Future of Families
Candidate Genes
Restricted Use Data Appendage

9-year Follow-Up Wave

Updated March 2023

Bendheim-Thoman
Center for Research on Child Wellbeing
Wallace Hall
Princeton University
Princeton, NJ 08544
http://crcw.princeton.edu

Columbia Population Research Center
1255 Amsterdam Avenue, Room 715
Columbia University
New York, NY 10027
http://cupop.columbia.edu/

Prepared by the staff at the Bendheim-Thoman Center for Research on Child Wellbeing (CRCW), and the Department of Molecular Biology, Princeton University. For more information about Future Families, please visit our web site at https://ffcws.princeton.edu/ or email ffdata@princeton.edu

Contents

1. DATA APPENDAGE OVERVIEW

The Future of Families Candidate Genes Restricted Use Data Appendage (ff_gen_9yr_res5.dta) contains genetic information on focal children from the Future of Families and Child Wellbeing Study (FFCWS) and their biological mothers. In order to obtain and process genetic information, saliva samples were provided by focal children and their biological mothers during in-home visit assessments at the 9-year follow-up wave. The goal of collecting genetic information was to allow researchers to test hypotheses about the relationships among genes, family and community environments, and child development. Two aspects of the Future of Families study design make it especially suitable for studying gene-environment interactions.

First, because non-marital births were oversampled, children in the study are disproportionately exposed to the kinds of family and community stresses that are hypothesized to interact with underlying, predisposing genetic variation. Second, because of the study's longitudinal design and its focus on family relationships and community contexts, the study provides substantial dataon children's cumulative exposure to stressful family and community environments.

This data appendage includes results from the genotyping of several candidate polymorphisms hypothesized to influence child development through their interactions with children's social environments. These candidate polymorphisms include serotonin transporters (5-HTTLPR and STin2), dopamine transporter (DAT1), dopamine D2 receptor (DRD2), dopamine D4 receptor (DRD4), catechol-O-methyltransferase (COMT), melanocortin 4 receptor (MC4R), transmembrane protein (TMEM18), brain-derived neurotropic factor (BDNF), fat mass and obesity-associated protein (FTO), and tryptophan hydroxylase (TPH1, TPH2). For information on how candidate polymorphisms were genotyped, please refer to *Quantification and Genotyping*. For general information about the structure, function, and hypothesized influences of the candidate polymorphisms on child development, please refer to *References and Suggested Readings*.

Please note: An earlier version of this data file previously contained a telomere length (TL) variable. In March 2023, this variable was added to the Public Use data and removed from this data appendage.

2. FILE LAYOUT

The file contains 4,898 observations (one per family) and is sorted by *idnum.*

3. VARIABLE NAMING CONVENTION

Genetic variable names are 7-8 characters long. The first 3 letters contain the variable prefix. The latter 4-5 characters contain the variable suffix, which includes an abbreviated description of the contextual data. The variable names are constructed as follows:

Position	Character	Indicates
1	g	Genetic component (all variables in appendage)
2	m	Mother's genetic component
	k	Child's genetic component
3	5	Fifth follow-up interview (9-year)
4+	**** (e.g., stin2)	Genetic component (e.g., STin 2 VNTR polymorphism)

For example, variable gk5stin2: the prefix (position 1-3) gk5 refers to the genetic component (g) corresponding to the child (k) obtained at the fifth follow-up interview (9-year) (5); the suffix stin2 (position 4+) refers to the STin 2 VNTR polymorphism.

4. DATA COLLECTION AND PROCESSING PROCEDURES

4.1 Overview

As part of the Year 9 follow-up wave, we attempted to collect saliva samples for genetic analysis from all focal children and biological mothers completing the in-home visit activities. In cases where a biological father or non-parental figure was the primary caregiver, or the biological mother was not present for the in-home visit, a saliva sample was collected from only the child. Families completing the home visit activities received a $65 payment to the primary caregiver and a $30 payment to the child. No additional remuneration was provided specifically for the contribution of saliva samples. Ultimately, 3,403 in-home visits were conducted; 2,884 unique child samples and 2,670 unique mother samples were collected.

4.2 Collection kit

Interviewers used the Oragene® DNA Self-Collection Kit to collect saliva samples from focal children and biological mothers during the Home Visit. The Self-Collection Kit is a repository for the collection, preservation, and transportation of saliva.

4.3 Process for administering

The respondents were instructed to spit into the container until the liquid portion reached a line on the interior of the container (the ideal volume of saliva to be collected was 2 ml). The container was then capped. In the process of screwing the cap onto the container, a liquid preservative was released. The container was then put into a small plastic biohazard bag that contained absorbent material if the container were to leak. The plastic bag was then put into a mailer.

In cases where the child had developmental or physical limitations prohibiting the interviewer from collecting the full sample by having the child spit into the collection kit, the child accessory kit was used. The child accessory kit contained a set of five saliva sponges used with the Oragene® self-collection kits. The saliva sponges were inserted into the child's mouth and moved around the upper and lower cheek pouches on both sides of the mouth to collect saliva. The sponges were stored inside the containers and then sealed as described above. Respondents were instructed to rinse their mouth out 5 minutes prior to the saliva sample collection. They were also provided with a packet of sugar and instructed to use ¼ tsp. if they were having difficulty stimulating saliva.

4.4 Transfer of samples to the biorepository/laboratory

As Westat received specimen containers from the field, they were inspected to make sure that samples did not contain personal identifiers and placed in a shipment box with other received collection kits. Until they were mailed, these boxes were secured in the locked field room and maintained at room temperature. On an approximately monthly basis during the field period, Westat shipped boxes of specimen containers at room temperature to the laboratory of Dr. Daniel Notterman, Co-Principal Investigator of the FFCWS, in the Department of Molecular Biology at Princeton University. A transmittal form containing the IDs of the enclosed containers was emailed to lab staff. The lab confirmed receipt of the boxes with Westat.

Saliva collection kits were shipped monthly from October 2007 through May 2010 to the Notterman laboratory at Princeton by FedEx from Westat. In total, 31 shipments were received. Upon receipt of the shipments, lab technicians used a barcode reader to inventory the individual samples. These data were imported into a Microsoft Access database where a full inventory of receipted samples is kept.

4.5 DNA purification and storage

Extraction was completed 1 to 2 weeks after receipt of samples from Westat. DNA was extracted from the entire sample using the Oragene® prepIT•L2P Laboratory Protocol for Manual Purification of DNA (DNA Genotek). Briefly, when samples were ready to be processed, they were incubated at 50°C in a water incubator for a minimum of 1 hour. The mixed Oragene®-DNA/saliva sample was transferred to a 15 ml centrifuge tube. A 1/25 ul volume portion of Oragene®-prepIT•L2P solution was added to the microcentrifuge tube and mixed by vortexing for a few seconds. The sample was incubated on ice for 10 minutes, then centrifuged at room temperature for 10 minutes at 3,500 rpm. The clear supernatant was carefully transferred with a pipet into a fresh centrifuge tube, avoiding the precipitate at the bottom of the tube. A volume of room temperature 100% ethanol equal to the volume of the supernatant was added to the supernatant and gently mixed by inversion 10 times. The sample was allowed to stand for 10 minutes at room air to allow the DNA to fully precipitate. The tube was then centrifuged for 10 minutes at room temperature at 3,500 rpm. The supernatant was decanted and discarded, taking care to avoid disturbing the DNA pellet. An ethanol wash consisting of 1 ml of 70% ethanol was added to the tube without disturbing the pellet. After standing at room temperature for 1 minute, the tube was gently swirled to completely remove the ethanol, taking care not to disturb the pellet. The pellet was air dried after which the DNA was rehydrated by adding 0.5 to 1.0 ml of TE solution (10 mM Tris-HCl, 1 mM EDTA, pH 7.5), vortexing the sample for 30 seconds (s), incubating it at room temperature, and transferring the rehydrated DNA to 3 x 1.7 ml microcentrifuge tubes for storage (2 tubes were stored at -80°C and one in the lab refrigerator at 4°C). DNA concentration was determined using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher).

5. QUANTIFICATION AND GENOTYPING

This section describes how candidate polymorphisms were genotyped. For general information about the structure, function, and hypothesized influences of candidate polymorphisms on child development, please refer to *References and Suggested Readings*.

5.1 Serotonin transporter - 5-HTTLPR

DNA samples were analyzed for the 5HTTLPR length polymorphism by polymerase chain reaction (PCR) followed by gel electrophoresis to distinguish the short (14 repeats, a 375 basepair (bp) fragment,) and long (16 repeats, a 419 bp fragment) allele forms. In some cases,

bands signifying the presence of alleles longer than L were noticed. These were designated X, XX, XXX and L+. Additional tests including these rare, longer alleles were conducted. In one instance, a band signifying the presence of an allele different than S but smaller than L was observed and designated S+. PCR was performed with the following primers: forward, 5′ATG CCA GCA CCT AAC CCCT AAT GT3′; reverse, 5′GGA CCG CAA GGT GGG CGG GA3′. PCR was carried out on a PTC-225 DNA engine (MJ Research), using the following cycling conditions: 2 minute (min) denaturing step at 95°C, followed by 35 cycles of 94°C for 30 seconds (s), 66°C for 30 s, and 72°C for 40 s, and a final extension phase of 72°C for 5 min. Reactions were performed in 1 × PCR Buffer (Denville Scientific), containing 1.5 mM $MgCl_2$, 500 nanograms (ng) of genomic DNA, 5 picomoles (pmol) of each primer, 0.3 mM deoxynucleotide triphosphates (dNTPs), and 1 U Taq polymerase (Denville Scientific). PCR products were separated on a 2.0% (weight/volume (w/v)) agarose gel (Denville Scientific) supplemented with ethidium bromide (0.03% (w/v)) and visualized by UV illumination.

5.2 Serotonin transporter - STin2

The STin 2 VNTR polymorphism is a 17-bp VNTR in the intron 2 region of the 5-HTT gene. The two main alleles of the STin2 are the 10 repeats and 12 repeats (31, 35). We again used PCR followed by gel electrophoresis to distinguish between repeats. Forward primer was 5′-GTC AGT ATC ACA GGC TGC GAG -3′; reverse primer was 5′-TGT TCC TAG TCT TAC GCC AGT G-3′. PCR was carried out on a PTC-225 DNA engine (MJ Research) with the following cycling conditions: 2-min denaturing step at 95°C, followed by 32 cycles of 95°C for 30 s, 63.2°C for 30 s (with touch-down −0.1°C) and 72°C for 30 s, and a final extension phase of 72°C for 5 min. Reactions were performed in 1 × PCR Buffer (Denville Scientific), containing 1.5 mM $MgCl_2$, 500 ng of genomic DNA, 5 pmol of each primer, 0.3 mM dNTPs, and 1 U Taq polymerase (Denville Scientific). PCR products were separated on a 2.0% (w/v) agarose gel (Denville Scientific) supplemented with ethidium bromide (0.03% (w/v)) and visualized by UV illumination.

5.3 Dopamine transporter - DAT1 dat1rs40184

The SNP rs40184 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 m, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using the TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number or C_2960969_10. Data were analyzed using Sequence Detection Systems v.2.3 for the 7900HT RT-PCR Machine by Applied Biosystems.

5.4 Dopamine D2 receptor - DRD2 rs1800497

PCR was performed for 3159 DNA with the following primers: forward, 5′-CCT TCC TGA GTG TCA TCA AC-3′; reverse, 5′- ACG GCT CCT TGC CCT CTA G-3′. PCR was carried out on a PTC-225 DNA engine (MJ Research), using the following cycling conditions: 5-min denaturing step at 95 °C, followed by 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 60 s, and a final extension phase of 72 °C for 5 min. Reactions were performed in 1 × PCR Buffer (Denville Scientific), containing 1.5 mM $MgCl_2$, 500 ng of genomic DNA, 5 pmol of eachprimer, 0.3 mM dNTPs, and 1 U Taq polymerase (Denville Scientific). PCR products were digested with the TaqI restriction endonuclease (NEB cat #R0149), which cuts the 236 bp amplicon into 124 bp and 112 bp fragments (if the reference G allele is present) and separated on a 3.0% (w/v) HiRes agarose gel (ISCBioExpress; cat #E3115) supplemented with ethidium bromide (0.03% (w/v)) and visualized by UV illumination. The SNP rs1800497 was genotyped

for the remaining 2,414 DNA samples as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_7486676. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.5 Dopamine D4 receptor - DRD4 rs1800955

The single nucleotide polymorphism (SNP) rs1800955 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_7470700_30. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.6 Dopamine D4 receptor - DRD4 VNTR

The variable number tandem repeat (VNTR) in exon 3 of the DRD4 gene position on chromosome 11p15.5 was amplified using specifically designed primers (forward, 5' AGGTGGCACGTCGCGCCAAGCTGCA; reverse, 5'TCTGCGGTGGAGTCTGGGGTGGGAG). The forward oligonucleotide was tagged with a 6-FAM fluorescent dye supplied from Integrated DNA technologies (IDT; Coralville, IA) while the HEX and PET fluorescently labeled forward primers were obtained from Life Technologies (Grand Island, NY). The reverse primer was unlabeled and obtained from IDT. PCR amplification was performed in a 10 µL reaction mixture containing 1 ng saliva DNA, 0.4 uM each primer, 1 x GC Enhancer and Amplitaq Gold 360 mix Life Technologies (Grand Island, NY). PCR conditions were 10 mins at 95°C, 35 cycles of 15 sec at 95°C, 15 sec at 68°C and 30 sec at 72°C; followed by 7 mins at 72°C. Following amplification, 0.33 ul each of the FAM-, VIC- and PET-labeled PCR products were combined. Following addition of the GeneScan 500 LIZ size standard and HiDi formamide, the mixture was loaded onto the Applied Biosystems 377 Genetic Analyzer (Foster City, CA). Genotypes were determined by capillary electrophoresis as follows:

***DRD4* repeat allele**	**Product size (bp)**
2R	218
3R	266
4R	314
5R	362
6R	410
7R	458
8R	506
9R	554
10R	602

5.7 Catechol-O-methyltransferase - COMT rs4680

The SNP rs4680 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays

from Applied Biosystems (Foster City, CA), assay number or C_25746809_50. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.8 Melanocortin 4 receptor - MC4R rs17782313

The SNP rs17782313 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number or C_32667060_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.9 Transmembrane protein - TMEM18 rs6548238

The SNP rs6548238 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® Polymerase Chain Reaction (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_29311887_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.10 Brain-derived neurotropic factor - BDNF rs4074134

The SNP rs4074134 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number or C_416535_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.11 Fat mass and obesity-associated - FTO rs9939609

The SNP rs9939609 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number or C_30090620_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.12 Tryptophan hydroxylase - TPH1 rs1800532

The SNP rs1800532 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_8940793_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.13 Tryptophan hydroxylase - TPH2 rs1386494

The SNP rs1386494 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_8872341_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

5.14 Tryptophan hydroxylase - TPH2 rs4570625

The SNP rs4570625 was genotyped as described by TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). TaqMan® PCR (denaturation: 95°C, 10 min, annealing: 95°C, 15 s; extension: 60°C, 60 s; for 40 cycles) was performed on genomic DNA from working aliquots using TaqMan® Genotyping Master Mix and SNP Genotyping Assays from Applied Biosystems (Foster City, CA), assay number C_226207_10. Data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems.

6. QUALITY AND QUALITY CONTROL

DNA quantification was performed using absorbance after extraction. Absorbance was measured using either the NanoDrop system or Bio-Tek Synergy2 multi-mode plate reader. The method listed for quantification was a part of Oragene® DNA extraction protocol previously referenced in the "Extraction and Storage" section of this report:

1. Measure absorbance at 320 nm, 280 nm and 260 nm.
2. Calculate corrected A280 and A260 values by subtracting the absorbance at 320 nm (A320) from the A280 and A260 values.
3. DNA concentration in ng/μL = corrected A260 × 10 (dilution factor) × 50 (conversion factor).
4. A260/A280 Ratio: Divide corrected A260 by corrected A280.

Equations were built in Microsoft Excel worksheets to produce needed corrections as listed below and all results were added to an Access database. A separate Access database lists the original concentration, A260, A280, 260/280, 260/320, 320 raw, corrected 260/280, and corrected DNA concentration for each sample.

Problem samples with lower corrected concentrations were also quantified using Invitrogen's Qubit dsDNA BR assay kit (formally Quant-iT dsDNA BR assay kit) using the Bio-Tek Synergy 2 multi-mode plate reader or Invitrogen's Qubit® 1.0 Fluorometer. The Qubit® Fluorometer can calculate the concentration of each sample using the follow method per protocol. A link for this manual is: *http://tools.invitrogen.com/content/sfs/manuals/mp32850.pdf*

To calculate the concentration of the sample, the following equation is used:

$$Concentration\ of\ sample = QF\ value \times \frac{200}{x}$$

where *QF value* = the value given by the Qubit® 2.0 Fluorometer*;* and
x = the number of microliters of sample added to the assay tube

This equation generates a result with the same units as the value given by the Qubit® 2.0 Fluorometer (i.e., if the Qubit® 2.0 Fluorometer gave a concentration in μg/mL, the result of the equation will be in μg/mL).

The general procedure for checking reliability was to run controls on all gels and TaqMan® SNP Genotyping Assays Protocol (Applied Biosystems, Foster City, CA). The genes that ran manually-loaded agarose gels (this would include the 5-HTTPLR and Stin2 genes) had a positive (sample with double bands) and negative (DNase free H_2O) control. After the gel was run, a lab technician reviewed the gel image and call base repeat length. Dr. Notterman made a final review of all gel images and repeat decisions.

For genotyping using TaqMan® SNP Genotyping Assays on the Applied Biosystems 7900HT machines for the gene call, (such as DAT1) additional controls were included. An NTC (No Template Control, using DNase free H_2O) and then three different positives (FAM, VIC, and both) were run with every plate. The data were analyzed using Sequence Detection Systems v.2.3 for 7900HT RT-PCR Machine by Applied Biosystems which combined the samples and NTC samples to create resultcalls and clusters. A lab technician reviewed allelic cluster results and manually checked the multicomponent amplification plot results on each sample that the program called "undetermined" to decide whether the sample needed to be repeated. The percent of repeats is different between samples run on gel and samples run using the SNP Genotyping. The 5- HTTLPR and Stin2 results were repeated as samples were extracted, so both lab technicians ran sample repeats over the two-year period the lab received samples.

7. MISSING VALUES

All missing values for genetic component variables are coded using similar conventions found in other FFCWS data files. More specifically, the following five missing codes are used.

-9 Not in wave

Family did not participate in 9-year in-home assessments

-7 N/A PCG not Bio Mother

Primary caregiver who participated in 9-year in-home assessment was not the focal child's biological mother. Therefore, according to study protocols, no saliva sample was collected from the primary caregiver. This code is applicable only to gm5* variables.

-5 Not collected

Saliva sample was not collected from respondent for reasons other than refusal.

-3 Missing/negative

Saliva sample was collected from respondent; however, results were negative.*

-1 Refused

Respondent refused to provide saliva sample for genetic quantification and genotyping.

Counts of missing codes

Missing code	**Mother sample**		**Child sample**	
	Percent	N	Percent	N
-9 Not in wave	30.5	1495	30.5	1495
-7 N/A PCG not Bio Mother	4.8	233	0	0
-5 Not collected	1.3	63	1.5	75
*-3 Missing/negative**	[0.0, 0.4]**	[0, 19]**	[0.1, 0.5]**	[3, 25]**
-1 Refused	1.3	437	9.1	444
Non-missing	[54.1, 54.5]**	[2651, 2670]**	[58.4, 58.8]	[2859, 2881]**

Total	100	4898	100	4898

**Note regarding missing/negative results:* Samples were repeated by a lab technician at least three times with different dilutions to improve chances of amplification. If band was faint after repeat, it was called "negative." Results with genetic mismatches were repeated to confirm results. The percent of repeats on the SNP protocol was generally less than 5% when using a 384 well plate. Repeats were run together for better statistical results and manually checked by a lab technician. If the multicomponent amplification plot had low fluorescence that made confirmation of a sample's result unclear, it was also called "negative". There are a few results called "negative" because the result changed multiple times over repeating and could not be confirmed. Results with genetic mismatches were repeated to confirm results.

** Counts vary across genetic component variables. Ranges across all variables are shown in brackets.

8. ACKNOWLEDGEMENTS

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) through grant NIH 7R01HD076592.

9. DATA DICTIONARY

Note: Variable names listed in the data dictionary exclude the 3 character length prefixes. Please refer to the *Variable Naming Convention* section for more information about variable prefixes.

Variable	Description
*saliva	Respondent contributed saliva sample for genetic quantification and genotyping
*5httlpr	5-HTTLPR genetic component
*stin2	Stin2 genetic component
*dat1rs40184	DAT1 rs40184 genetic component
*drd2rs1800497	DRD2 rs1800497 genetic component
*drd4rs1800955	DRD4 rs1800955 genetic component
*drd4VNTR	DRD4 VNTR genetic component
*comtrs4680	COMT rs4680 genetic component
*mc4rrs17782313	MC4R rs17782313 genetic component
*tmem18rs6548238	TMEM18 rs6548238 genetic component
*bdnfrs4074134	BDNF rs4074134 genetic component
*ftors9939609	FTO rs9939609 genetic component

*tph1rs1800532	TPH1 rs1800532 genetic component
*tph2rs1386494	TPH2 rs1386494 genetic component
*tph2rs4570625	TPH2 rs4570625 genetic component

10. REFERENCES AND SUGGESTED READINGS

Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. *Development and psychopathology*, *23*(01), 39-52.

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Pijlman, F. T., Mesman, J., & Juffer, F. (2008). Experimental evidence for differential susceptibility: Dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Developmental Psychology*, *44*(1), 293-300.

Belsky, J., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science*, *16*(6), 300-304.

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*(6), 885-908.

Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, *17*(2), 271-301.

Caspi, A., Hariri, A. R., Holmes, A., Uher, R., & Moffitt, T. E. (2010). Genetic sensitivity to the environment: The case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American Journal of Psychiatry*, *167*(5), 509-527.

Lee, D., Brooks-Gunn, J., McLanahan, S. S., Notterman, D., & Garfinkel, I. (2013). The Great Recession, genetic sensitivity, and maternal harsh parenting. *Proceedings of the National Academy of Sciences*, *110*(34), 13780-13784.

Freese, J., & Shostak, S. (2009). Genetics and social inquiry. *Annual Review of Sociology*, *35*, 107-128.

Kegel, C. A., Bus, A. G., & Van IJzendoorn, M. H. (2011). Differential susceptibility in early literacy instruction through computer games: The role of the dopamine D4 receptor gene (DRD4). *Mind, Brain, and Education*, *5*(2), 71-78.

Lachman, H. M., Papolos, D. F., Saito, T., Yu, Y. M., Szumlanski, C. L., & Weinshilboum, R. M. (1996). Human catechol-O-methyltransferase pharmacogenetics: Description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics and Genomics*,*6*(3), 243-250.

Mitchell, C., McLanahan, S., Brooks-Gunn, J., Garfinkel, I., Hobcraft, J., Notterman, D. (2013). Genetic differential sensitivity to social environments: Implications for research. *American Journal of Public Health*. *103*(S1), 102-110.

Mitchell, C., Notterman, D., Brooks-Gunn, J., Hobcraft, J., Garfinkel, I., Jaeger, K., & McLanahan, S. (2011). Role of mother's genes and environment in postpartum depression. *Proceedings of the National Academy of Sciences*, *108*(20), 8189-8193

Mitchell, C., McLanahan, S., Hobcraft, J., Brooks-Gunn, J., Garfinkel, I., & Notterman, D. (2015). Family Instability, Genes, and Children's Externalizing Behavior. *American Journal of Sociology*, 120(4).

Propper, C., Willoughby, M., Halpern, C. T., Carbone, M. A., & Cox, M. (2007). Parenting quality, DRD4, and the prediction of externalizing and internalizing behaviors in early childhood. *Developmental Psychobiology*, *49*(6), 619-632.

Stein, R. A. (2012). Epigenetics and environmental exposures. *Journal of Epidemiology and Community Health*, *66*(1), 8-13.

Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Mesman, J. (2008). Dopamine system genes associated with parenting in the context of daily hassles. *Genes, Brain and Behavior*, *7*(4), 403-410.

Young, S. E., Smolen, A., Corley, R. P., Krauter, K. S., DeFries, J. C., Crowley, T. J., & Hewitt, J. K. (2002). Dopamine transporter polymorphism associated with externalizing behavior problems in children. *American Journal of Medical Genetics*, *114*(2), 144-149.