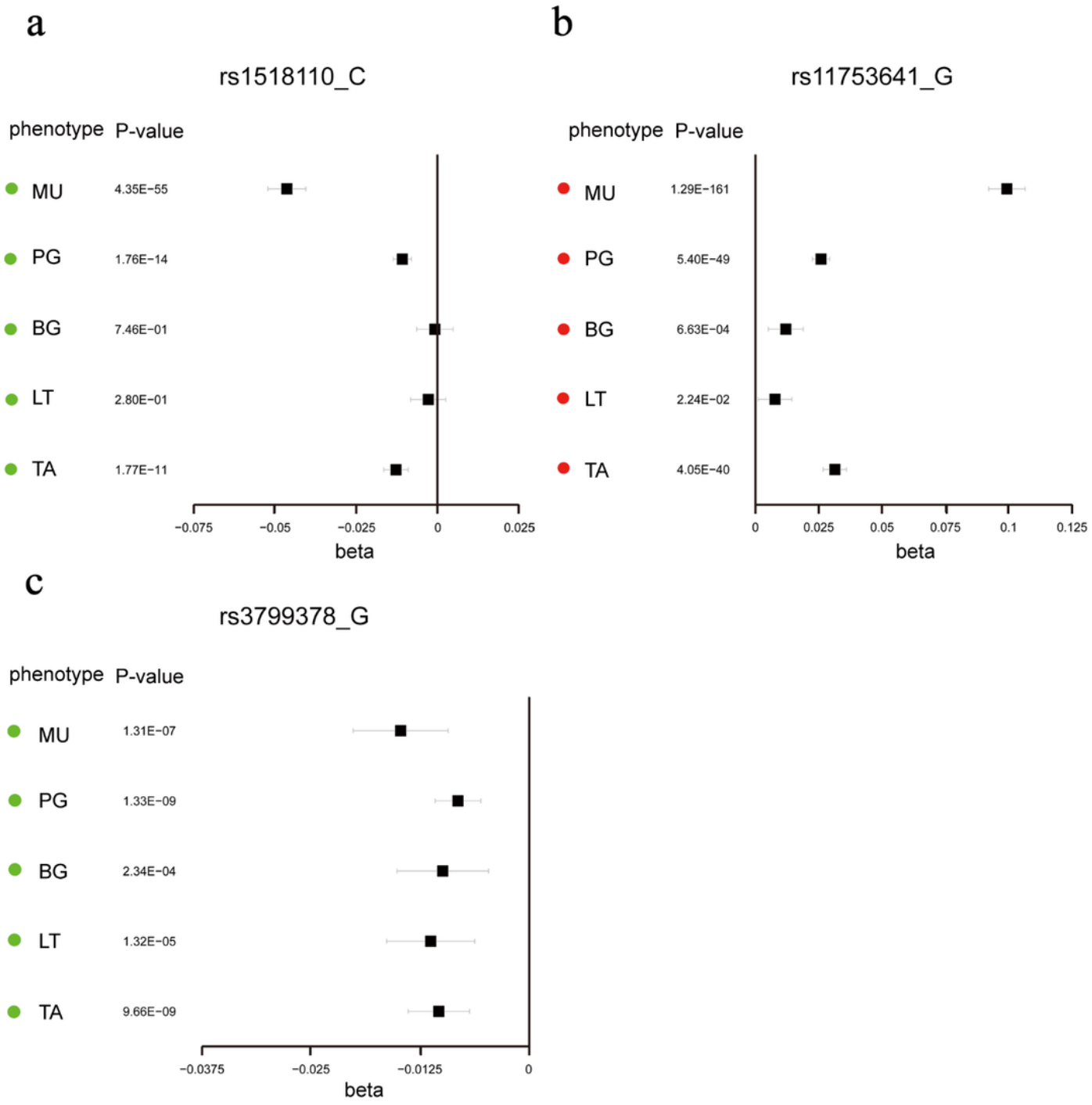

Figure 2

The forest plots of the index SNPs at the novel loci. The effect size, confidence interval, and effect direction of each index SNP at the novel loci are shown. The risk effect is shown in red, and the protective effect is shown in green. (a) SNP rs1518110; (b) SNP rs17753641; and (c) SNP rs3799378.

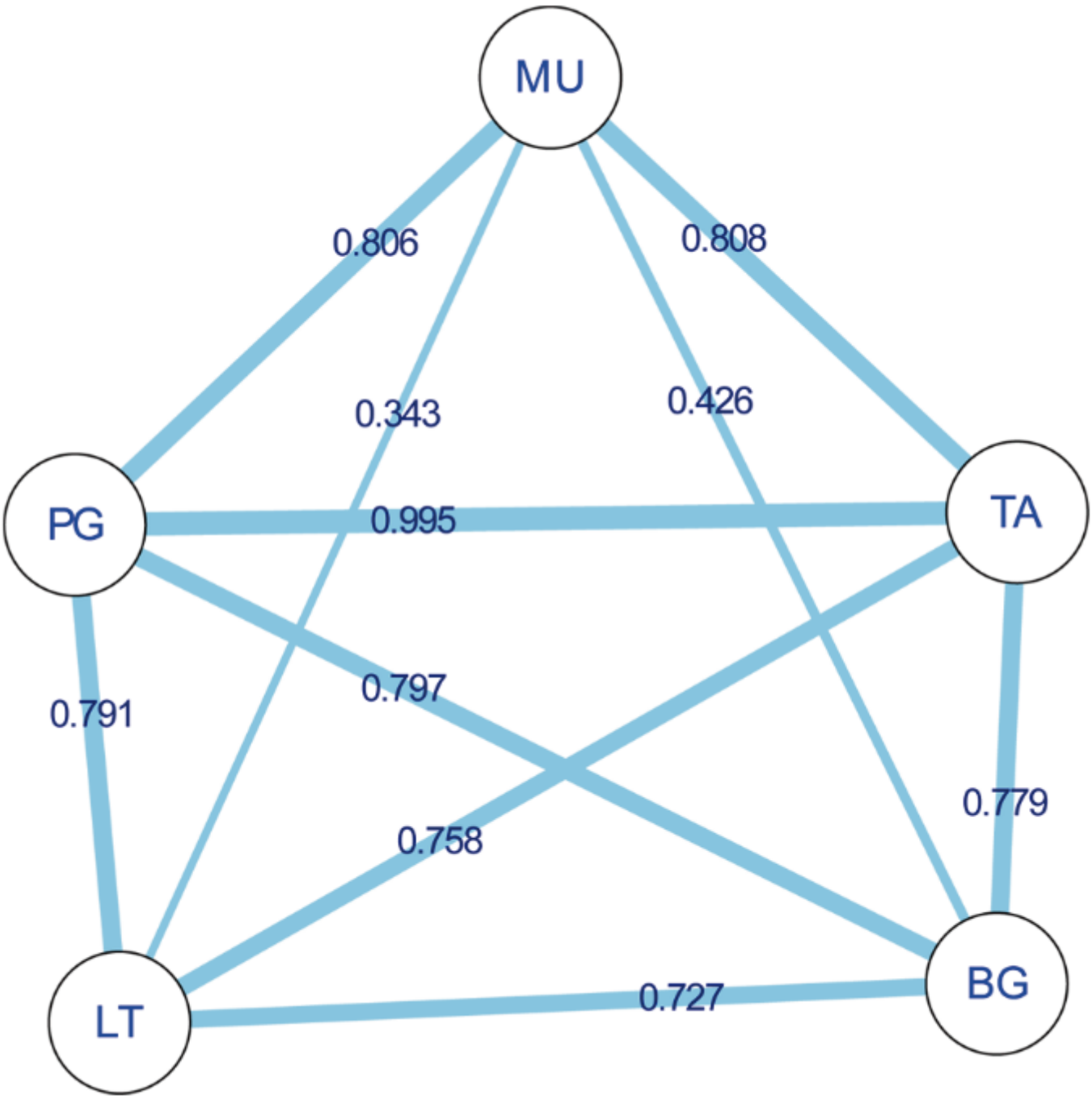

Figure 3

SNP-based genetic correlations between the five oral inflammatory traits. Each node represents a trait with edges indicating the strength of the pairwise correlations. The width of the edges is proportional to the strength of rg.

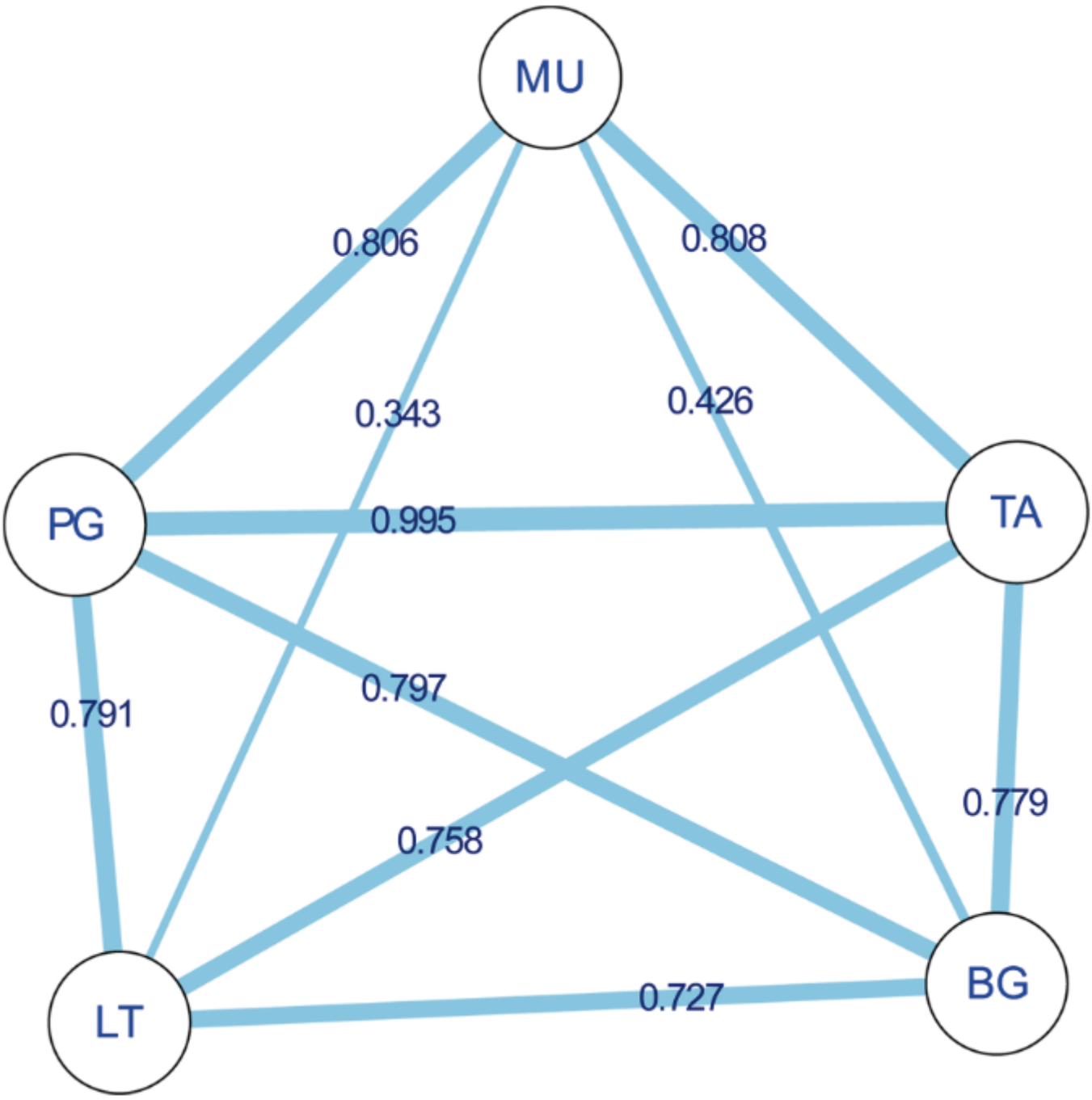

Figure 3

SNP-based genetic correlations between the five oral inflammatory traits. Each node represents a trait with edges indicating the strength of the pairwise correlations. The width of the edges is proportional to the strength of rg.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.docx
- AdditionalFile1.docx
- AdditionalFile2.docx
- AdditionalFile2.docx
- AdditionalFile3.xlsx
- AdditionalFile3.xlsx
- AdditionalFile4.docx
- AdditionalFile4.docx
- AdditionalFile5.docx
- AdditionalFile5.docx
- AdditionalFile6.xlsx
- AdditionalFile6.xlsx
- AdditionalFile7.docx
- AdditionalFile7.docx
- AdditionalFile8.xlsx
- AdditionalFile8.xlsx
- AdditionalFile9.xlsx
- AdditionalFile9.xlsx
- AdditionalFile10.xlsx
- AdditionalFile10.xlsx
- AdditionalFile11.docx
- AdditionalFile11.docx
- AdditionalFile12.xlsx
- AdditionalFile12.xlsx
- AdditionalFile13.xlsx
- AdditionalFile13.xlsx
- AdditionalFile14.docx
- AdditionalFile14.docx
- AdditionalFile15.xlsx
- AdditionalFile15.xlsx
- AdditionalFile16.docx
- AdditionalFile16.docx