Published in final edited form as:

Biol Psychiatry. 2018 July 15; 84(2): 148–159. doi:10.1016/j.biopsych.2017.11.010.

Reproducibility in Imaging Genetics: The Case of Threat-Related Amygdala Reactivity

Reut Avinun¹, Adam Nevo², Annchen R. Knodt¹, Maxwell L. Elliott¹, and Ahmad R. Hariri¹

¹Laboratory of NeuroGenetics, Department of Psychology & Neuroscience, Duke University, Durham, NC, USA

²Cardiothoracic Division, Department of Surgery, Duke University Medical Center, Durham, NC, USA

Abstract

BACKGROUND—Low reproducibility rates are a concern in most, if not all, scientific disciplines. In psychiatric genetics specifically, intermediate brain phenotypes more proximal to putative genetic effects were touted as a strategy leading to increased power and reproducibility. Here we attempt to replicate previously published associations between single nucleotide polymorphisms (SNPs) and threat-related amygdala reactivity, which represents a robust brain phenotype not only implicated in the pathophysiology of multiple disorders, but also as a biomarker of future risk.

METHODS—We conducted a literature search for published associations between SNPs and threat-related amygdala reactivity and found 37 unique findings. Our replication sample consisted of 1117 young adult volunteers (629 women, mean age 19.72±1.25 years) for whom both genetic and functional MRI data were available.

RESULTS—Of the 37 unique associations identified, only 3 replicated as previously reported. When exploratory analyses were conducted with different model parameters compared to the original findings, significant associations were identified for 28 additional studies: 8 of these were for a different contrast/laterality; 5 for a different sex and/or race/ethnicity; and 15 in the opposite direction as well as for a different contrast, laterality, sex, and/or race/ethnicity. No significant associations, regardless of model parameters, were detected for 6 studies. Notably, none of the significant associations survived correction for multiple comparisons.

CONCLUSIONS—We discuss these patterns of poor replication with regard to the general strategy of targeting intermediate brain phenotypes in genetic association studies as well as the growing importance of advancing the reproducibility of imaging genetics findings.

Corresponding Author: Reut Avinun, Ph.D., Laboratory of NeuroGenetics, Department of Psychology & Neuroscience, Duke University, Grey Building 2020 West Main St, Ste 0030 Durham, NC 27705. reut.avinun@duke.edu.

DISCLOSURES

The authors report no biomedical financial interests or potential conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Reproducibility; genes; single nucleotide polymorphism (SNP); functional MRI; amygdala; imaging genetics

Background

Candidate gene studies of complex behavioral phenotypes suffer from low reproducibility, a concern certainly not unique to this discipline (1–3). Considerable effort has been made to improve the reproducibility of behavioral genetics findings. Amongst a number of strategies, including the reduction of heterogeneity in clinical phenotypes (4), the adoption of dimensional symptom measures (5), and an increase in sample sizes (6), the use of intermediate systems-level brain phenotypes as targets of association studies has been advanced in the last nearly two decades (7–9). While there is no question that any genetic influence on behavior and related risk for mental illness is mediated through the brain, demonstrating this obligate pathway using available neuroimaging metrics has proved inconsistent at best (10–12).

Here, we seek to contribute to this literature by examining the reproducibility of published association studies of single nucleotide polymorphisms (SNPs) and threat-related amygdala reactivity using data from a large cohort of young adult volunteers. In doing so, we evaluate the observed patterns of reproducibility, and address possible barriers to replication in imaging genetics. This is especially germane as the opportunity steadily increases for large genome wide association studies of intermediate brain phenotypes (13–15) to generate many novel loci, especially SNPs, of likely unknown function.

We focus on the intermediate brain phenotype of threat-related amygdala reactivity for several reasons. First, threat-related amygdala reactivity has been consistently implicated in the pathophysiology of multiple psychiatric disorders, especially depression and anxiety (16), and is thus a leading candidate mechanism through which genetic risk for these disorders may manifest. Second, threat-related amygdala reactivity has emerged as a biomarker of future risk for mood and anxiety disorders (17–20), and thus reproducible genetic correlates of this phenotype may one day be clinically useful in preventing illness. Third, candidate gene studies of threat-related amygdala reactivity are amongst, if not, the most common in the imaging genetics literature, and thus this intermediate brain phenotype is primed for a systematic evaluation. Fourth, the amygdala represents an anatomically and functionally conserved cornerstone of translational neuroscience research poised to reveal the detailed mechanisms through which genetic variation impacts behavior and risk for psychopathology (21). Finally, our research group has published multiple candidate gene studies of this phenotype (22–28) and we have recently generated a large sample of overlapping SNP and amygdala reactivity data that affords significant power to evaluate reproducibility while varying a number of parameters (e.g., race/ethnicity, fMRI task/contrast, diagnosis) that could influence observed associations. Importantly, the size of our replication sample is larger than all included discovery studies.

Methods

Literature Search

In order to find genetic association studies of amygdala reactivity published by April 2017, we conducted a search in PubMed using the following terms: gene, polymorphism, amygdala, and imaging. This search was restricted to titles and abstracts of human studies. Additionally, we searched for relevant papers in the reference list of a recent review on the genetics of anxiety (29). Lastly, we used Google Scholar to search for studies published since 2015 by using the following terms: gene, amygdala, fMRI, and threat or anxiety. Only papers published in the English language were considered for inclusion. The search yielded 102 relevant papers, which included 107 findings. Of these there were 37 unique associations between SNPs and threat-related amygdala reactivity which we attempted to replicate (information regarding inclusion and exclusion criteria is provided in the supplementary information methods and Table S1). Notably, 9 of the 37 associations were based on earlier smaller subsamples of our current replication sample. In cases where more than one study reported on the same SNP, we chose the study that was most similar to ours in terms of sample characteristics and fMRI task. Study demographics and functional details regarding the genes and polymorphisms that have been associated with threat-related amygdala reactivity are presented in Tables S2 and S3.

Participants and Study Design

Data were derived from 1330 participants (762 women, mean age 19.70 ± 1.25 years) who successfully completed the Duke Neurogenetics Study (DNS), which assessed a range of behavioral and biological traits among young adult, university students. The DNS was approved by the Duke University School of Medicine Institutional Review Board, and all participants provided written informed consent prior to participation. Detailed recruitment and exclusion criteria are provided in the supplementary information. The current analyses were conducted on a subset of 1117 participants (629 women, mean age 19.72 ± 1.25 years) for whom there was also whole-genome SNP data available. Of the 1117 participants used in our analyses, 236 individuals had at least one DSM-IV diagnosis (further information regarding diagnoses can be found in the supplementary information). Based on self report, there were 509 non-Hispanic Caucasians, 122 African Americans, 296 Asians, 72 Latino/as, 2 Pacific Islanders, and 116 multiracial or other participants in this final sample.

Race/Ethnicity

Because self-reported race and ethnicity are not always an accurate reflection of genetic ancestry, an analysis of identity by state of whole-genome SNPs was performed in PLINK (30). The first four multidimensional scaling components were used as covariates to reduce possible confounding effects of ancestry. When data from non-Hispanic Caucasians and Asians, our two largest homogeneous subgroups, were analyzed separately, the first two multidimensional scaling components within each subgroup were used as covariates.

Genotyping

DNA was isolated from saliva derived from Oragene DNA self-collection kits (DNA Genotek) customized for 23 and Me (www.23andme.com). DNA extraction and genotyping were performed through 23 and Me by the National Genetics Institute (NGI), a CLIA-certified clinical laboratory and subsidiary of Laboratory Corporation of America. One of two different Illumina arrays with custom content was used to provide genome-wide SNP data, the Human Omni Express or Human Omni Express-24 (31–33). Further information regarding genotype imputation and proxy matching can be found in the supplementary information.

Hardy-Weinberg equilibrium was tested with each race/ethnic subgroup (supplementary table S4). The Asian group was split based on a K-means clustering analysis of its two multidimensional scaling components. After applying the FDR correction for multiple comparisons all SNPs were in Hardy-Weinberg equilibrium ($p > .001$).

Amygdala reactivity task

All participants completed an emotional face-processing task shown to consistently elicit robust amygdala reactivity in this and other samples (22, 24, 34, 35). For additional information on the task and preprocessing methods see the supplementary information.

Statistical analyses

Mplus version 7 (36) was used to conduct linear regression analyses with amygdala reactivity as the outcome, under the default option of maximum likelihood estimation with robust standard errors, which is robust to non-normality. Sex (1=males; 2=females), clinical diagnosis (0=no diagnosis; 1=one or more past or current diagnoses), MDS components, and age were included as covariates unless otherwise specified. Each SNP was coded based on the target association, so that higher values indicated an expected higher amygdala reactivity (i.e., 0=lowest; 1=intermediate; and 2=highest). In cases where higher reactivity was associated with homozygosity for one allele, this genotype was coded as 1 and all others were coded as 0. All analyses were conducted on BOLD parameter estimates extracted from functional clusters exhibiting a significant main effect of contrast in the right and left amygdala based on the entire sample. The mean of right and left values was used to test for associations with bilateral reactivity. Descriptive analyses, analysis of variance (ANOVA), and Fisher's Exact tests were conducted in SPSS version 19.

For 10 of the original reported associations, for which an exact replication attempt was feasible (e.g., same task contrast, race/ethnicity, diagnostic status), we further conducted voxel-wise regressions in SPM12. Here, we tested each of the 10 SNPs previously found to be significantly associated with amygdala reactivity to the contrast of angry and fearful > shapes in non-Hispanic Caucasians while controlling for sex, age, and 2 MDS components. These more lenient voxel-wise regressions were tested within our anatomically-defined amygdala ROIs at an uncorrected threshold of $p < 0.001$.

Defining a replication

A replication was considered successful when $p < .05$ and the association was exactly the same as previously reported, i.e., same allele/genotype associated with higher reactivity. Furthermore, a replication required the association to be present in individuals from the same ancestral population (i.e., Caucasians, Asians, or mixed race/ethnicity), sex, and diagnostic status in studies where a clinical sample was examined, as well as in the same hemisphere and with the same task contrast of the original finding. When such a replication was not feasible (e.g., the contrast used was not available in our data), we chose the most similar contrast that was available. For example, when an original report used the contrast of happy, sad, and angry > shapes we used angry, neutral, surprised, and fearful > shapes, or in the 3 cases of fear conditioning tasks we used all threat related contrasts.

When replication was unsuccessful, we conducted exploratory analyses using other threat-related contrasts (e.g., angry > shapes or fearful > neutral), testing left, right, and mean reactivity independently, including all participants while controlling for ancestry (in cases where the original finding was specific to Caucasians or Asians), excluding those with a diagnosis, and/or examining men and women separately. This was done in order to exhaust all possibilities and to enable other researchers to draw independent conclusions regarding reproducibility, as there is no clear definition of a successful replication in the literature.

When an original finding was replicated in the same direction, but with a different contrast and/or laterality, it was considered to be a “partial replication”, and when it was replicated for a different sex and/or race/ethnicity, it was categorized as significant with a “different population”. Associations in our sample that were in the opposite direction of the original finding (i.e., the other allele was associated with higher activation), were categorized as a separate group of “opposite direction” findings. Lastly, when a significant association was not achieved in any tested model, the replication attempt was categorized as unsuccessful.

Results

Descriptive statistics

Of the 37 associations tested, 22 were of variants with putative functionality based on prior studies of gene expression or transcription (supplementary table S2). Associations for 5 studies were based on fMRI data collected using 1.5T scanners (two were missing this information), 3 were based on non-face contrasts, and 23 were with unilateral reactivity (13 with the left hemisphere and 10 with the right). Of the 32 findings that included reports on race/ethnicity, 23 were based on Caucasian samples. Sample sizes ranged between 29 and 808, and participant mean age ranged between 11 and 46 years.

Replication analyses

The regression results for each replication attempt are presented in Table 1. Standardized estimates are reported to facilitate comparison of effect sizes across SNPs, but the p values are based on the unstandardized values. As shown in Figure 1, of the 37 original associations, only 3 were replicated: rs1391187 of the *FRAS1*-related extracellular matrix protein 3 (*FREM3*) gene (24), which was originally found in a subsample of the current

replication sample and should therefore be considered as a confirmation; rs17689966 of the corticotropin-releasing hormone/factor receptor 1 (CRHR1) gene, which was significant in panic disorder patients, and rs16944 of the interleukin-1b gene (37), which was significant in major depressive disorder patients.

Our exploratory analyses revealed significant associations for another 28 studies: 8 showed a significant association with a different contrast/laterality compared to the original finding; 5 were significant for a different sex and/or race/ethnicity; and 15 were significant in the opposite direction, so that the allele associated with higher reactivity in the original study was associated with lower reactivity in our sample, and were further different based on task contrast, laterality, sex, and/or race/ethnicity. There were no significant associations under any model parameters for the remaining 6 studies.

Finally, to confirm that our primary analytic method did not bias the findings, we also tested 10 original associations using a more lenient voxel-wise regression analysis. We found no evidence for replication even at an uncorrected threshold of $p < 0.001$.

Effect of study characteristics on reproducibility

To minimize the likelihood of empty cells, we used Fisher's Exact tests and created the following replication variable: 1=fully replicated and partially replicated with a different contrast/laterality ($N=11$); 2=not replicated or significant with a different population (ethnicity/gender; $N=11$); 3=significant in the opposite direction ($N=15$). There were no significant relationships between replication and a) laterality (0=unilateral, 1=bilateral; $p=.15$); b) if the SNP was genotyped directly (0=genotyped, 1=imputed/proxy; $p=.55$); or c) if the study was done on a subsample of the DNS (0=not DNS, 1=DNS; $p=.11$). After excluding previous DNS studies, which always included individuals with a clinical diagnosis, we also found that the inclusion of patients in the original study did not affect reproducibility (0=no patients; 1=included patients; $p=.23$). Nonetheless, it should be noted that other than a previous DNS study, 2 of the 3 fully replicated studies were based on patient samples. Additionally, we tested the effect of putative SNP functionality (0=no known functionality; 1=known functionality; $p=.053$), and found a trend for more replications (exact and different laterality/contrast) and more opposite direction findings for putatively functional SNPs.

Lastly, replication categories were not significantly different by sample size ($F(2, 34)=1.15$, $p=.33$; full replication and contrast/laterality: $M=251 \pm 275$; null and different population: $M=187 \pm 102$; opposite direction: $M=145 \pm 117$). However, we also tested the effect of sample size after excluding previous DNS studies, which were characterized by relatively large sample sizes. In this case, sample size did have a significant effect ($F(2, 25)=4.32$, $p=.024$; full replication and contrast/laterality: $M=43 \pm 16$; null and different population: $M=170 \pm 95$; opposite direction: $M=123 \pm 85$). As the variances were not homogenous based on the Levene's test ($F(2, 25)=4.63$, $p=.02$), the Games-Howell post-hoc test was used to compare the means of the sample sizes. Surprisingly the full replication and different contrast/laterality group was characterized by significantly smaller sample sizes compared with the null and different population group ($p=.017$) and the opposite direction group ($p=.011$). The

mean sample sizes of the studies in the two latter groups were not significantly different ($p=.50$).

Correction for Multiple Comparisons

For the 28 significant exploratory analyses (i.e., analyses with different model parameters than the original association) p values ranged between .01 and .49. None of these associations would survive correction for multiple comparisons, which necessitate correcting for 6 contrasts, 2 race/ethnicity groups, 3 hemisphere possibilities (left, right, mean), and 3 different sex groups (males, females, and combined sample). Even this is lenient, however, as comprehensive correction for all 37 SNPs requires a p value of $.05/37=.001$ for significance. In this case, none of the 37 SNPs would significantly be associated with amygdala reactivity.

Discussion

In the current study, we attempted to replicate 37 published associations between SNPs in candidate genes and threat-related amygdala reactivity, an intermediate brain phenotype implicated in the pathophysiology of and risk for mood and anxiety disorders. Generally, the replication rate was low with only 3 of the 37 SNPs exhibiting identical associations with amygdala reactivity in our sample. Exploratory analyses were conducted for the remaining 34 SNPs: 8 showed a significant association with a different contrast/laterality compared to the original finding; 5 were significant for a different sex and/or race/ethnicity; 15 were significant in the opposite direction, so that the allele associated with higher reactivity in the original study was associated with lower reactivity in our sample, and were further different based on task contrast, laterality, sex, and/or race/ethnicity; and 6 were not significant regardless of model parameters and the tested sample.

Interestingly, other than a study that was based on an earlier DNS subsample (24), the replicated associations were in patient samples (37, 38). Additionally, the only other SNP association derived from a sample of patients was also replicated in our sample, but with a different contrast (39). These findings suggest the possibility of higher reproducibility for imaging genetics findings derived from relatively homogenous samples. This interpretation is somewhat supported by our finding that studies with smaller samples were more likely to replicate or exhibit significant associations with different laterality and/or task contrast. However, as none of the associations would have survived multiple comparisons correction for the 37 SNPs tested, these patterns should be confirmed in further research before firm conclusions can be drawn.

Our finding of low reproducibility rates is in stark contrast to the critical, obligate nature of brain function in the emergence of both normal and abnormal behavior, as well as the more proximate position brain function holds to any molecular or cellular effects of genetic polymorphisms. Several factors may have influenced the low reproducibility we observed including small effect sizes, racial/ethnic as well as clinical heterogeneity between samples, different research paradigms, and different analytic methodologies. We briefly consider each below with an eye towards improving reproducibility.

Small effect sizes

Complex phenotypes such as brain function are likely affected by many common polymorphisms each with small effect sizes typically accounting for less than 1% of the phenotypic variance (40). While the genetic effects on intermediate phenotypes were expected to be larger, this does not appear to be the case with available brain function and structural volume phenotypes (41). For example, in a genome-wide association study of 7,612 participants from the UK Biobank, no significant associations were found with amygdala reactivity (42), suggesting that the effects of SNPs are very small and/or that epistasis and gene-environment interactions are at play. Consequently, detecting the effect of a particular SNP on quantitative phenotypes, including those derived from neuroimaging, is challenging and necessitates large sample sizes, the consideration of interacting factors, or a strong *a priori* hypothesis that might enable finding an effect with a relatively small target sample (e.g., specific patient samples with more homogeneous target phenotypes).

High heterogeneity between and within sampled populations

Allele frequencies of common genetic polymorphisms have a high degree of variability between populations. In the context of imaging genetics, the effects of such allele frequency differences may be exaggerated through interactions with cultural factors, which can contribute to variability in brain function including amygdala reactivity (43). Additional heterogeneity between samples may rise from difference in age range, sex distribution, and clinical status. Genetic influences (44) and brain function (45) change during development and therefore imaging genetics associations may be specific to a certain developmental window. Likewise, genetic effects (46, 47) and brain function (48) differ by sex. Lastly, the presence or absence of clinical disorder and accompanying pharmacologic treatment may affect reproducibility. For example, we were able to replicate some associations only when limiting analyses to subsamples with or without a clinical diagnosis.

Methodological differences

Even studies that attempt to answer similar questions often do so using different methods. In imaging genetics, such methodological differences include scanner hardware and software, DNA isolation and extraction, genotyping, behavioral activation task, image analyses, and the strategy for multiple comparisons correction. All of these can affect the findings and limit reproducibility. Neuroimaging results may further diverge between studies as a function of scan length (49), state mood (50), body-mass-index (51), as well as caffeine (52), nicotine (53), and alcohol consumption (54). Carefully documenting and, when possible, controlling for such variability may aid in improving reproducibility.

Defining a successful replication

One of the obstacles we encountered was how best to define what constitutes a successful replication. Does a successful replication necessitate a resampling of the same population? Use of the same analysis tools? Use of the same materials and tasks? Does laterality matter? If the finding was significant bilaterally in the replication, but it was only significant in the right amygdala in the target study, can it still be considered a successful replication? What if the result is significant but in a different direction? To our knowledge, there are no accepted

standards to guide these decisions. Low reproducibility rates can have a big impact on the field by creating an atmosphere of distrust and casting doubt on true findings. Consequently, devising reproducibility guidelines in each field should be a priority.

Increasing reproducibility

Several considerations can help improve reproducibility in imaging genetics, including 1) adopting a common core set of reliable neuroimaging assays (e.g., using the same tasks, contrasts, and preprocessing methods), as has been advanced for behavioral phenotypes (e.g., the PhenX toolkit, consensus measures for phenotypes and exposures; 55); 2) adopting similar extraction and genotyping protocols, and reporting on genotyping error-rates; 3) using the same strategy for multiple comparisons correction; 4) using larger samples; 5) limiting sample heterogeneity; 6) aggregating multiple SNPs, each responsible for only a small effect, to create polygenic scores (56). For such scores, SNPs could be aggregated based on genome-wide association studies of complex phenotypes (57, 58) or biologically relevant systems. To aid researchers in generating biologically informed polygenic scores based on the association studies reviewed herein, we provided brief details as to the putative functionality of each genetic variant in supplementary tables S2 and S3; 7) despite much criticism (59–61), application of gene-environment and gene-gene interaction analyses may still be beneficial to reproducibility in imaging genetics (59–61). For example, in the current analyses the direction of the effects of two SNPs (rs4680 and rs1800497) differed according to sex. The effect of common genetic variants on complex phenotypes is likely both subtle and dynamic, thus associations are unlikely to be constrained to direct, main effects. The interested reader is referred to Bogdan et al. (62) and to Carter et al., (63) for detailed reviews of these and other factors related to reproducibility in imaging genetics.

Limitations

Our study, of course, is not without limitations which can be addressed in future research. First, we only included studies that examined the main effects of SNPs on amygdala reactivity. It will be of interest to test whether gene-environment and gene-gene interaction effects or other forms of genetic variation (e.g., variable number tandem repeats) yield similar reproducibility rates. Second, our sample mostly comprised of healthy participants. We were not able to test for two associations that were found in clinical samples that are not well represented in the DNS (64, 65). It is possible that replication rates will be different when patient samples are examined separately. Third, there was high heterogeneity in the analytic methods used across the original studies, and most of them did not adopt the conservative approach we applied herein, which reduces Type I error and any correlation coefficient inflation that may result when an explanatory variable (i.e., genotype) is used to identify activation in voxel-based analyses (66). Additionally, the utility of a replication only under specific analytic methods is debatable, and even the most lenient voxel-wise analyses did not yield significant associations in our sample. Fourth, it is possible that some of the studies used additional or different covariates in their statistical analyses (e.g., added body mass index or scan time). Lastly, while our amygdala reactivity paradigm has been widely adopted and used extensively in imaging genetics, some of the target studies used different paradigms, such as fear conditioning or passive viewing of facial expressions, which may have further limited reproducibility.

Conclusions

The importance of replication lies in preventing false theories from taking root and biasing future research. From a practical standpoint, replicable findings will prevent the waste of valuable resources and allow better utility for the scientific community and general public (67). The low reproducibility rate found in the current study is consistent with the observation that the effects of common SNPs on intermediate neural phenotypes are small and comparable to those on other complex phenotypes (41). Notably, however, this does not mean we should abandon the use of intermediate phenotypes in genetics research. As others have suggested (41), the use of intermediate phenotypes is crucial to the understanding of the biological pathways through which genetic variation ultimately shapes behavior and risk for mental illness. This potential can be better realized by adopting uniform guidelines for imaging genetics research, such as utilizing standardized data collection and analysis; increasing homogeneity and size of samples; applying proper correction for multiple comparisons; controlling for possible confounders; and using more elaborate statistical models that better match the intricate nature of complex phenotypes (62). It is our hope that the current study will help hasten these processes and lead to further discoveries and reliable findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Duke Neurogenetics Study participants and the staff of the Laboratory of NeuroGenetics. The Duke Neurogenetics Study received support from Duke University as well as US-National Institutes of Health grants R01DA033369 and R01DA031579. RA, ARK, and ARH received further support from US-National Institutes of Health grant R01AG049789. MLE was supported by the National Science Foundation Graduate Research Fellowship under Grant No. NSF DGE-1644868.

References

1. Baker M. 2016; 1,500 scientists lift the lid on reproducibility. *Nature*. 533:452–454. [PubMed: 27225100]
2. Open Science Collaboration. 2015; Estimating the reproducibility of psychological science. *Science*. 349:aac4716. [PubMed: 26315443]
3. Prinz F, Schlange T, Asadullah K. 2011; Believe it or not: how much can we rely on published data on potential drug targets? *Nature reviews Drug discovery*. 10:712–712.
4. Luo X, Stavrakakis N, Penninx BW, Bosker FJ, Nolen WA, Boomsma DI, et al. 2016; Does refining the phenotype improve replication rates? A review and replication of candidate gene studies on major depressive disorder and chronic major depressive disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 171:215–236.
5. Prisciandaro J, Roberts J. 2009; A comparison of the predictive abilities of dimensional and categorical models of unipolar depression in the National Comorbidity Survey. *Psychol Med*. 39:1087–1096. [PubMed: 18845012]
6. Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, et al. 2013; A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 18:497–511. [PubMed: 22472876]
7. Hariri AR. 2009; The Neurobiology of Individual Differences in Complex Behavioral Traits. *Annu Rev Neurosci*. 32:225. [PubMed: 19400720]

8. Hariri AR, Weinberger DR. 2003; Imaging genomics. *Br Med Bull.* 65:259–270. [PubMed: 12697630]
9. Meyer-Lindenberg A, Weinberger DR. 2006; Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience.* 7:818–827. [PubMed: 16988657]
10. Munafo MR, Brown SM, Hariri AR. 2008; Serotonin transporter (5-HTTLPR) genotype and amygdala activation: a meta-analysis. *Biol Psychiatry.* 63:852–857. [PubMed: 17949693]
11. Bastiaansen JA, Servaas MN, Marsman JBC, Ormel J, Nolte IM, Riese H, et al. 2014; Filling the gap: relationship between the serotonin-transporter-linked polymorphic region and amygdala activation. *Psychol Sci.* 25:2058–2066. [PubMed: 25253281]
12. Murphy S, Norbury R, Godlewska B, Cowen P, Mannie Z, Harmer C, et al. 2013; The effect of the serotonin transporter polymorphism (5-HTTLPR) on amygdala function: a meta-analysis. *Mol Psychiatry.* 18:512–520. [PubMed: 22488255]
13. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, et al. 2014; The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav.* 8:153–182. [PubMed: 24399358]
14. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. 2015; UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12:e1001779. [PubMed: 25826379]
15. Holmes AJ, Hollinshead MO, O’Keefe TM, Petrov VI, Fariello GR, Wald LL, et al. 2015; Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific data.* 2:150031. [PubMed: 26175908]
16. Hariri, A. *Looking inside the disordered brain.* Sunderland, MA: Sinauer; 2015.
17. Mattson WI, Hyde LW, Shaw DS, Forbes EE, Monk CS. 2016; Clinical neuroprediction: Amygdala reactivity predicts depressive symptoms 2 years later. *Soc Cogn Affect Neurosci.* 11:892–898. [PubMed: 26865423]
18. Swartz JR, Knodt AR, Radtke SR, Hariri AR. 2015; A neural biomarker of psychological vulnerability to future life stress. *Neuron.* 85:505–511. [PubMed: 25654256]
19. McLaughlin KA, Busso DS, Duys A, Green JG, Alves S, Way M, et al. 2014; Amygdala response to negative stimuli predicts PTSD symptom onset following a terrorist attack. *Depress Anxiety.* 31:834–842. [PubMed: 24995938]
20. Admon R, Lubin G, Stern O, Rosenberg K, Sela L, Ben-Ami H, et al. 2009; Human vulnerability to stress depends on amygdala’s predisposition and hippocampal plasticity. *Proc Natl Acad Sci U S A.* 106:14120–14125. [PubMed: 19666562]
21. Hariri AR, Holmes A. 2015; Finding translation in stress research. *Nat Neurosci.* 18:1347–1352. [PubMed: 26404709]
22. Swartz JR, Prather AA, Di Iorio CR, Bogdan R, Hariri AR. 2016A Functional Interleukin-18 Haplotype Predicts Depression and Anxiety through Increased Threat-Related Amygdala Reactivity in Women but Not Men. *Neuropsychopharmacology.*
23. Swartz JR, Waller R, Bogdan R, Knodt AR, Sabhlok A, Hyde LW, et al. 2015A common polymorphism in a Williams syndrome gene predicts amygdala reactivity and extraversion in healthy adults. *Biol Psychiatry.*
24. Nikolova YS, Iruku SP, Lin C-W, Conley ED, Puralewski R, French B, et al. 2015FRAS1-related extracellular matrix 3 (FREM3) single-nucleotide polymorphism effects on gene expression, amygdala reactivity and perceptual processing speed: An accelerated aging pathway of depression risk. *Frontiers in psychology.* :6. [PubMed: 25759672]
25. Nelson EC, Agrawal A, Heath AC, Bogdan R, Sherva R, Zhang B, et al. 2016; Evidence of CNH3 involvement in opioid dependence. *Mol Psychiatry.* 21:608. [PubMed: 26239289]
26. Waller R, Corral-Frias NS, Vannucci B, Bogdan R, Knodt AR, Hariri AR, et al. 2016An oxytocin receptor polymorphism predicts amygdala reactivity and antisocial behavior in men. *Soc Cogn Affect Neurosci.* :nsw042.
27. Lohoff FW, Hodge R, Narasimhan S, Nall A, Ferraro TN, Mickey BJ, et al. 2014; Functional genetic variants in the vesicular monoamine transporter 1 modulate emotion processing. *Mol Psychiatry.* 19:129–139. [PubMed: 23337945]

28. Arloth J, Bogdan R, Weber P, Frishman G, Menke A, Wagner KV, et al. 2015; Genetic differences in the immediate transcriptome response to stress predict risk-related brain function and psychiatric disorders. *Neuron*. 86:1189–1202. [PubMed: 26050039]
29. Savage JE, Sawyers C, Roberson-Nay R, Hettema JM. 2016; The genetics of anxiety-related negative valence system traits. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 174:156–177.
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. 2007; PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*. 81:559–575. [PubMed: 17701901]
31. Do CB, Tung JY, Dorfman E, Kiefer AK, Drabant EM, Francke U, et al. 2011; Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet*. 7:e1002141. [PubMed: 21738487]
32. Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S, et al. 2010; Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet*. 6:e1000993. [PubMed: 20585627]
33. Tung JY, Do CB, Hinds DA, Kiefer AK, Macpherson JM, Chowdry AB, et al. 2011; Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS ONE*. 6:e23473. [PubMed: 21858135]
34. Zhou Z, Zhu G, Hariri AR, Enoch M-A, Scott D, Sinha R, et al. 2008; Genetic variation in human NPY expression affects stress response and emotion. *Nature*. 452:997–1001. [PubMed: 18385673]
35. Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, et al. 2002; Serotonin transporter genetic variation and the response of the human amygdala. *Science*. 297:400–403. [PubMed: 12130784]
36. Muthén, LK, Muthén, BO. *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén; 2007.
37. Baune BT, Dannlowski U, Domschke K, Janssen DG, Jordan MA, Ohrmann P, et al. 2010; The interleukin 1 beta (IL1B) gene is associated with failure to achieve remission and impaired emotion processing in major depression. *Biol Psychiatry*. 67:543–549. [PubMed: 20044070]
38. Weber H, Richter J, Straube B, Lueken U, Domschke K, Scharfner C, et al. 2016; Allelic variation in CRHR1 predisposes to panic disorder: evidence for biased fear processing. *Mol Psychiatry*. 21:813. [PubMed: 26324098]
39. Domschke K, Dannlowski U, Hohoff C, Ohrmann P, Bauer J, Kugel H, et al. 2010; Neuropeptide Y (NPY) gene: impact on emotional processing and treatment response in anxious depression. *Eur Neuropsychopharmacol*. 20:301–309. [PubMed: 19854625]
40. Plomin R, DeFries JC, Knopik VS, Neiderhiser JM. 2016; Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*. 11:3–23. [PubMed: 26817721]
41. Munafo MR, Flint J. 2014; The genetic architecture of psychophysiological phenotypes. *Psychophysiology*. 51:1331–1332. [PubMed: 25387716]
42. Elliott L, Sharp K, Alfaro-Almagro F, Douaud G, Miller K, Marchini J, et al. 2017 The genetic basis of human brain structure and function: 1,262 genome-wide associations found from 3,144 GWAS of multimodal brain imaging phenotypes from 9,707 UK Biobank participants. *bioRxiv*. : 178806.
43. Derntl B, Habel U, Robinson S, Windischberger C, Kryspin-Exner I, Gur RC, et al. 2009; Amygdala activation during recognition of emotions in a foreign ethnic group is associated with duration of stay. *Soc Neurosci*. 4:294–307. [PubMed: 19479638]
44. Casey B, Glatt CE, Tottenham N, Soliman F, Bath K, Amso D, et al. 2009; Brain-derived neurotrophic factor as a model system for examining gene by environment interactions across development. *Neuroscience*. 164:108–120. [PubMed: 19358879]
45. Blakemore SJ, Choudhury S. 2006; Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*. 47:296–312. [PubMed: 16492261]
46. Mirowska-Guzel D, Mach A, Gromadzka G, Czlonkowski A, Czlonkowska A. 2008; BDNF A196G and C270T gene polymorphisms and susceptibility to multiple sclerosis in the polish population. Gender differences. *J Neuroimmunol*. 193:170–172. [PubMed: 18061279]

47. Shalev I, Lerer E, Israel S, Uzevovsky F, Gritsenko I, Mankuta D, et al. 2009; BDNF Val66Met polymorphism is associated with HPA axis reactivity to psychological stress characterized by genotype and gender interactions. *Psychoneuroendocrinology*. 34:382–388. [PubMed: 18990498]
48. Cosgrove KP, Mazure CM, Staley JK. 2007; Evolving knowledge of sex differences in brain structure, function, and chemistry. *Biol Psychiatry*. 62:847–855. [PubMed: 17544382]
49. Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, Kirk GR, et al. 2013; The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *NeuroImage*. 83:550–558. [PubMed: 23747458]
50. Harrison BJ, Pujol J, Ortiz H, Fornito A, Pantelis C, Yucel M. 2008; Modulation of brain resting-state networks by sad mood induction. *PLoS One*. 3:e1794. [PubMed: 18350136]
51. Willeumier KC, Taylor DV, Amen DG. 2011; Elevated BMI is associated with decreased blood flow in the prefrontal cortex using SPECT imaging in healthy adults. *Obesity*. 19:1095–1097. [PubMed: 21311507]
52. Perthen JE, Lansing AE, Liao J, Liu TT, Buxton RB. 2008; Caffeine-induced uncoupling of cerebral blood flow and oxygen metabolism: a calibrated BOLD fMRI study. *NeuroImage*. 40:237–247. [PubMed: 18191583]
53. Karama S, Ducharme S, Corley J, Chouinard-Decorte F, Starr JM, Wardlaw JM, et al. 2015; Cigarette smoking and thinning of the brain's cortex. *Mol Psychiatry*. 20:778. [PubMed: 25666755]
54. Luchtman M, Jachau K, Tempelmann C, Bernarding J. 2010; Alcohol induced region-dependent alterations of hemodynamic response: implications for the statistical interpretation of pharmacological fMRI studies. *Experimental Brain Research*. 204:1–10. [PubMed: 20502888]
55. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. 2011; The PhenX Toolkit: get the most from your measures. *Am J Epidemiol*. 174:253–260. [PubMed: 21749974]
56. Bogdan R, Hyde L, Hariri A. 2013; A neurogenetics approach to understanding individual differences in brain, behavior, and risk for psychopathology. *Mol Psychiatry*. 18:288–299. [PubMed: 22614291]
57. Whalley H, Pampmeyer M, Sprooten E, Romaniuk L, Blackwood D, Glahn D, et al. 2012; The influence of polygenic risk for bipolar disorder on neural activation assessed using fMRI. *Translational Psychiatry*. 2:e130. [PubMed: 22760554]
58. Holmes AJ, Lee PH, Hollinshead MO, Bakst L, Roffman JL, Smoller JW, et al. 2012; Individual differences in amygdala-medial prefrontal anatomy link negative affect, impaired social functioning, and polygenic depression risk. *J Neurosci*. 32:18087–18100. [PubMed: 23238724]
59. Duncan LE, Keller MC. 2011; A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry. *The American journal of psychiatry*. 168:1041. [PubMed: 21890791]
60. Eaves L, Verhulst B. 2014; Problems and pit-falls in testing for G×E and epistasis in candidate gene studies of human behavior. *Behav Genet*. 44:578. [PubMed: 25195167]
61. Keller MC. 2014; Genex environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol Psychiatry*. 75:18–24. [PubMed: 24135711]
62. Bogdan R, Salmeron B, Carey C, Agrawal A, Calhoun V, Garavan H, et al. 2017; Imaging Genetics and Genomics in Psychiatry: A Critical Review of Progress and Potential. *Biol Psychiatry*. 82:165–175. [PubMed: 28283186]
63. Carter CS, Bearden CE, Bullmore ET, Geschwind DH, Glahn DC, Gur RE, et al. 2017; Enhancing the Informativeness and Replicability of Imaging Genomics Studies. *Biol Psychiatry*. 82:157–164. [PubMed: 27793332]
64. Stevens JS, Almli LM, Fani N, Gutman DA, Bradley B, Norrholm SD, et al. 2014; PACAP receptor gene polymorphism impacts fear responses in the amygdala and hippocampus. *Proc Natl Acad Sci U S A*. 111:3158–3163. [PubMed: 24516127]
65. Haram M, Bettella F, Brandt CL, Quintana DS, Nerhus M, Bjella T, et al. 2016; Contribution of oxytocin receptor polymorphisms to amygdala activation in schizophrenia spectrum disorders. *British Journal of Psychiatry Open*. 2:353–358. [PubMed: 27847593]

66. Viviani R. 2010; Unbiased ROI selection in neuroimaging studies of individual differences. *Neuroimage*. 50:184–189. [PubMed: 19900563]
67. Arrowsmith J. 2011; Trial watch: Phase II failures: 2008–2010. *Nat Rev Drug Discov*. 10:328–329.
68. Chang H, Hoshina N, Zhang C, Ma Y, Cao H, Wang Y, et al. 2017 The protocadherin 17 gene affects cognition, personality, amygdala structure and function, synapse development and risk of major mood disorders. *Mol Psychiatry*.
69. Fakra E, Hyde LW, Gorka A, Fisher PM, Munoz KE, Kimak M, et al. 2009; Effects of HTR1A C (– 1019) G on amygdala reactivity and trait anxiety. *Arch Gen Psychiatry*. 66:33–40. [PubMed: 19124686]
70. Dannlowski U, Kugel H, Franke F, Stuhrmann A, Hohoff C, Zwanzger P, et al. 2011; Neuropeptide-S (NPS) receptor genotype modulates basolateral amygdala responsiveness to aversive stimuli. *Neuropsychopharmacology*. 36:1879–1885. [PubMed: 21525857]
71. Smoller JW, Paulus MP, Fagerness JA, Purcell S, Yamaki LH, Hirshfeld-Becker D, et al. 2008; Influence of RGS2 on anxiety-related temperament, personality, and brain function. *Arch Gen Psychiatry*. 65:298–308. [PubMed: 18316676]
72. Smoller JW, Gallagher PJ, Duncan LE, McGrath LM, Haddad SA, Holmes AJ, et al. 2014; The human ortholog of acid-sensing ion channel gene ASIC1a is associated with panic disorder and amygdala structure and function. *Biol Psychiatry*. 76:902–910. [PubMed: 24529281]
73. Hariri AR, Gorka A, Hyde LW, Kimak M, Halder I, Ducci F, et al. 2009; Divergent effects of genetic variation in endocannabinoid signaling on human threat-and reward-related brain function. *Biol Psychiatry*. 66:9–16. [PubMed: 19103437]
74. Bogdan R, Williamson DE, Hariri AR. 2012; Mineralocorticoid receptor Iso/Val (rs5522) genotype moderates the association between previous childhood emotional neglect and amygdala reactivity. *Am J Psychiatry*. 169:515–522. [PubMed: 22407082]
75. Neumann SA, Brown SM, Ferrell RE, Flory JD, Manuck SB, Hariri AR. 2006; Human choline transporter gene variation is associated with corticolimbic reactivity and autonomic-cholinergic function. *Biol Psychiatry*. 60:1155–1162. [PubMed: 16876130]
76. Holz NE, Buchmann AF, Boecker R, Blomeyer D, Baumeister S, Wolf I, et al. 2015; Role of FKBP5 in emotion processing: results on amygdala activity, connectivity and volume. *Brain Struct Funct*. 220:1355–1368. [PubMed: 24756342]
77. Tost H, Kolachana B, Hakimi S, Lemaitre H, Verchinski BA, Mattay VS, et al. 2010; A common allele in the oxytocin receptor gene (OXTR) impacts prosocial temperament and human hypothalamic-limbic structure and function. *Proc Natl Acad Sci U S A*. 107:13936–13941. [PubMed: 20647384]
78. Redlich R, Stacey D, Opel N, Grotegerd D, Dohm K, Kugel H, et al. 2015; Evidence of an IFN- γ by early life stress interaction in the regulation of amygdala reactivity to emotional stimuli. *Psychoneuroendocrinology*. 62:166–173. [PubMed: 26313134]
79. Joeyen-Waldorf J, Nikolova YS, Edgar N, Walsh C, Kota R, Lewis DA, et al. 2012; Adenylate cyclase 7 is implicated in the biology of depression and modulation of affective neural circuitry. *Biol Psychiatry*. 71:627–632. [PubMed: 22264442]
80. Tesli M, Skatun KC, Ousdal OT, Brown AA, Thoresen C, Agartz I, et al. 2013; CACNA1C risk variant and amygdala activity in bipolar disorder, schizophrenia and healthy controls. *PLoS ONE*. 8:e56970. [PubMed: 23437284]
81. Ousdal OT, Brown AA, Jensen J, Nakstad PH, Melle I, Agartz I, et al. 2012; Associations between variants near a monoaminergic pathways gene (PHOX2B) and amygdala reactivity: a genome-wide functional imaging study. *Twin Res Hum Genet*. 15:273–285. [PubMed: 22856363]
82. Blasi G, Bianco LL, Taurisano P, Gelao B, Romano R, Fazio L, et al. 2009; Functional variation of the dopamine D2 receptor gene is associated with emotional control as well as brain activity and connectivity during emotion processing in humans. *J Neurosci*. 29:14812–14819. [PubMed: 19940176]
83. Kilpatrick LA, Labus JS, Coveleskie K, Hammer C, Rappold G, Tillisch K, et al. 2011; The HTR3A polymorphism c.-42C> T is associated with amygdala responsiveness in patients with irritable bowel syndrome. *Gastroenterology*. 140:1943–1951. [PubMed: 21420406]

84. Lee B-T, Lee H-Y, Han C, Pae C-U, Tae WS, Lee M-S, et al. 2011; DRD2/ANKK1 TaqI A polymorphism affects corticostriatal activity in response to negative affective facial stimuli. *Behav Brain Res.* 223:36–41. [PubMed: 21510979]
85. Canli T, Congdon E, Gutknecht L, Constable R, Lesch K. 2005; Amygdala responsiveness is modulated by tryptophan hydroxylase-2 gene variation. *J Neural Transm.* 112:1479–1485. [PubMed: 16245070]
86. Woudstra S, Bochdanovits Z, Van Tol M, Veltman D, Zitman F, Van Buchem M, et al. 2012; Piccolo genotype modulates neural correlates of emotion processing but not executive functioning. *Translational Psychiatry.* 2:e99. [PubMed: 22832909]
87. Wellman CL, Camp M, Jones VM, MacPherson KP, Ihne J, Fitzgerald P, et al. 2013; Convergent effects of mouse Pet-1 deletion and human PET-1 variation on amygdala fear and threat processing. *Exp Neurol.* 250:260–269. [PubMed: 24100022]
88. Marusak HA, Furman DJ, Kuruvadi N, Shattuck DW, Joshi SH, Joshi AA, et al. 2015; Amygdala responses to salient social cues vary with oxytocin receptor genotype in youth. *Neuropsychologia.* 79:1–9. [PubMed: 26477647]
89. Liu X, Akula N, Skup M, Brotman MA, Leibenluft E, McMahon FJ. 2010; A genome-wide association study of amygdala activation in youths with and without bipolar disorder. *J Am Acad Child Adolesc Psychiatry.* 49:33–41. [PubMed: 20215924]
90. Domschke K, Baune BT, Havlik L, Stuhrmann A, Suslow T, Kugel H, et al. 2012; Catechol-O-methyltransferase gene variation: impact on amygdala response to aversive stimuli. *Neuroimage.* 60:2222–2229. [PubMed: 22387174]
91. Ridder S, Treutlein J, Nees F, Lang S, Diener S, Wessa M, et al. 2012; Brain activation during fear conditioning in humans depends on genetic variations related to functioning of the hypothalamic–pituitary–adrenal axis: first evidence from two independent subsamples. *Psychol Med.* 42:2325–2335. [PubMed: 22410078]
92. Soliman F, Glatt CE, Bath KG, Levita L, Jones RM, Pattwell SS, et al. 2010; A genetic variant BDNF polymorphism alters extinction learning in both mouse and human. *Science.* 327:863–866. [PubMed: 20075215]

Figure 1.

Reproducibility in imaging genetics of threat-related amygdala reactivity.

Note. Yes - full replication; Different contrast or laterality - partial replication. Significant with a change in fMRI task contrast and/or hemispheric laterality; Different population: association was significant in a different gender and/or ethnicity; No - was not replicated in any tested model ; and opposite direction - the allele/genotype that was previously associated with higher reactivity was associated with lower reactivity in our sample (most of these associations also differed in contrast, laterality, and/or ethnicity).

Table 1

Replication results of 37 unique associations between single nucleotide polymorphisms and threat-related amygdala reactivity.

Target Study	Sex (ethnicity) ^a	Scanner	Significant contrast	Gene	Allele/Genotype (laterality) [*]	If replicated	B and SE
Chang et al., 2017 (68)	F+M (1)	3T	angry and fearful > shapes	PCDH17 (rs9537793)	G (right)	Different contrast and direction	N=509, $\beta=-.05$, SE=.04, $p=.27$. But significant in fear > shapes (N=509, $\beta=-.08$, SE=.04, $p=.048$).
Fakra et al., 2009 (69)	F+M (1)	3T	angry and fearful > shapes	HTR1A (rs6295, Used proxy rs358527)	C (minor allele, Bilateral)	Different contrast and direction	N=506, $\beta=-.06$, SE=.04, $p=.18$. But significant in angry > shapes ($\beta=-.10$, SE=.04, $p=.019$).
Dannowski et al., 2011 (70)	F+M (1)	3T	angry and fearful > shapes	NPSR1 (rs324981, Used proxy rs324987)	T (minor allele, right)	Different contrast, laterality, and direction	N=509, $\beta=-.003$, SE=.04, $p=.94$. But mean amygdala reactivity significant for the whole sample in angry and fearful > neutral (N=1117, $\beta=-.06$, SE=.03, $p=.037$).
Smoller et al., 2008 (71)	F+M (N/A)	1.5T+3T	happy, sad, and angry > shapes	Rgs2 (rs4606, Used proxy rs3767488)	G (minor allele, left)	Different gender and direction	Not significant in faces > shapes ^b (N=1111, $\beta=.008$, SE=.03, $p=.79$) or in angry > shapes (N=1111, $\beta=-.01$, SE=.03, $p=.73$). Angry > shapes was significant in females without a diagnosis in various ethnicities (N=510, $\beta=-.11$, SE=.05, $p=.023$) and in Caucasians (N=213, $\beta=-.12$, SE=.06, $p=.048$).
Smoller et al., 2014 (72)	F+M (1)	3T	angry and fearful > shapes	ASIC1/ACCN2 (rs10875995)	C (bilateral)	Different contrast and ethnicity	N=509, $\beta=.01$, SE=.04, $p=.79$. Significant in the whole sample in angry and fearful > neutral (N=1117, $\beta=.07$, SE=.03, $p=.031$).
Haerri et al., 2009 (73)	F+M (1)	3T	angry and fearful > shapes	FAAH (rs324420)	CC (bilateral)	Different ethnicity and direction	Not significant in Caucasians only (N=509, $\beta=-.06$, SE=.04, $p=.17$), but

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Bogdan et al., 2012 (74)	F+M (3)	3T	angry and fearful > shapes	NR3C2 (rs5522)	C (right)	Different contrast, laterality, gender, and direction.	significant in the whole sample (N=1117, $\beta = -.06$, SE=.03, p=.039). In Caucasians, angry > shapes (N=509, $\beta = -.086$, SE=.04, p=.042).
Neumann et al., 2006 (75)	F+M (1)	3T	angry and fearful > shapes	CHT1 (rs333229, Imputed)	GG (right)	Different contrast, ethnicity, and direction	N=1117, $\beta = -.04$, SE=.03, p=.15. Significant in males, bilateral amygdala angry and fearful > neutral (N=488, $\beta = -.09$, SE=.047, p=.049).
Holz et al., 2015 (76)	F+M (1)	3T	angry and fearful > shapes	FKBP5 (rs1360780)	T (left)	Different contrast, laterality, and direction	N=492, $\beta = -.03$, SE=.04, p=.53. But significant in the whole sample in angry > shapes (N=1082, $\beta = -.06$, SE=.03, p=.045).
Tost et al., 2010 (77)	F+M (1)	3T	angry and fearful > shapes	OXTR (rs53576)	G (bilateral)	No	N=509, $\beta = -.03$, SE=.04, p=.41. But significant bilaterally in fearful > neutral (N=509, $\beta = -.09$, SE=.04, p=.03).
Redlich et al., 2015 (78)	F+M (1)	3T	angry and fearful > shapes	IFN- γ (rs1861494)	C (right)	Different contrast and gender	N=509, $\beta = -.01$, SE=.04, p=.81. Significant in females, angry and fearful > neutral (N=268, $\beta = .12$, SE=.05, p=.03).
Joeyen-Waldorf et al., 2012 (79)	F+M (1)	3T	angry and fearful > shapes	ADCY7 (rs1064448)	T (left and bilateral)	No	LEFT: N=508, $\beta = .008$, SE=.04, p=.85; BIL/ATERAL: $\beta = .000$, SE=.04, p=.99.
Tesli et al., 2013 (80)	F+M (1)	1.5T	angry and fearful > shapes	CACNA1C (rs1006737)	A (left)	No	N=509, $\beta = -.01$, SE=.04, p=.76. Also separately tested BD patients, various ethnicities (N=32, $\beta = -.30$, SE=.21, p=.16).
Ousdal et al., 2012 (81)	F+M (N/A)	1.5T	angry and fearful > shapes	PHOX2B (rs10014254, Imputed)	T (left)	Different contrast, laterality, and direction	N=507, $\beta = -.04$, SE=.04, p=.30. Significant in the

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Blasi et al., 2009 (82)	F+M (1)	3T	Greater amygdala reactivity during implicit, but not explicit, processing. angry, fearful, happy, and neutral > fixation crosshair	DRD2 (rs1076560)	GG (left)	Different direction	right amygdala, angry and fearful > neutral (N=507, $\beta=-.10$, SE=.04, p=.017).
Kilpatrick et al., 2011 (83)	F (N/A)	3T	Greater amygdala response to angry and fearful > implicit baseline and shapes > implicit baseline, but not in angry and fearful > shapes	HTR3A (rs1062613)	CC (bilateral in shapes and right in faces)	Different contrast and direction	In Caucasians, not significant in females in angry and fearful > shapes (N=268, $\beta=-.056$, SE=.06, p=.37). Significant in females, and in both sexes in faces > shapes (N=509, $\beta=-.09$, SE=.04, p=.04). In the whole sample significant in left amygdala angry > neutral (N=1117, $\beta=-.06$, SE=.03, p=.04).
Lee et al., 2011 (84)	F (2)	1.5T	angry and sad > neutral	DRD2/ANKK1 (rs1800497)	Homozygotes higher than heterozygotes (bilateral)	Different contrast	Not significant in angry > neutral (N=151, $\beta=-.05$, SE=.08, p=.53). Significant in females (N=151, $\beta=-.17$, SE=.08, p=.038) and in males (N=122, $\beta=-.21$, SE=.08, p=.01) in fearful > shapes.
Canli et al., 2005 (85)	F+M (N/A)	3T	fear > neutral, sad > neutral; and happy > neutral	TPH2 (rs4570625)	T (bilateral)	Different contrast	Caucasians. Not significant in fearful > neutral (N=509, $\beta=-.02$, SE=.04, p=.67). Significant in angry and fearful > shapes (N=509, $\beta=-.09$, SE=.04, p=.040).
Woudstra et al., 2012 (86)	F+M (1)	3T	angry and fearful > neutral	PCLO (rs2522833)	C (left)	Different contrast and ethnicity	N=509, $\beta=-.02$, SE=.04, p=.56. But significant in MDDs, bilaterally, various

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Zhou et al., 2008 (34)	F+M (1)	3T	angry and fearful > shapes	NPY (Haplotypes of 6 SNPs: rs3037354, rs17149106, rs16147, rs16139, rs5573, rs5574)	Low expression diplotypes (right)	Different contrast, ethnicity, and gender	ethnicities in angry and fearful > shapes (N=63, $\beta=34$, SE=12, $p=.017$) and in the whole sample in angry > shapes (N=1117, $\beta=.06$, SE=.03, $p=.043$).
Domschke et al., 2010 (39)	F+M (1)	3T	masked angry > neutral	NPY (rs16147)	C (bilateral)	Different contrast	N=199, $\beta=.05$, SE=.07, $p=.48$. Significant in females, various ethnicities, fear > shapes (N=196, $\beta=.15$, SE=.07, $p=.04$).
Baune et al., 2010 (37)	F+M (1)	3T	angry, sad, and happy > neutral	IL-1 β (rs16944)	A (left)	Yes	In MDD patients, N=27, $\beta=.11$, SE=.19, $p=.57$. Significant in fearful > neutral (N=27, $\beta=.39$, SE=.13, $p=.001$). A trend in anger > neutral in the whole sample (N=1110, $\beta=-.06$, SE=.03, $p=.050$).
Wellman et al., 2013(87)	F+M (2)	3T	angry and fearful > shapes	PET1 (rs860573)	A (AA very rare so combined with AG; bilateral)	No	In MDD patients, angry and fearful > neutral N=27, $\beta=.50$, SE=.16, $p=.006$. Also MDDs of various ethnicities (N=63, $\beta=.41$, SE=.12, $p=.001$).
Nelson et al., 2016 (25)	F+M (1)	3T	blunted amygdala habituation	CNIH3 (rs10799590)	GG (A allele was associated with blunted habituation, which is a related to resilience. right)	Different contrast and laterality	Not significant in habituation (N=509, $\beta=-.01$, SE=.04, $p=.72$), but significant bilaterally in fear > neutral in Caucasians (N=509, $\beta=.10$, SE=.04, $p=.025$) and in the whole sample (N=1117, $\beta=.08$, SE=.03, $p=.016$).
Waller et al., 2016 (26)	M (1)	3T	Angry > neutral	OXTR (rs1042778)	TT (only in males, right)	Different ethnicity, genotype, and coding	N=241, $\beta=.08$, SE=.07, $p=.29$. Significant in males without a diagnosis in various ethnicities with

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Swartz et al., 2015 (23)	F+M (3)	3T	angry and fearful > shapes	GTF21 (rs13227433, Imputed)	C (bilateral)	Different contrast and laterality	additive coding, T as risk (N=374, $\beta=-.14$, SE=.06, $p=.028$). N=1102, $\beta=-.03$, SE=.03, $p=.38$. Significant in the left amygdala in angry and fearful > neutral (N=1102, $\beta=-.06$, SE=.03, $p=.041$).
Swartz et al., 2016 (22)	F (1)	3T	angry and fearful > shapes	IL-18 (rs187238, used proxy rs795467, and rs1946518)	CG haplotype (CC in rs187238 and GG in rs1946518). Only in females, left centromedial.	Different gender, laterality, and direction	N=268, $\beta=-.05$, SE=.06, $p=.40$. Significant in males and females bilaterally, whole amygdala, in angry and fearful > shapes (N=509, $\beta=-.09$, SE=.04, $p=.036$).
Lohoff et al., 2014 (27)	F+M (3)	3T	angry, fearful, surprised, and neutral > shapes	VMAT1 (rs1390938, Imputed)	A (left)	No	N=1110, $\beta=-.01$, SE=.03, $p=.64$. Trend in participants without a diagnosis, various ethnicities (N=880, $\beta=-.06$, SE=.04, $p=.07$).
Nikolova et al., 2015 (24)	F+M (1)	3T	angry, fearful, surprised, and neutral > shapes	FREM3 (rs7676614)	G (bilateral)	Different laterality	N=509. Not significant in both hemispheres ($\beta=-.07$, SE=.04, $p=.10$), but significant in the left ($\beta=-.10$, SE=.046, $p=.029$).
Nikolova et al., 2015 (24)	F+M (1)	3T	angry, fearful, surprised, and neutral > shapes	FREM3 (rs1391187, Imputed)	A (bilateral)	Yes (confirmed)	N=508, $\beta=-.14$, SE=.04, $p=.002$.
Arloth et al., 2015 (28)	F+M (3)	3T	angry and fearful > neutral (was not significant in the angry and fearful > shapes contrast)	19 SNPs that are associated with the genetic regulation of GR stimulated gene expression	Left centromedial	Different contrast and laterality	N=1117, $\beta=-.01$, SE=.03, $p=.71$. Significant in right centromedial amygdala, angry and fearful > shapes (N=1117, $\beta=-.06$, SE=.03, $p=.043$).
Marusak et al., 2015 (88)	F+M (3)	3T	neutral > shapes (not with fear, angry, or happy).	OXTR (rs2254298)	A (bilateral)	Different contrast, laterality, and direction	Trend in neutral > shapes (N=1117, $\beta=-.06$, SE=.03, $p=.056$). Significant in the right amygdala in angry and fearful > neutral (N=1117, $\beta=-$

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Liu et al., 2010 (89)	F+M (mostly 1)	3T	Perceived threat > nose width (hostile > nose width)	DOK5 (rs2023454, Used proxy rs1699924)	GG (major allele, Right)	No	-06, SE=-03, p=-041). angry and fearful > shapes (N=509, $\beta=-02$, SE=-04, p=.69).
Domschke et al., 2012 (90)	F+M (1)	3T	angry and fearful > shapes	COMT (rs4680)	Val (left, Driven mostly by females)	Different contrast and direction	N=509, $\beta=-003$, SE=-04, p=.94. Also not significant in females (N=268, $\beta=-.10$, SE=-06, p=.12). Significant in females, left amygdala, fear > neutral (N=268, $\beta=-.12$, SE=-06, p=.04), and in males, right amygdala, angry and fearful > shapes (N=241, $\beta=-.14$, SE=-066, p=.03). In females, various ethnicities, angry and fearful > shapes was significant (N=629, $B=-.11$, SE=-04, p=.01).
Findings with a non-face contrast							
Ridder et al., 2012 (91)	F+M (N/A)	1.5T	CS+ (danger signal) > CS- (safety signal) during early acquisition	NR3C1 (Carriers of 3 or more minor alleles vs all others: rs33389, rs41423247 used proxy rs853180, rs4986593)	0-no minor alleles; 1-one or more minor alleles (left)	Different direction	Not significant in angry and fearful > shapes (N=509, $\beta=-02$, SE=-04, p=.64), but significant angry and fearful > neutral in Caucasians and in the whole sample (N=509, $\beta=-.09$, SE=-04, p=.03; N=1117, $\beta=-.06$, SE=-03, p=.043, respectively).
Soliman et al., 2010 (92)	F+M (3)	3T	Greater activation during fear conditioning extinction	BDNF (rs6265)	T allele (Met allele, Left)	Different laterality	Not significant in angry and fearful > shapes (N=1117, $\beta=-04$, SE=-04, p=.21). Significant bilaterally in angry > shapes (N=1117, $\beta=-09$, SE=-035, p=.015). In Caucasians, significant in the right amygdala angry and

Target Study	Sex (ethnicity) [^]	Scanner	Significant contrast	Gene	Allele/Genotype (laterality)*	If replicated	B and SE
Weber et al., 2015 (38)	F+M (1)	N/A	CS- > baseline\fixation cross (safety signal presentation)	CRHR1 (rs17689966, Imputed)	G (right)	Yes	fearful > shapes (N=509, $\beta=-.11$, SE=.05, $p=.040$). In panic disorder patients, not significant in angry and fearful > shapes (N=22, $\beta=-.24$, SE=.17, $p=.13$). Significant in fearful > shapes (N=22, $\beta=-.38$, SE=.13, $p=.004$). Also significant bilaterally in females in fearful > neutral (N=262, $\beta=-.14$, SE=.06, $p=.022$).

[^] Ethnicity: 1-Caucasians; 2-Asians; 3-Various ethnicities; F-females; M-males; CS-conditional stimulus.

* The allele or genotype that was associated with higher amygdala reactivity.

[^] Our faces > shapes contrasts translates to angry, fearful, surprised, and neutral > shapes.