Title
Genome-wide scan of dental fear and anxiety nominates novel genes

Authors
Yuqiao Zhou1,14, Daniel W. McNeil 2,3,14, Simon Haworth4,5, Tom Dudding4,5, Jonathan M. Chernus1,14, Chenxing Liu1,14, Dongjing Liu1,14, Casey D. Wright2,14, Jamey Brumbaugh2,14, Cameron L. Randall6,14, Robert J. Weyant7,14, Richard J. Crout8,14, Betsy Foxman9,14, Steven Reis10, Nicolas J. Timpson4,11, Mary L. Marazita1,10,12,13,14, John R. Shaffer1,12,13,14

Affiliations
1. Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh
2. Department of Psychology, West Virginia University, Morgantown, WV
3. Department of Dental Public Health and Professional Practice, School of Dentistry, West Virginia University, Morgantown, WV
4. Medical Research Council Integrative Epidemiology Unit, Bristol Medical School, Bristol, BS8 2BN, UK.
5. Bristol Dental School, University of Bristol, Bristol, BS1 2LY, UK.
6. Department of Oral Health Sciences, University of Washington School of Dentistry, Seattle, WA
7. Department of Dental Public Health, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA
8. Department of Periodontics, School of Dentistry, West Virginia University, Morgantown, WV
9. Center for Molecular and Clinical Epidemiology of Infectious Diseases, Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan
10. The Clinical and Translational Science Institute, University of Pittsburgh, Pittsburgh, PA
11. Avon Longitudinal Study of Parents and Children, University of Bristol, Bristol, BS8 2BN
12. Center for Craniofacial and Dental Genetics, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA
13. Department of Oral and Craniofacial Sciences, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA
14. Center for Oral Health Research in Appalachia (COHRA)

Author contact information:
Yuqiao Zhou
yuz80@pitt.edu

Daniel W. McNeil
dmcneil@wvu.edu

Simon Haworth
simon.haworth@bristol.ac.uk
Conflicts of Interest
All authors report no conflicts of interest.

Acknowledgements
The authors express their gratitude to the participants of the COHRA1, COHRA2, Dental SCORE and ALSPAC studies for their contributions, as well as the investigators and staff involved in the data collection, cleaning, and management. This work was supported by NIH grants U01-DE018903, R01-DE014899, X01-HG009878, R01-HL089292, and Pennsylvania Department of Health grant ME-02-384. ALSPAC receives core funding from the UK Medical Research Council, the Wellcome Trust, and the University of Bristol. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the paper.

Author Contributions
Y. Zhou, J.R. Shaffer, D.W. McNeil, S. Haworth, T. Dudding, N. Timpson, M.L. Marazita contributed to conception, design and data acquisition, data analysis, and critically revised the manuscript. Y. Zhou drafted the manuscript. J.M. Chernus, C. Liu, and D. Liu, C.L. Randall contributed to data analysis, interpretation and critically revised the manuscript. C. Wright and J. Brumbaugh contributed to data interpretation, critically revised the manuscript. R.J. Crout, R.J. Weyant, B. Foxman, S. Reis contributed to conception and data acquisition, critically revised the manuscript. All authors gave final approval and agree to be accountable for all aspects of the work.

Abstract

Objectives: Dental care-related fear and anxiety (DFA) is prevalent, affects oral healthcare utilization, and is related to poor oral health and decreased quality of life. In addition to learned and cultural factors, genetics is hypothesized to contribute to DFA. Therefore, we performed a genome-wide association study to identify genetic variants contributing to DFA. Methods: Adult and adolescent participants were from four cohorts (three from the US-based Center for Oral Health Research in Appalachia, n=1144, 1164, and 535; and the UK-based Avon Longitudinal Study of Parents and Children [ALSPAC], n=2078). Two self-report instruments were used to assess DFA: the Dental Fear Survey (US cohorts) and Corah’s Dental Anxiety Scale (ALSPAC). Genome-wide scans were performed for the DFA total scores and subscale scores (avoidance, physiological arousal, fear of dental treatment-specific stimuli), adjusting for age, sex, educational attainment, recruitment site, and genetic ancestry. Results across cohorts were combined using meta-analysis. Results: Heritability estimates for DFA total and subscale scores were similar across cohorts and ranged from
23 to 59%. The meta-analysis revealed three significant (p<5E-8) associations between genetic loci and two DFA subscales: physiological arousal and avoidance. Nearby genes included NTSRI (p=3.05E-8), DMRTA1 (p=4.40E-8) and FAM84A (p=7.72E-9). Of these, NTSRI, which was associated with the avoidance subscale, mediates neurotensin function and its deficiency may lead to altered fear memory in mice. Gene enrichment analyses indicated that loci associated with the DFA total score and physiological arousal subscale score were enriched for genes associated with severe and persistent mental health (e.g., schizophrenia) and neurocognitive (e.g., autism) disorders. **Conclusions:** Heritability analysis indicated that DFA is partly explained by genetic factors, and our association results suggested shared genetic underpinnings with other psychological conditions.

Keywords: Genetics, Heredity, GWAS, Meta-analysis, Genomics, Behavioral Science
Introduction

Dental fear and anxiety (DFA) is a prevalent problem that affects dental treatment-seeking behavior and diminishes oral health and quality of life (McNeil and Randall 2014). Approximately 45% of American adults report at least moderate levels of dental fear with 5-10% reporting that they avoid dental care as a result. Both fear and anxiety are involved in distress about dental treatment, and have extreme manifestations in phobia (McNeil and Randall 2014) (See Appendix for further information). Like other fears and anxieties, DFA is characterized by behavioral, cognitive, and physiological manifestations of distress (McNeil and Randall 2014). Fear of pain has been demonstrated to be a major component of DFA (McNeil and Berryman 1989; Randall et al. 2017a), and is partially genetically determined (Randall et al. 2017b; Ray et al. 2010). There are multiple pathways (Carter et al. 2014) that can lead to the onset and continuation of DFA, which can change over the lifespan (Thomson et al. 2009). In addition to learning (e.g., classical conditioning related to pain or other negative dental experiences, operant conditioning involved in avoidance of treatment) and social/cultural factors such as familial transmission (e.g., modeling and other social learning; (McNeil et al. 2019), genetic risk factors are hypothesized to contribute to the predisposition, precipitation, and perpetuation of specific phobias, such as DFA (Randall et al. 2016).

In support of this hypothesis, genetically modified animal models have identified multiple genes that contribute to behaviors relevant to anxiety and fear. For example, mouse models for anxiety and fear have implicated: Reln, Bdnf, SLC6A4 (Singewald and Holmes 2019); Syngap1 (Ozkan et al. 2014); AdipoR2 (Zhang et al. 2017); Ntsr1 (Yamada et al. 2010); and others. Some of these genes play critical roles in neuronal migration and excitability (Reln, Syngap1 and AdipoR2), while others encode proteins that act as neuromodulators in the nervous system, such as dopamine (Ntsr1) and serotonergic (Bdnf, SLC6A4) systems. While these experiments support the role of genetic risk factors in complex anxiety behaviors, there is no specific animal model for DFA.

In humans, prior evidence also indicates a genetic contribution to fear and anxiety (Levey et al. 2020), including distress about oral healthcare. Indeed, quantitative assessments of DFA obtained from self-report instruments are about 30% heritable (Randall et al. 2017a), although minimal work has addressed the role of specific genes in the etiology of DFA. One recent candidate gene study suggested that variants in MC1R may predict greater levels of dental fear (Randall et al. 2016). Furthermore, other work by our group has nominated candidate genes associated with pain-related fear and anxiety (Randall et al. 2017b), which also may have implications for DFA. However, no genome-wide association studies (GWAS) have been reported specifically for DFA. Therefore, we performed a GWAS to identify genetic variants contributing to DFA.

Methods
The project is reported in accordance to the STROBE guidelines for cohort studies and received ethical approval by the pertinent Institutional Review Boards and Ethics Committees (details in Appendix).

**Overview of study design and cohorts**

Four cohorts (Center for Oral Health Research in Appalachia 1 [COHRA1], n=1144; COHRA2, n=1164; Dental SCORE, n=535; Avon Longitudinal Study of Parents and Children [ALSPAC], n=2078) were included in the analysis:

1. **COHRA1** is a population-based cohort comprising northern Appalachian families (Polk et al. 2008). For analyses presented here, only participants aged 13 years and older were included. Specifically, there were 124 adolescents, ages 13-17 years old; the remainder of the sample were adults, ages 18 - 74. There were 723 (63.2%) females and 421 (36.8%) males. The analysis sample comprised a subset of 886 unrelated participants and another 258 participants were related to one or more individuals in the unrelated set. The number of unique households was 700.

2. **COHRA2** is a longitudinal cohort that included adult women from West Virginia and southwest Pennsylvania who were in their first and second trimesters of pregnancy, following them and their babies for up to seven years (Neiswanger et al. 2015). The current GWAS included the mothers only. Mean age was 29.1 (SD = 5.4), range 18 – 47 years.

3. **Dental SCORE** is a cross-sectional cohort of unrelated individuals who participated in the same data collection protocol as COHRA1 (Bambs et al. 2011). Mean age was 63.1 (SD = 7.4), range 45-75 years. There were 356 (66.5%) females and 179 (33.5%) males.

4. **ALSPAC** is a longitudinal birth cohort which recruited pregnant women living near Bristol, UK with an estimated delivery date between 1991 and 1992. Since then, participating mothers and their children have been followed up with clinical assessments and questionnaires and is ongoing (Boyd et al. 2013). The current GWAS included adolescents in ALSPAC who attended a clinic at age 17.5 years during which the DFA assessment was performed. There were 1268 (61%) females and 810 (39%) males. Those participants answered questions about dental health and attitudes towards dental care.

**Dental fear assessments**

In COHRA1, COHRA2 and Dental SCORE, the Dental Fear Survey (DFS) was used to assess dental fear, quantified as a total score and three subscale scores: behavioral avoidance (hereafter referred to as “avoidance”), fear of specific dental stimuli (“specific stimuli”) and physiological arousal associated with dental treatment (“physiological arousal”). Responses to individual items are based on a 5-point Likert-type scale. DFA total and subscale scores and were generated by summing the Likert-like points across multiple items (i.e., all items for the total score and specific items for each subscale) and were treated as continuous variables. These variables were analyzed using standard statistical models for quantitative traits. The
distribution of DFA (total and subscale) scores in the examined cohorts are shown in the Appendix Figures S1 and S2. The total score ranges from 20 to 100, with higher scores indicating greater DFA. The DFS has established psychometric properties and has been widely used internationally. In ALSPAC, Corah's Dental Anxiety Scale (DAS) (Corah et al. 1978) was adapted for use in the study to measure DFA. Total scores were generated by summing responses across all four domains (A-D), producing values between 4 and 20; additionally, a physiological arousal score was generated by summing domains B, C, and D (see Appendix for details).

**Genotyping**

Participants were genotyped on Illumina genotyping arrays and quality control (QC) processing was performed as described in the Appendix. Imputation was performed using the Michigan Imputation Server Haplotype Reference Consortium (HRC) Phase 1 panel (Das et al. 2016), except for COHRA1, in which we used the 1000 Genomes project phase 1 version 2 release (Consortium 2015) (Appendix Table S4). Variants were excluded if they had low imputation quality (info score below 0.4 for COHRA1, r² below 0.3 for COHRA2, Dental SCORE and ALSPAC), departed from Hardy–Weinberg equilibrium (p<0.0001), or had minor allele frequencies<1%. Descriptions of genotyping, quality control and imputation for each cohort are provided in the Appendix.

**GWAS and meta-analysis**

GWAS were performed for the DFA total score and the three subscale scores, while adjusting for age, sex, educational attainment, income (COHRA1 and COHRA2 only), recruitment site, genetic principal components (COHRA2 and Dental SCORE only), and other study-specific covariates as described in Appendix Table S4. Linear regression was used to test each single nucleotide polymorphism (SNP) for association under the additive genetic model. Analysis used PLINK for COHRA2 and Dental SCORE because these samples included only unrelated individuals. To assess population structure in COHRA2 and Dental SCORE cohorts, we performed principal component analysis (PCA) using subsets of uncorrelated SNPs. Based on the scatterplots of the principal components (PCs) and scree plots, we determined that population structure was captured in six PCs of ancestry. Mixed models as implemented in EMMAX and BOLT-LMM were used for COHRA1 and ALSPAC, respectively, to model the variance due to the kinship (comprising both the family relatedness and population structure) in the samples, while adjusting for pertinent covariates. Stouffer’s p-value based meta-analysis was carried out using METAL (Willer et al. 2010).

Overlapping phenotypes among the four cohorts were meta-analyzed via two sets of meta-analyses: US three-cohort (COHRA1, COHRA2 and Dental SCORE) meta-analyses of the total score and avoidance, physiological arousal, and specific stimuli subscales, and US and UK four-cohort meta-analyses (COHRA1,
COHRA2, Dental SCORE and ALSPAC of the total score and physiological arousal subscale. The intersection comprising 6.4 million SNPs across the four cohorts went into the meta-analyses.

GWAS results were visualized using composite Manhattan plots and Q-Q plots. We used the conventional threshold of \( p < 5 \times 10^{-8} \) (i.e., Bonferroni correction for 1 million tests) for genome-wide statistical significance. Because we performed 20 GWAS (6 meta-analyses, 14 individual cohort ones), study-wide statistical significance was determined at \( p < 2.5 \times 10^{-9} \) (i.e. \( p < 5 \times 10^{-8} \) divided by 20). We also reported “suggestive” evidence of association of \( p < 1 \times 10^{-6} \).

*Gene ontology enrichment analyses*

To explore the biological processes and phenotypes associated with the identified GWAS loci, we performed Genomic Regions Enrichment of Annotations (GREAT) (McLean et al. 2010) analysis, using independent SNPs with \( p < 1 \times 10^{-6} \) in the meta-analyses. In addition, to explore the known functional effects of our most significant SNPs as reported by public databases, we used Functional Mapping and Annotation of genetic associations (FUMA) (Watanabe et al. 2017), with a \( p < 1 \times 10^{-6} \) threshold for lead SNPs. We included only SNPs deemed approximately independent of each other (i.e., \( r^2 < 0.6 \))

*Heritability estimates*

Heritability within the COHRA1, COHRA2 and Dental SCORE cohorts was assessed using the GREML method, implemented in the GCTA software package (Yang et al. 2011), using participant-level phenotype data and a genetic relatedness matrix estimated from common genetic variants (with effect allele frequency >1%). Family-based heritability estimates for the COHRA1 cohort were published elsewhere (Randall et al. 2017a). Heritability of ALSPAC and that of the meta-analysis were assessed using SumHer (Speed and Balding 2019), which uses GWAS summary statistics to estimate SNP heritability using the linkage disequilibrium-adjusted kinship method (LDAK). The reference linkage disequilibrium (LD) scores were taken from HapMap3 (Consortium 2010) reference data.

*Results*

*Heritability of DFA*

Distributions of DFA total score and subscales are shown in Supplemental Figures S1-2. As expected for these population-based samples, the majority of participants had low to moderate levels of DFA. A minority of participants had high levels of DFA, including 11.3% (95% confidence interval [CI] 9.5%-13.3) of COHRA1, 9.5% (95% CI 8.0-11.4%) of COHRA2, 6.4% (95% CI 4.4-8.8%) of Dental SCORE, and 7.8% (95% CI 6.7-9.0%) of ALSPAC, defined as total scores above 60 (US cohorts) or 12 (ALSPAC). We estimated SNP-based heritability of the DFA total score and subscales using the GREML method in COHRA1, COHRA2, and Dental SCORE. Heritability estimates ranged from 0 to 0.28, 0 to 0.19, and 0.23 to 0.67 in COHRA1, COHRA2, and Dental SCORE, respectively (Table 1). As expected, point estimates of SNP-based heritability
were lower than family-based heritability in COHRA1 (Randall et al. 2017a), although the 95% CIs overlapped between the two approaches. Likewise, CIs of SNP-based heritability estimates largely overlapped across cohorts. Heritability of DFA in ALSPAC was estimated using summary statistics via SumHer (Speed and Balding 2019) to be 0.37 (SD 0.49) for the total score and 0.26 (SD 0.54) for physiological arousal. SumHer-based heritability estimates from the meta-analysis summary statistics produced estimates ranging from 0.23 to 0.59, with large standard deviations (Table 1).

Table 1. Heritability estimates of dental fear total score and sub-scales.

<table>
<thead>
<tr>
<th></th>
<th>family-based h² (CI) (Randall et al. 2017a)</th>
<th>SNP-based h² (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SOLAR (CI)</td>
<td>COHRA1 (n=1370)</td>
</tr>
<tr>
<td><strong>Total score</strong></td>
<td>0.30 (0.14-0.46)</td>
<td>0.13 (0.00-0.37)</td>
</tr>
<tr>
<td><strong>Physiological arousal</strong></td>
<td>0.14 (0.00-0.3)</td>
<td>&lt;0.01 (0.00-0.12)</td>
</tr>
<tr>
<td><strong>Specific stimuli</strong></td>
<td>0.36 (0.2-0.52)</td>
<td>0.28 (0.14-0.52)</td>
</tr>
<tr>
<td><strong>Avoidance</strong></td>
<td>0.22 (0.06-0.38)</td>
<td>0.05 (0.0-0.21)</td>
</tr>
</tbody>
</table>

Note.

h²: heritability estimate  
CI: 95% confidence interval  
SD: standard deviation  
*Analyses were performed separately by genotyping batches; the majority of the COHRA1 sample was genotyped and analyzed together and the remainder of the COHRA1 sample was genotyped and analyzed with COHRA2, here denoted “COHRA2”

GWAS meta-analysis results

The two sets of meta-analyses included up to 6.4 million variants in up to 4921 participants aged 13-75. Results for GWAS meta-analysis across three US cohorts are shown as a composite Manhattan plot in Figure 1A for four DFA phenotypes (total score, avoidance, physiological arousal, and specific stimuli). Results from the meta-analysis across the three US cohorts plus one UK cohort are shown in Figure 1B, for the total score and physiological arousal. Results for individual cohorts are shown in Supplemental Figures S3-6. The genomic inflation factors ranged from 0.93 to 1.02 across scans, suggesting little evidence of systematic inflation of test statistics.

US three-cohort meta-analysis

We observed two genome-wide significant associations for physiological arousal and avoidance, one at chromosome 9p21.3 (lead SNP rs62572327; each copy of the C allele results in an increase [beta coefficient] of 1.72 to 2.33 points on the DFS subscale score; p=3.05E-8 for physiological arousal) near the gene DMRTA1, and the other at chromosome 20q13.33 (lead SNP rs8124973; beta ranges from 0.79 to 1.71 for the A allele;
$p=4.40E-8$ for avoidance) near $NTSR1$ and $SLCO4A1$ (Figure 1A). rs8124973 shows chromatin marks of promoter and enhancer activity in multiple tissues and cells, including the brain. These two loci are depicted in Figure 2A and 2B regional association plots, and the details for the lead SNPs are shown in Table 2. In addition to the genome-wide significant associations, 12 suggestive associations were observed (Table 2) including the genomic regions that harbor $PDE10A$ (lead SNP rs9356408; $p=6.38E-8$ for avoidance; Table 2 and Figure 2C), and $PAFAHIBI1$ (lead SNP rs8077138, $p=5.6E-7$ for physiological arousal; Table 2 and Figure 2D).

**US and UK four-cohort meta-analysis**

One genome-wide significant association was observed for physiological arousal at a region on chromosome 2p24.3 (lead SNP rs11694373; beta ranges from 1.41 to 2.26 for the C allele; $p=7.72E-9$) near the gene $FAM84A$ (Figure 1B). The lead SNP is located within a long intergenic non-protein coding RNA ($LINC00276$).

A suggestive association signal for total score was observed at chromosome 12p13.33 (lead SNP rs725652; $p=5.05E-8$; Figure 2E), within 96 kb of $ADIPOR2$, $CACNA2D4$ and $CACNA1C$. Other suggestive associations were near the genes $SYNGAP1$ (lead SNP rs1800838, an eQTL for that gene, $p=9.1E-7$; Figure 2F), $COL13A1$ (lead SNP rs77245622, $p=8.32E-7$), and $NEDD1$ (lead SNP rs11108685, an eQTL for $NEDD1$, $p=4.53E-7$). A list of all suggestive associations is available in Table 3.

Table 2. US three-cohort (n=2843) meta-analysis results for loci showing genome-wide significant or suggestive evidence of association with DFA phenotypes.

<table>
<thead>
<tr>
<th>Locus</th>
<th>BP</th>
<th>SNP</th>
<th>MAF</th>
<th>$p$(Total score)</th>
<th>$p$(Physiological Arousal)</th>
<th>$p$(Avoidance)</th>
<th>$p$(Specific Stimuli)</th>
<th>Gene nearby and biological relevance</th>
<th>SNP annotation (HaploReg v4 (Ward and Kellis 2016) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p32.3</td>
<td>52866 066</td>
<td>rs7556402</td>
<td>0.011 -0.03</td>
<td>5.27E-7</td>
<td>3E-4</td>
<td>2.58E-5</td>
<td>7.14E-7</td>
<td>$BTF3L4$, $ORC1$</td>
<td>The leading SNP is in a region showing chromatin marks for promoter and enhancer activity in multiple tissues including brain</td>
</tr>
<tr>
<td>2p25.1</td>
<td>75221 74</td>
<td>rs76719058</td>
<td>0.016 -0.047</td>
<td>0.00000 15</td>
<td>0.0011</td>
<td>0.0016</td>
<td>0.035</td>
<td>$LOC100506274$</td>
<td></td>
</tr>
<tr>
<td>2p24.3</td>
<td>14502 100</td>
<td>rs11694373</td>
<td>0.019 -0.032</td>
<td>3.03E-6</td>
<td>1.27E-7</td>
<td>3.20E-5</td>
<td>3.13E-4</td>
<td>$FAM84A$: Encodes Neurologic Sensory Protein 1, a candidate gene for schizophrenia (Sundararajan et al. 2018)</td>
<td>The leading SNP is in a region showing chromatin marks for enhancer activity in colon mucosa and intestine cell</td>
</tr>
<tr>
<td>2q14.2</td>
<td>12182 6568</td>
<td>rs11901120</td>
<td>0.01-0.052</td>
<td>1.10E-3</td>
<td>1.63E-2</td>
<td>6.97E-8</td>
<td>6.13E-2</td>
<td>$GLI2$</td>
<td>The leading SNP is in a region showing chromatin marks for promoter and enhancer in multiple brain tissues</td>
</tr>
</tbody>
</table>
Reported SNPs were mapped to genome build GRCh37/hg19.

### Table 3

**Table 3. US and UK (n=4921) four-cohort meta-analysis results for loci showing genome-wide significant or suggestive evidence of association with dental fear phenotypes.**

<table>
<thead>
<tr>
<th>Locus</th>
<th>BP</th>
<th>SNP</th>
<th>MAF</th>
<th>(p) (Physiological Arousal)</th>
<th>(p) (Total score)</th>
<th>Gene nearby and biological relevance</th>
<th>SNP annotation (HaploReg v4 (Ward and Kellis 2016))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2p16.3</td>
<td>49814095</td>
<td>rs113974660</td>
<td>0.01-0.014</td>
<td>1.46E-5</td>
<td>2.58E-7</td>
<td>1.05E-3</td>
<td>NRXN1</td>
</tr>
<tr>
<td>3p12.2</td>
<td>81096082</td>
<td>rs2639278</td>
<td>0.33-0.43</td>
<td>0.0006-8</td>
<td>0.0077-7</td>
<td>9.68E-7</td>
<td>Intergenic</td>
</tr>
<tr>
<td>6q27</td>
<td>166408316</td>
<td>rs9356408</td>
<td>0.04-0.05</td>
<td>1.07E-5</td>
<td>1.12E-3</td>
<td>6.38E-8</td>
<td>5.25E-3</td>
</tr>
<tr>
<td>9p21.3</td>
<td>22538241</td>
<td>rs62572327</td>
<td>0.014-0.025</td>
<td>6.27E-7</td>
<td>3.05E-8</td>
<td>5.58E-5</td>
<td>2.78E-5</td>
</tr>
<tr>
<td>9p34.3</td>
<td>138654419</td>
<td>rs71508811</td>
<td>0.01-0.022</td>
<td>1.10E-5</td>
<td>4.19E-5</td>
<td>7.72E-7</td>
<td>2.07E-3</td>
</tr>
<tr>
<td>17p13.3</td>
<td>2558393</td>
<td>rs8077138</td>
<td>0.252-0.261</td>
<td>4.55E-6</td>
<td>5.63E-7</td>
<td>5.95E-5</td>
<td>4.15E-4</td>
</tr>
<tr>
<td>17p21.31</td>
<td>43277758</td>
<td>rs11870697</td>
<td>0.2-0.26</td>
<td>3.26E-7</td>
<td>2.24E-4</td>
<td>1.58E-6</td>
<td>2.69E-6</td>
</tr>
<tr>
<td>20q13.3</td>
<td>61308731</td>
<td>rs8124973</td>
<td>0.13-0.18</td>
<td>5.94E-7</td>
<td>1.47E-6</td>
<td>4.4E-8</td>
<td>6.04E-4</td>
</tr>
<tr>
<td>20p12.1</td>
<td>16921399</td>
<td>rs73086099</td>
<td>0.016-0.022</td>
<td>3.74E-6</td>
<td>9.027E-7</td>
<td>1.19E-6</td>
<td>0.002-4</td>
</tr>
<tr>
<td>Xq26.2</td>
<td>133593711</td>
<td>rs5978007</td>
<td>0.011-0.03</td>
<td>1.54E-5</td>
<td>2.45E-2</td>
<td>1.53E-7</td>
<td>3.72E-1</td>
</tr>
</tbody>
</table>

Note. Abbreviations: BP, base pair (variant position); MAF, minor allele frequency. Reported SNPs were mapped to genome build GRCh37/hg19.
<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position</th>
<th>rsID</th>
<th>Minor Allele Frequency</th>
<th>Beta</th>
<th>p-Value</th>
<th>Gene and Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>2p24.3</td>
<td>14502100</td>
<td>rs11694373</td>
<td>0.016-0.02</td>
<td>7.72E-9</td>
<td>1.46E-6</td>
<td><em>FAM84A</em>: encodes Neurologic Sensory Protein 1, highly expressed in colon cancer cells (Kobayashi et al. 2006); candidate gene for schizophrenia (Sundararajan et al. 2018)</td>
</tr>
<tr>
<td>2q11.2</td>
<td>99802191</td>
<td>rs116519747</td>
<td>0.02-0.032</td>
<td>7.65E-7</td>
<td>6.7E-6</td>
<td><em>KIAA1211L</em></td>
</tr>
<tr>
<td>6p21.32</td>
<td>33272541</td>
<td>rs1800838</td>
<td>0.01-0.016</td>
<td>5.39E-6</td>
<td>9.2E-7</td>
<td><em>SYNGAP1</em> regulates synaptic plasticity and neuronal homeostasis. Mice with <em>SYNGAP1</em> knocked out have deficits in fear memory, decreased anxiety-like behavior, and increased locomotor activity. Allelic variants of this gene are associated with intellectual disability and autism spectrum disorder (Ozkan et al. 2014)</td>
</tr>
<tr>
<td>10q22.1</td>
<td>71891900</td>
<td>rs77245622</td>
<td>0.36-0.49</td>
<td>3.74E-5</td>
<td>8.32E-7</td>
<td><em>COL13A1</em></td>
</tr>
<tr>
<td>12p13.33</td>
<td>1994572</td>
<td>rs725652</td>
<td>0.012-0.034</td>
<td>3.18E-7</td>
<td>5.05E-8</td>
<td><em>AdipoR2</em>: Receptor for ADIPOQ, an essential hormone secreted by adipocytes that regulates glucose and lipid metabolism; plays a role in neuropathic pain perception and inflammatory pain; knockout mice exhibit enhanced contextual fear responses (Zhang et al. 2017); <em>CACNA2D4</em>, <em>CACNA1C</em>: encode members of the voltage-dependent calcium channel.</td>
</tr>
<tr>
<td>12q23.1</td>
<td>97156844</td>
<td>rs11108685</td>
<td>0.012-0.02</td>
<td>5.17E-7</td>
<td>4.53E-7</td>
<td><em>NEDD1</em>: Neural precursor cell expressed gene</td>
</tr>
</tbody>
</table>

Note: Abbreviations: BP, base pair (variant position); MAF, minor allele frequency. Reported SNPs were mapped to genome build GRCh37/hg19.

In addition to associations observed in the meta-analyses, we also scrutinized loci that were significantly ($p<5E-8$) associated with DFA phenotypes in individual cohorts. In total, 10 genome-wide significant associations were observed in individual cohorts but not reaching the meta-analysis suggestive level (Figures S1-4, Table S3). Two of these showed some evidence of replication in a second cohort. For the COHRA1 cohort, the lead SNP, rs118016777 (beta=4.22, p=3.02E-8 for physiological arousal), had the same direction of effect and a nominally significant $p$ in COHRA2 (beta=1.31, $p=0.048$). This SNP is located in a region harboring *NPTXR* (Figure 2G, Table S3). Though not showing any evidence of replication, the lead SNP for an association observed in COHRA2, rs78589876 (beta=3.67, $p=8E-9$), is located near the biologically plausible gene, *TRPC1*, which regulates transmission of sensory signals (Figure 2H, Table S3). Manhattan plots showing the individual cohort results are available in supplemental figures S3-S6.

Gene ontology enrichment analyses and functional annotation
To explore the biological processes and phenotypes associated with the identified GWAS loci, we performed enrichment analysis in GREAT, using SNPs with p<1E-6 in the meta-analyses. For the 85 independent lead SNPs across all four phenotypes combined, we detected significant enrichment for calcium ion transport (Figure S7A) and bradykinin receptors and G protein-coupled receptors (GPCRs) (Figure S7B). We then performed functional annotation of associated variants and candidate gene annotation using FUMA. For the four-cohort meta-analysis, FUMA gene-set analysis revealed loci associated with total score are enriched for GWAS catalog reported gene-sets for psychiatric disorders such as autism spectrum disorder or schizophrenia (p=9.35E-8, Fig S7D)

Discussion

We performed the first GWAS of DFA, identifying associations in and around several genes, including some with known biological roles in fear memory and anxiety traits from animal models. In particular, the two sets of meta-analyses revealed three genome-wide significant associations for two DFA subscales, physiological arousal and avoidance. Though the GWAS study design, itself, does not indicate which specific gene accounts for an observed association, genes near the significant loci included DMRTA1, NTSR1 and FAM84A. Note, the associated SNPs were not actually located within these genes, which is consistent with results for most GWAS of complex human traits, and suggests a regulatory role rather than an impact on protein structure, although the exact causal variants have not been determined. Of these genes, NTSR1, which was associated with the avoidance subscale, is especially interesting in the context of fear and anxiety. NTSR1 mediates the multiple functions of neurotensin, such as hypotension and antinociception. Ntsr1-deficient mice may have emotional disorders, such as those involving fear memory (Yamauchi et al. 2007). DMRTA1 is a cis-regulated candidate gene located within a known QTL for contextual fear memory in mice (Carhuatanta et al. 2014). FAM84A encodes Neurologic Sensory Protein 1, which is a candidate gene for schizophrenia (Sundararajan et al. 2018). Although these associations did not meet the study-wide significance threshold accounting for the number of individual cohorts by phenotypes (i.e. p<2.5E-9), this threshold is overly conservative since the total scores and subscales are all highly correlated with each other, and the meta-analyses are not independent of the individual cohorts.

A few genes near the suggestive or genome-wide significant loci showed corroborating evidence in the context of fear, anxiety, and stress in animal models. For example, an association for total score (lead SNP rs725652, p=5.05E-8, US and UK four-cohort meta-analysis) was observed at the genomic region harboring ADIPOR2. ADIPOR2 plays a role in neuropathic pain perception and has been shown in knockout mice to influence contextual fear memories (Zhang et al. 2017). This finding coincides with prior work demonstrating the importance of memory of fear, anxiety, and pain among highly fearful/anxious dental patients (Kyle et al. 2016). A suggestive association for total score was identified at an eQTL for SYNGAP1 (lead SNP rs1800838,
SYNGAP1 regulates synaptic plasticity and has been shown in knockout mice to cause a deficit in fear memory and decreased anxiety-like behavior (Ozkan et al. 2014).

We scrutinized 6 previously implicated (Randall et al. 2016) exonic SNPs in MC1R for any evidence of association. In contrast to previous work, we did not see significant associations between these MC1R variants and DFA in any of our cohorts or meta-analyses after adjusting for multiple testing. However, we did see nominal evidence of association ($p<0.05$) for two of the SNPs (rs1805007 and rs11547464) in COHRA2, ALSPAC, and the US three-cohort meta-analysis, suggesting a need for future studies to clarify the role of MC1R (results not shown).

Gene enrichment analysis indicated that genomic loci associated with DFA are enriched for biological processes including bradykinin receptors (Figure S5), which are associated with anxiety disorders and other pathophysiologic conditions such as inflammation, trauma, burns, shock, and allergy (Rouhiainen et al. 2019). We also detected enrichment in G protein-coupled receptors (GPCRs) which have been linked to regulation of fear and anxiety (Catapano and Manji 2007). FUMA gene-set analysis revealed that the loci reaching the suggestive level in our genome scans were enriched for genes associated with severe and persistent mental health (schizophrenia) and neurocognitive (autism) conditions (adjusted $p=1.82E-9$).

Previously our group and others have found that in addition to learning and social/environmental factors, genetic influences are important in the etiology of DFA (Randall et al. 2017a; Ray et al. 2010). DFA phenotypes were previously estimated to be 14%-30% heritable from family data in COHRA1 (Randall et al. 2017a) (Table 1). Here we found that the SNP-based heritability estimates in COHRA1 (0-28%) were lower than previously reported family-based estimates (but with overlapping CIs). The lower heritability estimates are not surprising given that the SNP-based methods used in the present analysis only reflect the additive heritability due to common variants, thus potentially underestimating the heritability relative to family studies (Docherty et al. 2016). Furthermore, this result suggests that the genetic architecture of DFA may include other types of variants, such as low frequency, rare, copy number, and/or structural variants, as well as gene-gene and gene-environment interactions. Such genetic contributions were not investigated in this study of common variants, and future research is needed to explore them. Interestingly, among individual cohorts, heritability estimates were highest for Dental SCORE, which is also the oldest cohort, although the wide CIs overlapped across studies. Notably, SumHer produced heritability estimates with large standard deviations, which can be interpreted to mean that there is uncertainty in the heritability estimate. A potential explanation for the uncertainty is that our sample size (under 5000) was too small to obtain precise estimates in the SumHer analysis. Indeed, the small sample sizes are a limitation for all of heritability estimates reported here. In addition, heterogeneity may exist in the genetic architecture of DFA across strata of age, geography, and
environmental exposures, which could explain the diminished heritability estimate across the combined samples compared to the heritability estimates within each study.

While our results suggested a meaningful genetic component of DFA, and we nominate some specific loci for follow-up, our study did not address whether genetics alone can lead to high DFA scores, nor what environmental exposures influence DFA and how they may interact with genetics. As a complex trait, there are many factors contributing to the etiology of DFA besides genetics, such as learning via direct experiences (e.g., classical conditioning, respondent conditioning) and social learning, sex and gender, age, and cultural factors (McNeil & Randall, 2014). Moreover, DFA can change over time (Thomson et al. 2019), highlighting the importance of environmental influences. The core of DFA can be related to many different antecedents (e.g., fear of pain, claustrophobia/being closed-in in dental settings [McNeil & Berryman, 1989]), of which generalized anxiety is one. Analyses presented in the Appendix explore the relationship between DFA and general anxiety in our samples. The association between DFA and general anxiety across the samples was weak, however, consistent with prior studies, although some individuals who have generalized anxiety likely also experience DFA. In addition, gene-environment interactions – that is, when the effect of an environmental exposure depends on the underlying genetics, or equivalently, when the manifestation of a genetic liability depends on the environmental context – may also play a role. Although not within the scope of our study, gene-environment interactions are hypothesized to be an important contributor to the “missing heritability” for complex traits, and identifying and modeling gene-environment interactions remain an interesting area that needs to be further investigated.

Some study limitations should be noted. First, while both instruments used to assess dental fear are validated, DFA traits were measured using self-report instruments. Measurement of complex psychological or behavioral phenotypes is stronger when additional assessment tools are used (e.g., combined use of behavioral or psychophysiological recording, clinician-rated measures, and/or physiological assessment).

Secondly, the sample size was modest for a GWAS meta-analysis, and therefore many true associations may have gone undetected due to low power to detect variants of small effect and/or low frequency. Thus, additional studies and larger sample sizes are necessary. Moreover, we studied DFA in population-based samples, which include only a minority of participants with high levels of DFA, reflective of the phenotype distribution in the underlying population. Therefore, the genetic associations reported here represent mean differences across the full range of the phenotype, and are not necessarily driven by individuals with high DFA. Other approaches, such as case-control or extreme phenotype study designs may be more apt to detect variants specifically leading to high dental fear. Lastly, our study focused solely on the discovery of independently acting genetic variants on DFA, and by design did not address other contributors, such as prior negative dental experiences, social learning, or comorbidities such as anxiety disorders or oral disease. For example, one could hypothesize that having extensive dental caries may be related to the development
of DFA. We intentionally avoided including these factors in our analysis model, as they themselves may be partly controlled by genetics, so their inclusion could prevent the discovery of genetic loci comprising the shared genetic architecture of DFA and other traits/comorbidities. This analysis decision reflects the proximal goal of the study – identification of DFA-associated genetic variants – and does not imply that environmental exposures and comorbidities are unimportant. Instead, such relationships will need to be further explored in further analyses. However, it is explicable to note that the exclusion of important environmental sources of variance and comorbidities from our analysis model would not lead to false positive genetic associations.

In conclusion, we shed light on the complex genetic architecture of DFA and nominated novel genetic loci for follow-up in future investigations. The findings related to genes involved in fear memories are consistent with other literature on the importance of memory of dental experiences in dental fear and anxiety (MeNeil et al. 2011). Understanding the genetic contributions to DFA may lead to improved interventions, which may in turn enhance treatment utilization and oral health status generally.

FIGURE LEGENDS

Figure 1. Manhattan plots for US three-cohort meta-analysis and UK four-cohort meta-analysis
Combined Manhattan plot showing US three-cohort (A) and US and UK four-cohort (B) meta-analysis results for total fear score and subscales. For US three-cohort meta-analysis, results for four phenotypes are plotted while for US and UK four-cohort meta-analysis, two phenotypes are represented. The red solid line indicates the genome-wide significance threshold ($p=5E-8$) and the blue dotted line indicates the suggestive significance threshold ($p=1E-6$). Highlighted in different colors are the genetic variants that surpassed the suggestive significance level. Gene symbols or cytogenic location are marked above the variants. The y-axis shows the $-\log_{10}$-transformed $p$, indicating significance of the association for each SNP across the genome (x-axis; organized by physical position on each chromosome).

Figure 2. LocusZoom (Pruim et al. 2010) plots showing the significant and suggestive associations (left y-axis; log_{10}-transformed p-values) with dental fear traits (A) and (B) genome-wide significant associations observed in the US three-cohort meta-analysis; (C) and (D) suggestive associations observed in US three-cohort meta-analysis; (E) and (F) suggestive significant associations observed in US and UK four-cohort meta-analysis; (G) genome-wide significant association for COHRA1 and (H) genome-wide significant associations for COHRA2. Each point depicts the significance ($-\log_{10}$-transformed $p$; left y-axis) of a SNP. Square indicates genotyped SNPs, and circle indicates imputed SNPs.
The shading of the points represents the linkage disequilibrium (i.e., correlation; $r^2$) of the SNP with the lead SNP. The blue recombination rate overlay (right $y$-axis) provides information about the LD-block structure of region. Positions of known genes are shown below the plot.

References


