

**Biomarkers**  
**in Wave III of the Add Health Study**

The Add Health Biomarker Team

## **The Add Health Biomarker Team**

Myron S. Cohen	University of North Carolina at Chapel Hill
Qinghua Feng	University of Washington
Francesca A. Florey	University of North Carolina at Chapel Hill
Carol A. Ford	University of North Carolina at Chapel Hill
Kathleen Mullan Harris	University of North Carolina at Chapel Hill
John K. Hewitt	Institute for Behavioral Genetics, University of Colorado
Marcia M. Hobbs	University of North Carolina at Chapel Hill
King K. Holmes	University of Washington
Nancy B. Kiviat	University of Washington
Lisa E. Manhart	University of Washington
William C. Miller	University of North Carolina at Chapel Hill
Martina Morris	University of Washington
John L. Schmitz	University of North Carolina at Chapel Hill
Andrew Smolen	Institute for Behavioral Genetics, University of Colorado
Joyce W. Tabor	University of North Carolina at Chapel Hill
Patricia A. Totten	University of Washington
J. Richard Udry	University of North Carolina at Chapel Hill

Edited by  
Lisa E. Manhart

## **Table of contents**

I.	Introduction.....	4
II.	Eligibility for Biomarker Samples.....	6
III.	STI and HIV Testing: Field, Transportation and Notification Procedures.....	8
IV.	Laboratory Testing for Sexually Transmitted Infections.....	17
V.	Test Considerations in the Estimation of HIV and STI Prevalence.....	31
VI.	DNA Sampling, Genotyping, and Determination of Zygosity.....	42

## **Chapter I - Introduction**

One of the many unique features of Wave III of the Add Health Study was the collection of biological samples. These biological samples permitted the identification of individuals with sexually transmitted infections [STI] (including HIV), and genotype ascertainment for pairs of full-siblings or twins who resided in the same households. The STI testing allows for analyses of individual, household, family, and environmental risk factors for laboratory-confirmed sexually transmitted infections (versus self-report), and the genetic sample facilitates analyses that differentiate between parental, social, and genetic influence, as well as the extent to which genetic differences in neurotransmitter function are associated with a wide range of behaviors.

The inclusion of these biomarker data requires special considerations in the analysis of Wave III Add Health data. Thus, the purpose of this monograph is to outline relevant procedures, design, and sampling schemes used in the collection of biomarker data, and to serve as a user's guide for its analysis and interpretation. The monograph is intended to supplement existing descriptions of the Add Health study, rather than to replace them. Therefore, please refer to the web pages describing the Add Health Study design for more extensive detail on the study ([www.cpc.unc.edu/addhealth](http://www.cpc.unc.edu/addhealth)) and the sampling weights necessary to work with the data ([www.cpc.unc.edu/addhealth/codebooks/wave3](http://www.cpc.unc.edu/addhealth/codebooks/wave3)).

Issues that require special consideration include sample design (e.g., who was selected for each type of biomarker test), specimen collection, laboratory methods, and laboratory

test performance. Each of these themes is described in separate chapters to this monograph, but should be viewed as complementary to each other.

## **Chapter II – Eligibility for Biomarker Samples**

Add Health Wave III data collection was designed to collect oral mucosal transudate (OMT) samples for HIV testing and urine samples for STI testing from all original Add Health respondents (n=20,745) who consented to provide these specimens. Original respondents who were full siblings or twins (n=3787) were also asked to provide a saliva sample for DNA analysis. Biological samples for HIV and STI testing were also collected from a special sample of recruited partners who consented to provide these specimens. The partner sample at Wave III was designed to include partnerships representing the full spectrum of relationship intimacy and commitment. The total sample was expected to be 1,500 partners, consisting of approximately one third married, one third cohabiting and one third dating partners.

The partner interview was identical to the main respondent interview with the exception of questionnaire sections that used preloaded information from earlier wave data. All Add Health respondents at Wave III were 18 years old or older at the time of data collection.

Interviews were conducted nationwide, including Alaska and Hawaii. Respondents who were overseas for the duration of the field work or who were in the Armed Forces and deployed overseas for the duration of the field work, were not eligible for a Wave III interview. Interviews were conducted mostly in-home, but some interviews were conducted in school settings, work places, and other public places.

## DNA Sample

Original respondents who were identified to be full siblings or twins at earlier waves (3,787) were asked to provide a saliva sample for DNA analysis. However, the recruited partners of these respondents were not asked to provide a DNA sample.

## HPV/*Mycoplasma genitalium* Sample

The University of Washington (UW) requested 4,200 urine specimens from original female Add Health respondents. In order to meet this quota, 7,000 original female Add Health respondents were randomly flagged for potential recruitment into this sample. The UW also requested 1,000 urine specimens from original male Add Health respondents. In order to meet this quota, 2,000 original male Add Health respondents were randomly flagged for potential recruitment into this sample. Frozen aliquots of urine samples were shipped to the UW on dry ice and processed in the HPV and *M. genitalium* laboratories respectively.

## **Chapter III – STI and HIV Testing: Field, Transportation and Notification**

### **Procedures**

Carol A. Ford, M.D.  
Associate Professor of Pediatrics and Internal Medicine  
Adolescent Medicine Program  
School of Medicine  
University of North Carolina at Chapel Hill

Francesca A. Florey, M.S.  
Carolina Population Center  
University of North Carolina at Chapel Hill

Joyce Tabor, M.S.  
Carolina Population Center  
University of North Carolina at Chapel Hill

### **Introduction**

The Add Health Wave III survey was a massive data collection endeavor that spanned the continental United States, Hawaii, and Alaska. Between August 2001 and April 2002, a total of 15,197 original Add Health respondents and 1,507 of their married, cohabiting, or dating partners were interviewed in their homes or another location of their choice by a staff of 422 interviewers. While it was anticipated that most interviews would be collected in in-home settings, to the extent that permission was obtained by local authorities and prison officials, and confidentiality of the interview could be assured, respondents and partners were interviewed and provided biological specimens in military bases, penitentiary facilities, and other institutions. All Wave III respondents 18 years of age or older were asked to provide a urine specimen for *Chlamydia trachomatis* (CT), *Neisseria gonorrhoeae* (GC), and other experimental STI testing, and an oral mucosal transudate (OMT) specimen for *Human Immunodeficiency Virus Type-1* (HIV-1) testing.

This section summarizes field, transportation, and notification procedures related to these biospecimens.

### **Informed Consent**

Wave III used a multiple-step consent process. Before the start of the interview, the interviewer described the Wave III interview and obtained consent for participation. During this process, respondents were informed that they would be asked to provide urine and OMT specimens at the end of the interview and, at the time they were asked to provide these specimens, they would be asked to sign additional consent forms. Completion of the interview was not linked with a requirement to provide biospecimens. At the end of the interview and in the following order, respondents were asked to independently consent to and provide a: 1) *OMT Specimen for HIV Test*; and 2) *Urine Specimen for STI Tests*.

Interviewers provided respondents with the following information regarding the collection of HIV/STI biospecimens:

- The OMT specimen would be tested for the presence of HIV antibodies.
- The urine specimen would be tested for two curable STIs—chlamydia and gonorrhea.

The urine samples would also be tested for other STIs but, because those tests have not yet been approved by the FDA as reliable using urine specimens, they would only be used for exploratory purposes and their results would not be made available. Add Health would not use urine samples to test for drug use.

- Add Health would not report the results for these tests to anyone else, and the security procedures make it impossible for others to know the respondent's HIV or STI status.
- If they provided the HIV OMT sample, they would receive an incentive of \$10.
- If they provided the urine specimen, they would receive an incentive of \$10.
- A toll-free telephone number to call four weeks after collection date to obtain results of the urine and HIV testing and to receive counseling.
- A toll-free telephone number to obtain additional information about the study, as well as the number of the UNC IRB contact.
- The toll-free 24-hour Centers for Disease Control and Prevention National STD and AIDS Hotline telephone number for general questions about HIV and STIs.

Respondents were also provided verbal and written information about CT, GC, and HIV. Respondents were told that most who are infected do not experience symptoms, and provided with information on transmission, benefits of being tested and learning results, and the meaning of a positive and negative test result. Respondents were told that persons notified of a positive test would be asked to go see a doctor or nurse. They also received information explaining that Add Health testing did not include testing for all STIs, and was not a substitute for regular medical care (including yearly Pap smears for sexually active women).

Respondents received written sheets to be given to a health care provider if they had a positive test. This sheet provided information on interpretation of test results, recommended repeat serologic testing for positive HIV tests, and provided guidance and

references for management and treatment issues based on national guidelines [1]. Health care professionals were also referred to their local health department, infectious disease specialists in their region, and given the Centers for Disease Control and Prevention National STD and AIDS Hotline telephone number if they had questions.

For respondents who were incarcerated at the time of the Add Health interview, biospecimen collection was only conducted if local IRB permission was obtained, prison administration agreed to all privacy and confidentiality requirements in the Add Health study protocol, the respondent was guaranteed unmonitored access to telephone calls to get his/her results, the respondent had access to adequate health services to re-test for, and/or treat STIs or HIV, and the respondent was given access to a private place to provide the urine sample. Special informed consent forms as well as HIV and STI information sheets were designed for respondents who were incarcerated at the time of the interview. There were no differences in incentive payments offered to incarcerated and non-incarcerated respondents.

### **Field Collection**

OMT specimens were collected for HIV testing using the Orasure® Oral HIV-1 Antibody Testing System. Respondents were instructed to place the pad of the collection device between the lower cheek and gum, to rub back and forth until moist, and to keep in place for a minimum of 2 minutes (timed by interviewer). The collection device was then placed in the vial supplied with the kit, and placed in a plastic bag.

Urine specimens were tested for CT and GC testing using Ligase Chain Reaction (LCR™) amplification technology in the Abbott LCx® Probe System. The assay required that 15-20 cc of first stream urine be collected in a plastic, preservative-free, sterile urine specimen collection cup from respondents who have not urinated within one hour prior to collection, and that the specimen be cooled immediately to 2-8°C until processed. Respondents received instructions about collection procedures from interviewers who had received extensive training (including training by Abbott Laboratory representatives), and samples were collected in a 30 cc cup with a black line marked at 15 cc. After securing the lid of the collection cup, specimens were placed in a plastic bag, and then inside of a styrofoam cooler with two ice packs. If respondents asked to be excused to urinate before the end of the interview, consent and specimen collection occurred at that time to avoid situations where respondents would not be eligible to provide a specimen at the end of the interview because they had voided within the previous one-hour.

A paper requisition was completed by the field interviewer and shipped along with the biospecimens. A bar code label, identical to that used for labeling the urine and OraSure specimens, was affixed to the requisition. The date and time of sample collection, and interviewer identification number were also recorded on the requisition.

### **Transportation**

STI/HIV tests were packaged and mailed to the University of North Carolina at Chapel Hill. Add Health packaging was approved by Federal Express as having met the

minimum packaging requirement for urine and saliva diagnostic specimens. Packaging was also in compliance with OSHA standards for non-infectious, diagnostic specimens and biological requirements. These requirements specify that packaging must include a primary receptacle, absorbent and cushioning material, a water-tight secondary container, other packaging (shipping container), and clinical specimens outer labels. This package was placed in a FedEx Diagnostic Envelope.

Our goal was to have samples arrive at the laboratory within 96 hours of collection, and received by 10:00 a.m. for processing. After collection and packaging, field interviewers were required to place the packaged box inside of their soft-side cooler until they could get to a manned Federal Express office or until Federal Express could collect the package. If the field interviewer could not get to a Federal Express office before the end of the day, or if the Federal Express office was closed (as in the case of holidays or Sundays), field interviewers were asked to keep the packaged containers in their refrigerators until the following day.

A unique challenge faced by the project was the closing of the Federal Express service and airports in the week following the September 11, 2001 attack on the World Trade Center and the Pentagon. During this time, field interviewers were instructed to keep all the specimens that were collected in their refrigerator and to ship as soon as airports re-opened and the Federal Express service re-initiated. Despite these shipping delays, the project team decided to maintain the field work operation active, to the extent possible.

## **Notification Issues**

Research subjects who have participated in the Add Health Study are protected by a Certificate of Confidentiality issued by the Department of Health and Human Services in accordance with the provisions of section 301 (d) of the Public Health Service Act (42 U.S.C.,241(d)). Under the terms of the Confidentiality Certificate, UNC-CH, the Principal Investigator, and research staff are authorized to protect the privacy of the individuals who are the subjects of the Add Health project by withholding their identifying characteristics from all persons not connected with the conduct of the study.

Collection of biological specimens and testing for STIs was handled in a manner consistent with the Add Health Security Management Protocol. In order to remove the link between the biospecimen sample and the respondent's identifying information, a specimen identification number was pre-assigned for each Add Health respondent by the Add Health Security Manager in Canada. These specimen identification numbers were referred to as the UID, which was unique and different from the respondent's questionnaire identification number. The UID for each case was pre-loaded into the laptop computers used in the field interview process. At the time of biospecimen collection, the link between the respondent identification number and the UID was broken. The biospecimen sample was labeled using a barcode label that contained only the UID, and the connection between the respondent identification information and the UID was erased in the electronic files.

This protocol results in anonymous testing because HIV and STI results are never linked to a respondent's identifying information. At the end of specimen collection, respondents were told that the results of tests would be available by calling a toll-free number one month after the date of collection, and that they would be the only one who could access their results. To be sure no one else could receive their results, they were asked to select a password which they directly entered into the computer. All respondents were given a card with the toll-free number, their UID, and a hand-written date that represented when their results should be available. Respondents were told they would not be able to get their results without this information, and to keep the card in a safe place.

The UID and password information for each Add Health respondent was transmitted electronically to Research Triangle Institute (RTI). RTI transmitted this information to the American Social Health Association (ASHA) who maintained a toll-free telephone results and counseling service for the duration of the study. This information was stored by ASHA until such time that they received STI and HIV results from the UNC Labs for these cases (linked by the UID).

The ASHA results and counseling telephone service was open from 3:00 PM to 11:00 PM Eastern time five days a week. Three telephone counselors were available at all times. Respondents who called were asked to provide their UID and personal password. Once the counselor confirmed that the information was correct, the counselor provided the respondent with the available test results. A respondent was able to refuse to receive one or all of the results at any time during the call. The respondent was permitted to call

and receive his/her results twice. The ASHA counselor also provided counseling, referral information, and the National CDC AIDS and STD Hotline telephone number [2]. All respondents with a positive test were told to seek health care. Respondents were counseled to be re-tested at a health care facility if: 1) their Add Health HIV test was positive; 2) any test result was indeterminate; 3) any procedural problem occurred during specimen collection, transportation, or testing; 4) they suspected test results were wrong; or 5) they could not get test results because of a lost UID or forgotten password. Respondents who did not call for positive results had a second opportunity to learn results if contacted during follow-up HIV and STI notification interviews.

## **REFERENCES**

1. Centers for Disease Control and Prevention. 1998 Guidelines for Treatment of Sexually Transmitted Disease. *MMWR*. 1998;47:1-118.
2. Centers for Disease Control and Prevention. Revised Guidelines for HIV Counseling, Testing, and Referral and Revised Recommendations for HIV Screening of Pregnant Women. *MMWR*. 2001;50.

## **Chapter IV: Laboratory Testing for Sexually Transmitted Infections**

### **Human Immunodeficiency Virus Type 1 (HIV-1) testing**

John L. Schmitz, Ph.D.  
Department of Pathology and Laboratory Medicine  
University of North Carolina

Wave III respondents who provided consent were tested for the presence of antibodies to HIV-1 in oral mucosal transudate (OMT) specimens collected with the OraSure<sup>®</sup> HIV-1 Oral specimen collection device (Epitope, Beaverton, OR).

Upon receipt in the laboratory, OMT samples were inspected for the presence of a bar code label encoding the subjects' UID number and a corresponding requisition with the same UID bar code label affixed. After verification of the UID and collection date, the tip of the OraSure vial was broken off, the vial placed into a polypropylene centrifuge tube labeled with the corresponding UID and the sample centrifuged for 15 minutes at 600 x g. After centrifugation, the OraSure<sup>®</sup> collection device was discarded and the sample volume was verified to be at least 750µl. Samples with less than 750µl were not tested as per the manufacturer's instructions. The centrifuged tubes were then capped and stored at room temperature until testing for HIV-1 antibodies was conducted, typically on the next day's run.

OMT samples were tested for the presence of HIV-1 antibodies using the Oral Fluid Vironostika<sup>®</sup> HIV-1 Microelisa System (Organon Teknika Corporation, Durham, NC) following the manufacturer's instructions. Each run of microelisas contained 3 negative

and 1 positive controls. The direct manual method was used to dilute controls and specimens directly in the wells of the microtiter plate that were coated with inactivated HIV-1 antigen. After mixing, the plates were covered with an adhesive plate sealer and incubated at 37°C for 90 to 100 minutes. The plates were then washed four times with an automated plate washer (ELP-40 Microplate Strip Washer, BIO-TEK, Instruments, Inc., Winooski, VT), followed by the addition of appropriately diluted peroxidase-conjugated goat anti-human immunoglobulin. The plates were covered and incubated at 37°C for 30 to 35 minutes followed by 4 washes. One hundred and fifty microliters of ABTS (2,2'-azino-di-[3-ethylbenzthiazoline-6-sulfonate]) substrate was then added and the plates incubated for 10 to 12 minutes uncovered. Finally, 150µl of stop solution was added. The plates were mixed and the absorbance of each well determined with an automated microelisa plate reader (ELx800, BIO-TEK Instruments, Inc., Winooski, VT) at 405nm.

The microelisa plate reader automatically determined assay acceptability and the positive cutoff value based on the negative and positive control absorbance values. OMT samples with an absorbance value above the positive cutoff (reactive samples) were retested in duplicate wells of the microelisa assay on the next run of the test. If either or both of the duplicate wells gave a positive result the sample was considered repeatedly reactive and a confirmatory Western immunoblot assay was performed.

To confirm repeatedly reactive microelisa results, we used the OraSure<sup>®</sup> HIV-1 Western Blot Kit (Organon Teknika Corporation, Durham, NC), testing samples according to the manufacturer's instructions. Each run of blots contained one negative, one weak positive

and one strong positive control sera. The on-strip dilution method was used to add controls and repeatedly reactive samples to the troughs of trays containing nitrocellulose strips with preblotted HIV-1 proteins. After control and sample addition, the trays were covered and incubated for 3 hours at room temperature while rotating at 60 revolutions per minute (rpm). Following incubation, the diluted controls and samples were aspirated and 2ml of sample buffer were added to each trough of the tray. Sample buffer was immediately aspirated. This procedure was repeated 2 more times, followed by two five-minute washes in sample buffer. Next, an alkaline phosphatase conjugated, goat anti-human IgG (heavy and light chains) F(ab')<sub>2</sub> fragment was added to each trough of the tray, which was then covered and incubated for 45 minutes at room temperature with rotation at 60rpm. The previously described wash procedure was then performed followed by a 5-minute wash in deionized water. One milliliter of substrate (BCIP/NBT) was added to each trough of the tray, which was then mixed by hand and incubated 10 minutes at room temperature. Color development was stopped by rinsing the strips 3 times in deionized water, followed by a 5-minute wash in deionized water. The strips were then air-dried and banding patterns recorded on worksheets. Criteria for acceptable runs as defined by the manufacturer were reviewed for each run. If the criteria were not met, the entire run was repeated.

Western blots were interpreted according to the following criteria: A blot was defined as negative if no bands were present. A blot was defined as positive if 2 or 3 bands corresponding to the p24, p41 or gp120/160 were present. Any banding pattern not meeting this criterion was defined as indeterminate.

OMT specimens with a negative ELISA result on initial testing were reported as seronegative for antibodies to HIV-1 as were specimens that were reactive on the initial run but for which both duplicate repeat tests were negative. In addition, specimens repeatedly reactive by ELISA but demonstrating no bands by Western blot testing were also reported as seronegative. OMT specimens that were repeatedly reactive by ELISA and met the criteria for a positive Western blot were reported as seropositive for antibodies to HIV-1. OMT specimens that were repeatedly reactive by ELISA but did not demonstrate a banding pattern meeting the criteria for a positive Western blot were reported as indeterminate.

HIV-1 antibody testing by the OraSure<sup>®</sup> method was only conducted on specimens that were properly labeled (i.e., labeled with a bar-coded UID label on the transport vial and a corresponding UID barcode label on the accompanying requisition). Any samples that were received with incomplete or no labeling were discarded and not tested.

Additionally, as per the manufacturer's instructions, OraSure<sup>®</sup> samples were required to be maintained at 4 – 37°C and tested within 21 days of collection. Samples that were received beyond the 21-day time limit were not tested. Specimen temperature was maintained by the inclusion of ice packs in the packing containers used for shipping. All OraSure<sup>®</sup> samples were discarded after testing was completed.

## ***Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (GC) Testing**

John L. Schmitz, Ph.D.  
Department of Pathology and Laboratory Medicine  
University of North Carolina

Wave III respondents who provided consent were tested for the presence of CT and GC DNA in urine specimens. Urine was collected and shipped to UNC laboratories as described in Chapter III. Urine sample requirements for CT and GC LCR testing were more restrictive than for HIV-1 antibody testing and are described below:

### **Criteria for urine collection.**

- First void sample, collected at least 1 hour after the respondent had previously urinated.
- 20ml maximum volume collected.
- Urine samples were placed in coolers with ice packs immediately after collection and shipped in packaging containing 2 ice packs.
- Samples were received in the laboratory and processed within 4 days of collection.
- Samples were shipped on the day of collection or the following day by overnight express to arrive in the laboratory by 10:00am. Samples collected on Friday evenings or Saturdays were held at refrigerated temperature until shipment on the following Monday. These samples arrived at the laboratory in sufficient time to meet the 4-day sample-processing requirement.
- All samples, except those with improper labeling, were tested.

Upon receipt in the laboratory, urine specimens were inspected for the presence of a bar code label encoding the subject's UID number. This UID number must have matched the UID bar code label affixed to the requisition accompanying the specimen. After verification of the UID, volume of urine received, determination that the sample was cold and had been received within 4 days of collection, the urine was processed for CT and GC testing. Additional aliquots of urine were processed for long-term storage and further testing for *Trichomonas vaginalis*, *Mycoplasma genitalium* and human papillomavirus.

CT and GC testing was performed by Ligase Chain Reaction (LCR) testing (Abbott LCx Probe System, Abbott Park, IL). A one-milliliter aliquot of urine was processed with the Urine Specimen Preparation Kit (Abbott LCx Probe System, Abbott Park, IL) following the manufacturer's instructions. Briefly, 1ml of well-mixed urine was added to a microfuge tube and centrifuged at 9,000 x g for 15 minutes at room temperature. The supernatant was aspirated and 1ml of LCx Urine Specimen Resuspension Buffer was added to the remaining pellet. The microfuge tubes were then vortexed to dislodge the pellet and heated at 97°C for 15 minutes, pulse-centrifuged and tested immediately or on the next days run. For amplification, 100µl of processed urine was added to individual LCx Chlamydia and LCx Gonorrhoeae Amplification Vials. Along with the respondent samples, 2 negative controls, 2 calibrators, and 1 positive control were prepared. Vials were placed into thermocyclers and amplification was carried out under assay-defined parameters. After amplification, controls, calibrators and respondent samples were placed into carousels for detection on the LCx microparticle enzyme immunoassay (MEIA) detection instrument.

In the presence of amplified CT and/or GC DNA in the respondent, 2 pairs of pathogen specific probes bind and are ligated. Probes (ligated and unligated) are captured onto microparticles via an antibody to the capture hapten on one end of the ligated probe. After a wash step, bound, ligated probes (only produced in samples with CT or GC DNA) are specifically identified by reaction with a second, alkaline phosphatase conjugate antibody that binds to the detection hapten on the other end of the ligated probe.

Unligated but bound probes are not detected as they have no detection hapten on the other end of the probe. The presence of bound antibody is detected by adding the substrate (4-Methylumbelliferyl Phosphate), which results in the generation of a fluorescent signal detected by the LCx instrument.

To determine whether a sample contains CT and/or GC DNA, the LCx detection instrument determines a cutoff value for each run based on calibrator rates. A signal-to-cutoff (S/CO) ratio is then determined for each sample in the assay run. The following criteria were used to determine the final results of each respondent sample (for both CT and GC) based on the S/CO ratio, and a summary of results is presented below.

Initial S/CO <0.80: Negative for CT and/or GC

Initial S/CO ≥0.80: All samples with this S/CO were repeat tested

Repeat S/CO ≥1.00: Positive for CT and/or GC

Repeat S/CO <1.00: Negative for CT and/or GC

Laboratory result	<i>Chlamydia trachomatis</i>	<i>Neisseria gonorrhoeae</i>
Positive	645	57
Negative	12,668	12,339
No results	699	1,616
Refused	1,185	1,185

### ***Trichomonas vaginalis* testing**

Marcia M. Hobbs, Ph.D.  
Research Associate Professor of Medicine, Microbiology & Immunology  
University of North Carolina

Urines were tested using an in-house PCR-ELISA that detects *Trichomonas vaginalis* DNA. This *T. vaginalis* PCR-ELISA is not FDA approved and is currently used for research purposes only. However, formal validation studies of the performance of the assay compared to wet mount and culture are published in the *Journal of Clinical Microbiology* [1,2] and described in the next section. The PCR primers used in this assay, TVK3 and TVK7, specifically amplify a 312-bp sequence from repetitive DNA in the *T. vaginalis* genome. PCR with these primers is negative with human DNA, other organisms found in the genitourinary tract and other *Trichomonas* species [3]. PCR products were detected using an ELISA format with a nucleotide probe (TVK) that binds a specific DNA sequence in the amplified product from positive specimens. Results were determined from absorbance values that ranged from 0 to 3.5. Based on the published ROC analysis of the test in urine from women, an ELISA absorbance cutoff of 3.0 was used to determine a positive result for specimens from female subjects. The published ROC analysis of the test in urine from men established 2.0 as the cutoff for male subjects. Negative test results were coded 0, and positive test results were coded 1.

Using a modification of the published assay procedures [1,2] Add Health urine specimens were tested using the specifics outlined below. Urine specimens were received and logged into the computer system in the UNC Hospitals Clinical Microbiology/Immunology Laboratories. A 1ml aliquot of each sample was labeled with the UID

number, and samples were transported to the Sexually Transmitted Disease Cooperative Research Center (STD CRC) Microbiology Core Laboratory for *T. vaginalis* testing. Samples were prepared within 2 days of receipt using the Amplicor CT/NG Urine Specimen Prep kit (Roche Diagnostic Systems, Indianapolis, IN) according to the manufacturer's instructions. Urine preps were stored at  $-70^{\circ}\text{C}$  until PCR-ELISA testing.

For PCR, 50  $\mu\text{l}$  of thawed urine prep was used as template in a reaction containing 40 pmol each of primers TVK3 and TVK7 (digoxigenin [DIG] labeled), 200 mM each dATP, dCTP, dGTP and dUTP, 2.5 U *Taq* polymerase (Gibco BRL, Grand Island, NY) 4mM  $\text{MgCl}_2$  and 1.0 U AmpErase (uracil N-glycosylase; Applied Biosystems, Foster City, CA) in 1x PCR buffer (Gibco BRL) in a final volume of 0.1 ml. PCR consisted of an initial 5-min incubation at  $95^{\circ}\text{C}$  followed by 40 cycles of denaturation at  $90^{\circ}\text{C}$  for 1 min, annealing at  $60^{\circ}\text{C}$  for 30 sec, and extension at  $72^{\circ}\text{C}$  for 2 min. Purified *T. vaginalis* DNA and sterile water were used as positive and negative controls with every batch of PCR tests.

PCR products were detected using the PCR ELISA DIG detection kit (Roche Diagnostic Systems) according to the manufacturer's instructions. For ELISA, 60  $\mu\text{l}$  of PCR product was incubated with 50 ng of biotinylated TVK probe in hybridization buffer in a final volume of 1.0ml; 200  $\mu\text{l}$  of the hybridization solution was pipetted into streptavidin-coated wells of the kit microtiter plate and incubated at  $37^{\circ}\text{C}$  for 2 h. with gentle rocking. Plates were washed 5 times in a microtiter plate washer containing kit washing solution, and 200  $\mu\text{l}$  of anti-DIG POD solution (diluted as per kit instructions) was added to each

well and incubated at 37°C for 30 min. with gentle rocking. Plates were washed as above, and 200 µl of ABTS substrate solution was added to each well and incubated at 37°C in the dark for 1 h. with gentle rocking. Absorbance was read at 405nm in a spectrophotometric plate reader with a linear range up to 3.5; absorbance of the kit negative control read at 492 nm was subtracted from sample readings as directed in the kit instructions. Tests were considered valid if the absorbance of the positive control was  $\geq 2.0$  and the negative control was  $< 0.2$ . Duplicate values for individual specimens were averaged and, based on the published ROC analyses of the test in urine, ELISA absorbance values (Abs) of 3.0 for specimens from women and 2.0 for specimens from men were used as cutoffs. Indeterminate results were not possible, and positive results were not verified by re-testing. Below is a summary of results.

Table 2: Laboratory testing for *Trichomonas vaginalis*

	Positive	Negative	No results	Refused
<i>Trichomonas vaginalis</i>	340	12,869	803	1,185

### Human papillomavirus (HPV) testing

Nancy B. Kiviat, M.D.  
 Professor of Pathology  
 University of Washington

Qinghua Feng, Ph.D.  
 Research Scientist  
 University of Washington

HPV test result data will be made available in the second release of Wave III data scheduled for the Fall of 2003.

Five ml aliquots of urine were flagged and shipped to the University of Washington and selected for HPV testing as described in Chapter 2. Upon receipt of the shipment, sample volume, condition, and receipt date were recorded. An additional internal lab number was assigned and used for DNA preparation, PCR amplification and dot blot detection. All samples were catalogued, thawed, and two 2.5 ml aliquots prepared. If there were less than 5 ml in the tube, the amount of sample was divided evenly into two aliquots. One aliquot was retained in the HPV laboratory for HPV testing and the other was re-frozen and forwarded to the *M. genitalium* laboratory for *M. genitalium* testing.

An experimental, non-commercial polymerase chain reaction (PCR)/Dot blot test was used to identify the presence of genital HPV DNA in the urine specimen, irrespective of HPV type. The urine specimen was thawed, aliquoted and spun to collect the urine pellet. The urine pellet was resuspended in 600 µl STM (storage and transfer media from Digene) and digested with 20 µg/ml protease K. Genomic HPV DNA was isolated from 100 µl of the digested urine pellet using QIAamp blood DNA mini column. DNA from each sample was amplified using HPV L1 consensus primers MY09/MY11/HMB01 and beta-globin primers PC04/GH20 [4]. The amplification mixture contained 6 mM MgCl<sub>2</sub>, 1x PCR buffer II, 7.5 U Gold Tag DNA polymerase (AmpliTaq Gold, Roche Molecular Diagnostics, Cat. No. N808-0249), 200 µM dNTPs, 50 pmole MY09 and MY11 primers, 10 pmole HMB01, PC04 and GH20 primers and 50-100 ng template DNA in 50 µl. The DNA was amplified using the following program: 95°C, 9 min; 95°C, 20 sec, 55°C, 30 sec, 72°C, 30 sec for 40 cycles; 72°C, 10 min. Ten microliter PCR products were then

dotted onto two nylon filters and probed with biotin-labeled beta-globin and HPV generic probes respectively. The dot blot results were read by two persons independently and any discrepancies resolved by a third individual's interpretation.

Beta-globin, a housekeeping gene, was used to monitor sample sufficiency and PCR amplification. When present, it indicates that sufficient cellular DNA is present and that amplification is occurring (e.g., the sample is not inhibited), and, when absent, it identified samples that were insufficient for analysis. Thus, recorded results indicated whether the specimen was positive for beta-globin and HPV. If the beta-globin result was negative, the HPV result was inconclusive and those samples were excluded. No re-testing of positive samples was performed.

### ***Mycoplasma genitalium* (MG) testing**

Patricia A. Totten, Ph.D.  
Research Associate Professor  
University of Washington

Although a general description of methods for *M. genitalium* PCR testing is presented here, more details will be provided with publication of the development and validation of this test (SM Dutro, JK Hebb, CA Garin, JP Hughes, GE Kenny and PA Totten, submitted). Data on *M. genitalium* test results will also be available in the second release of Wave III data scheduled for Fall 2003.

Samples to be tested for *M. genitalium* were received first in the HPV laboratory, thawed, aliquoted, and frozen at  $-80^{\circ}$  until preparation. After thawing, samples were processed using the AMPLICOR<sup>®</sup> CT/NG specimen preparation kit, according to manufacturer's directions (Roche Diagnostics Corporation, Indianapolis, IN, "urine procedure"). To test the subset of samples for *M. genitalium*, we used an *M. genitalium*-specific PCR assay targeting the MgPa genes as described [5] with modifications for the detection of PCR products using a microwell plate rather than a Southern blot based assay (SM Dutro, JK Hebb, CA Garin, JP Hughes, GE Kenny and PA Totten, submitted). This experimental, non-commercial assay is not FDA approved. Briefly, the PCR assay targets the MgPa genes of *M. genitalium* using the previously described primer sets, 5'-TGAAACCTTAACCCCTTGG-3' and 5'-AGGGGTTTTCCATTTTGC-3' [5], except that the primers were biotinylated in the current assay to allow attachment to streptavidin-coated plates. *M. genitalium*-specific PCR products were allowed to hybridize to a digoxigenin-labeled probe DNA, which was subsequently detected using the PCR ELISA DIG detection kit (Roche Diagnostic Systems). Cutoffs, determined by ROC analysis, were used to score tests as positive or negative (SM Dutro, JK Hebb, CA Garin, JP Hughes, GE Kenny and PA Totten, submitted). All positive and equivocal specimens were confirmed as positive if they repeated with one positive or two equivocal results upon repeat in duplicate using the same MgPa PCR assay. All negative specimens were not retested and reported as negative. Positive controls and negative controls for the whole procedure, including sample preparation (*M. genitalium* whole cells and sample preparation reagents alone) as well as the PCR assay alone (*M. genitalium* genomic DNA

and sterile water) were included in every batch of PCR assays and required to have the expected result for a valid run.

## REFERENCES

1. Kaydos SC, Swygard H, Wise SL, Sena AC, Leone PA, Miller WC, Cohen MS, Hobbs MM. Development and validation of a PCR-based enzyme-linked immunosorbent assay with urine for use in clinical research settings to detect *Trichomonas vaginalis* in women. *J Clin Microbiol* 2002;40(1):89-95
2. Kaydos-Daniels SC, Miller WC, Hoffman I, Banda T, Dzinyemba W, Martinson F, Cohen MS, Hobbs MM. Validation of a urine-based PCR-enzyme-linked immunosorbent assay for use in clinical research settings to detect *Trichomonas vaginalis* in men. *J Clin Microbiol* 2003;41(1):318-23
3. Kengne P, Veas F, Vidal N, Rey JL, Cuny G. *Trichomonas vaginalis*: repeated DNA target for highly sensitive and specific polymerase chain reaction diagnosis. 1994, *Cell Mol Biol* 40:819-831.
4. Manos, MM. The use of polymerase chain reaction amplification for the detection of genital human papillomavirus. In: Furth M., Greaves, M., eds. *Molecular Diagnostics of Human Cancer*, volume 7, Cold Spring, New York: Cold Spring Harbor Laboratory Press, 1989:209-214.
5. Totten PA, Schwartz MA, Sjostrom KE, Kenny GE, Handsfield HH, Weiss JB, Whittington WL. Association of *Mycoplasma genitalium* with nongonococcal urethritis in heterosexual men. *J Infect Dis* 2001;183(2):269-276.

## **Chapter V: Test Considerations in the Estimation of HIV and STI Prevalence**

William C. Miller, M.D., Ph.D., MPH  
Assistant Professor, Departments of Epidemiology and Medicine  
University of North Carolina

Marcia M. Hobbs, Ph.D.  
Research Associate Professor of Medicine and Microbiology & Immunology  
University of North Carolina

Nancy B. Kiviat, M.D.  
Professor, Department of Pathology  
University of Washington

Qinghua Feng, Ph.D.  
Research Scientist, Department of Pathology  
University of Washington

Patricia A. Totten, Ph.D.  
Research Associate Professor, Department of Medicine  
University of Washington

### **Test characteristics, measured prevalence, and estimated true prevalence**

The tests used for the identification of HIV and sexually transmitted infections are imperfect, with sensitivities and specificities less than 1. Estimation of the prevalence of these infections and associations with various potential risk factors will be affected by the test performance. The magnitude of these effects will depend on the performance characteristics of the tests and the prevalence of the disease in the population. Previous estimates of test performance can be used to guide analyses of the potential effects of the use of these imperfect tests.

The results of previous studies assessing test performance must be applied cautiously to the Add Health population. The study population of Wave III of Add Health is community-based, rather than clinic-based. Consequently, the prevalence of HIV

infection and most STIs is expected to be relatively low, certainly lower than most clinic populations. Furthermore, the majority of infections will be asymptomatic. In contrast, the majority of the tests used in the Add Health study were evaluated in clinic-based populations, frequently in symptomatic persons. The presence of symptoms is often related to the organism burden. Sensitivity of the diagnostic tests tends to be lower with lower organism burden.

The combination of low prevalence and imperfect diagnostic tests leads to more substantial measurement error than would be encountered in higher prevalence populations. The magnitude of the measurement error can be estimated using a simple formula that relies on estimates of the sensitivity and specificity.

$$\text{True Prevalence Estimate} = \frac{\text{Measured Prevalence} + \text{Specificity} - 1}{\text{Sensitivity} + \text{Specificity} - 1}$$

For example, given sensitivity of 90%, specificity of 99.5%, and measured prevalence of 1%, the prevalence estimate would be:

$$\text{True Prevalence estimate} = (0.01 + 0.995 - 1) / (0.90 + 0.995 - 1) = 0.0056 \text{ or } 0.56\%$$

Generally, in the range of the sensitivities and specificities of the tests used in this study and the expected prevalences of the STIs, the specificity has the greatest effect on the

measured prevalence (Figures 1 & 2). One important feature to note is that the measured prevalence cannot be below a certain level with a given specificity. Thus, it may be possible to argue that a test performed with higher specificity than previously estimated because of the measured prevalence. This possibility highlights the fact that the actual performance of a diagnostic test is never certain within the context of a given research study.

### **Test Characteristics**

The tests used in Add Health have all been validated in other populations. However, the generalizability of the validation study populations, which are typically clinic-based, to the largely asymptomatic general population of 18-26 year olds in Add Health varies. Furthermore, the methodological issues related to test validation for some of these assays are substantial, leading to further uncertainty in the test characteristics.

The diagnostic tests used for the identification of *Chlamydia trachomatis* (CT), *Neisseria gonorrhoeae* (GC), *Trichomonas vaginalis* (TV), human papillomavirus (HPV) and *Mycoplasma genitalium* (MG) are nucleic acid amplification tests (NAAT). Assessing the performance of NAAT is problematic because no perfect reference standard exists and, in many cases, the new NAAT tests are considerably more sensitive than the traditional reference standard, culture. Various methods have been used to overcome the reference test bias, but none of the commonly used methods are completely satisfactory [1]. Thus, the performance characteristics of NAATs used in Add Health are uncertain.

Still, the results of previous studies provide a reasonable starting place for considering the test performance.

*CT LCR Assay:* The package insert for the Abbott LCx Assay reports a sensitivity of 94.1% and specificity of 95.2% in urine for asymptomatic women. In asymptomatic men, the sensitivity and specificity are reported as 90.0% and 97.9%, respectively. These estimates of the specificity are likely to be underestimates due to poor sensitivity of the reference standard.

In a recent meta-analysis, four studies of the performance of urine-based LCR for CT in asymptomatic men and women were identified [2]. The sensitivity ranged from 87.5 – 96%. The specificities were all reported as 100%, although these estimates were based on discrepant analysis, a biased procedure known to lead to overestimation of test performance. A recent large study in men that used a combination of tests as the reference standard estimated the sensitivity as 80.4 – 89.3% and specificity as 95.5 – 99.6% [3]. A similar study in women revealed sensitivity of 80.8% – 83.4% and specificity of 97.8% – 98.9% [4]. The imperfect sensitivity of the combined reference standards of these studies may result in a slight underestimation of specificity.

*GC LCR Assay:* The package insert for the Abbott LCx Assay reports a sensitivity of 100.0% and specificity of 98.5% in urine for asymptomatic women. In asymptomatic men, the sensitivity and specificity are reported as 100.0% and 100.0% respectively. However, these results are based on very small numbers of subjects. The overall

sensitivity and specificity in combining symptomatic and asymptomatic men and women are calculated as 94.9% and 99.9%. These estimates of the specificity are likely to be underestimates due to the imperfect sensitivity of the reference standard.

In a systematic review of the performance of LCR for GC, the reported sensitivity in women was 94.3% - 100% and the specificity was 98.7% - 99.7% before discrepant analysis [5]. In men, the sensitivity was 98.0% - 100% and specificity 98.6 – 100%. All studies reported specificities of 100% after discrepant analysis. In the Add Health population, all GC tests with S/CO of  $\geq 0.8$  were repeated and confirmed positive only if repeat amplification yielded a S/CO of  $\geq 1.0$ . This procedure should improve specificity and reduce the number of false positives.

*TV PCR Assay:* The TV PCR assay is available for research purposes only. In the only evaluation studies to date, the assay performed with a sensitivity of 90.8% and adjusted specificity of 93.4% in women [6]. In men, the sensitivity was 88.9% and the adjusted specificity was 94.5% [7]. The sensitivities and specificities of the assay by symptom status are listed in Table 1. These specificities were based on an algebraic adjustment of the observed specificities. The specificities are undoubtedly lower than actual performance, even after adjustment, because of the very poor sensitivity of culture, the reference standard.

Table 1. Performance of *T. vaginalis* PCR-ELISA compared to combined reference of wet mount and culture

Subject group		% Sensitivity (95% CI)	Adjusted % Specificity <sup>a</sup>
Women	Symptomatic <sup>b</sup>	89.5 (65.5, 98.2)	90.7
	Asymptomatic	91.8 (84.1, 96.2)	95.3
Men	Symptomatic <sup>c</sup>	87.2 (73.6, 94.7)	97.5
	Asymptomatic	90.9 (74.5, 97.6)	92.7

<sup>a</sup> Adjusted specificity using the formula of Staquet *et al.* (J. Chron. Dis. 34:599-610; 1981); assumes wet mount /culture has sensitivity of 70% for women and 50% for men.

<sup>b</sup> Abnormal vaginal discharge.

<sup>c</sup> Urethritis  $\geq 4$  wbc/hpf on urethral gram stain.

The limit of detection for *T. vaginalis* PCR-ELISA is 8 organisms per PCR (160 organisms per ml of urine). These data are from urine specimens with volumes ranging from 2 to 100ml (mean 31.8 ml); the test is slightly more sensitive in specimens  $\leq 20$ ml for women (100.0%; 95% CI: 82.2, 100.0) and  $\leq 30$  ml for men (92.75; 95% CI: 79.0, 98.1).

*HPV PCR/dot blot hybridization assay:* The analytic limit of detection for this assay is one HPV molecule per sample. For women, the assay is less sensitive in urine samples than it is in cervical swab samples because of the lower number of cervical epithelial cells and the greater amount of PCR inhibitors in urine. To address this, Qiagen columns were used to purify DNA and eliminate inhibitors. The assay may be less sensitive in men, because fewer HPV-containing epithelial cells are shed into urine, but remains specific. The dot blot detection procedure is more specific than the FDA-approved

Digene hybrid capture assay, and the generic probe can detect a majority of the 40 anogenital HPV types.

*MgPa PCR assay for M. genitalium:* The MgPa PCR assay is a new, experimental assay available for research purposes only. The analytical sensitivity (limit of detection) is 1 (with 38% confidence) and 15 (with 95% confidence) genome copies of *M. genitalium*. Since culture is extremely difficult for this fastidious organism, and alternative types of detection have not been developed, the sensitivity and specificity of the *M. genitalium* PCR has not been determined. However, analysis of a subset of PCR positive and negative samples using another *M. genitalium*-specific PCR test or by sequencing the PCR product is consistent with high specificity of this assay [8, 9]. Nevertheless, the relative sensitivity and specificity of the *M. genitalium* PCR test using different specimen types, both from men (urethra vs. urine) and women (vagina vs. cervix vs. urine) requires further study.

*HIV Orasure Assay:* The only published assessment of the HIV Orasure Assay was conducted among 3570 subjects from a variety of settings (blood banks, clinics, and others) [10]. Among 673 HIV positive persons, 665 were identified as positive by the Orasure assay (sensitivity=0.988). The specificity was 0.9986, based on 2880 negative EIA results, 13 EIA positive/WB negative, and 4 EIA positive/WB indeterminate results among a population of 2897 HIV negative persons.

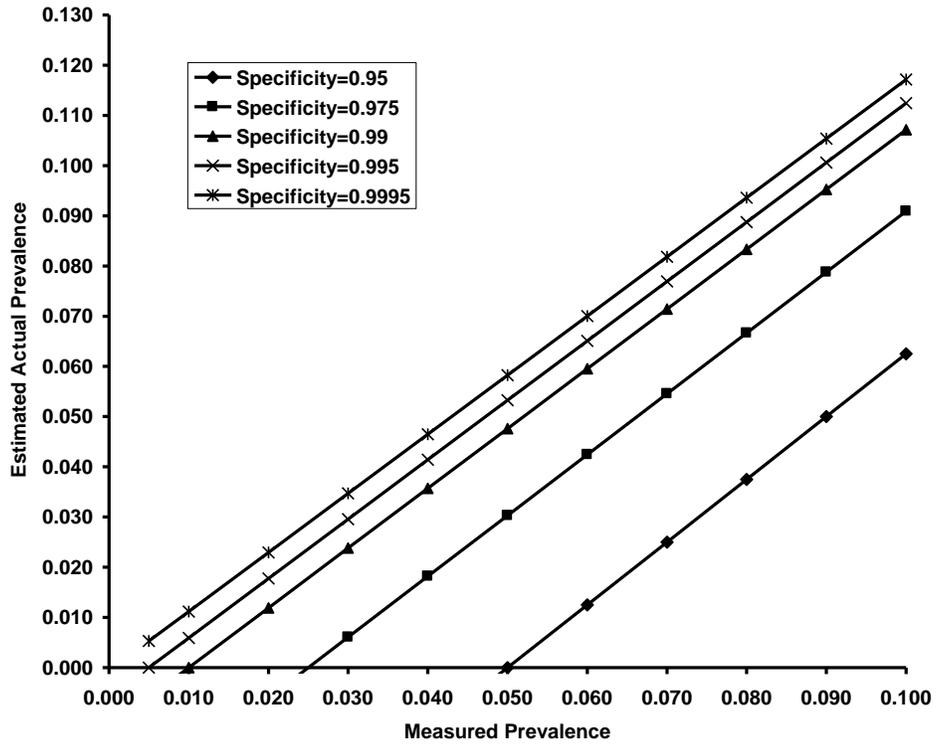


Figure 1. The relationship between measured prevalence and estimated actual prevalence with varying specificity. Sensitivity is held constant at 0.85.

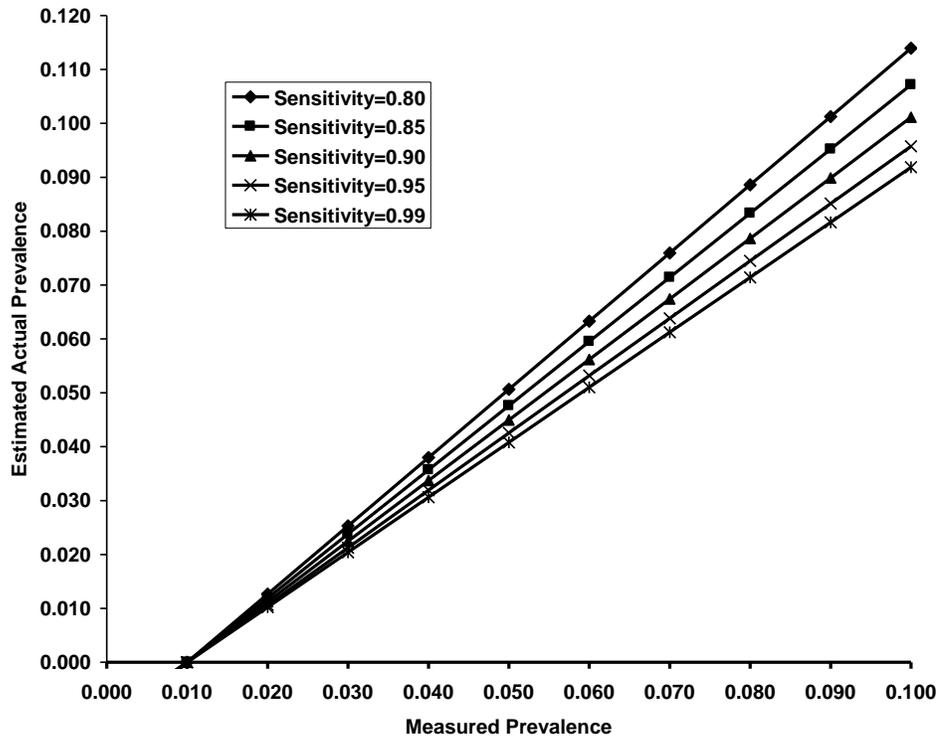


Figure 2. The relationship between measured prevalence and estimated actual prevalence with varying sensitivity. Specificity is held constant at 0.85.

## REFERENCES

1. Miller, WC. Can we do better than discrepant analysis for new diagnostic test evaluation? *Clin Infect Dis* 1998;27(5):1186-93.
2. Watson EJ, Templeton A, Russell I, Paavonen J, Mardh PA, Stary A, Pederson BS. The accuracy and efficacy of screening tests for *Chlamydia trachomatis*: a systematic review. *J Med Microbiol* 2002;51(12):1021-31.
3. Johnson RE, Green TA, Schachter J, Jones RB, Hook EW 3rd, Black CM, Martin DH, St Louis ME, Stamm WE. Evaluation of nucleic acid amplification tests as reference tests for *Chlamydia trachomatis* infections in asymptomatic men. *J Med Microbiol* 2002;51(12):1021-31.
4. Black CM, et al. Head-to-head multicenter comparison of DNA probe and nucleic acid amplification tests for *Chlamydia trachomatis* infection in women performed with an improved reference standard. *J. Clin. Microbiol.* 2002;40:3757-3763.
5. Koumans EH, Johnson RE, Knapp JS, St Louis ME. Laboratory testing for *Neisseria gonorrhoeae* by recently introduced nonculture tests: a performance review with clinical and public health considerations. *Clin Infect Dis* 1998;27(5):1171-80.
6. Kaydos SC, Swygard H, Wise SL, Sena AC, Leone PA, Miller WC, Cohen MS, Hobbs MM. Development and validation of a PCR-based enzyme-linked immunosorbent assay with urine for use in clinical research settings to detect *Trichomonas vaginalis* in women. *J Clin Microbiol* 2002;40(1):89-95
7. Kaydos-Daniels SC, Miller WC, Hoffman I, Banda T, Dzinyemba W, Martinson F, Cohen MS, Hobbs MM. Validation of a urine-based PCR-enzyme-linked

immunosorbent assay for use in clinical research settings to detect *Trichomonas vaginalis* in men. J Clin Microbiol 2003;41(1):318-23.

8. Cohen CR, Manhart LE, Bukusi EA, Astete S, Brunham RC, Holmes KK, Sinei SK, Bwayo JJ, Totten PA. Association between *Mycoplasma genitalium* and acute endometritis. Lancet 2002;359(9308):765-6.
9. Manhart LE, Critchlow CW, Holmes KK, Dutro SM, Eschenbach DA, Stevens CE, Totten PA. Mucopurulent cervicitis and *Mycoplasma genitalium*. J Infect Dis 2003;187(4):650-7.
10. Gallo D, George JR, Fitchen JH, Goldstein AS, Hindahl MS. Evaluation of a system using oral mucosal transudate for HIV-1 antibody screening and confirmatory testing. OraSure HIV Clinical Trials Group. JAMA 1997;277(3):254-8.

## **Chapter VI – DNA sampling, genotyping, and determination of zygosity.**

Andrew Smolen, PhD  
Institute for Behavioral Genetics  
University of Colorado  
Boulder, CO 80309

John K. Hewitt, PhD  
Institute for Behavioral Genetics  
University of Colorado  
Boulder, CO 80309

### **Introduction**

Respondents identified as full siblings or twins at earlier waves (n=3,787) were asked to supply saliva specimens for DNA analysis as described in Chapter 2, and researchers at the University of Colorado are conducting genetic analyses on these biospecimens. Data generated from these analyses will enable behavior genetic studies of adolescent health to move from variance decomposition into anonymous components, to the testing of specific hypotheses about the influence of individual genes and their expression in the context of environmental circumstances. Beyond the specific candidate genes identified for study, DNA collection in conjunction with the rich phenotypic and environmental data gathered in the Add Health project creates an important research resource for future studies.

Saliva samples were collected from full siblings or twins to genotype DNA for seven 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 initially targeted candidates are the dopamine transporter (DAT1), the dopamine D4 receptor (DRD4), the serotonin transporter (5HTT), monoamine oxidase A (MAOA), monoamine oxidase B (MAOB), the dopamine D2 receptor (DRD2), and the dopamine D5 receptor (DRD5). The laboratory methods for the three primary candidates (DAT1, DRD4, 5HTT) are described below. Genotyping for 11 neutral markers for the purposes of zygosity determination will also be conducted where needed.

### **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°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  $\mu$ l 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  $\mu$ l 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  $\mu$ l 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  $\mu$ l 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  $\mu$ g/ml), and an aliquot was diluted to a concentration of 20 ng/ $\mu$ l or less for a working sample. The average yield of DNA was  $58 \pm 1$   $\mu$ g. DNA samples were stored at 4°C and transported to the Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado for genotyping.

**Genotyping.** Initial screening of the samples included the analysis of three polymorphic candidate genes: the dopamine transporter, the serotonin transporter and the dopamine D4 receptor. Additional genes identified are the dopamine D2 receptor gene, the monoamine oxidase A promoter, and the CYP2A6 gene.

**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 Vandenberg 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 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 Variable Number Tandem Repeat (VNTR) in the 5' regulatory region of the gene [15] which has been shown to affect expression [16]. The MAOA gene product is primarily responsible for the degradation of dopamine, serotonin, and norepinephrine and is located in the nigro-striatal pathway of the brain [17,18]. A genotype by environment interaction has been reported for this polymorphism. This finding has been both supported [20] and questioned [21,22].

The MAOA-uVNTR polymorphism was assayed by a minor modification [21] of a published method [16]. 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' [16]. PCR reactions contained 1 µl of DNA (20 ng or less), 1.8 mM MgCl<sub>2</sub>, 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 (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 [13] as detailed previously [14] and analyzed in an ABI 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.

**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 SNPs, di-nucleotide repeats, and restriction endonuclease sites in coding and non-coding regions [23]. 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 [24,23]. We redesigned the assay as a SNP assay using the Applied Biosystem's "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 [25] was performed using the fluorogenic 5' nuclease (Taqman®, Applied Biosystems, Foster City, CA) method using reagents (VIC™ and FAM™ labeled probes and TaqMan® Universal PCR Master Mix without AMPerase® UNG) obtained from Applied Biosystems (ABI). Unlabeled forward and reverse primers were purchased from ABI or Integrated DNA Technologies (Coralville, IA). All

reactions were performed in an ABI Prism® 7000 Sequence Detection System using the allelic discrimination mode as described by Livak [26] and in the accompanying instrument documentation. Methods used were as described in the package inserts from ABI, except that the total reaction volume was decrease 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-CCTGCCTTGACCAGC-NFQMGB-3'; Probe 2: 5'- FAM-CTGCCTCGACCAGC-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 95°C followed by 40 cycles of 92°C for 15 sec and 60°C 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.

**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 [27,28]. 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 [29] that has been associated in epidemiological studies with reducing the number of cigarettes smoked [30,31,32]. The CYP2A6\*2 assay [25] was performed using the fluorogenic 5' nuclease (Taqman®, Applied Biosystems, Foster City, CA) method using reagents (VIC™ and FAM™ labeled probes and TaqMan® Universal PCR Master Mix without AMPerase® UNG) obtained from Applied Biosystems (ABI). Unlabeled forward and reverse primers were purchased from ABI or Integrated DNA Technologies (Coralville, IA). All reactions were performed in an ABI Prism® 7000 Sequence Detection System using the allelic discrimination mode as described by Livak [26] 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](http://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'-TCGTCCTGGGTGTTTTTCCT-3'; Probe 1: 5'- FAM-TCGACGCCCTCCGGGGC-TAMRA-3'; Probe 2: 5'- VIC-TCGACGCCCACCGGGGC-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 95°C followed by 40 cycles of 92°C for 15 sec and 60°C 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.

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

After amplification, 1  $\mu$ l of PCR product was combined with 2  $\mu$ L of loading buffer containing size standard (Genescan 2500 TAMRA<sup>®</sup>, PE-Biosystems), and 0.8  $\mu$ l 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.

We have tested the validity of allele calls for the three genotyping protocols described above, using preamplified DNA compared with genomic DNA for 315 individuals (1890 allele calls) from non-AddHealth 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].

**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  $\mu$ M of each of the four dNTP's, 2.5 mM MgCl<sub>2</sub>, 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  $\mu$ l. 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°C for 10 min (to activate the Gold polymerase); 35 cycles of 94°C for 10 sec, 55°C for 90 sec, and 72°C for 90 sec; a final elongation at 72°C for 8 min; and a hold at 4°C when finished. PCR products are stored at -20°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**

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. *Amer J Hum Genet* 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. *Behav Genet* 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. Vandenberg, 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). *Genet Epidemiol* 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. *Amer J Med. Genet* 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<sub>4</sub> (DRD4) gene. *Hum Mol Genet* 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. *J Neurochem* 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. Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucl Acids Res*

- 1992;19: 4008.
14. 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.
  15. 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.
  16. Sabol S.Z., Hu, S. and Hamer, D. A functional polymorphism in the monoamine oxidase a gene promoter. *Hum. Genet* 1998;103: 273-279.
  17. 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 Res* 1991;19: 4537-4541.
  18. Shih, J.C., Chen, K., and Ridd, M.J. Monoamine oxidase: From genes to behavior. *Annu Rev Neurosci* 1999;22: 197-217.
  19. Caspi A., McClay, J., Moffitt, T.E., Mill, J., Martin, J., Craig, I.W., Taylor, A., and Poulton, R. Role of genotype in the cycle of violence in maltreated children. *Science* 2002;297: 851-854.
  20. 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. *Arch Gen. Psychiatr* 2004;61: 738-744.
  21. Haberstick, B.C., Lessem, J.M., Hopfer, C.J., Smolen, A., Ehringer, M.A.,

- Timberlake, D., and Hewitt, J.K. MAOA genotype and antisocial behaviors in the presence of childhood and adolescent maltreatment. *Am. J. Hum. Genet* (Neuropsych. Genet.) 2005; In Press.
22. Young, S.E., Smolen, A, Hewitt, J.K, Haberstick, B.C, Stallings, M.C, Corley, R.P, and Crowley, T.J. Interaction between MAOA genotype and maltreatment in the risk for conduct disorder: failure to replicate. 2005; under review.
23. Noble, E. P. D2 dopamine receptor gene in psychiatric and neurologic disorders and its phenotypes. *Amer. J. Med. Genet* 2003;116B: 103-125.
24. 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. *Amer. J. Med. Genet.* 1998;81: 257-267.
25. Haberstick, B.C. and Smolen, A. Genotyping of three single nucleotide polymorphisms following whole genome preamplification of DNA collected from buccal cells. *Behav. Genet* 2004;34: 541-547.
26. Livak, K.J. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genetic Analysis: Biomolecular Engineering* 1999;14: 143-149.
27. Raunio, H., Rautio, A., Gullsten, H., and Pelkonen, O. Polymorphisms of CYP2A6 and its practical consequences. *J. Clin. Pharmacol* 2001;52: 357-363.
28. Oscarson, M. Genetic polymorphisms in the cytochrome P450 2A6 (CYP2A6) gene: Implications for interindividual differences in nicotine metabolism. *Drug Metab. Disp* 2001;29: 91-95.
29. Tricker, A.R. Nicotine metabolism, human drug metabolism, polymorphisms, and smoking behaviour. *Toxicology* 2003;183: 151-173.

30. 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.
31. Pianezza, M.L., Sellers, E.M., and Tyndale, R.F. Nicotine metabolism defect reduces smoking. *Nature* 1998;25: 750-751.
32. Tyndale, R.F. and Sellers, E.M. Variable CYP2A6-mediated nicotine metabolism alters smoking behaviour and risk. *Drug Metab. Disp* 2001;29: 548-552.