Supplementary Materials

for "Male-specific association reveals contribution of MRAS to painful temporomandibular disorder"

Supplementary Methods

Replication Cohorts

1) The Study of Health in Pomerania (SHIP)[2] is a cross-sectional study of a representative sample of the population of Pomerania, Germany. Subjects were aged 20-81 years, and included 51% females. Subjects reported symptoms by questionnaire regarding pain in the TMJ and facial muscles to assess presence and frequency of pain. During a clinical exam, the examiner inquired about pain or discomfort upon palpation of masticatory tissues, including TM joints (dorsocranial and lateral) palpated at 2 kg/cm^2, and masseter, temporalis, and medial pterygoid palpated at 1 kg/cm^2. Pain or discomfort during jaw movement, range of motion, and joint sounds were also assessed. TMD was defined as pain or discomfort during examination procedures in at least one muscle or TM joint. Subjects with both phenotype and genotype data totaled 3,651, including 607 cases (17%) and 3,044 controls.

2) The Northern Finland Birth Cohort (NFBC)[10] is a cohort study of all births during 1966 in the Oulu and Lapland provinces of northern Finland. An assessment for TMD was performed at the 46-year follow-up time-point. Subjects (52% female) reported symptoms by questionnaire using the following questions: a) "Do you experience temple, temporomandibular joint, face, or jaw pain once a week or more often?" b) "Do you experience pain once a week or more often while opening your mouth wide?" A clinical exam determined the presence of examiner-evoked pain in three or more temporomandibular muscles and/or joints. Palpation sites included the temporalis (1 kg force), masseter (1 kg force), lateral TMJs (0.5 kg force), and TMJs around the pole (1 kg force). For each palpation, participants reported the presence of pain as yes or no. Of 1,940 subjects that completed the TMD questionnaire and examination, 161 (11%) had TMD defined as positive responses to both questionnaire items and pain in response to one or more palpations.

3) The Brazilian case-control study (SPB) enrolled a community-based sample of females between the age of 18-44 years living in Piracicaba, São Paulo, Brazil. Pain history was determined by asking if subjects had pain in the head, face, jaw, or in front of the ears in the last 3 months. A study examiner manually palpated lateral and posterior aspects of both TM joints (0.45 kg) and subjects reported whether or not each palpation evoked pain. Evoked pain was also assessed following unassisted jaw opening, maximum assisted jaw opening, and right and left excursion and protrusion of the jaw. TMD case classification was determined by pain for at least 3 months and pain on examination in at least one TM joint. Of 636 subjects, 144 (22%) were classified as TMD cases and 492 were controls.

4) The OPPERA II Chronic TMD Replication case-control study was designed to confirm findings from the initial OPPPERA GWAS in an independent cohort. Potential subjects were recruited by telephone screening of 166,062 phone numbers listed in counties surrounding the four OPPERA recruitment sites. Classification of cases and controls were based on symptoms reported in a telephone interview, using responses to a brief screening questionnaire with high sensitivity and specificity for correctly classifying examiner-verified TMD status.[5] Subjects were classified as chronic facial pain cases if they reported pain in one or more craniofacial locations (excluding toothache or ear infection) that occurred for 5 days or more per month for at least six of the preceding 12 months. Controls reported no such facial pain in the preceding six months; controls also did not endorse the use of an occlusal splint or previous TMD diagnosis. Saliva samples for DNA genotyping were obtained using Oragene collection tubes (DNA Genotek Inc., Kanata, Ontario) mailed to the OPPERA Data Coordinating Center. DNA from all subjects was genotyped using the Affymetrix Axiom Precision Medicine Research Array by McGill University and Génome Québec Innovation Centre. Of 2,430 eligible subjects who completed the screening interview, 1,342 subjects (66% female, age 18-74) returned complete phenotype and genotype information and were included in the replication analysis, including 444 chronic TMD cases and 898 TMD-free controls.

5) The Complex Persistent Pain Conditions (CPPC): Unique and Shared Pathways of Vulnerability study included participants enrolled in a case control study of overlapping pain conditions conducted at UNC Chapel Hill. Subjects were aged 18-64, and included both sexes (86% female) and major ethnic and racial groups (68% non-Hispanic white). All subjects had at least one of four index CPPCs (episodic migraine, irritable bowel syndrome, fibromyalgia, or vulvar vestibulitis), or were otherwise healthy controls with none of these conditions. TMD status was determined by RDC exam,[4] using the same protocols as the OPPERA study.[9] Subjects with a positive diagnosis of TMD (regardless of any other painful comorbidities) were included as cases in the replication analysis (n = 266), while controls were subjects without TMD and without any of the index CPPCs (n = 224). DNA from all subjects was genotyped using the Affymetrix Axiom Precision Medicine Research Array by McGill University and Génome Québec Innovation Centre.

6) The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) is a population based cohort of 16,415 participants, aged 18 to 74 years, enrolled between 2008 and 2011. Subjects were self-identified Hispanic/Latino participants recruited at four U.S. community sites (the Bronx, New York; San Diego, California; Miami, Florida; and Chicago, Illinois). Details of the design, recruitment, and implementation of HCHS/SOL have been published.[8; 13] The Oral Health Questionnaire asked two questions about pain currently or in the past 12 months in a) the face and b) the jaw. To be classified as an orofacial pain case for this analysis, participants had to report pain in either the face OR in the jaw; controls reported no such pain. Participants were genotyped on the HCHS Custom 15041502 B3 array (Illumina Omni2.5M+custom content). Quality control was conducted as previously described;[3; 7] genotypes were cleaned and imputed at the

University of Washington following the same procedures as described for the OPPERA genotyping.

7) The UK Biobank[1] (UKB) dataset used in this analysis came from 503,325 residents of the United Kingdom recruited between 2006-2010. They were registered with the National Health Service, aged between 40-69 years and living less than 25 miles from a study center. At their initial assessment visit, subjects gave informed consent and completed questionnaires that included the question: "Have you had facial pains for more than 3 months?" with response options 'yes', 'no', 'do not know' and 'prefer not to answer'. Cases (n = 1,320) were defined as subjects who answered 'yes' for the facial pain question; controls (n = 135,853) were defined as all others, excluding those that reported having headaches for more than 3 months in response to a separate question. Genotyping was performed using the UK Biobank Axiom Array. This analysis was limited to available genotypes at the interim release.

Supplementary Table 1: Racial/ethnic composition of the OPPERA GWAS discovery cohort

Ethnicity	Number of unrelated study subjects
White	1,859
Black/African-American	701
Asian	223
Hispanic	159
Other/Multiple	47
Not Stated	29
Native American	8
Pacific Islander	4
Total	**3,030**

Supplementary Table 2: SNP quality filters

	Filter	SNPs lost	SNPs kept
1	none - all SNP probes		2,567,845
2	genotype scanning technical filters (CIDR)	13,713	2,554,132
3	missing rate >=2%	17,838	2,536,294
4	>1 discordant calls in 69 study duplicates	614	2,535,680
5	>1 Mendelian error in 9 HapMap trios	278	2,535,402
6	HWE p-value $<10^{-4}$ in either European or African subjects	3,790	2,531,612
7	sex difference in allele freq >=0.2 for autosomes/XY in white or black subjects	489	2,531,123
8	sex difference in heterozygosity >0.3 for autosomes/XY in white or black subjects	151	2,530,972
9	positional duplicates	42,460	2,488,512
10	MAF = 0	151,434	2,337,078
11	MAF <0.01	505,612	1,831,466
12	quality filter (rows 2 - 8)	1.44 %	
13	composite filter (quality rows 2 - 8 and informativeness rows 9-10)	8.99 %	
14	filter for quality and informativeness and MAF (rows 2-11)	28.68 %	

NOTE: The MAF < 0.01 entry here is for illustration only. The final MAF filtering is based on MAF and sample size considerations

Supplementary Table 3: Expected number of true discoveries among the first 75 top results

True total number of loci	Sample size: 1000 cases + 1000 controls	Sample size: 1,082 cases + 2,144 controls	Combined sample
A. for minor allele frequency 0.35			
25	2.9	6.6	16.3
50	5.7	12.8	31.3
75	8.5	18.6	44.6
100	11.1	24.0	55.6
125	13.7	29.1	63.6
150	16.1	33.8	68.8
B. for minor allele frequency 0.15			
25	0.8	2.0	6.8
50	1.6	4.0	13.2
75	2.4	5.9	19.3
100	3.2	7.8	24.9
125	4.0	9.6	30.1
150	4.7	11.4	34.9

Supplementary Table 3 shows the number of true associations expected to be discovered among the top 75 GWAS hits, i.e. 75 SNPs with the smallest p-values as a function of the total number of susceptibility loci covered by the GWAS. The numbers of expected discoveries are given for three study sizes: (1) a hypothetical reference sample of 1000 cases and 1000 controls; (2) our discovery sample of 1,082 cases and 2,144 controls; (3) a proposed meta-analysis of the discovery sample with an additional 1,000 cases and 1,000 controls from a hypothetical replication cohort. Assuming the number of truly associated SNPs to be 25, MAF=0.15, and sample sizes of the discovery cohort (1,082 cases and 2,144 controls), there is 80% chance that one or more signals in the discovery GWAS will score among the top 41 observed associations when they are ordered by p-value. For 50 such loci, the number of top results decreases to 10. For 75 such loci, the number of top results is 5.

Power calculations were carried out by the method of Kuo and Zaykin.[6], evaluating a range of numbers for the total number of susceptibility loci. This procedure was informative because of the multiple-testing context of GWA studies. Just as the family-wise error rates are commonly defined in terms of the probability of making "at least one" (or "at least K") incorrect rejections, power should be similarly defined in terms of correctly rejecting "at least K" hypotheses of no association. In studies that involve multiple tests of significance, such as GWA studies, power to detect one or more truly associated variants increases with the total number of associated variants. Thus, the total number of tests, as well as the number of truly associated loci should affect power calculations.

Statistical power was assessed assuming a homogeneous sample; however, stratified analyses to account for existing racial heterogeneity may cause a reduction in power. While we acknowledge this possibility, the effect of adjusting for additional covariates in models with a binary outcome can result in either decreased or increased power.[15] In case-control studies, a substantial reduction in power due to adjustment for covariates is expected for rare diseases, while for diseases with high prevalence, power may increase instead.[11] Furthermore, in terms of analytical tools, we chose the PCA method to adjust for population stratification, as it has been shown that that PCA-based methods are superior to other comparable methods.[14]

Supplementary Table 4: Race-stratified and mixed model association analyses

Because the OPPERA study was intentionally designed to recruit a diverse cohort, we considered the possibility that population stratification affected association findings, even after accounting for ancestry in the regression models. For SNPs with a significant association in the regression models, we subdivided the sample into homogeneous racial groups representing European- and African-American ancestry (Supplementary Table 4a). For rs5862730, which has comparable minor allele frequencies in both ancestry groups, we observed associations with TMD that were consistent in magnitude and direction in both subgroups. For rs10092633, the association among African-American females was in the same direction and of greater magnitude (OR=3.8) than the overall model. (As noted previously, rs10092633 was not observed among European-American females.) The trio of chromosome 3 SNPs associated with TMD in males showed strong association in both racial strata (OR > 2.2 in European-Americans, OR > 3.8 in African-Americans), although with much lower MAF in African-Americans (MAF = 0.02). Further analysis using a univariate logistic mixed model approach that incorporates the relatedness matrix to account for population and sample stratification also supported rs5862730 and the SNPs on chromosome 3 (male-specific) as significant, but did not show association for rs10092633 (Supplementary Table 4b).

					European-American only				African-American only				Meta-analysis			
Group	rsID	Chr	Position	EA	EAF	OR	SE	P	EAF	OR	SE	P	OR	P	Q	I
all	rs5862730	4	146211844	D	0.30	1.37	0.08	2.00E-04	0.32	1.69	0.15	5.20E-04	1.44	7.39E-07	0.22	32.67
fem	rs5862730	4	146211844	D	0.30	1.34	0.97	2.60E-03	0.32	2.26	0.20	4.27E-05	2.21	4.72E-05	0.60	0
fem	rs10092633	8	41123732	A	0.00	NA	NA	NA	0.14	3.80	0.29	3.76E-06	NA	NA	NA	NA
male	rs34612513	3	137541085	A	0.11	2.18	0.25	1.60E-03	0.02	5.32	0.71	1.80E-02	2.40	1.70E-04	0.23	29.33
male	rs28865059	3	137687399	C	0.12	2.26	0.24	7.70E-04	0.02	3.83	0.71	5.80E-02	2.39	1.50E-04	0.48	0
male	rs13078961	3	137687685	C	0.12	2.26	0.24	7.80E-04	0.02	3.83	0.71	5.80E-02	2.39	1.50E-04	0.48	0

Supplementary Table 4a Race-stratified association and meta-analysis. All SNPs exceeding the threshold for statistical significance in either the full cohort or sex-stratified analyses were examined for association within racial strata. A fixed-effects meta-analysis was used to combine stratum-specific effects. Race clusters corresponding to European-American and African-American ancestry were identified by principal components analysis to select homogeneous groups for race-stratified association tests. European-only analyses included 1627 subjects, of which 1129 were female and 498 were male. African-American-only analyses included 668 subjects, of which 367 were female and 301 were male. Subjects that did not cluster with these two ancestry groups (n = 735) were not included in race-stratified analyses. Totals differ from Supplementary Table 1 due to discrepancies between self-reported race and ancestry identified by PCA. rsID: SNP name; Chr: chromosome; EA: effect allele; EAF: effect allele frequency; OR: odds ratio; SE: standard error; P: *P*-value; D: deletion allele of an insertion/deletion polymorphism; Q: p-value for Cochrane's Q statistic; I: I^2 heterogeneity index. Chromosome and position are from GRCh37/hg19 (build 37).

Group	rsID	Chr	Position	EA	EAF	N	P	beta	se	l_remle	p_wald
all	rs5862730	4	146211844	D	0.33	3030	2.82×10^{-8}	-0.07	0.01	0.07	5.06×10^{-8}
fem	rs5862730	4	146211844	D	0.33	1956	1.70×10^{-8}	-0.09	0.02	0.11	1.90×10^{-8}
fem	rs10092633	8	41123732	A	0.03	1956	2.91×10^{-8}	0.61	0.33	0.08	6.57×10^{-2}
male	rs34612513	3	137541085	A	0.08	1074	1.49×10^{-8}	-0.20	0.03	0.04	4.35×10^{-10}
male	rs28865059	3	137687399	C	0.09	1074	2.21×10^{-8}	-0.19	0.03	0.04	8.42×10^{-10}
male	rs13078961	3	137687685	C	0.09	1074	2.22×10^{-8}	-0.19	0.03	0.04	8.44×10^{-10}

Supplementary Table 4b Univariate logistic mixed model analysis. All SNPs exceeding the threshold for statistical significance in either the full cohort or sex-stratified analyses were examined for association using a univariate mixed model approach using the GEMMA[16] software package. This method incorporates the relatedness matrix among study individuals to account for the population stratification and sample structure in the mixed-race OPPERA cohort. rsID: SNP name; Chr: chromosome; EA: effect allele; EAF: effect allele frequency; P: SNP *P*-value from discovery logistic regression; D: deletion allele of an insertion/deletion polymorphism; beta: beta estimate from the ulm model; se: standard error for the beta estimate; l_remle: remle estimate for lambda; p_wald: *P*-value from Wald test. Chromosome and position are from GRCh37/hg19 (build 37).

Supplementary Table 5. Significant DRG *cis*-eQTLs among SNPs associated with TMD.

SNP	Gene Name	Probe Name	Probe Type	Stratum	Beta	P-value	FDR
a)							
rs34612513	*MRAS*	PSR03013533	Exon 9	M+F	-0.66	1.06×10^{-6}	7.63×10^{-4}
rs13078961	*MRAS*	PSR03013533	Exon 9	M+F	-0.51	2.43×10^{-5}	5.83×10^{-3}
rs28865059	*MRAS*	PSR03013533	Exon 9	M+F	-0.51	2.43×10^{-5}	5.83×10^{-3}
b)							
rs13078961	*MRAS*	PSR03013527	Exon 8	M	-0.55	2.25×10^{-4}	4.73×10^{-3}
rs13078961	*MRAS*	PSR03013533	Exon 9	M	-0.56	8.68×10^{-4}	5.35×10^{-3}
rs13078961	*MRAS*	PSR03013523	Exon 5	M	-0.50	9.41×10^{-4}	5.35×10^{-3}
rs13078961	*MRAS*	PSR03013525	Exon 7	M	-0.48	1.02×10^{-3}	5.35×10^{-3}
rs13078961	*MRAS*	PSR03013534	Exon 9	M	-0.36	3.99×10^{-3}	1.65×10^{-2}
rs13078961	*MRAS*	PSR03013532	Exon 9	M	-0.40	4.71×10^{-3}	1.65×10^{-2}
rs13078961	*MRAS*	PSR03013524	Exon 6	M	-0.35	6.98×10^{-3}	2.09×10^{-2}

c) [accompanying XLS file]

SNPs significantly associated with TMD were evaluated for association with gene expression levels in human DRG tissue. **a)** Five selected TMD SNPs were tested against all *cis*-acting gene-level and exon-level probes. Separate analyses were performed on population strata including all subjects, males only, and females only. Correction for multiple testing was performed using the Benjamini-Hochberg (FDR) procedure, and statistical significance was set at 1%, i.e. FDR $P < 1 \times 10^{-2}$. eQTL discovery and FDR values were obtained from matrix_eQTL.[12] Significant results are presented. **b)** Replicated TMD SNP rs13078961 were tested against all *cis*-acting gene-level and exon-level probes of the *MRAS* gene in males and females separately. Here, since exon-level probe mRNA levels of the same gene are highly correlated, we assess statistical significance using FDR 10%. Only significant results are presented. **c)** Association between five selected TMD SNPs and *trans*-acting gene expression levels in human DRGs.

Supplementary Table 6. Gene Ontology (GO) biological pathway analyses by Ingenuity for *trans*-acting eQTLs in human DRGs. A total of 1,005 eQTL identifiers associated with male-specific TMD SNPs at FDR level 1% were mapped to Ingenuity ® Pathway Analysis (IPA) entities corresponding to a total of 503 genes/proteins. The IPA Core Analysis procedure identified "Inflammatory Response" and "Immunological Disease" (*P* value ranges: 7.63×10^{-6} - 2.21×10^{-29} and 8.98×10^{-6} - 5.44×10^{-29}, respectively) as the two top categories among diseases and biofunctions. The top canonical pathways included "Primary Immunodeficiency Signaling" and "B Cell/T cell Receptor Signaling" (*P* = 1.18×10^{-16} and 4.80×10^{-12}, respectively). The top upstream regulators included the TCR member CD3 and the myeloid-related transcription factor SPI-1 and as well as major cytokines IL4, IL10, and IL2 (*P* value range for all: 1.98×10^{-10} - 2.04×10^{-17}).

Top Canonical Pathways

Name	p-value	Overlap
Primary Immunodeficiency Signaling	1.18×10^{-16}	33.3% 16/48
B Cell Receptor Signaling	4.87×10^{-12}	11.6% 22/190
P13K Signaling in B Lymphocytes	2.80×10^{-10}	12.8% 17/133
T Cell Receptor Signaling	1.42×10^{-7}	11.4% 13/114
p70S6K Signaling	1.87×10^{-7}	10.2% 14/137

Top Upstream Regulators

Upstream Regulator	p-value overlap	Predicted Activation
IL4	2.04×10^{-17}	
lipopolysaccharide	4.92×10^{-15}	
IL2	9.98×10^{-13}	
SPI1	5.79×10^{-12}	
CD3	1.98×10^{-10}	

Top Diseases and Biological Functions

Diseases and Disorders

Name	p-value	#Molecules
Inflammatory Response	7.63×10^{-6} - 2.21×10^{-29}	154
Immunological Disease	8.98×10^{-6} - 5.44×10^{-29}	185
Hematological Disease	8.98×10^{-6} - 2.70×10^{-19}	140
Cancer	8.98×10^{-6} - 3.41×10^{-17}	373
Organismal Injury and Abnormalities	8.98×10^{-6} - 3.41×10^{-17}	373

Supplementary Table 7. Gene Ontology (GO) biological pathway analyses by Pathway Studio for *trans*-acting eQTL genes in human DRGs. Pathway analyses were performed using Pathway Studio®, reporting GO biological process pathways. Similarly to Ingenuity, pathway analyses using Pathway Studio on the eGenes associated with male-specific TMD SNPs at FDR level 1% at gene- (N=610) and exon-level (N=895) revealed multiple significant pathways related to the immune system; B cell receptor signaling pathway (GO:0050853, $P = 6.60 \times 10^{-17}$ and 8.89×10^{-18}) and immune response (GO:0006955, $P = 1.15 \times 10^{-13}$ and 6.53×10^{-21}) were of highest significance. Statistical significance for a pathway was assessed by assuming an experiment-wide alpha = 0.05, an estimated number of unrelated GO pathways N = 10,000, and a Bonferroni correction for multiple testing of $0.05/10{,}000 = 5 \times 10^{-6}$. NS: not significant after correction for multiple testing.

GO Biological Process Pathway	Gene-level P value	Exon-level P value
B cell receptor signaling pathway	6.60×10^{-17}	8.89×10^{-18}
immune response	1.15×10^{-13}	6.53×10^{-21}
innate immune response	2.40×10^{-13}	5.70×10^{-15}
Fc-epsilon receptor signaling pathway	2.57×10^{-13}	9.42×10^{-14}
immune system process	5.28×10^{-13}	5.85×10^{-13}
positive regulation of GTPase activity	2.80×10^{-12}	2.72×10^{-15}
T cell costimulation	1.35×10^{-11}	2.79×10^{-12}
Fc-gamma receptor signaling pathway involved in phagocytosis	2.64×10^{-11}	1.58×10^{-11}
intracellular signal transduction	2.01×10^{-10}	3.85×10^{-16}
regulation of immune response	1.23×10^{-9}	3.87×10^{-11}
phosphorylation	4.88×10^{-9}	1.18×10^{-6}
T cell receptor signaling pathway	1.14×10^{-8}	2.23×10^{-16}
regulation of small GTPase mediated signal transduction	2.60×10^{-8}	5.49×10^{-9}
positive regulation of B cell differentiation	2.68×10^{-8}	NS
protein autophosphorylation	4.41×10^{-8}	NS
peptide antigen transport	1.42×10^{-7}	NS
protein phosphorylation	1.75×10^{-7}	NS
complement activation, classical pathway	1.90×10^{-7}	5.31×10^{-9}

receptor-mediated endocytosis	2.99×10^{-7}	NSS
positive regulation of T cell proliferation	3.08×10^{-7}	1.06×10^{-8}
complement activation	5.48×10^{-7}	7.33×10^{-8}
small GTPase mediated signal transduction	5.65×10^{-7}	9.12×10^{-7}
platelet activation	1.11×10^{-6}	NSS
signal transduction by protein phosphorylation	1.29×10^{-6}	3.77×10^{-6}
cell surface receptor signaling pathway	1.86×10^{-6}	1.85×10^{-10}
blood coagulation	2.16×10^{-6}	NS
T cell differentiation	2.65×10^{-6}	1.42×10^{-10}
positive regulation of T cell receptor signaling pathway	4.75×10^{-6}	NS
positive regulation of gamma-delta T cell differentiation	4.75×10^{-6}	7.63×10^{-7}
B cell activation	NS	9.38×10^{-9}
T cell activation	NS	1.78×10^{-10}
axon guidance	NS	1.95×10^{-9}
cell adhesion	NS	6.13×10^{-14}
immunological synapse formation	NS	4.33×10^{-6}
negative thymic T cell selection	NS	2.82×10^{-8}
peptidyl-tyrosine autophosphorylation	NS	6.32×10^{-7}
positive regulation of neutrophil chemotaxis	NS	3.19×10^{-6}
positive regulation of signal transduction	NS	8.02×10^{-8}
signal transduction	NS	4.81×10^{-9}
transmembrane receptor protein tyrosine kinase signaling pathway	NS	2.18×10^{-7}

Supplementary Table 8. Replication of SNP rs13078961 as an eQTL for *MRAS* in human blood. The eQTL analysis evaluated association between rs13078961 genotype and *MRAS* mRNA level measured by gene- and exon-level probes. Statistical significance defined as $P < 0.05$. Only statistically significant associations are presented.

Gene Name	Probe Name	Probe Site	Stratum	Beta	*P* value
MRAS	PSR03013534	Exon 9	M+F	-0.21	3.76×10^{-2}
MRAS	PSR03013534	Exon 9	M	-0.39	2.17×10^{-2}
MRAS	PSR03013522	Exon 5	M	-0.38	2.62×10^{-2}

Supplementary Table 9. Association of *MRAS* mRNA expression in human blood with TMD caseness. Logistic regression was used to model the effect of *MRAS* mRNA level, measured by gene- and exon-specific probes, on TMD status. Statistical significance was defined as $P < 0.05$ indicating difference in expression levels by TMD status. Only significant associations are presented.

| Gene Name	Probe Name	Probe Type	Stratum	Average Level	Beta	P-value
MRAS	TC03000744	gene	M+F	5.27	-0.04	4.75×10^{-2}
MRAS	PSR03013530	Exon 9	M	9.79	-0.27	1.70×10^{-2}

Supplementary Figures

Supplementary Figure 1: *P*-value distributions for the test of association with TMD. a. Quantile-quantile (QQ) plot showing only genotyped SNPs passing quality control filters. **b.** QQ plot showing imputed SNPs passing filters.

a.
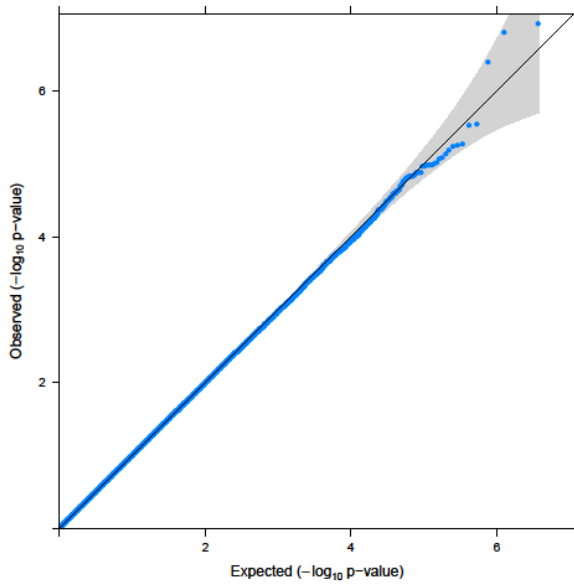
b.
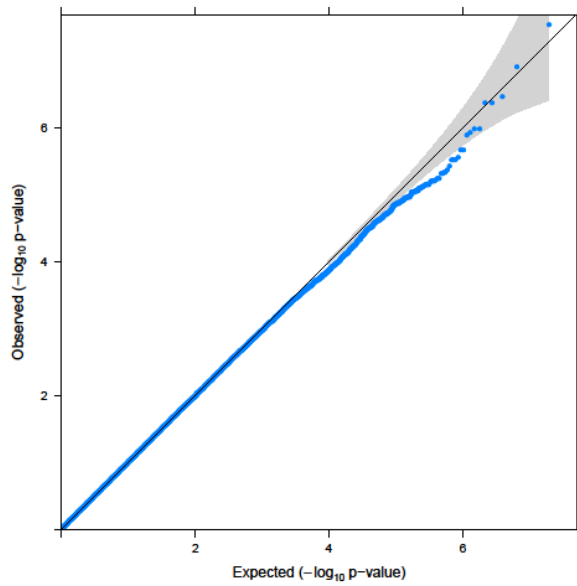
References

[1] Allen NE, Sudlow C, Peakman T, Collins R. UK biobank data: come and get it. Sci Transl Med 2014;6(224):224ed224.

[2] Bernhardt O, Gesch D, Schwahn C, Bitter K, Mundt T, Mack F, Kocher T, Meyer G, Hensel E, John U. Signs of temporomandibular disorders in tinnitus patients and in a population-based group of volunteers: results of the Study of Health in Pomerania. J Oral Rehabil 2004;31(4):311-319.

[3] Conomos MP, Laurie CA, Stilp AM, Gogarten SM, McHugh CP, Nelson SC, Sofer T, Fernandez-Rhodes L, Justice AE, Graff M, Young KL, Seyerle AA, Avery CL, Taylor KD, Rotter JI, Talavera GA, Daviglus ML, Wassertheil-Smoller S, Schneiderman N, Heiss G, Kaplan RC, Franceschini N, Reiner AP, Shaffer JR, Barr RG, Kerr KF, Browning SR, Browning BL, Weir BS, Aviles-Santa ML, Papanicolaou GJ, Lumley T, Szpiro AA, North KE, Rice K, Thornton TA, Laurie CC. Genetic diversity and association studies in US Hispanic/Latino populations: Applications in the Hispanic Community Health Study/Study of Latinos. Am J Hum Genet 2016;98(1):165-184.

[4] Dworkin S, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. J Craniomandib Disord 1992;6(4):301-355.

[5] Gonzalez YM, Schiffman E, Gordon SM, Seago B, Truelove EL, Slade G, Ohrbach R. Development of a brief and effective temporomandibular disorder pain screening questionnaire: reliability and validity. J Am Dent Assoc 2011;142(10):1183-1191.

[6] Kuo C-L, Zaykin DV. Novel rank-based approaches for discovery and replication in genome wide association studies. Genetics 2011;189(1):329-340.

[7] Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, Boehm F, Caporaso NE, Cornelis MC, Edenberg HJ, Gabriel SB, Harris EL, Hu FB, Jacobs KB, Kraft P, Landi MT, Lumley T, Manolio TA, McHugh C, Painter I, Paschall J, Rice JP, Rice KM, Zheng X, Weir BS, Investigators ftG. Quality control and quality assurance in genotypic data for genome-wide association studies. Genet Epidemiol 2010;34(6):591-602.

[8] Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, Criqui MH, Elder JP. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20(8):642-649.

[9] Ohrbach R, Fillingim RB, Mulkey F, Gonzalez Y, Gordon S, Gremillion H, Lim P-F, Ribeiro-Dasilva M, Greenspan JD, Knott C, Maixner W, Slade G. Clinical findings and pain symptoms as potential risk factors for chronic TMD: descriptive data and empirically identified domains from the OPPERA Case-Control Study. J Pain 2011;12(11, Supplement):T27-T45.

[10] Pelkonen ES, Mäki PH, Kyllönen MA, Miettunen JA, Taanila AM, Sipilä KK. Pain-related symptoms of temporomandibular disorders in the offspring of antenatally depressed mothers and depressed parents: a 31-year follow-up of the Northern Finland Birth Cohort 1966. Eur J Pain 2013;17(7):1048-1057.

[11] Pirinen M, Donnelly P, Spencer CCA. Including known covariates can reduce power to detect genetic effects in case-control studies. Nat Genet 2012;44(8):848-851.

[12] Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 2012;28(10):1353-1358.

[13] Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Schneiderman N, Raij L, Talavera G, Allison M, Lavange L, Chambless LE, Heiss G. Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20(8):629-641.

[14] Wu C, DeWan A, Hoh J, Wang Z. A comparison of association methods correcting for population stratification in case-control studies. Ann Hum Genet 2011;75(3):418-427.

[15] Xing G, Xing C. Adjusting for covariates in logistic regression models. Genet Epidemiol 2010;34(7):769-771.

[16] Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet 2012;44(7):821-824.