DSDR | DATA SHARING FOR DEMOGRAPHIC RESEARCH

ICPSR 27031

# National Longitudinal Study of Adolescent Health (Add Health), 1994-2008: Wave III DNA [Restricted Use]

Kathleen Mullan Harris
*University of North Carolina-Chapel Hill*

J. Richard Udry
*University of North Carolina-Chapel Hill*

Original Add Health Wave III DNA Data Codebook

Inter-university Consortium for Political and Social Research
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

# Terms of Use

The terms of use for this study can be found at:
http://www.icpsr.umich.edu/cocoon/ICPSR/TERMS/27031.xml

# Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

NOTICE
WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

# *National Longitudinal Study of Adolescent Health*

# *Wave III*
# *DNA Code Book*

Andrew Smolen, John K. Hewitt
Institute for Behavior Genetics, University of Colorado at Boulder

❖ ❖ ❖ ❖ ❖

Carolina Population Center
University of North Carolina at Chapel Hill

January 2009

This research is from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu).

# WAVE III DNA CODEBOOK

At Wave III, saliva samples were collected from full siblings or twins to genotype DNA for five candidate polymorphisms. These candidate genes have been reported to be associated with individual differences in behavior related to mental health; are reported to be functional, exonic, in promoter regions, or affect gene expression; are expressed in the brain; and have *prima facie* involvement in neurotransmission. The candidates are the dopamine transporter (DAT1), the dopamine D4 receptor (DRD4), the serotonin transporter (5HTT), monoamine oxidase A (MAOA), and the dopamine D2 receptor (DRD2). Genotyping for 11 neutral markers for the purposes of zygosity determination was conducted where needed. Also available for these Wave III respondents are genotypes for CYP2A6, rs2304297, rs892413, rs4950, and rs13280604.

## Field collection methods

Respondents who were selected for this sample were asked to read and sign separate informed consent forms and there was no incentive payment for providing a DNA specimen. The respondent was asked to insert one sterile cytology brush into the mouth and rub the cheeks and gums for 20 seconds to collect buccal cells. The tip of the brush was placed in a 2 ml screw cap tube containing 200 µl of lysis buffer (1% isopropyl alcohol [v/v] in 50 mM Tris-HCl, 1 mM EDTA and 1% sodium dodecyl sulfate, pH 8.0). The subjects then swished their mouths with 10 ml of 4% sucrose vigorously for 30 sec, and the contents were discharged into a 50 ml conical test tube which was sealed with parafilm ("wash 1"). This was followed by a second mouth wash done identically ("wash 2"). The tubes were labeled and packaged for shipment with ice packs to maintain a temperature of 4$^{o}$C until received in the University of Arizona laboratory for DNA extraction. Specimens were later transferred to the University of Colorado for genetic typing and analysis.

## Laboratory methods

**DNA Samples**. DNA samples were prepared in the laboratory of Dr. David Rowe, University of Arizona. Genomic DNA was isolated from buccal cells using a modification of published methods.[1-4] The brush and mouthwash samples were prepared separately, and combined at a later stage. On day one, 1 ml of lysis buffer (6 M guanidine-HCl, 100 mM Tris-HCl and 10 mM EDTA, pH 7.5) and 25 µl of proteinase K (10mg/ml) were added to each 2 ml tube containing the swab. Tubes were placed in a rotator in a 55°C incubator overnight. The "wash 1" and "wash 2" samples were combined and centrifuged at 1,800 rev/min for 10 min at room temperature. The supernatant was discarded, and 1 ml of lysis buffer was added to the resulting pellet. The resuspended pellet was transferred to a fresh 2 ml tube, and 25 µl of proteinase K (10mg/ml) was added. Samples were placed in a rotating incubator overnight at 55°C.

On day two, the brush heads were removed from their tubes, and 200 µl of binding matrix (10 mM sodium acetate and 0.1g/ml diatomaceous earth [Sigma] in lysis buffer) was added to each brush and wash tube. The tubes were placed on a rotator at room temperature for 15 min, and centrifuged at maximum speed for 2 min. The supernatant

fluid was discarded, and 1 ml of wash buffer (50% ethanol [v/v] in 400 mM sodium chloride, 20 mM Tris-HCl and 2 mM EDTA, pH 7.5) was added to the pellets in each tube. The tubes were placed on a rotator at room temperature for 15 min, and centrifuged at maximum speed for 2 min again. The supernatant fluid was discarded and the resulting pellets were vacuum dried overnight.

On day three, 200 µl of elution buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.8) was added to each dried pellet. The tubes were placed in a rotating incubator at 55 for 30 min, and centrifuged at maximum speed for 2 min. The supernatant fluids from each individual's brush and wash tubes were combined and stored in a single 0.5 ml tube. The yield of DNA was quantified by absorbance at 260 nm (1 O.D. = 50 µg/ml), and an aliquot was diluted to a concentration of 20 ng/µl or less for a working sample. The average yield of DNA was 58 ± 1 µg. DNA samples were stored at 4°C and transported to the Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado for genotyping.

**Dopamine transporter (DAT1, locus symbol: SLC6A3).** The DAT1 gene, which maps to 5p15.3[5] contains a 40 bp Variable Number Tandem Repeat (VNTR) polymorphism in the 3' untranslated region. This VNTR ranges from 3 to 11 copies, and has been reported to effect translation of the DAT protein in human striatum.[6] The 9-repeat (440 bp) and 10-repeat (480 bp) polymorphisms are the two most common alleles in Caucasian, Hispanic and African American populations.[7] The assay was a modification of the method of Vandenbergh et al.[5] The primer sequences were: forward, 5'-TGTGGTGTAGGGAACGGCCTGAG-3' (fluorescently labeled), and reverse: 5'-CTTCCTGGAGGTCACGGCTCAAGG-3'.

**Dopamine D4 Receptor (DRD4).** The DRD4 gene, which maps to 11p15.5, contains a 48 bp VNTR in the third exon.[8] This VNTR can consist of 2 to 11 repeats, although 4 and 7 are most common. This polymorphism results in a variation in the third cytoplasmic loop of the receptor protein, which has been shown to affect the function of the D4 receptor in vivo.[8] The assay used was a modification of the method of Sander et al.[9] The primer sequences[10] were: forward, 5'-AGGACCCTCATGGCCTTG-3' (fluorescently labeled), and reverse, 5'-GCGACTACGTGGTCTACTCG-3'. This method results in polymerase chain reaction (PCR) product of (in bp): 379, 427, 475 (4x), 523, 571, 619 (7x), 667, 715, 763 and 811.

**Serotonin transporter (5HTT, locus symbol SLC6A4).** The 5HTT gene, which maps to 17q11.1-17q12, contains a 44 bp VNTR in the 5' regulatory region of the gene.[11] The VNTR in the promoter appears to be associated with variations in transcriptional activity: the long variant (528 bp) has approximately three times the basal activity of the shorter promoter (484 bp) with the deletion.[12] The assay used was a modification of the method of Lesch et al.[12] The primer sequences were: forward, 5'-GGCGTTGCCGCTCTGAATGC-3' (fluorescently labeled), and reverse, 5'-GAGGGACTGAGCTGGACAACCAC-3.' These primer sequences yield products of 484 or 528 bp.

**Monoamine Oxidase A promoter** (MAOA-uVNTR). The MAOA gene, which maps to Xp11.3-11.4, contains a 30 bp VNTR in the 5' regulatory region of the gene,[13] which has been shown to affect expression.[14] The MAOA gene product is primarily responsible for the degradation of dopamine, serotonin, and norepinephrine and is located in the nigro-striatial pathway of the brain.[15,16] A genotype by environment interaction has been reported for this polymorphism. This finding has been both supported[17] and questioned.[18,19]

The MAOA-uVNTR polymorphism was assayed by a minor modification 18 of a published method.[14] Primer sequences for the 30bp VNTR in the promoter region of the MAOA open reading frame were: forward, 5'ACAGCCTGACCG-TGGAGAAG-3' (fluorescently labeled), and reverse, 5'-GAACGTGACGCTCCATTCGGA-3'.[14] PCR reactions contained 1 µl of DNA (20 ng or less), 1.8 mM $MgCl_2$, 180 µM deoxynucleotides, with 7'-deaza-2'-deoxyGTP (Roche Applied Science, Indianapolis, IN) substituted for one-half of the dGTP, 200 nM forward and reverse primers (Integrated DNA Technologies (IDT), Coralville, IA) and 1 unit of AmpliTaq Gold® polymerase (PE-Biosystems, Foster City, CA), in a total volume of 20 µL. Amplification was performed using a modified version of touchdown PCR[20] as detailed previously[21] and analyzed in an Applied Biosystems (ABI, Foster City, CA) PRISM® 3100 Genetic Analyzer using protocols supplied by the company. Products of this reaction included five possible fragment sizes that included 291, 321, 336, 351, and 381bps (2, 3, 3.5, 4 and 5 repeats, respectively). Genotypes were scored independently by two individuals.

**CYP2A6*2 SNP (dbSNP reference: rs1801272; LocusID: 1548).** The CYP2A6 gene maps to 19q12-13.2 and has been found to contain several genomic alterations that include synonymous and nonsynonymous nucleotide changes and genetic deletions.[22,23] One nonsynonymous change is a T→A nucleotide substitution at codon 60 in exon 3, that results in a substitution of leucine for histidine. This produces a catalytically inactive protein product[24] that has been associated in epidemiological studies with reducing the number of cigarettes smoked.[25,26,27] The CYP2A6*2 assay[28] was performed using the fluorogenic 5'nuclease (TaqMan®, ABI) method using reagents (VIC™ and FAM™ labeled probes and TaqMan® Universal PCR Master Mix without AMPerase® UNG) obtained from ABI. Unlabeled forward and reverse primers were purchased from ABI or IDT. All reactions were performed in an ABI PRISM® 7000 Sequence Detection System using the allelic discrimination mode as described by Livak[29] and in the accompanying instrument documentation. Methods used were as described in the package inserts from ABI, except that the total reaction volume was decreased to 20 µl. Reactions contained 1 µl of DNA (approximately 20 ng of DNA), 8.5 µl of water, 0.5 µl of 40x primer/probe mix and 10 µl of 2x TaqMan® Universal Master Mix. Primer and probe sequences were obtained from the SNP500Cancer Database (//snp500cancer.nci.nih.gov). Final primer and probe concentrations were 900 mM and 200 mM, respectively. The sequences of primers and probes were: Forward Primer: 5'-AGGAGGCGGGCTTCCT-3'; Reverse Primer: 5'-TCGTCCTGGGTGTTTTCCT-3'; Probe 1: 5'- FAM-TCGACGCCC***T***CCGGGGC-TAMRA-3'; Probe 2: 5'- VIC-TCGACGCCC***A***CCGGGGC-TAMRA-3'. Probe 1 (FAM/TAMRA) anneals to the A-containing SNP (histidine-containing, inactive enzyme) and probe 2 (VIC/TAMRA) anneals to the T-containing SNP (leucine-containing, active enzyme). Cycling parameters

used were a 10 min hold at 95ºC followed by 40 cycles of 92ºC for 15 sec and 60ºC for 60 sec. Each 96 well plate included six non-template controls, four samples homozygous for the A SNP, four samples homozygous for the T SNP and two heterozygous samples. Genotypes were scored independently by two individuals.

**DRD2 TaqIA (dbSNP reference: rs1800497, LocusID: 1813).** The gene encoding the dopamine D2 receptor maps to 11q23 and contains a number of genomic variations that include single nucleotide polymorphisms (SNPs), di-nucleotide repeats, and restriction endonuclease sites in coding and non-coding regions.[30] The dopamine D2 receptor gene (DRD2) has a polymorphic TaqI restriction endonuclease site about 2500 bp downstream (3' untranslated region) from the coding region of the gene. This site is designated the TaqIA site to distinguish it from another TaqI restriction site (TaqIB) located elsewhere in the gene. The A1 allele has a point mutation C→ T (TCGA to TTGA), which eliminates the TaqI site. This is not a functional polymorphism as far as we know. This polymorphism is generally assayed by incubating the full-length PCR product (304 bp) with TaqI enzyme. Amplicons that do not contain the restriction site are not cleaved (TaqIA-1 allele). Those that contain the restriction site are cleaved to 178 and 126 bp fragments (TaqIA-2 allele). It has been reported that the TaqIA-1 allele is associated with severe alcoholism.[30,31] We redesigned the assay as a SNP assay using the ABI "TaqMan® Assays-by-Design™ for SNP Genotyping Service." Primer and probe sequences are given in Table I. For this assay the T-containing probe 1 is equivalent to the TaqIA-1 allele, and the C-containing probe 2 is equivalent to the TaqIA-2 allele.

The DRD2 TaqIA assay[28] was performed using the fluorogenic 5'nuclease (TaqMan®, ABI) method using reagents (VIC™ and FAM™ labeled probes and TaqMan® Universal PCR Master Mix without AMPerase® UNG) obtained from ABI. Unlabeled forward and reverse primers were purchased from ABI or IDT. All reactions were performed in an ABI PRISM® 7000 Sequence Detection System using the allelic discrimination mode as described by Livak[29] and in the accompanying instrument documentation. Methods used were as described in the package inserts from ABI, except that the total reaction volume was decreased to 20 µl. Reactions contained 1 µl of DNA (approximately 20 ng of DNA), 8.5 µl of water, 0.5 µl of 40x primer/probe mix and 10 µl of 2x TaqMan® Universal Master Mix. Final primer and probe concentrations were 900 mM and 200 mM, respectively. The sequences of primers and probes were: Forward Primer: 5'-GTGCAGCTCACTCCATCCT-3'; Reverse Primer: 5'-GCAACACAGCCATCCTCAAAG-3'; Probe 1: 5'- VIC-CCTGCCT***T***GACCAGC-NFQMGB-3'; Probe 2: 5'- FAM-CTGCCT***C***GACCAGC-NFQMGB-3'. Probe 1 (VIC/NFQMGB) anneals to the "T" form ("A" on the opposite strand, TaqIA-1, no restriction site, 304 bp fragment) and probe 2 (FAM/NFQMGB) anneals to the "C" form ("G" on the opposite strand, TaqIA-2, containing the restriction site, 178 bp restriction fragment). Cycling parameters used were a 10 min hold at 95ºC followed by 40 cycles of 92ºC for 15 sec and 60ºC for 60 sec. Each 96 well plate included six non-template controls, four samples homozygous for the A SNP, four samples homozygous for the G SNP and two heterozygous samples. Genotypes were scored independently by two individuals.

**Amplification of target sequences by Polymerase Chain Reaction**. PCR reactions contained 1 µl of genomic DNA (20 ng or less), 10% dimethylsulfoxide (DMSO, Hybra-Max® grade, Sigma, St. Louis, MO), 1.8 mM $MgCl_2$, 180 µM deoxynucleotides, with 7'-deaza-2'-deoxyGTP (Roche Applied Science) substituted for one-half of the dGTP, forward and reverse primers (DRD4, 480 nM, DAT1, 400 nM, 5HTT, 600 nM, obtained from IDT) and 1 unit of AmpliTaq Gold® polymerase (PE-Biosystems), and 1x PCR II buffer (PE-Biosystems) in a total volume of 20 µl. Amplification was performed using touchdown PCR[13]. A 95$^o$C incubation for 10 min was followed by two cycles of 95$^o$C for 30 sec, 65$^o$C for 30 sec, and 72$^o$C for 60 sec. The annealing temperature was decreased every two cycles from 65$^o$C to 57$^o$C in 2$^o$C increments (10 cycles total), and a final thirty cycles of 95$^o$C for 30 sec, 65$^o$C for 30 sec, and 72$^o$C for 60 sec and a final 30 min incubation at 72$^o$C.

After amplification, 1 µl of PCR product was combined with 2 µl of loading buffer containing size standard (Genescan 2500 TAMRA®, PE-Biosystems), and 0.8 µl was loaded into each of 48 lanes of a 12-cm gel. PCR products were electrophoresed through a 4.25% polyacrylamide gel under denaturing conditions (6M urea) with Perkin Elmer ABI PRISM® 377 DNA sequencer using protocols supplied by the company. Allele sizes were scored by two investigators independently, and inconsistencies were reviewed and rerun if necessary.

The validity of allele calls for the three genotyping protocols described above were tested using preamplified DNA compared with genomic DNA for 315 individuals (1890 allele calls) from non–Add Health samples. Allele calls agreed in 98.7% of cases between the two methods, with essentially equivalent, and low, rates of missing genotypes, incorrect calls, and unresolved inconsistencies.[14]

**Neuronal nicotinic cholinergic receptors (nAChRs).** These ligand-gated ion channel receptors are expressed on cell bodies and nerve terminal within the brain where they modulate the release of neurotransmitters such as glutamate, GABA, serotonin, norepinephrine, and dopamine (Wonnacott, 1997). nAChRs are normally activated by acetylcholine, but nicotine and related compounds also affect the function of these receptors, often at very low concentrations. nAChRs have a pentameric structure and are composed of distinct combinations of 11 neuronal subunits (α-2 - α-7, α-9 - α-10, β2 – β4). The nAChRs can be homomeric (composed of the same subunit combination) or heteromeric (composed of different subunit combinations). For a more thorough introduction and review of the nAChRs, see Dani (2001), Dani and Bertrand (2007), Gotti, Zoli and Clememti (2006).

A total of two SNPs have been genotyped in the **nAChR alpha-6 subunit** and included: rs2304297 and rs892413. The gene encoding this receptor subunit (CHRNA6, GeneID: 8973) is located on chromosome 8 at position 8p11.21 and spans 16.01 kilo-bases. SNP rs2304297 alleles are: 2 = C, 3 = G. SNP rs892413 alleles are: 1 = A, 2 = C.

A total of two SNPs have been genotyped in the **nAChR beta-3 subunit** and included: rs4950 and rs13280604. The gene encoding this receptor subunit (CHRNB3, GeneID: 1142) is located on chromosome 8 at position 8p11.2 and spans 39.99 kilo-bases. SNP rs4950 alleles are: 3 = G, 1 = A. SNP rs13280604 alleles are: 3 = G, 1 = A.

Genomic DNA was isolated as described earlier and subsequently preamplified using the whole genome **Primer Extension Preamplification** (**PEP**) method described by Zhang et al. (1992). Each 20 µL reaction contained approximately 20 ng of genomic DNA, 2.5 mM $MgCl_2$, 200 µM dNTP's, 13.33 µM primers (random 15mer, Operon Technologies, Alameda, CA), 10% DMSO and 2.5 units of Taq Polymerase (Promega, Madison, WI). Amplification was performed using 55 cycles of 94$^o$C for 60 sec, 37$^o$C for 2 min, 55$^o$C for 4 min and a final 20 min incubation at 72$^o$C. Following preamplification, 100 µl of water was added to each PEP sample, and 1 µl of the diluted PEP reaction was used as the source of DNA. The quality of the data obtained using these methods has been shown to be reliable for genotyping (Anchordoquy, 2003). TaqMan® assays for allelic discrimination (ABI) were used to determine SNP genotypes, as described for the TaqIA locus earlier, using ABI PRISM® 7000 and 7900 instruments.

**Determination of Zygosity.** Zygosity of twin subjects in the sample will be determined by genotyping 11 highly informative microsatellite markers, which we have been using for the last six years. The PCR conditions will include 200 µM of each of the four dNTPs, 2.5 mM $MgCl_2$, 200 nM forward primer (fluorescently labeled), 200 nM reverse primer, approximately 20 ng of genomic DNA, 1 unit of AmpliTaq Gold polymerase and 1x PCR II buffer in a total volume of 20 µl. Microsatellite loci used for zygosity determination are: D1S1679, D2S1384, D3S1766, D4S1627, D6S1277, D7S1808, D8S1119, D9S301, D10S1208, D13S796, D15S652 and D20S481. Three multiplexed PCR reactions containing four primer pairs each will be run for each individual. Tubes are capped and placed in a Perkin Elmer 9600 thermocycler. The amplification is run with the following conditions: 95$^o$C for 10 min (to activate the Gold polymerase); 35 cycles of 94$^o$C for 10 sec, 55$^o$C for 90 sec, and 72$^o$C for 90 sec; a final elongation at 72$^o$C for 8 min; and a hold at 4$^o$C when finished. PCR products are stored at -20$^o$C until analyzed. Electrophoresis will be performed as described above. To control for potential genotyping errors, any analysis that is questionable is repeated. This includes poor amplification (small peaks), potential problems with gel quality, size standards or software problems such as lane tracking or analysis programs (Genscan or Genotyper), all of which are routine.

**REFERENCES CITED**


1. Lench, N., Stanier, P., and Williamson, R. Simple non-invasive method to obtain DNA for gene analysis. *Lancet* 1988:1356-1358.
2. Meulenbelt, I., Droog, S., Trommelen, G.J., Boomsma, D.I., and Slagboom, P.E. High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *American Journal of Human Genetics* 1995;57: 1252-1254.
3. Spitz, E., Moutier, R., Reed, T., Busnel, M., C., Marchaland, C., Roubertoux, P.L., and Carlier, M. Comparative diagnoses of twin zygosity by SSLP variant analysis, questionnaire, and dermatoglyphic analysis. *Behavior Genetics* 1996;26: 55-64.
4. Freeman, B., Powell, J., Ball, D., Hill, L., Craig, I., and Plowmin, R. DNA by mail: An inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. *Behavior Genetics* 1997;27: 251-257.
5. Vandenbergh, D.J., Perisco, A.M., Hawkins, A.L., Griffin, C.A., Li X., Jabs E.W.,

and Uhl G.R. Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. *Genomics* 1992;14: 1104-1106.
6. Heinz, A., Goldman, D., Jones, D.W., Palmour, R., Hommer, D., Gorey, J.G., Lee, K.S., Linnoila, M., and Weinberger, D.R. Genotype influences in vivo dopamine transporter availability in human striatum. *Neuropsychopharmacology* 2000;22: 133-139.
7. Doucette-Stamm, L.A., Blakey, D.J., Tian, J., Mockus, S., and Mao, J.I. Population genetic study of human dopamine transporter gene (DAT1). *Genetic Epidemiology* 1995;12: 303-308.
8. Van Tol, H.H., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik H.B., and Jovanovic, V. Multiple dopamine D4 receptor variants in the human population. *Nature* 1992;358: 149-152.
9. Sander, T., Harms, H., Dufeu, P., Kuhn, S., Rommelspacher, H., and Schmidt, L.G. Dopamine D4 receptor Exon III alleles and variation of novelty seeking in alcoholics. *American Journal of Medical Genetics* 1997;74: 483-487.
10. Lichter, J.B., Barr, C.L., Kennedy, J.L., Van Tol, H.H.M., Kidd, K.K., and Livak, K.J. A hypervariable segment in the human dopamine receptor $D_4$ (DRD4) gene. *Human Molecular Genetics* 1993;2: 767-773.
11. Heils, A., Teufel, A., Petri, S., Stober, G., Riederer, P., Bengel, D., and Lesch, K.P. Allelic variation of the human serotonin transporter gene expression. *Journal of Neurochemistry* 1996;66: 2621-2624.
12. Lesch, K.P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S., Benjamin, J., Muller, C.R., Hamer, D.H., and Murphy, D.L. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science*, 1996;274: 1527-1531.
13. Samochowiec, J., Lesch, K.P., Rottmann, M., Smolka, M., Syagailo, Y.V., Okladnova, O., Rommelspacher, H., Winterer, G., Schmidt, L.G., and Sander, T. Association of a regulatory polymorphism in the promoter region of the monoamine oxidase a gene with antisocial alcoholism. *Psychiatry Research* 1999;86: 67-72.
14. Sabol S.Z., Hu, S., and Hamer, D. A functional polymorphism in the monoamine oxidase a gene promoter. *Human Genetics* 1998;103: 273-279.
15. Chen, Z., Hotamisligil, G.S., Huang, J., Wen, L., Ezzeddine, D., Aydin-Muderrisoglu, N., Powell, J.F., Huan, R.H., Breakefield, X.O., Craig, I.W., and Hsu, Y.P. Structure of the human gene for monoamine oxidase type A. *Nucleic Acids Research* 1991;19: 4537-4541.
16. Shih, J.C., Chen, K., and Ridd, M.J. Monoamine oxidase: From genes to behavior. *Annual Review of Neuroscience* 1999;22: 197-217.
17. Foley, D.L., Eaves, L.J., Wormley, B., Silberg, J.L., Maes, H.H., Kuhn, J., and Riley, B. Childhood adversity, monoamine oxidase A genotype, and risk for conduct disorder. *Archives of General Psychiatry* 2004;61: 738-744.
18. Haberstick, B.C., Lessem, J.M., Hopfer, C.J., Smolen, A., Ehringer, M.A., Timberlake, D., and Hewitt, J.K. Monoamine oxidase A (MAOA) and antisocial behaviors in the presence of childhood and adolescent maltreatment. *American Journal of Medical Genetics Part B* (*Neuropsychiatric Genetics*) 2005;135B(1): 59-64.
19. Young, S.E., Smolen, A., Hewitt, J.K., Haberstick, B.C., Stallings, M.C., Corley, R.P., and Crowley, T.J. Interaction between MAO-A genotype and maltreatment in

the risk for conduct disorder: failure to confirm in adolescent patients. *American Journal of Psychiatry* 2006;163: 1019-1025.
20. Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., and Mattick, J.S. "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* 1992;19: 4008.
21. Anchordoquy, H.C., McGeary, C., Liu, L., Krauter, K.S., and Smolen, A. Genotyping of three candidate genes after whole-genome preamplification of DNA collected from buccal cells. *Behavior Genetics* 2003;33: 73-78.
22. Raunio, H., Rautio, A., Gullsten, H., and Pelkonen, O. Polymorphisms of CYP2A6 and its practical consequences. *Journal of Clinical Pharmacology* 2001;52: 357-363.
23. Oscarson, M. Genetic polymorphisms in the cytochrome P450 2A6 (CYP2A6) gene: Implications for interindividual differences in nicotine metabolism. *Drug Metabolism and Disposition* 2001;29: 91-95.
24. Tricker, A.R. Nicotine metabolism, human drug metabolism, polymorphisms, and smoking behaviour. *Toxicology* 2003;183**:** 151-173.
25. London, S.J., Idle, J.R., Daly, A.K., and Coetzee, G.A. (1999). Genetic variation of CYP2A6, smoking, and risk of cancer. *Lancet* 1999;353: 898-899.
26. Pianezza, M.L., Sellers, E.M., and Tyndale, R.F. Nicotine metabolism defect reduces smoking. *Nature* 1998;25: 750-751.
27. Tyndale, R.F., and Sellers, E.M. Variable CYP2A6-mediated nicotine metabolism alters smoking behaviour and risk. *Drug Metabolism and Disposition* 2001;29: 548-552.
28. Haberstick, B.C., and Smolen, A. Genotyping of three single nucleotide polymorphisms following whole genome preamplification of DNA collected from buccal cells. *Behavior Genetics* 2004;34: 541-547.
29. Livak, K.J. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genetic Analysis: Biomolecular Engineering* 1999;14: 143-149.
30. Noble, E.P. D2 dopamine receptor gene in psychiatric and neurologic disorders and its phenotypes. *American Journal of Medical Genetics* 2003;116B: 103-125.
31. Noble, E.P., Ozkaragoz, T.Z., Ritchie, T.L., Zhang, X., Belin, T.R., and Sparks, R.S. D2 and D4 dopamine receptor polymorphisms and personality. *American Journal of Medical Genetics* 1998;81: 257-267.

**OTHER REFERENCES**

Dani, J.A. (2001). Overview of nicotinic receptors and their roles in the central nervous system. *Biological Psychiatry* 49:166-174.

Dani, J.A., and Bertrand, D. (2007). Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annual Review of Pharmacology* 47:699-729.

Wonnacott, S. (1997) Presynaptic nicotinic ACh receptors. *Trends Neuroscience* 20:92-98.

Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W., and Arnheim, N. (1992). Whole Genome Amplification from a Single Cell: Implications for Genetic Analysis. *Proceedings of the National Academy of Science* (USA) 89:5847-5851.

# Wave III DNA

Wave III DNA

| **aid** | | Str 8 | Respondent Identifier<br>NOTE: Smallest 5 and largest 5 values are displayed. |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1 | 0.0% | 11316754 | |
| 1 | 0.0% | 12571478 | |
| 1 | 0.0% | 14614574 | |
| 1 | 0.0% | 14714942 | |
| 1 | 0.0% | 15574600 | |
| 2564 | 99.6% | 17608172-99718344 | NOTE: Range of values omitted from display |
| 1 | 0.0% | 99718845 | |
| 1 | 0.0% | 99718944 | |
| 1 | 0.0% | 99719278 | |
| 1 | 0.0% | 99719938 | |
| 1 | 0.0% | 99886994 | |

| **dat1a** | | Str 3 | Dopamine Transporter-DAT1-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1 | 0.0% | 6R | 6R |
| 13 | 0.5% | 7R | 7R |
| 20 | 0.8% | 8R | 8R |
| 973 | 37.8% | 9R | 9R |
| 1555 | 60.4% | 10R | 10R |
| 12 | 0.5% | | Missing |

| **dat1b** | | Str 3 | Dopamine Transporter-DAT1-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 5 | 0.2% | 8R | 8R |
| 126 | 4.9% | 9R | 9R |
| 2386 | 92.7% | 10R | 10R |
| 45 | 1.7% | 11R | 11R |
| 12 | 0.5% | | Missing |

| **drd4a** | | Str 2 | Dopamine Receptor-DRD4-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 412 | 16.0% | 2R | 2R |
| 136 | 5.3% | 3R | 3R |
| 1865 | 72.5% | 4R | 4R |
| 15 | 0.6% | 5R | 5R |
| 5 | 0.2% | 6R | 6R |
| 126 | 4.9% | 7R | 7R |
| 15 | 0.6% | | Missing |

| **drd4b** | | Str 3 | Dopamine Receptor-DRD4-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 31 | 1.2% | 2R | 2R |
| 16 | 0.6% | 3R | 3R |
| 1449 | 56.3% | 4R | 4R |
| 56 | 2.2% | 5R | 5R |
| 34 | 1.3% | 6R | 6R |
| 914 | 35.5% | 7R | 7R |
| 48 | 1.9% | 8R | 8R |
| 3 | 0.1% | 9R | 9R |

| | | | |
|---|---|---|---|
| 8 | 0.3% | 10R | 10R |
| 15 | 0.6% | | Missing |

| **httlpra** | | Str 1 | Serotonin Transporter-5HTTLPR-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 862 | 33.5% | L | L |
| 1696 | 65.9% | S | S |
| 16 | 0.6% | | Missing |

| **httlprb** | | Str 1 | Serotonin Transporter-5HTTLPR-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2049 | 79.6% | L | L |
| 509 | 19.8% | S | S |
| 16 | 0.6% | | Missing |

| **maoa_va** | | Str 4 | Monoamine Oxidase A-uVNTR-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 45 | 1.7% | 2R | 2R |
| 28 | 1.1% | 3.5R | 3.5R |
| 1283 | 49.8% | 3R | 3R |
| 1167 | 45.3% | 4R | 4R |
| 16 | 0.6% | 5R | 5R |
| 35 | 1.4% | | Missing |

| **maoa_vb** | | Str 4 | Monoamine Oxidase A-uVNTR-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 14 | 0.5% | 2R | 2R |
| 17 | 0.7% | 3.5R | 3.5R |
| 733 | 28.5% | 3R | 3R |

| | | | |
|---|---|---|---|
| 1730 | 67.2% | 4R | 4R |
| 45 | 1.7% | 5R | 5R |
| 35 | 1.4% | | Missing |

| **cyp2a6a** | | Num 1 | Cytochrome P450 2A6-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2555 | 99.3% | 1 | T |
| 19 | 0.7% | . | Missing |

| **cyp2a6b** | | Num 1 | Cytochrome P450 2A6-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2437 | 94.7% | 1 | T |
| 118 | 4.6% | 2 | A |
| 19 | 0.7% | . | Missing |

| **drd2a** | | Str 2 | Dopamine D2 Receptor-TaqIA-Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 201 | 7.8% | A1 | A1 |
| 2356 | 91.5% | A2 | A2 |
| 17 | 0.7% | | Missing |

| **drd2b** | | Str 2 | Dopamine D2 Receptor-TaqIA-Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1162 | 45.1% | A1 | A1 |
| 1395 | 54.2% | A2 | A2 |
| 17 | 0.7% | | Missing |

| **s000001a** | | Num 1 | rs2304297 - Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |

| | | | |
|---|---|---|---|
| 2017 | 78.4% | 2 | C |
| 351 | 13.6% | 3 | G |
| 206 | 8.0% | . | Missing |

| **s000001b** | | Num 1 | rs2304297 - Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1087 | 42.2% | 2 | C |
| 1281 | 49.8% | 3 | G |
| 206 | 8.0% | . | Missing |

| **s000002a** | | Num 1 | rs892413 - Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2127 | 82.6% | 1 | A |
| 319 | 12.4% | 2 | C |
| 128 | 5.0% | . | Missing |

| **s000002b** | | Num 1 | rs892413 - Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1218 | 47.3% | 1 | A |
| 1228 | 47.7% | 2 | C |
| 128 | 5.0% | . | Missing |

| **s000003a** | | Num 1 | rs4950 - Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1172 | 45.5% | 1 | A |
| 1245 | 48.4% | 3 | G |
| 157 | 6.1% | . | Missing |

| **s000003b** | | Num 1 | rs4950 - Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2088 | 81.1% | 1 | A |
| 329 | 12.8% | 3 | G |
| 157 | 6.1% | . | Missing |

| **s000004a** | | Num 1 | rs13280604 - Allele A |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1119 | 43.5% | 1 | A |
| 1179 | 45.8% | 3 | G |
| 276 | 10.7% | . | Missing |

| **s000004b** | | Num 1 | rs13280604 - Allele B |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 1942 | 75.4% | 1 | A |
| 356 | 13.8% | 3 | G |
| 276 | 10.7% | . | Missing |

| **dna2** | | Num 1 | DNA Based Zygosity |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 308 | 12.0% | 1 | Monozygotic (MZ) |
| 412 | 16.0% | 2 | Dizygotic (DZ) |
| 1854 | 72.0% | . | Missing |

| **concord** | | Str 5 | Concordance of 12 Zygosity Markers |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 2 | 0.1% | 0/12 | 0 of 12 Markers |
| 14 | 0.5% | 1/12 | 1 of 12 Markers |
| 2 | 0.1% | 2/11 | 2 of 11 Markers |

| | | | |
|---|---|---|---|
| 36 | 1.4% | 2/12 | 2 of 12 Markers |
| 2 | 0.1% | 2/8 | 2 of 8 Markers |
| 6 | 0.2% | 3/11 | 3 of 11 Markers |
| 62 | 2.4% | 3/12 | 3 of 12 Markers |
| 6 | 0.2% | 3/8 | 3 of 8 Markers |
| 8 | 0.3% | 4/11 | 4 of 11 Markers |
| 78 | 3.0% | 4/12 | 4 of 12 Markers |
| 78 | 3.0% | 5/12 | 5 of 12 Markers |
| 66 | 2.6% | 6/12 | 6 of 12 Markers |
| 4 | 0.2% | 7/11 | 7 of 11 Markers |
| 36 | 1.4% | 7/12 | 7 of 12 Markers |
| 12 | 0.5% | 8/12 | 8 of 12 Markers |
| 38 | 1.5% | 11/11 | 11 of 11 Markers |
| 2 | 0.1% | 11/12 | 11 of 12 Markers |
| 268 | 10.4% | 12/12 | 12 of 12 Markers |
| 1854 | 72.0% | | Missing |

| **zygchg** | | Num 1 | Zygosity Changed-Based on Saliva Test |
|---|---|---|---|
| Frequency | Percent | Value | Label |
| 654 | 25.4% | 0 | No Change |
| 66 | 2.6% | 1 | Change |
| 1854 | 72.0% | . | Missing |