



# Infant inhibited temperament in primates predicts adult behavior, is heritable, and is associated with anxiety-relevant genetic variation

Andrew S. Fox<sup>1,2</sup> · Ronald A. Harris<sup>3</sup> · Laura Del Rosso<sup>2</sup> · Muthuswamy Raveendran<sup>3</sup> · Shawn Kamboj<sup>1</sup> · Erin L. Kinnally<sup>1,2</sup> · John P. Capitanio<sup>1,2</sup> · Jeffrey Rogers<sup>3</sup>

Received: 30 October 2020 / Revised: 23 April 2021 / Accepted: 4 May 2021  
© The Author(s) 2021. This article is published with open access

## Abstract

An anxious or inhibited temperament (IT) early in life is a major risk factor for the later development of stress-related psychopathology. Starting in infancy, nonhuman primates, like humans, begin to reveal their temperament when exposed to novel situations. Here, in Study 1 we demonstrate this infant IT predicts adult behavior. Specifically, in over 600 monkeys, we found that individuals scored as inhibited during infancy were more likely to refuse treats offered by potentially-threatening human experimenters as adults. In Study 2, using a sample of over 4000 monkeys from a large multi-generational family pedigree, we demonstrate that infant IT is partially heritable. The data revealed infant IT to reflect a co-inherited substrate that manifests across multiple latent variables. Finally, in Study 3 we performed whole-genome sequencing in 106 monkeys to identify IT-associated single-nucleotide variations (SNVs). Results demonstrated a genome-wide significant SNV near *CTNNA2*, suggesting a molecular target worthy of additional investigation. Moreover, we observed lower *p* values in genes implicated in human association studies of neuroticism and depression. Together, these data demonstrate the utility of our model of infant inhibited temperament in the rhesus monkey to facilitate discovery of genes that are relevant to the long-term inherited risk to develop anxiety and depressive disorders.

## Introduction

An anxious temperament (AT) or inhibited temperament (IT) early in life is a major risk factor for the later development of anxiety and depressive disorders [1–4]. Children

with an extreme IT early in life are characterized by increased behavioral inhibition and anxiety in novel and potentially threatening contexts. In humans, IT reflects an inhibited nature or disposition that can define how an individual approaches novel situations throughout their life [5–7]. Although IT often emerges during the second year of life around the time that a child develops the ability to behaviorally cope with threat [8, 9] it can be preceded by increased reactivity in infancy [5]. Temperament is partially inherited, and in human samples the heritability estimates range from 20 to 50% [10–16]. Because IT is heritable, and a risk-factor for the later development of stress-related psychopathology, IT likely reflects an underlying genetic disposition and/or mediates the relationship between genetic variation and psychopathology.

To understand which molecular components contribute to anxiety and depressive disorders, researchers have begun to perform genome-wide association studies (GWAS) with dispositional anxiety and psychopathology in humans. Large-scale efforts for anxiety disorders [17, 18], depressive disorders [19, 20], and neuroticism [21] are all ongoing. Polygenic risk-scores and genetic correlation analyses, suggest each of these disorders partially overlap in their

These authors contributed equally: Andrew S. Fox, Ronald A. Harris

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-021-01156-4>.

✉ Andrew S. Fox  
dfox@ucdavis.edu

✉ Jeffrey Rogers  
jr13@bcm.edu

<sup>1</sup> Department of Psychology, University of California, Davis, CA, USA

<sup>2</sup> California National Primate Research Center, University of California, Davis, CA, USA

<sup>3</sup> Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

genetic substrates [17, 21]. Although some studies have identified genes with significant single-nucleotide variants (SNVs; e.g., in *CRHR1*, *ESR1*, *NTRK2*, etc.) [17, 21, 22], the specific SNVs influencing psychopathology may not be the same in individuals with different genetic backgrounds and/or who were exposed to different environmental stressors.

Rhesus macaques provide a powerful and well-validated animal model to study the genetic and molecular bases of early-life dispositional anxiety because of their similarity to humans. Next to humans, rhesus macaques are the most widely geographically distributed primates in the world, thereby exhibiting outstanding ecological flexibility and adaptability [23, 24]. As Old World monkeys, rhesus macaques are phylogenetically close to humans (with a common ancestor ~10–12 million years closer to humans than marmosets and ~50 million years closer than mice). This recent evolutionary divergence between rhesus macaques and humans is reflected in similarities in their genomes, transcript expression, brain circuits, and resulting socio-emotional behavior [25–28]. Thus, the rhesus monkey model provides a unique opportunity for translational research that supports and extends the insights gained from human genetic studies.

Capitanio et al. have developed and implemented a rhesus monkey model of infant IT, as part of their standardized BioBehavioral Assessment (BBA) that has been applied to over 4000 animals. Infant IT is characterized by below-average expression of latent variables which emerge from behavioral observations spread over 2 days, termed “Activity” and “Emotionality”, both of which are indicative of behavioral inhibition [29] (see methods for details). Notably, molecular studies of dispositional anxiety in macaques are concordant with human GWAS studies, hinting at evolutionarily-conserved mechanisms (e.g., CRH [30, 31], and neurotrophic pathways [4, 32, 33]. Infant IT among macaques provides an excellent model for the study of the human risk for anxiety and depressive disorders [4, 34].

The rhesus monkey model of IT provides numerous advantages that facilitate the identification of genes and molecules that contribute to the risk for psychopathology. Rhesus populations have higher levels of genetic variation and lower linkage disequilibrium than do equivalent sample sizes from human populations, likely due to the historic population bottleneck experienced by ancient human populations that gave rise to modern humans [26, 35–38]. Importantly, this genetic variation is demonstrated in the 853 rhesus macaque whole-genome sequences that Dr. Rogers et al. have collected from US primate centers, which has identified >85 million SNVs, including 408,496 missense variants, 9921 stop codons gained and 7918 splice acceptor or donor variants [39]. Moreover, the breeding strategy at primate centers results in large multi-

generational pedigrees in which each animal has many close and more distant relatives, and this will enrich the population for putatively “rare” variation, thus increasing statistical power in heritability and genetic associations analyses [40].

Here, we demonstrate that infant IT in macaques can be used as an endophenotype to complement human studies of psychopathology, and gain insight into the nature of inherited anxiety. We demonstrate that infant IT is associated with lasting behavioral changes (study 1); show that infant IT is heritable (study 2); and begin to study the genetic variation that underlie this early-life risk-factor (study 3).

## Method

### Methods overview

All studies were performed in rhesus macaques (*Macaca mulatta*) in accordance with the federal guidelines of animal use and care and with the approval of the University of California, Davis Institutional Animal Care and Use Committee. Primary analyses across all studies included animals that underwent BBA-testing during infancy (3–4 months of age). Subsets of animals were selected for Food Retrieval Task testing in Study 1 ( $n = 679$ ; 59M/620F), heritability analyses in Study 2 ( $n = 4433$ ; 2019M/2414F), and whole-genome sequencing in Study 3 ( $n = 106$ ; 49M/57F). The only animals that did not undergo BBA-testing were a subset of female animals that underwent test-retest analyses in Study 1 (>2 tests:  $n = 649$ ; >3 tests:  $n = 288$ ; 4 tests:  $n = 88$ ). Additional detailed methods can be found in the *Supplementary Methods* as well as previous publications [29].

### Assessment of infant IT (Studies 1–3)

The methods for scoring IT in the BBA program have been previously described [29, 34]. Animals (90–120 days old) were relocated for a 25 h testing period (see supplementary methods). IT is defined based on four factors that emerged from factor analyses in a subset of several hundred animals [29]. Animals were considered to be “inhibited” if their scores were below the mean for both “Activity” and “Emotionality” across 2 days of testing, otherwise they were classified as “not inhibited.” The “Activity” factor includes time locomoting; time NOT hanging from the top or side of the cage; rate of environmental exploration; and whether or not the animals ate food\*, drank water\*, or were seen crouching in the cage\* (\* = dichotomized due to rarity). The “Emotionality” factor includes rate of cooing; rate of barking; and whether the animals scratched\*, displayed threats\*, or lipsmacked\* (\* = dichotomized due to rarity).

## Food retrieval task

The Food Retrieval task was administered at ~6:00 am on the day after relocation, prior to morning health and husbandry, by a technician that was blind to infant IT scores. To ensure there was no familiarity with the testing context, the Food Retrieval Task was performed in a different location than BBA-testing. The technician stood in front of the cage, then approached the animal and hand presented a food treat for 5 s, taking care to avert her eyes from the monkey (a stopwatch with an audible beep was used for timing). Treat retrieval was recorded. If the animal did not accept the treat, the technician placed the treat on the forage board and stepped back from the cage, averted her eyes, waited 5 s, and again recorded whether or not the treat was retrieved. Three trials were run consecutively for each animal. Because humans in such close proximity can be perceived as threatening, the Food Retrieval task sets up a potential conflict for the animals between a fear of the human versus attraction to a favored treat.

To estimate the relationship between infant IT and refusal to take food in the food retrieval task, we used logistic regressions. To estimate test-retest stability of the dichotomous reach variable, we used chi-square tests [41, 42]. All statistical analyses were implemented in Python (version 3.7.3), statsmodels (version 0.10.0; <https://www.statsmodels.org/stable/index.html> [43]; was used for regression analyses, and (Pingouin; <https://pingouin-stats.org/> [44]; was used to perform chi-squared tests.

## Heritability of IT

For this study, we analyzed variation among 4433 infants assessed through the BBA program between May of 2001 and January of 2017. The total sample consisted of 407 inhibited females, 403 inhibited males, 2007 non-inhibited females, and 1616 non-inhibited males. As in our previous work [45–48], all heritability and co-heritability estimates were performed using SOLAR-Eclipse (<http://solar-eclipse-genetics.org/>). Prior to heritability estimation phenotype variables were normalized using an inverse normal transformation. All heritability analyses controlled for sex. Because all animals were assessed between 3 and 4 months of age, we did not include Age or Age-squared as covariates in heritability analyses.

## Methods for whole-genome sequencing and mapping

Blood samples were collected from 36 inhibited animals and 70 non-inhibited animals. DNA was extracted and sequenced at the Human Genome Sequencing Center,

Baylor College of Medicine using either the Illumina HiSeq 2000 or Illumina HiSeq X Ten system. WGS sequence data for the 106 animals are publicly available through the NCBI SRA (<https://www.ncbi.nlm.nih.gov/biosample/?term=Bio+Behavior+Assessment>). Paired end reads were aligned to the rhesus Mmul\_10 reference using BWA mem with an average mapped sequence depth of 33.66X across the samples. The GATK v. 4.1.2.0 [49] pipeline was used to identify single-nucleotide variants (SNVs) and insertions/deletions (indels) smaller than 7bp. Variant Effect Predictor [50] was used to annotate variants based on merged Ensembl and RefSeq gene models.

IT-related variants were analyzed using FaST-LMM [51], which implements a linear mixed model that takes potential relatedness into account, controlling for sex. Sequence variants of interest were further examined by lifting the rhesus positions over to the orthologous human position and performing CADD analysis [52].

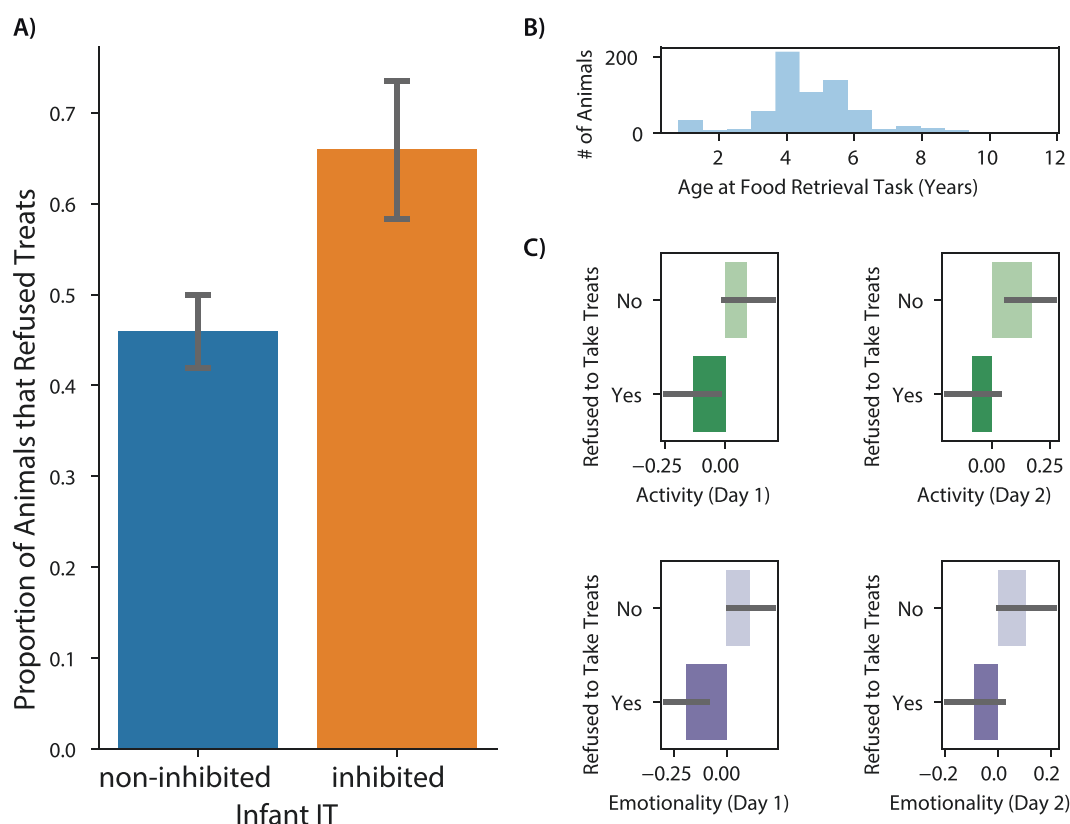
Permutation tests were used to compare IT-related associations to relevant gene lists extracted from published human genome-wide gene-association studies (GWAS). First, the minimum  $p$  value for each gene was computed. Then the average minimum  $p$  value for IT-associations for each gene in the target-list was computed and stored. Finally, for each analysis we performed 10,000 permutations with a similarly sized set of randomly selected genes, and determined the average  $p$  value of those gene-sets. The  $p$  value was computed as the proportion of permutations that resulted in a lower  $p$  value than the target gene-set.

## Results

### Study 1: Infant IT predicts later-life behavior

To begin, we determined whether infant IT reflects an animal's stable disposition throughout their lifespan. We assessed behavioral inhibition during a Food Retrieval Task, when adult animals ( $n = 679$ ) were offered a treat by a human experimenter. We hypothesized that inhibited animals would forgo reward in the presence of this potentially threatening human. Logistic regressions demonstrated animals with high levels of IT at 3–4 months of age are less willing to take a treat years later ( $z = 3.248$ ,  $p = 0.001$ ; Fig. 1).

IT is a composite measure of two latent factors, across 2 days, i.e., Day 1 and Day 2 Activity & Emotionality. To be sure each factor of IT was contributing to the underlying temperamental variable, we assessed each variable as a predictor of treat refusal. Analyses revealed each measure, Activity and Emotionality, on each day to be predictive of treat refusal (all  $p$ 's < 0.014). Supplementary analyses



**Fig. 1** Infant Inhibited Temperament (IT) are more likely to refuse treats from a potentially-threatening human intruder. **A** Proportion of animals that refused treats in the Food Retrieval Task in relation to infant IT. **B** A histogram of ages for animals when tested in the Treat

Refusal Task, on average >4 years after assessment of infant IT. **C** The factors that contribute to IT, i.e., Day 1 and 2 Activity and Emotionality, are each related to the propensity to refuse treats. Error-bars reflect bootstrapped 95% confidence interval.

revealed no significant effects of sex ( $p = 0.209$ ); significant effects of IT on treat refusal in both males ( $t = 2.710$ ,  $p = 0.007$ ) and females ( $t = 3.193$ ,  $p = 0.001$ ); an effect of age on treat refusal, such that older animals were less likely to refuse a treat ( $z = -5.419$ ,  $p < 0.001$ ); and that IT remained significantly associated with treat refusal while controlling for age ( $z = 3.785$ ,  $p < 0.001$ ) (see *Supplementary Results* for additional descriptions and consideration of IT as a continuous variable). Interestingly, although animals have multiple opportunities to retrieve treats, the Food Retrieval Task typically results in an all-or-nothing result, with only ~17% (49/292) of animals who did not retrieve a treat on the first trial going on to retrieve any treat. Unsurprisingly, we obtained similar results when examining first-treat refusal, for IT ( $z = 3.248$ ,  $p = 0.001$ ), age ( $z = -6.064$ ,  $p < 0.001$ ), and IT controlling for age ( $2.935$ ,  $p = 0.003$ ). Together these data suggest that IT as assessed in this protocol is a trait-like measure, which is susceptible to change with experience, but does remain detectably consistent within an animal across contexts as they mature.

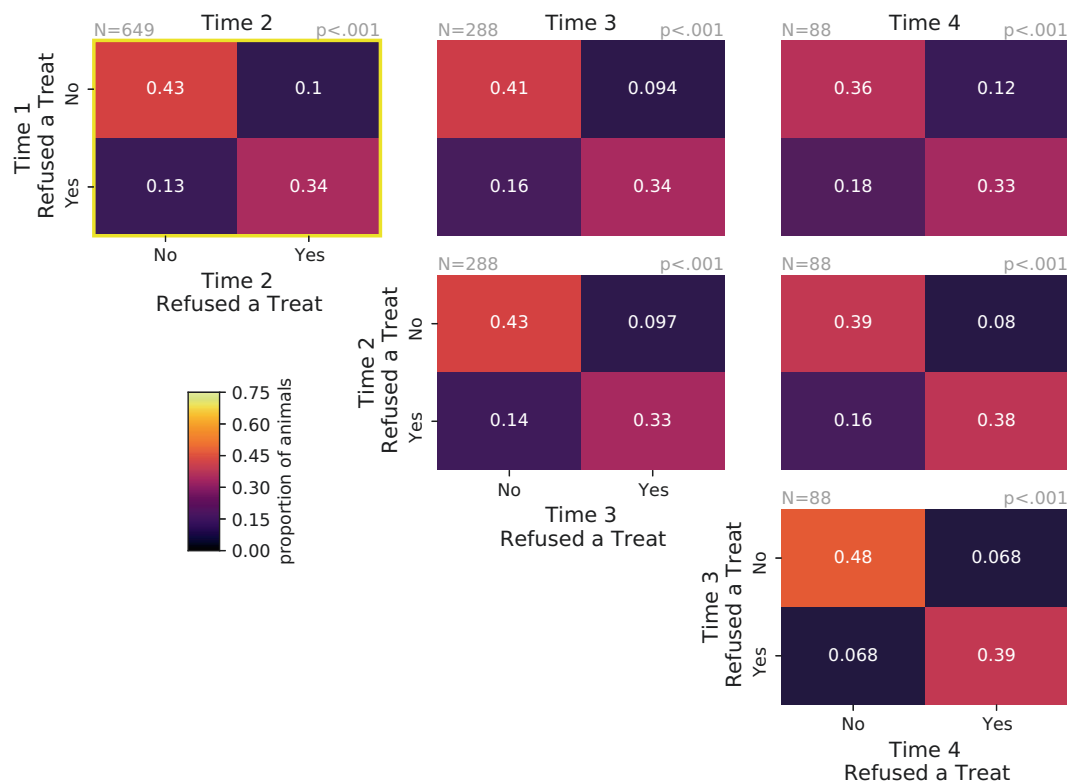
We next evaluated treat-refusal as a stable assessment. In a separate group of 649 female animals, we performed the food retrieval task multiple times, longitudinally. Animals

received 2 ( $N = 649$ ) to 4 ( $N = 88$ ) assessments, on average 1.8 years apart. Chi-squared tests revealed treat-inhibition to be stable between each assessment ( $p$ 's  $< 0.001$ ; Fig. 2), such that 69% of animals responded consistently during the 4th assessment on average 5.24 years later (range = 1.9–9.2 years). Again, treat refusal tended to be all-or-nothing, with 82% of animals that refused the first treat refusing all treats ( $\chi^2 = 430.68$ ,  $p < 0.000001$ ), and first-treat retrieval showing similar stability ( $p$ 's  $< 0.001$ ).

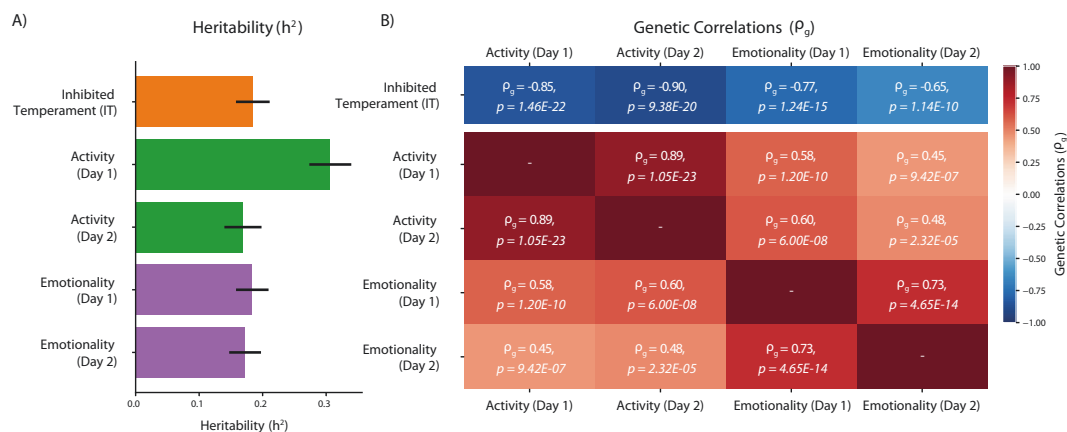
Together, these data demonstrate infant IT reflects life-long behavioral inhibition, that reflects varied behaviors across novel contexts.

## Study 2: Infant IT is heritable

We next examined the heritability of IT, and its contributing factors, across 4433 animals that were part of a large multi-generational pedigree. Results showed that IT is significantly heritable with  $h^2 = 0.19$  ( $\text{std}_{h^2} = 0.027$ ;  $p = 3.7\text{e-}27$ ; Fig. 3A). Results also found IT's latent factors, i.e., Day 1 and Day 2 Activity and Emotionality, to be significantly heritable ( $p$ 's  $< 1.0\text{e-}16$ ) with  $h^2$  values ranging from 0.17 to 0.30. To test the extent to which the latent factors



**Fig. 2** Treat refusal in the Food Retrieval Task is stable over multiple assessments. Contingency tables with the proportion of animals that refused a treat across assessments. The N and  $p$  value for the chi-squared test are above each contingency table in gray.



**Fig. 3** Inhibited Temperament (IT) and its latent factors are heritable, and genetically correlated. Inhibited Temperament (IT) is heritable, as are the contributing latent factors Activity and Emotionality across both days of testing (A) error-bars reflect standard

deviation of the heritability estimate). In addition, IT shows a significant negative genetic correlation with each of latent factors, which in turn are positively correlated with each other (B).

contributing to IT resulted from overlapping genetic variation, we performed genetic correlation analyses. Genetic correlation analyses test the extent to which covariation in phenotypes are associated with covariation in estimated genetic variation, as estimated by relatedness. Results demonstrated all latent factors that comprise IT were significantly genetically correlated with each other ( $\rho_g$

estimates from 0.45 to 0.89;  $p$ 's  $< 0.0001$ ) and with IT ( $\rho_g$  estimates 0.65–0.90,  $p$ 's  $< 1 \times 10^{-9}$ ; Fig. 3B). These data suggest a partially shared genetic substrate that contributes to these latent factors.

Together, the data from Study 2 demonstrate infant IT is significantly heritable, making it an ideal starting point to identify genetic variation. Here, we focus genetic analyses



on IT which is likely to identify genetic variation that contribute to multiple latent variables, and are less likely to be specifically related to an individual feature of IT (e.g., propensity toward locomotion).

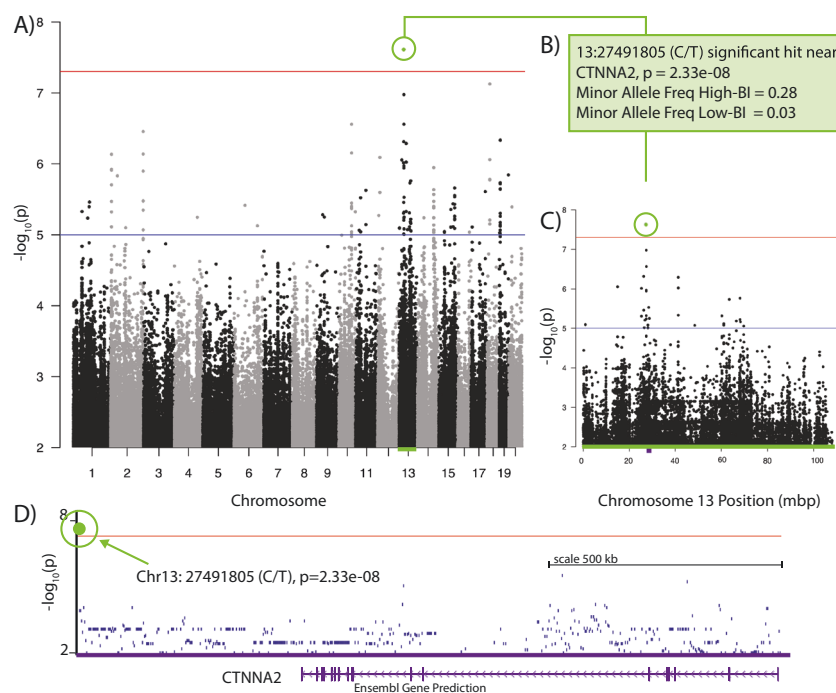
### Study 3: GWAS of IT

In this study, we surveyed BBA subjects for DNA sequence polymorphisms that might reasonably be hypothesized to influence IT, and determined whether the results were consistent with relevant human GWAS studies. Across the 106 individuals, we identified 53,030,128 SNVs and 6,435,882 indels using whole-genome sequencing, and used FaST-LMM to identify IT-related SNVs, taking relatedness into account. We identified a single SNV (13:27491805:C:T) that exceeded a significance threshold  $p < 5 \times 10^{-8}$  which is a standard threshold for GWAS genome wide significance (Table S1, Fig. 4). No indels met this threshold. Two additional nearby SNVs had the third and fourth lowest  $p$  values in the dataset, but failed to reach the genome wide significance threshold (Table S1, Fig. 4B), including 13:27444729:G:C, 47,076bp upstream and 13:27493293:T:A, which was 1477bp downstream of the top SNV. All of these SNVs are intergenic, but

13:27491805:C:T is 473,868 bp downstream of the 3' end of Catenin Alpha-2 (*CTNNA2*). However, although follow-up analyses revealed 4 indels and 43 SNVs in *CTNNA2*, we were unable to identify variation that was both related to IT ( $p < 0.05$ ) and predicted to be functional (i.e., CADD > 10, see supplementary results, Tables S2–S3).

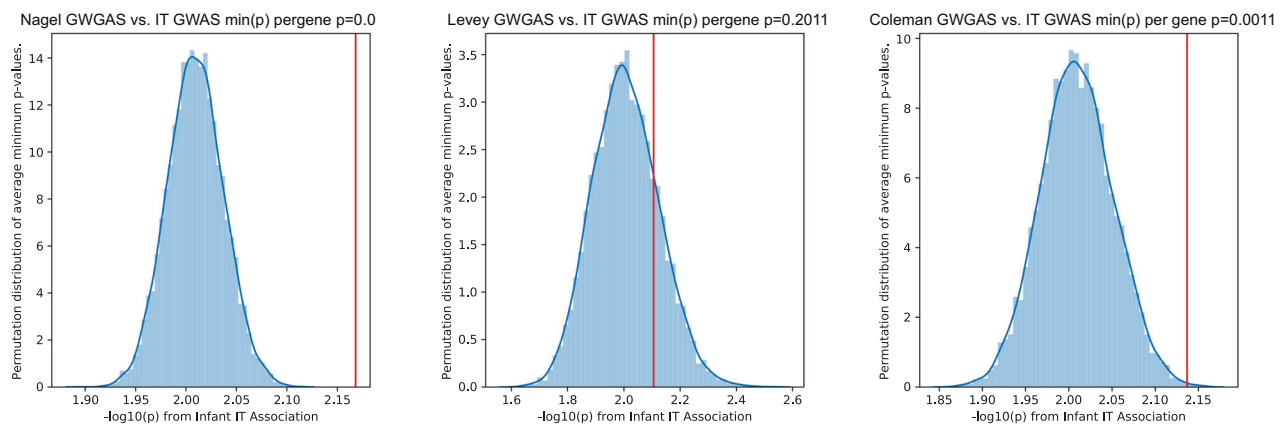
In addition to our genome-wide significant hit in *CTNNA2*, we relaxed the formal genome-wide significance threshold and explored variation in other genes that were marginally significantly ( $p < 0.01$ ) associated with IT (Table S4), but did not reach genome-wide significance. Interestingly, our results revealed uncorrected associations with genetic variation in genes we have previously implicated in inhibition, including an 3' UTR variant in *NTRK3* ( $p = 0.005$ ) and an intronic variant in *PRKCD* ( $p = 0.007$ ) [31, 33, 53]. Such associations must be considered less meaningful but can provide indications of genetic effects that deserve further study.

Using permutation analyses, we tested for IT-related enrichment in genes relevant to stress-related psychopathology to determine if further genetic studies of monkey IT are likely to identify evolutionarily-conserved genes and pathways that are relevant to human psychopathology. Because species differences prevent base-pair by base-pair



**Fig. 4** A SNV on Chromosome 13 near *CTNNA2* is associated with IT. **A** Genome-wide Manhattan plot of BioBehavioral Assessment (BBA) associations with SNVs. The x-axis is the chromosomes and y-axis is the  $-\log_{10}$  of the FaST-LMM  $p$  value. The genome-wide significance threshold  $p < 5 \times 10^{-8}$  is represented by the red line. The significant SNV is highlighted in green. **B** A description of the significant SNV. **C** Chromosome 13 Manhattan plot of BioBehavioral

Assessment (BBA) associations with SNVs. The x-axis is the chromosomes and y-axis is the  $-\log_{10}$  of the FaST-LMM  $p$  value. The genome-wide significance threshold  $p < 5 \times 10^{-8}$  is represented by the red line. The significant SNV is highlighted in green. **D** UCSC Genome Browser view of SNVs with FaST-LMM  $-\log_{10} p$  values in the vicinity of *CTNNA2*. The genome-wide significance threshold  $p < 5 \times 10^{-8}$  is represented by the red line.



**Fig. 5** Permutation analyses of the minimum  $p$  value for each gene from human GWGAS studies. Permutation tests revealed the average minimum  $p$  value from genes found in human GWGAS studies of Neuroticism ( $p < 0.001$ ) and Major Depressive Disorder ( $p = 0.011$ )

comparisons, we examined gene-level enrichment based on GWGAS from studies of human Neuroticism (547 genes; Nagel) [21], Anxiety Disorders (31 genes; Levey) [17], and Depressive Disorders (251 genes; Coleman) [22]. We examined the minimum IT-related  $p$  value for target genes deemed significant in published GWGAS studies. Permutation analyses found the average minimum  $p$  value to be significantly lower in target gene lists as compared to randomly selected genes, for neuroticism (Nagel:  $p < 0.0001$ ) and depressive disorders (Coleman:  $p = 0.0018$ ), but not Anxiety Disorders (Levey:  $p = 0.21$ ; though this list only contained 31 genes) (Fig. 5). Interestingly, genes that were significant in the human GWGASs and the current analyses ( $p < 0.005$ , uncorrected) include *CTNNA2*, *PRKCD*, *NTRK3* and *ESR1*, highlighting the potential for identifying evolutionarily-conserved mechanisms. These data highlight the promise of our nonhuman primate approach, and suggest that additional data will identify additional IT-related genetic signals.

## Discussion

An early-life IT represents one of the strongest known risk-factors for the later development of psychopathology. Here, we demonstrate infant IT in rhesus macaques to be a heritable endophenotype that reflects stable context-independent behavior, and demonstrates promise for identifying molecules that contribute to stress-related psychopathology in humans. The data suggest infant IT is a potentially continuous measure that reflects underlying biological processes that manifest across different behaviors in different contexts (see *Supplementary Results*). These data mirror the work done in humans, in which children classified as behaviorally inhibited tend to remain inhibited

was significantly lower than the average minimum  $p$  value in equally sized random subsets of genes. This same analysis did not reach significance for a GWGAS of Anxiety Disorders ( $p = 0.201$ ).

as they mature, and are at a substantial risk to develop stress-related psychopathology [1–3, 54–57].

The extremely large number of phenotyped animals ( $N > 4000$ ) that are part of a large multi-generational pedigree, provides compelling evidence for heritability, and a unique opportunity for further study of the specific SNVs that contribute to IT. Though generally consistent with our previous behavioral genetics studies of adolescent/adult rhesus macaques [30, 34, 36, 46, 58–60], previous heritability estimates tended to be slightly higher (e.g., ~25–40% heritable: [30, 45–48]. The analyses we describe here include over 4000 animals and found infant IT to be ~19% heritable (see Fig. 3). It is possible that these heritability estimates will increase as animals mature, as has been observed in twin studies of human anxiety [61, 62]. Like with human anxiety and depressive disorders, infant IT was associated with the family structure, such that inhibited animals are more likely to be related to inhibited animals.

Our preliminary genome-wide association analyses revealed a promising hit near the *CTNNA2* gene in chromosome 13, which encodes Catenin alpha-2. This is of particular interest because *CTNNA2* has been previously implicated in human association studies, including a GWGAS of anxiety disorders, and encodes a neuron-specific catenin that is important for cell-to-cell adhesion and synaptic plasticity, is expressed throughout cortical and subcortical structures, and is implicated in numerous studies of IT-related phenotypes [17, 63–70]. The association analysis presented here does not constitute evidence for a definitive association between *CTNNA2* and infant IT in the rhesus monkey. That said, together with findings in humans, these data contribute to the rationale for further study of *CTNNA2* in anxiety-related behavior in animal models, and highlight a potential molecular mechanism that may drive stable anxiety across the lifespan. More generally, our

results are consistent with many molecules contributing to IT, reflecting disruption across many of the brain regions and cell-types that are thought to play a role in anxiety and threat-processing (See Supplementary discussion) [30, 32, 33, 46, 47, 53].

Moreover, these data suggest that further study of infant ITs genetics may identify molecules that are relevant to human psychopathology that can be targeted for mechanistic study. In this sample of 106 rhesus macaques, we identified IT-associations that were enriched for genes identified in GWAS studies of ~449k subjects phenotyped for neuroticism and ~624k subjects phenotyped for major depressive disorder in humans (including ~185k patients). We did not find significant enrichment for the genes implicated in a GWAS of ~200k subjects phenotyped for anxiety disorders. This is likely because this study only identified 31 genes and will likely increase as the number of anxiety disorder genes goes up. Interestingly, although not significantly enriched for average  $p$  value, the *CTNNA2* gene was identified in the million veterans analysis of anxiety disorders, suggesting *CTNNA2* is a relevant gene, like *CRHR1* [30], that is shared between human anxiety disorders and rhesus monkey infant-IT. Together, these data highlight the utility of using our rhesus monkey model of infant IT to identify molecules that may be relevant to human psychopathology, and suggest there is a shared molecular substrate that is conserved across primates.

## Conclusion

Elucidating the genes and genetic mechanisms that predispose individuals to stress-related psychopathology can inform our understanding of psychopathology, and guide the development of biologically-informed treatments. Understanding of disease mechanisms and developing optimal treatments requires human epidemiological insights, as well as well-validated animal models. Future studies will be needed to overcome the limitations of the current preliminary GWAS study, and increase the number of subjects to obtain additional statistical power. Importantly, with the power afforded by a pedigree-based whole-genome sequencing approach in a genetically variable species, non-human primate studies may be able to identify genes and specific genetic variants that would otherwise take hundreds of thousands of subjects using a traditional GWAS approach [26, 36, 38, 40]. Here, we help provide support for a translational model that can support and extend the insights gained from human genetic studies to identify the mechanisms that underlie stress-related psychopathology.

**Acknowledgements** We would like to thank the staff at the California National Primate Research Center, Harsha Doddapaneni, Donna

Muzny and the HGSC data production team, Dr William Mason for helping to develop the Food Retrieval Task, as well as H. Sompura. This work was supported by NIH grants OD010962 to JPC, R01MH121735 to ASF, OD011173 to JR, and the California National Primate Research Center (P51OD011107).

**Author contributions** FAS: Conceptualization, Formal Analysis, Methodology, Software, Visualization, Writing—original draft (lead); HRA: Data curation, Formal Analysis, Software, Visualization, Writing—review & editing; LDR: Conceptualization (Food Retrieval), Investigation (BBA and Food Retrieval) (lead), Project Administration (BBA program) (lead), Supervision (BBA) (lead), Writing—review and editing; MR: Data curation, Formal analysis, Project administration; KS: Writing—original draft; ELK: Resources (provided DNA), Writing—review and editing. CJP: Conceptualization (BBA program), Data curation (BBA), Methodology (BBA), Resources (BBA data), Writing—review and editing; RJ: Conceptualization, Supervision, Writing—review and editing.

## Compliance with ethical standards

**Conflict of interest** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Biederman J, Hirshfeld-Becker DR, Rosenbaum JF, Hérot C, Friedman D, Snidman N, et al. Further evidence of association between behavioral inhibition and social anxiety in children. *Am J Psychiatry*. 2001;158:1673–9.
2. Clauss JA, Blackford JU. Behavioral inhibition and risk for developing social anxiety disorder: a meta-analytic study. *J Am Acad Child Adolesc Psychiatry*. 2012;51:1066–75.e1.
3. Fox NA, Henderson HA, Marshall PJ, Nichols KE, Ghera MM. Behavioral inhibition: linking biology and behavior within a developmental framework. *Annu Rev Psychol*. 2005;56:235–62.
4. Fox AS, Kalin NH. A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. *Am J Psychiatry*. 2014;171:1162–73.
5. Kagan J. Temperament and the reactions to unfamiliarity. *Child Dev*. 1997;68:139–43.
6. Kagan J, Reznick JS, Snidman N. Biological bases of childhood shyness. *Science*. 1988;240:167–71.
7. Zentner M, Shiner R. Fifty years of progress in temperament research: a synthesis of major themes, findings, challenges, and a look forward. In M. Zentner & R. L. Shiner editors, *Handb Temperam*. The Guilford Press (p. 673–700).



8. Scarr S, Salapatek P. Patterns of fear development during infancy. *Merrill-Palmer Q Behav Dev.* 1970;16:53–90.
9. Schaffer HR. The onset of fear of strangers and the incongruity hypothesis. *J Child Psychol Psychiatry.* 1966;7:95–106.
10. Dilalla LF, Kagan J, Reznick JS. Genetic etiology of behavioral inhibition among 2-year-old children. *Infant Behav Dev.* 1994;17:405–12.
11. Dunn J, Plomin R. Determinants of maternal behaviour towards 3-year-old siblings. *Br J Dev Psychol.* 1986;4:127–37.
12. Goldsmith HH, Gottesman II. Origins of variation in behavioral style: a longitudinal study of temperament in young twins. *Child Dev.* 1981;52:91–103.
13. Robinson JL, Kagan J, Reznick JS, Corley R. The heritability of inhibited and uninhibited behavior: a twin study. *Dev Psychol.* 1992;28:1030–7.
14. Battaglia M, Bajo S, Strambi LF, Brambilla F, Castronovo C, Vanni G, et al. Physiological and behavioral responses to minor stressors in offspring of patients with panic disorder. *J Psychiatr Res.* 1997;31:365–76.
15. Rosenbaum JF, Biederman J, Gersten M, Hirshfeld DR, Meminger SR, Herman JB, et al. Behavioral inhibition in children of parents with panic disorder and agoraphobia. A controlled study. *Arch Gen Psychiatry.* 1988;45:463–70.
16. Rosenbaum JF, Biederman J, Hirshfeld-Becker DR, Kagan J, Snidman N, Friedman D, et al. A controlled study of behavioral inhibition in children of parents with panic disorder and depression. *Am J Psychiatry.* 2000;157:2002–10.
17. Levey DF, Gelernter J, Polimanti R, Zhou H, Cheng Z, Aslan M, et al. Reproducible Genetic Risk Loci for Anxiety: results From ~200,000 Participants in the Million Veteran Program. *Am J Psychiatry.* 2020;177:223–32.
18. Purves KL, Coleman JRI, Meier SM, Rayner C, Davis KAS, Cheesman R, et al. A major role for common genetic variation in anxiety disorders. *Mol Psychiatry.* 2019. 20 November 2019. <https://doi.org/10.1038/s41380-019-0559-1>.
19. Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci.* 2019;22:343–52.
20. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry.* 2013;18:497–511.
21. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet.* 2018;50:920–7.
22. Coleman JRI, Gaspar HA, Bryois J. Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Breen G. The Genetics of the Mood Disorder Spectrum: genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biol Psychiatry.* 2020;88:169–84.
23. Campbell CJ. Primates in Perspective. In: Campbell C. J., Fuentes A., MacKinnon K. C., Bearder S. K., & Stumpf R., editors. *Int. Encycl. Primatol., Primates in Perspective.* (2 ed.) Hoboken, NJ, USA. John Wiley & Sons, Oxford University Press Inc.; 2010. p. 1.
24. Richard AF, Goldstein SJ, Dewar RE. Weed macaques: the evolutionary implications of macaque feeding ecology. *Int J Primatol.* 1989;10:569.
25. Bakken TE, Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, et al. A comprehensive transcriptional map of primate brain development. *Nature* 2016;535:367–75.
26. Rhesus Macaque Genome Sequencing and Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science.* 2007;316:222–34.
27. Rogers J, Gibbs RA. Comparative primate genomics: emerging patterns of genome content and dynamics. *Nat Rev Genet.* 2014;15:347–59.
28. Wallis JD. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat Neurosci.* 2011;15:13–19.
29. Golub MS, Hogrefe CE, Widaman KF, Capitanio JP. Iron deficiency anemia and affective response in rhesus monkey infants. *Dev Psychobiol.* 2009;51:47–59.
30. Rogers J, Raveendran M, Fawcett GL, Fox AS, Shelton SE, Oler JA, et al. CRHR1 genotypes, neural circuits and the diathesis for anxiety and depression. *Mol Psychiatry.* 2013;18:700–7.
31. Kalin NH, Fox AS, Kovner R, Riedel MK, Fekete EM, Roseboom PH, et al. Overexpressing Corticotropin-Releasing Factor in the Primate Amygdala Increases Anxious Temperament and Alters Its Neural Circuit. *Biol Psychiatry.* 2016;80:345–55.
32. Fox AS, Oler JA, Shelton SE, Nanda SA, Davidson RJ, Roseboom PH, et al. Central amygdala nucleus (Ce) gene expression linked to increased trait-like Ce metabolism and anxious temperament in young primates. *Proc Natl Acad Sci USA.* 2012;109:18108–13.
33. Fox AS, Souzaiaia T, Oler JA, Kovner R, Kim JMH, Nguyen J, et al. Dorsal Amygdala Neurotrophin-3 Decreases Anxious Temperament in Primates. *Biol Psychiatry.* 2019;86:881–9.
34. Capitanio JP. Naturally Occurring Nonhuman Primate Models of Psychosocial Processes. *ILAR J.* 2017;58:226–34.
35. Bimber BN, Ramakrishnan R, Cervera-Juanes R, Madhira R, Peterson SM, Norgren RB, et al. Whole Genome Sequencing Predicts Novel Human Disease Models in Rhesus Macaques. *Genomics.* 2017;109:214–20.
36. Fawcett GL, Raveendran M, Deiros DR, Chen D, Yu F, Harris RA, et al. Characterization of single-nucleotide variation in Indian-origin rhesus macaques (*Macaca mulatta*). *BMC Genom.* 2011;12:311.
37. Rogers J, Garcia R, Shelledy W, Kaplan J, Arya A, Johnson Z, et al. An initial genetic linkage map of the rhesus macaque (*Macaca mulatta*) genome using human microsatellite loci. *Genomics.* 2006;87:30–8.
38. Xue C, Raveendran M, Harris RA, Fawcett GL, Liu X, White S, et al. The population genomics of rhesus macaques (*Macaca mulatta*) based on whole-genome sequences. *Genome Res.* 2016;26:1651–62.
39. Warren WC, Harris RA, Haukness M, Fiddes IT, Murali SC, Fernandes J, et al. Sequence diversity analyses of an improved rhesus macaque genome enhance its biomedical utility. *Science.* 2020;370:eabc6617. <https://doi.org/10.1126/science.abc6617>.
40. Glahn DC, Nimgaonkar VL, Raventós H, Contreras J, McIntosh AM, Thomson PA, et al. Rediscovering the value of families for psychiatric genetics research. *Mol Psychiatry.* 2019;24:523–35.
41. Cressie N, Read TRC. Multinomial Goodness-of-Fit Tests. *J R Stat Soc Ser B Methodol.* 1984;46:440–64.
42. Yates F. Contingency Tables Involving Small Numbers and the  $\chi^2$  Test. *Suppl. J R Stat Soc.* 1934;1:217–35.
43. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. *Proceedings of the 9th Python in Science Conference.* 2010.
44. Vallat R. Pingouin: statistics in Python. *J Open Source Softw.* 2018;3:1026.
45. Tromp DPM, Williams LE, Fox AS, Oler JA, Roseboom PH, Rogers GM, et al. Altered Uncinate Fasciculus Microstructure in Childhood Anxiety Disorders in Boys But Not Girls. *Am J Psychiatry.* 2019;176:208–16.

46. Fox AS, Oler JA, Shackman AJ, Shelton SE, Raveendran M, McKay DR, et al. Intergenerational neural mediators of early-life anxious temperament. *Proc Natl Acad Sci USA*. 2015;112:9118–22.
47. Oler JA, Fox AS, Shelton SE, Rogers J, Dyer TD, Davidson RJ, et al. Amygdalar and hippocampal substrates of anxious temperament differ in their heritability. *Nature*. 2010;466:864–8.
48. Fox AS, Oler JA, Birn RM, Shackman AJ, Alexander AL, Kalin NH. Functional Connectivity within the Primate Extended Amygdala Is Heritable and Associated with Early-Life Anxious Temperament. *J Soc Neurosci*. 2018;38:7611–21.
49. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma*. 2013;43:11.10.1–11.10.33.
50. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17:122.
51. Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. FaST linear mixed models for genome-wide association studies. *Nat Methods*. 2011;8:833–5.
52. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47:D886–94.
53. Kovner R, Souaiaia T, Fox AS, French DA, Goss CE, Roseboom PH, et al. Transcriptional Profiling of Primate Central Nucleus of the Amygdala Neurons to Understand the Molecular Underpinnings of Early-Life Anxious Temperament. *Biol Psychiatry*. 2020:S000632232031595X.
54. Beesdo K, Bittner A, Pine DS, Stein MB, Höfler M, Lieb R, et al. Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Arch Gen Psychiatry*. 2007;64:903–12.
55. Chronis-Tuscano A, Degnan KA, Pine DS, Perez-Edgar K, Henderson HA, Diaz Y, et al. Stable early maternal report of behavioral inhibition predicts lifetime social anxiety disorder in adolescence. *J Am Acad Child Adolesc Psychiatry*. 2009;48:928–35.
56. Essex MJ, Klein MH, Slattery MJ, Goldsmith HH, Kalin NH. Early risk factors and developmental pathways to chronic high inhibition and social anxiety disorder in adolescence. *Am J Psychiatry*. 2010;167:40–6.
57. Hirshfeld-Becker DR, Biederman J, Henin A, Faraone SV, Davis S, Harrington K, et al. Behavioral inhibition in preschool children at risk is a specific predictor of middle childhood social anxiety: a five-year follow-up. *J Dev Behav Pediatr JDBP*. 2007;28:225–33.
58. Williamson DE, Coleman K, Bacanu S-A, Devlin BJ, Rogers J, Ryan ND, et al. Heritability of fearful-anxious endophenotypes in infant rhesus macaques: a preliminary report. *Biol Psychiatry*. 2003;53:284–91.
59. Coyne SP, Lindell SG, Clemente J, Barr CS, Parker KJ, Maestripietri D. Dopamine D4 receptor genotype variation in free-ranging rhesus macaques and its association with juvenile behavior. *Behav Brain Res*. 2015;292:50–5.
60. Kinnally EL, Karere GM, Lyons LA, Mendoza SP, Mason WA, Capitanio JP. Serotonin pathway gene-gene and gene-environment interactions influence behavioral stress response in infant rhesus macaques. *Dev Psychopathol*. 2010;22:35–44.
61. Bergen SE, Gardner CO, Kendler KS. Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: a meta-analysis. *Twin Res Hum Genet J Int Soc Twin Stud*. 2007;10:423–33.
62. Waszczuk MA, Zavos HMS, Gregory AM, Eley TC. The phenotypic and genetic structure of depression and anxiety disorder symptoms in childhood, adolescence, and young adulthood. *JAMA Psychiatry*. 2014;71:905–16.
63. Park C, Falls W, Finger JH, Longo-Guess CM, Ackerman SL. Deletion in *Catna2*, encoding  $\alpha$ N-catenin, causes cerebellar and hippocampal lamination defects and impaired startle modulation. *Nat Genet*. 2002;31:279–84.
64. Ehlers CL, Gizer IR, Bizon C, Slutske W, Peng Q, Schork NJ, et al. Single Nucleotide Polymorphisms in the REG-CTNNA2 region of chromosome 2 and NEIL3 associated with impulsivity in a Native American sample. *Genes Brain Behav*. 2016;15:568–77.
65. Terracciano A, Esko T, Sutin AR, de Moor MHM, Meirelles O, Zhu G, et al. Meta-analysis of genome-wide association studies identifies common variants in CTNNA2 associated with excitement-seeking. *Transl Psychiatry*. 2011;1:e49.
66. Edwards AC, Aliev F, Bierut LJ, Bucholz KK, Edenberg H, Hesselbrock V, et al. Genome-wide association study of comorbid depressive syndrome and alcohol dependence. *Psychiatr Genet*. 2012;22:31–41.
67. Sun Y, Chang S, Liu Z, Zhang L, Wang F, Yue W, et al. Identification of novel risk loci with shared effects on alcoholism, heroin, and methamphetamine dependence. *Mol Psychiatry*. 2019;26:1152–61.
68. Boraska V, Franklin CS, Floyd JAB, Thornton LM, Huckins LM, Southam L, et al. A genome-wide association study of anorexia nervosa. *Mol Psychiatry*. 2014;19:1085–94.
69. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, Ferguson A, et al. Genome-wide association study of multisite chronic pain in UK Biobank. *PLoS Genet*. 2019;15:e1008164.
70. Schaffer AE, Breuss MW, Caglayan AO, Al-Sanaa N, Al-Abdulwahed HY, Kaymakçalan H, et al. Biallelic loss of human CTNNA2, encoding  $\alpha$ N-catenin, leads to ARP2/3 complex overactivity and disordered cortical neuronal migration. *Nat Genet*. 2018;50:1093–101.

# Supplemental Materials

## Infant Inhibited Temperament in Primates Predicts Adult Behavior, is Heritable, and is Associated with Anxiety-Relevant Genetic Variation

### Authors

\*Andrew S. Fox<sup>1,2+</sup>, \*Ronald A. Harris<sup>3</sup>, Laura Del Rosso<sup>2</sup>, Muthuswamy Raveendran<sup>3</sup>, Shawn Kamboj S<sup>1</sup>, Erin L. Kinnally<sup>1,2</sup>, John P. Capitanio<sup>1,2</sup>, Jeffrey Rogers<sup>3+</sup>

### Affiliations

1. Department of Psychology, University of California, Davis
2. California National Primate Research Center, University of California, Davis
3. Human Genome Sequencing Center and Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas

\* These two authors contributed equally to this work.

+ Correspondence should be directed to Andrew Fox ([dfox@ucdavis.edu](mailto:dfox@ucdavis.edu)) and Jeffrey Rogers ([jr13@bcm.edu](mailto:jr13@bcm.edu))

### Manuscript Information

#### **Supplemental Information:**

*Supplemental Text:* 3446

*Supplementary Tables:* 4

#### **Supplemental References**

**Running title:** Infant Inhibited Temperament in Primates

# Supplemental Methods

## Subjects: Overview

All subjects were rhesus macaques (*Macaca mulatta*). Primary analyses across all studies included animals that underwent BBA testing during infancy (3-4 months of age). Subsets of animals were selected for Food Retrieval Task testing in Study 1 (n=679; 59M/620F), heritability analyses in Study 2 (n=4433; 2019M/2414F), and whole-genome sequencing in Study 3 (n=106; 49M/57F). The only animals that did not undergo BBA-testing were a subset of female animals that underwent test-retest analyses (n=649F) in Study 1. Details on each group can be found below. All studies were performed in accordance with the federal guidelines of animal use and care and with the approval of the University of California, Davis Institutional Animal Care and Use Committee.

## Assessment of Infant IT (Studies 1-3)

Across all studies, inhibited temperament status in infant rhesus macaques was identified using the California National Primate Research Center's (CNPRC) BioBehavioral Assessment (BBA) program. This program, which has been described in detail elsewhere (29, 34) comprises a series of standardized behavioral and physiological assessments, conducted over a 25-hr long period when infant macaques ranged from 90-120 days of age. Assessed animals were from all four rearing environments at CNPRC: field cages, corn cribs, indoor mother reared colony, or nursery reared colony. Field cages enclosures (approximately half an acre each) house approximately 50–200 animals each. Corn cribs are roughly 400 square foot outdoor enclosures that house approximately 15–30 animals each. Field cage and corn crib reared animals were raised in their respective outdoor enclosure with their biological or foster mother. Indoor mother reared and nursery reared animals were raised indoors with no social group. Indoor mother reared animals were housed in a cage with their biological or foster mother, and at most one additional adult female and infant macaque pair. Nursery reared animals were individually housed until 3 weeks of age, at which time they were given visual access to an infant of the same age, and eventually paired with this peer at 5 weeks.

Subjects are separated from their mothers, arrive in the testing area at 0900, and are housed in standard laboratory cages (0.58 × 0.66 × 0.81 m, Lab Products, Maywood, NJ). Animals are tested in cohorts of 5-8 animals. Beginning at 0915h on Day 1, a trained observer records the behavior of each animal for five minutes using a predetermined random order and a standard ethogram for this species (see Golub et al., 2009). This assessment is repeated at 0700h the next day (Day 2). At 1000h on Day 2, infants are returned to their mothers, who had been housed in a separate area of the facility, and allowed to nurse for an hour before all animals are returned to their home cages.

As described by Golub et al. (2009) (29), behavioral data from these assessments were subjected to an exploratory factor analysis using a sample of several hundred animals. Once a factor structure was identified, a confirmatory factor analysis, involving several hundred different animals was conducted to confirm the structure. These data revealed a two factor structure that was replicable: Activity (proportion of time spent locomoting; proportion of time spent not hanging from the side of the cage, rate of environmental exploration, and dichotomous codes for whether the animal ate food, drank water, or crouched) and Emotionality (rate of cooing, rate of barking, and dichotomous codes for whether the animal scratched, displayed threats, or lipsmacked) (definitions of all behaviors are found in Golub et al. (2009) (29). For each year, four scales were created, Activity and Emotionality for Day 1, and separately for Day 2, and values for each were z-scored with mean of 0 and standard deviation of 1. Animals were considered “inhibited” if their scores were below the mean for all four scales, and were otherwise classified as “not inhibited.”

The BBA program has assessed over 5017 rhesus macaques, who are seasonal breeders, with ~78% of animals born March-May. Using the full sample, and after controlling for rearing history, birth month has a small but significant effect on the extent to which an animal is inhibited, accounting for approximately 0.3% of the variance in IT, with animals born earlier in the season slightly more likely to be BI than animals born later.

## Study 1:

### **Subjects: IT and Food Retrieval task (Study 1a)**

The subjects were 679 (n=59 male) primarily female macaques that were relocated from an outdoor enclosure (either field cages or corn cribs) for various reasons (project harvest, shipping, cage relocation etc). No animals were paired. Animals were relocated by early afternoon, and the Food Retrieval test was administered at approximately 0600h the next day, prior to morning health and husbandry. This approach minimizes any effects of separation and relocation on the animals, by leveraging separations that are a part of normal animal and veterinary care.

### **Subjects: Test-Retest of Food Retrieval task (Study 1b)**

Test-retest was done on a separate group of adult female animals, many of whom did not undergo infant BBA testing. Subjects were 649 female subjects that underwent multiple food retrieval tests (at least 2 tests: n=649; at least 3 tests: n=288; 4 tests: n=88). The method for repeated assessment was identical to the initial testing.

7

### **Food Retrieval Task**

The Food Retrieval task was administered at approximately 6:00 am on the day following the animal's relocation, and prior to morning health and husbandry, by a technician that was blind to infant IT scores. In all cases, the Food Retrieval Task was performed in a different location than BBA testing, to ensure there was no familiarity with the testing context. The technician stood in front of the animal's cage, then approached the animal and hand presented a food treat (Condition 1) for a period of five seconds, taking care to avert her eyes from the monkey (a stopwatch with an audible beep was used for timing). If the treat was taken during the five seconds, this was recorded. Any behavioral responses directed toward the technician were also recorded (data not shown). If the animal did not accept the treat, the technician took a step back, recorded the animal's position in the cage, stepped forward, placed the treat on the forage board and stepped back from the cage (Condition 2), averted her eyes, waited five seconds, and again recorded whether or not the treat was retrieved as well as any behavioral response directed toward the technician. Three such trials were run consecutively for each animal. Because humans in such close proximity can be perceived as threatening, the Food Retrieval task sets up a potential conflict for the animals between a fear of the human versus attraction to a favored treat.

### **Statistical Analyses: Food Retrieval Task**

To estimate the relationship between infant AT and refusal to take food in the food retrieval task, we used logistic regressions. To estimate test-retest stability of the dichotomous reach variable, we used chi-square tests (41, 42). All statistical analyses were implemented in Python (version 3.7.3), statsmodels (version 0.10.0; <https://www.statsmodels.org/stable/index.html>; (43) was used for regression analyses, and (Pingoin; <https://pingouin-stats.org/>; (44) was used to perform chi-squared tests.



## Study 2:

### **Subjects: Heritability of IT**

For this study, we analyzed variation among 4433 infants (3-4 months of age) that were assessed for infant-IT as a part of the BBA program between May of 2001 and January of 2017. Parentage was assessed through breeding, blood typing, DNA. Since 1999 parentage was established with microsatellite markers(71, 72); before that, a combination of approaches were taken, including blood typing and breeding records. The total sample consisted of 407 inhibited females, 403 inhibited males, 2007 non-inhibited females, and 1616 non-inhibited males.

### **Statistical Analyses: Heritability Estimation**

All heritability and bivariate heritability analyses were performed using SOLAR-Eclipse (<http://solar-eclipse-genetics.org/>). In brief, heritability is estimated using maximum likelihood, as the proportion of total of variance in an N-by-N pair-wise phenotype-covariance matrix based on the relatedness between animals (73, 74). Prior to heritability estimation phenotype variables were normalized using an inverse normal transformation. All heritability analyses controlled for sex. Because all animals were assessed between 3-4 months of age, we did not include Age or Age-squared as covariates in heritability analyses.

Genetic correlations ( $\rho_g$ ) were estimated using bivariate heritability analyses in SOLAR-Eclipse(46, 75, 76). Bivariate heritability analyses are performed using methods similar to the heritability analyses detailed above, with a covariance matrix that represents both traits and their interaction. Significance was determined by comparing the full model to a second model in which the  $\rho_g$  parameter was fixed to be 0.

## Study 3:

### **Subjects: Whole-Genome-Sequencing**

A total of 106 animals assessed for infant IT, 36 inhibited animals and 70 non-inhibited animals, were selected for whole-genome sequencing (n=49 male, n=57 female). Animal selection was based on the following criteria: field cage born and reared; reared by biological mother (ie, foster-reared animals excluded); minimally related (kinship coefficient < 0.03125); DNA available in our bank (minimal DNA for animals born in 2005-2008, so those animals were excluded).

### **Methods for genome sequencing and mapping**

Rhesus macaque DNA samples (n=106) provided by Dr. John Capitanio and Dr. Erin Kinnally were sequenced at the Human Genome Sequencing Center, Baylor College of Medicine using either the Illumina HiSeq 2000 or Illumina HiSeq X Ten system. WGS sequence data for the 106 animals are publicly available through the NCBI SRA (<https://www.ncbi.nlm.nih.gov/biosample/?term=Bio+Behavior+Assessment>). The paired end reads were aligned to the rhesus Mmul\_10 reference genome assembly using BWA mem with an average mapped sequence depth of 33.66X across the samples. The GATK v. 4.1.2.0 (49) pipeline was used to identify single nucleotide variants (SNVs) and insertions/deletions (indels) smaller than 7bp. Variant Effect Predictor (VEP) (50) was used to annotate variants based on merged Ensembl and RefSeq gene models.

### **Statistical Analyses: Genome-Wide Associations**

Variants were analyzed for association using FaST-LMM (51) which implements a linear mixed model that takes into account potential relatedness among samples. FaST-LMM was run using the BioBehavioral Assessment (BBA) of inhibited (n=36) or not inhibited (n=70) as the phenotype and sex as a covariate.

Sequence variants of interest were further examined by lifting the rhesus positions over to the orthologous human position and performing CADD (52) analysis which predicts the functional impact of variants. CADD integrates annotations from multiple genomic resources and calculates a score that predicts the functional impact of a variant. A CADD score of 10 indicates that the variant is predicted to be among the 10% most functional variants across the entire genome. A CADD score of 20 means that the variant falls in the top 1% most functional.

### **Statistical Analyses: Permutation Analyses to Compare with Published Genome-wide gene-association studies (GWGAS)**

We compared our results to 3 genome-wide gene-association studies (GWGAS) on human Neuroticism (547 genes; (21), Anxiety Disorders (31 genes; (17), and Depressive Disorders (251 genes; (22). A list of relevant genes was extracted from each published GWGAS study. To perform the permutation analysis, we first computed the minimum p-value for each gene. We computed the average minimum p-value in our IT-association, in the genes reported in each list. Then, for each analysis we performed 10,000 permutations with a similarly sized set of random genes, and determined the average p-value of those gene-sets. The p-value was computed as the proportion of permutations that resulted in a lower p-value than the target gene-set.

## **Supplemental Results**

### **Exploratory analyses for IT and behavioral inhibition in the Food Retrieval Task:**

Exploratory logistic regressions were performed to examine potentially confounding variables of sex or age during Food Retrieval Task. There were no significant effects of sex ( $p=.209$ ). Additionally, IT was significantly associated with treat refusal in both males ( $t=2.710$ ,  $p=0.007$ ) and females ( $t=3.193$ ,  $p=0.001$ ), separately. There was substantial variation in the age at which animals were exposed to the Food Retrieval Test (Fig 1). Logistic regressions found that age was significantly correlated with treat refusal, such that older animals were less likely to refuse a treat ( $z=-5.419$ ,  $p<.001$ ). Importantly, IT remained significantly associated with treat refusal when entered simultaneously with age in a logistic regression ( $z=3.785$ ,  $p<.001$ ).

Interestingly, although animals have multiple opportunities to retrieve treats, the Food Retrieval Task typically results in an all-or-nothing, result, with only ~17% (49/292) of animals who did not retrieve a treat on the first trial going on to retrieve any treat. Unsurprisingly, our main results were maintained when restricting Food Retrieval Task data to refusal of the first treat, for IT ( $z=3.248$ ,  $p=0.001$ ), age ( $z=-6.064$ ,  $p<.001$ ), and IT controlling for age (2.935,  $p=0.003$ ). Together these data suggest that IT as assessed in this protocol is a trait-like measure, which is susceptible to change with experience, but does remain detectably consistent within an animal across contexts as they grow up.

Variance was similar for inhibited and non-inhibited animals in the for overall treat refusal (inhibited/non-inhibited variance =.23/.25) and first treat refusal (inhibited/non-inhibited variance =.21/.25) groups.

### **Exploration of IT as a continuous measure:**

The measure of infant IT that is the focus of this manuscript has previously been conceptualized as a categorical measure (29, 77), following the J. Kagan's seminal work in

humans (78). We sought to explore the utility of a continuous measure of IT, which would better characterize the infants in this study. To this end, we calculated a continuous measure of infant IT by taking the average of z-scored Activity and Emotionality across the two test days. Results held when using continuous infant IT as a predictor, such that it predicted treat refusal ( $z=4.980, p<0.001$ ), when controlling for age ( $z=4.938, p<0.001$ ), as well as in both females ( $z=4.436, p<0.001$ ) and males ( $z=2.225, p=0.026$ ). Importantly, when entered simultaneously into a regression with the categorical measure of IT, the continuous measure of IT remained significant ( $z=3.045, p=0.002$ ). Similarly, hierarchical logistic regressions revealed that the continuous IT accounts for additional variance after accounting for the categorical measure ( $X^2=9.47, p=0.002$ ), the reverse is not true ( $X^2=0.82, p=0.37$ ). Finally, heritability analyses revealed the continuous measure of IT to have a slightly higher point-estimate of heritability than the categorical measure (22% vs. 19% heritable). Taken together, these data suggest that, if anything, the continuous measure of IT performs slightly better than the categorical measure in predicting that future studies may be better suited to use the continuous measure.

### Genome-wide significant hit and *CTNNA2*:

We searched for predicted functional variants by reciprocally lifting over each SNV to the orthologous human position and annotating each SNV with CADD information including the calculated CADD PHRED Score (Table S2). The genome-wide significant hit, 13:27491805:C:T is in an open chromatin region based on Ensembl annotations. This could have regulatory implications, but the low CADD score suggests it may not. The nearby SNV 13:27444729:G:C is annotated as intergenic, but it does have a higher CADD score (CADD=4.436). The human nucleotide orthologous to 13:27493293:T:A could not be identified with existing information and tools.

Because the variant exhibiting genome-wide significant association with IT was near *CTNNA2*, we further examined SNVs and indels in that gene without imposing a FaST-LMM p-value threshold. There were 4 indels in the UTRs of *CTNNA2*, but the most significant FaST-LMM p-value was 0.14. There were 43 SNVs in *CTNNA2* including 28 UTR variants, 14 synonymous variants and a single missense variant. The 13:28743267:C:T missense variant (Table S3) has a high CADD score of 28.5 suggesting it may have functional significance, but the p-value ( $p=0.48$ ) suggests it is unlikely to have an association with IT, and the allele frequency difference between inhibited and non-inhibited animals is small. Thus, the genome-wide significant result near *CTNNA2* seems to affect IT via an unknown mechanism.

## Supplementary Discussion

Our preliminary genome-wide association analyses revealed a promising hit near the *CTNNA2* gene in chromosome 13. *CTNNA2* encodes catenin alpha-2 (also called cadherin-associated protein, alpha 2). Catenin alpha-2 is a neuron-specific catenin that is important for cell-to-cell adhesion and synaptic plasticity, and is expressed throughout cortical and subcortical structures (brain-map.org). Using a pedigree-based analysis in humans, researchers showed that biallelic loss of *CTNNA2* results in severe deficits in neuronal migration accompanied by intellectual impairment and autism-like features (70). The only study to our knowledge to examine the mechanistic role of *CTNNA2* in fear/anxiety-related behavior found that *CTNNA2* knockout mice have a deficiency in fear-potentiated startle (63). Despite the limited understanding surrounding the function of *CTNNA2* in anxiety, numerous discovery-based

analyses point to its relevance for psychopathology. Most excitingly Levey et al., (2019)(17) found SNPs in human *CTNNA2* to be associated with anxiety disorders in their GWAS (1/31 genes tested above). Consistent with a shared genetic substrate between anxiety, depression, and addiction, GWAS studies in humans have also implicated variation in the *CTNNA2* gene in impulsivity (64), excitement seeking (65), depression with comorbid substance abuse (66), substance dependence (67), anorexia(68), and chronic pain (69). As noted in the main manuscript, the association analysis presented here does not constitute evidence for a definitive association between *CTNNA2* and infant IT in the rhesus monkey. That said, together with findings in humans, these data contribute to the rationale for further study of *CTNNA2* in anxiety-related behavior in animal models, and highlight a potential molecular mechanism that may drive stable anxiety across the lifespan.

In addition to our finding near the *CTNNA2* gene, the data point to the potential involvement of other molecular systems in infant IT. Studies in humans and nonhuman primates have identified an evolutionarily conserved distributed neural circuit that underlies anxiety and anxiety-like behavior. Importantly, this circuit includes cortical regions (e.g. orbital prefrontal cortex and insular regions [OPro/Al]), extended amygdala regions (e.g. central nucleus of the amygdala [Ce], and bed nucleus of the stria terminalis [BST]), as well as regions of the brainstem (e.g. periaqueductal gray [PAG]). In nonhuman primates, we have demonstrated that metabolism and connectivity across these circuits is heritable, providing a neural substrate that can mediate the effects of genes on anxiety(46, 47). We predict that the heritability and genetic effects are mediated by alterations within these circuits. For example, though it did not survive multiple comparison correction, we identified IT-related variation in *ESR1* and *NTRK2*, which have been implicated in human GWAS studies, as well as *NTRK3* and *PRKCD*, which have been mechanistically implicated in anxiety-like behavior in animal models. These findings should be interpreted cautiously, as they did not reach formal levels of significance, but suggest a possible convergence with human genetics studies and rhesus RNA-sequencing studies that should be explored with additional statistical power.

Consistent with the results presented here, mechanistic studies in nonhuman primates and rodents have implicated specific molecules and cell-types within these regions that contribute to these distributed alterations in anxiety-related brain function. For example, we have demonstrated that expression of *NTRK3* is associated with anxiety-related responding (32, 33). Moreover, experimental manipulation of *NTRK3*-signalling in the dorsal amygdala region is sufficient to increase anxiety-related behavior in adolescent rhesus macaques(33). The data presented here, hint at additional support for the role of *NTRK3* in the risk to develop anxiety and depressive disorders by implicating a 3' variant as potentially contributing to infant IT. Along similar lines, rodent studies have identified cells expressing *PKC $\delta$*  in the lateral Ce (CeL) that, when optogenetically stimulated, result in decreased freezing to a cue (79). Moreover, these neurons seem to be required for fear learning in the BLA(80), and required for the effects of benzodiazepines(81). These data are supported by RNA-sequencing studies in the nonhuman primate where Kovner and colleagues identified expression of the gene that encodes *PKC $\delta$* ,

*PRKCD*, to be associated with behavioral inhibition in response to potential threat (53).

Interestingly, Kovner et al., demonstrated Ce PKC $\delta$ -expressing cells project to BST, suggesting a potential mechanism for the heritability of Ce-BST functional connectivity(48, 53). That said, again, these findings should be interpreted cautiously, and considered as modest evidence worthy of further investigation, rather than proof of these specific SNVs or genes as relevant to IT.

There are likely thousands of genes(82) and near infinite SNVs that contribute to anxiety and depressive disorders. Identifying any single gene that is causally related to a phenotype implicitly implicates related molecules in the disorder. For example, when an individual non-synonymous SNV implicates a particular G-protein coupled receptor, this suggests the molecules that bind that receptor, regulate that receptor, or mediate its intracellular effects also may play a role in that disorder. For example, identifying BI<sup>+</sup>-related polymorphism in the intracellular loop of the *CRHR1* gene(30) implicates CRH ligand (31), and intracellular signaling kinases that interact with CRHR1 in the expression of anxiety. The cascade of implicated molecules does not stop there. Genes involved in causing CRH to be released and other receptors on CRH-expressing cells can also be considered likely to play a role in the risk to develop stress-related psychopathology. Although each of these implicitly related molecules is unlikely to be the cause of psychopathology in any given patient, that does not mean that they cannot be useful for treatment targets.



## Supplemental References

71. Andrade MCR, Penedo MCT, Ward T, Silva VF, Bertolini LR, Roberts JA, et al. Determination of genetic status in a closed colony of rhesus monkeys (*Macaca mulatta*). *Primates J Primatol.* 2004;45:183–186.
72. Kanthaswamy S, von Dollen A, Kurushima JD, Alminas O, Rogers J, Ferguson B, et al. Microsatellite markers for standardized genetic management of captive colonies of rhesus macaques (*Macaca mulatta*). *Am J Primatol.* 2006;68:73–95.
73. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet.* 1998;62:1198–1211.
74. Blangero J, Diego VP, Dyer TD, Almeida M, Peralta J, Kent JW, et al. A Kernel of Truth: Statistical Advances in Polygenic Variance Component Models for Complex Human Pedigrees. *Adv Genet.* 2013;81:1–31.
75. Williams JT, Van Eerdewegh P, Almasy L, Blangero J. Joint Multipoint Linkage Analysis of Multivariate Qualitative and Quantitative Traits. I. Likelihood Formulation and Simulation Results. *Am J Hum Genet.* 1999;65:1134–1147.
76. Almasy L, Dyer TD, Blangero J. Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. *Genet Epidemiol.* 1997;14:953–958.
77. Capitanio JP. Behavioral Inhibition in Nonhuman Primates: The Elephant in the Room. In: Pérez-Edgar K, Fox NA, editors. *Behav. Inhib.*, Cham: Springer International Publishing; 2018. p. 17–33.
78. Kagan J. Galen's prophecy: temperament in human nature. New York: Westview Press; 1994.
79. Haubensak W, Kunwar P, Cai H, Cioocchi S, Wall N, Ponnusamy R, et al. Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature.* 2010;468:270–276.
80. Yu K, Ahrens S, Zhang X, Schiff H, Ramakrishnan C, Fenno L, et al. The central amygdala controls learning in the lateral amygdala. *Nat Neurosci.* 2017;20:1680–1685.
81. Griessner J, Pasieka M, Böhm V, Grössl F, Kaczanowska J, Pliota P, et al. Central amygdala circuit dynamics underlying the benzodiazepine anxiolytic effect. *Mol Psychiatry.* 2018:1–11.
82. Flint J, Greenspan RJ, Kendler KS. How genes influence behavior. Second edition. Oxford ; New York: Oxford University Press; 2020.

## Supplemental Tables

SNV ID	Ref Allele	Alt Allele	FaST-LMM p-value	Inhibited Alt AF	Not inhibited Alt AF
13:27491805:C:T	C	T	2.33e-08	0.2778	0.02899
13:27444729:G:C	G	C	1.06e-07	0.4028	0.08696
13:27493293:T:A	T	A	2.76e-07	0.2571	0.02899

**Table S1.** SNVs of interest based on FaST-LMM p-value including the allele nucleotides and alternate allele frequency by BBA status.

SNV ID	Human (hg38)	Ensembl Consequence	CADD Phred Score
13:27491805:C:T	2:81106490	Regulatory – Open Chromatin	0.344
13:27444729:G:C	2:81148697	Intergenic	4.436

**Table S2.** SNVs lifted over to the human genome and CADD results

SNV ID	Consequence	Ref Allele	Alt Allele	CADD Phred Score	FaST-LMM p-value	Inhibited Alt AF	Not inhibited Alt AF
13:28743267:C:T	missense	C	T	28.5	0.48	0	0.0072

**Table S3.** *CTNNA2* functional SNVs