TECHNICAL PAPER 66

PCR detection of Y-chromosome sequences in vaginal fluid: Preliminary studies of a potential biomarker for sexual behavior.

J.M. Zenilman, J. Yuenger, N. Galai, C.F. Turner, S.M. Rogers

Reference Citation

Zenilman JM, Yuenger J, Galai N, Turner CF, Rogers SM. PCR detection of Y-chromosome sequences in vaginal fluid: Preliminary studies of a potential biomarker for sexual behavior. (Published in Sexually Transmitted Diseases, in press.).
Polymerase Chain Reaction Detection of Y Chromosome Sequences in Vaginal Fluid: Preliminary Studies of a Potential Biomarker for Sexual Behavior

JONATHAN M. ZENILMAN, MD,*† JEFFREY YUENGER,* NOYA GALAI,† CHARLES F. TURNER, PHD,‡§ AND SUSAN M ROGERS, PHD¶

Background: Self-reported measures of sexual behavior are subject to nontrivial reporting biases.

Objective: The objective of this study was to develop a behavioral biomarker of recent sexual activity among females that is inexpensive, easily administered, and can be used in low sexually transmitted disease prevalence populations.

Methods: We developed a polymerase chain reaction (PCR) assay to detect Y chromosome (Yc) fragments. The Yc primers were developed against a 200-basepair (bp) microsatellite repeat sequence, which is unique to the male genome. A standard PCR technique was used. Assay sensitivity was determined quantitatively using donated semen samples. To assess longevity of detectability, we recruited female subjects in monogamous relationships. Seventeen subjects had unprotected intercourse followed by 3 weeks of abstinence from vaginal intercourse. Self-administered vaginal swabs (SAVS) were collected every other day. In addition to the swabs, subjects kept daily sexual diaries. Swabs were processed by semiquantitative PCR, and Yc decay curves were determined for each subject. The half-life of Yc in vaginal fluid was calculated on the collection of individual decay curves by a random-effects regression model approach.

Results: The sensitivity of our Yc-PCR assay was determined to be 5 copies of Yc. In the longevity studies, Yc was detectable in SAVIS up to 15 postcoital days (PCD). Mean Yc DNA concentration in SAVIS eluate followed an exponential decay pattern for each subject. Mean concentrations were 66.7 ng/mL at PCD-1, 20.6 ng/mL at PCD-7, and 4.5 ng/mL at PCD-15. The estimated half-life for Yc clearance was 3.83 days.

Conclusion: The swab-based Yc-DNA PCR assay can detect coitus in women for a 2-week retrospective period. This can be used to validate sexual behavior-reporting and condom use in women and promises to be a useful tool in sexual behavior research.

CONSISTENT CONDOM USE IS a central behavioral strategy to control the spread of HIV and other sexually transmitted diseases (STDs).1–3. Condom promotion interventions have drawn on repetitive and intervention models to increase consistent condom use.11,12 A key methodologic weakness is that measurement of condom use is almost always dependent on self-reports.13–16 In settings such as STD clinics, self-reported condom use is biased toward overreporting, presumably as a result of contextual factors.17,18 Biologic approaches such as using incident STDs are costly and require establishing a laboratory and clinical infrastructure. A National Institute of Mental Health consensus conference16 has therefore recommended that STD biomarkers be used only in phase 3 clinical trials of behavioral interventions, and in close collaboration with laboratory experts and clinicians.16 Nevertheless, even in high-incidence settings such as STD clinics, only a minority of subjects with risky behavior gets infected with an STD. For example, in our previous Baltimore studies, 3-month STD incidence rates of 15% to 20% were observed.17,19 Therefore, if STD biomarkers were solely used as measures of “safer sex behavior,” 80% to 85% of individuals would be classified as low risk.

The intrinsic genetic difference between the male and female genome offers an opportunity to develop a sexual exposure biomarker, because Y chromosome (Yc) sequences are unique to males. Screening for the Y chromosome in vaginal fluid could provide a highly sensitive validity marker, because sperm and other semen-derived cellular material are deposited in the vagina during unprotected vaginal intercourse.

We adapted forensic protocols for Y chromosome detection in vaginal fluid to develop an easily used biomarker of recent sexual intercourse. We provide estimates of assay sensitivity, and we have measured the duration that Yc can be detected in the vagina after unprotected intercourse under controlled circumstances.
Development of the Y Chromosome Polymerase Chain Reaction Assay

We adapted a PCR assay for Y chromosome DNA components from forensic science protocols. The primers were specific for microsatellite repeats unique to the X and Y chromosomes. The Y primers were: Y3: 5'-GGTTTACACCTCGGGAGG and Y4: 5'-ACAAAAGGGTCATTGCTGTAGG; the X primers were X3: 5'-TATTGGAATCTCCGAGGA and X4: 5'-TTCTACACTAACAGGTTTCA. Primers were synthesized at the Johns Hopkins Genetics Core Laboratory. Y-specific primers amplify a 200-basepair (bp) sequence of the α-satellite region of the Y chromosome. X primers amplify a 157-bp sequence of the α-satellite region of the X chromosome and were used as controls.

Polymerase Chain Reaction Amplification

During initial development, PCR reactions were conducted exclusively by a female laboratory worker to prevent possible Y chromosome contamination in a sterile or clean PCR hood. PCR amplification of 9 μL of the DNA extract was performed using a modification of previously described procedures.23 The samples were amplified using a Perkin Elmer Cetus thermocycler in a final volume of 55 μL containing 45 μL PCR Supermix (Life Technologies) and either 0.4 μmol/L X primer or 0.8 μmol/L Y primer. The thermocycler profile was 3 cycles at 94°C for 2 minutes, 60°C for 2 minutes, 72°C for 2 minutes, followed by 25 cycles of 94°C for 1 minute, 60°C for 1 minute, and 72°C for 1 minute.

The X and Y amplifications were run in separate reaction tubes. The higher concentration (0.8 μM) allowed us to detect spermatozoa in concentrations of over 1:2000 spermatozoa to epithelial cells. However, this high concentration also occasionally resulted in the amplification of larger DNA fragments because we amplified a microsatellite repeat region. Controls for the PCR reaction included a control to detect Y chromosome contamination. A male DNA sample was used as a positive control for the initial experimentation.

Electrophoresis and Detection of Amplified DNA

The amplified products were separated on precast 5% polycrylamide mini gels (Bio-Rad) in tris borate EDTA (TBE) buffer. A total of 9 μL of PCR product and 1 μL loading buffer (20% w/v Ficoll 400, 0.1 mol/L Na2EDTA at pH = 8, 1% w/v SDS, and 0.25% w/v bromophenol blue) were added to each lane of the gel. Y bands were detected at 200 bp and X bands at 157 bp. On each gel was included a 50-bp DNA (Life Technologies) marker and a low-mass DNA ladder (Life Technologies) to generate the standard curve for spot density measurements for the Y bands. The electrophoresis running time was 1 hour at 70 V. The gels were stained with 0.5 μg/mL ethidium bromide for 3 minutes and destained with distilled water for 3 minutes. Gels were visualized, and spot densities of the Y chromosome DNA fragments were calculated and recorded using an Alpha Imager 2.200 (Alpha Innotech).

Phase 1: Determining the Sensitivity of the Assay

Semenal fluid was obtained from volunteers (courtesy Kevin Whalley, PhD, Department of Biophysics, Johns Hopkins University) under a protocol approved by The Johns Hopkins University Institutional Review Board. Visual cell counts were determined using a Neubauer hemocytometer and was estimated as 127.3 × 10^6 sperm/mL. One microliter of seminal fluid was used in the differential DNA extraction, and extracted sperm DNA was serially diluted as previously described.22 The DNA was extracted by boiling 9 minutes in a solution of 5% Chelex 100 and 1 μL used for PCR amplification. The Yc assay was run in duplicate on the dilutions to determine limit of detectability.

Phase 2: Clinical Characterization of the Assay in Sexually Active Women

Our initial experiment was to determine the longevity of detectable Y chromosome components in vaginal fluid after unprotected sexual intercourse. We recruited 21 heterosexual adult female volunteers between the ages of 19 to 37 who were in monogamous relationships, not pregnant, and who were currently using nonbarrier hormonal methods of contraception. Volunteers were informed of the study goals and tasks for data and specimen collection before consenting to participate. All study participants signed a written informed consent. Study procedures were approved by the Institutional Review Boards at The Johns Hopkins University School of Medicine and Research Triangle Institute.

Nineteen of the 21 female volunteers were instructed to engage in unprotected sexual intercourse and then refrain from further vaginal intercourse for a 21-day period. There were 2 persons who were currently sexually abstinent (>1 month) and were negative controls. All women were trained and given an instruction booklet that we had previously developed for collecting self-administered vaginal swabs (SAVS). On the postcoital morning (day 1) and every other day thereafter (ie, days 3, 5, . . . up to day 21) a vaginal fluid sample was obtained by self-administered swab. Collected swabs were stored at 4°C until delivery to the laboratory.

All subjects kept detailed daily diaries, which included prompts for penile/vaginal intercourse, the use of condoms, diaphragms, spermicide, lubricants, and other contraceptives, receptive oral sex, digital penetration by male partner, douching, menses, and the use of tampons or pads. Subjects were intensively counseled by study staff that the project objective was assay validation and of the importance of reporting validity on the self-reported behaviors. Subjects were encouraged to report protocol deviations and were advised that they would still receive their reimbursement for participation even if deviations were reported.

DNA Extraction From the Vaginal Swabs

Vaginal fluid components were removed from the SAVS by placing in 0.5 mL sterile water, incubating at room temperature for 5 minutes, and rotating vigorously. The specimen was centrifuged for 1 minute at 10,000 g and all but 200 μL of supernatant was removed and then vortexed to resuspend the pellet.

The resulting specimen was extracted for Y chromosome DNA by a modified 2-step differential extraction technique.23 Briefly, 2 μL of proteinase-K (10 mg/mL) was added to the sample and incubated at 56°C for 1 hour. The sample was centrifuged for 3 minutes at 10,000 g, the supernatant removed, and the pellet was resuspended in 0.5 mL wash buffer containing 10 mmol/L Tris-HCl 7.5, 10 mmol/L EDTA, 50 mmol/L NaCl, and 2% SDS. The pellet was vortexed then centrifuged for 3 minutes at 10,000 g. The supernatant was discarded and the washing was repeated 2 more times. The pellet was washed a final time in 1 mL distilled water, then vortexed, and centrifuged for 3 minutes. The pellet was then resuspended in 150 μL of 5% (w/v) Chelex 100 (100–200 mesh, sodium form, biotechnology grade; Bio Rad), 2 μL of proteinase-K (10 mg/mL), and 7 μL of 1 mol/L DTT. The sample was incubated for 1 hour at 56°C. After incubation, the sample was vortexed for 5 to 10 seconds, centrifuged for 5 to 10 seconds at 10,000 g, and boiled for 8 minutes. The supernatant was then used for DNA amplification using the procedure described here.
subject-specific effects. The vectors \( (b_{0i}, b_{1i}) \) are assumed to be independently distributed as multivariate normal with mean 0 and arbitrary variance–covariance matrix. The errors \( \{e_{it}\} \) are modeled as independent and identically distributed normal random variables with constant variance \( \sigma^2 \). These models are also known as hierarchical linear models.

When the reading of YcDNA was below the detectable level, we assigned these data points the value of 0.5 so that the log can be defined. The parameters were estimated using the restricted maximum likelihood (REML) method with SAS proc mixed program (SAS Institute, Cary, NC). Ninety-five percent confidence intervals were provided for the parameter estimates. Applying the mixed model described here will give not only estimates of the group-average decline curve and the estimated half-life, but also the estimated distribution of the decline parameters among individuals in the group.

**Results**

**Sensitivity of the Assay**

To determine the selectivity of the differential DNA extraction, an extraction was performed in which epithelial cells from a buccal swab of a female were mixed with sperm. Six tubes were filled with 1 mL of TE buffer containing approximately \( 5 \times 10^5 \) epithelial cells.

TE buffer containing 2560, 1280, 640, 320, 160, and 80 sperm (by definition, half would contain the X chromosome and half the Y chromosome) were added, respectively, into each tube. This created a ratio of Y chromosome sperm to epithelial cell from 1:195 to 1:6250. The cellular mixtures were subjected to differential DNA extraction. The microcortals of the DNA products were used in a 100 \( \mu \)L PCR amplification using 0.8 \( \mu \)M primer concentration. The PCR products were run on PAGE minigels and stained with ethidium bromide. Clear Y bands were evident down to a dilution of 1:6250 (a maximum of 4 y–chromosome-containing sperm).

**Longevity of Detectable Yc DNA in Vaginal Fluid**

The experimental design was semiquantitative because it is difficult to standardize the quantity of semen present in the vagina or the amount of sample collected. Two of our 19 subjects were excluded because they reported multiple coital episodes in the diaries. In our 17 experimental subjects, the mean DNA content was 66.7 ng on day 1, which decreased to 20.6 ng by day 7, 4.5 ng by day 15, and 2.3 ng on day 17 (Fig. 1). The predicted line, based on hierarchical linear model of ln(YcDNA), had a mean intercept of 2.99 with standard error of 0.39 (95% confidence interval [CI], 2.24–3.76) and the slope was \(-0.181 \) with a standard error of 0.022 (95% CI, \(-0.22 --0.14 \)). The model also showed that there was significant variability in the initial values (intercept) among the women but a nonsignificant variability in the slopes. This finding supports the notion that the decline of YcDNA has a similar rate for different women, whereas the main differences are the result of the initial quantity present. The estimated half-life was 3.83 days (95% CI, 3.10–5.02). No YcDNA was detected in the samples obtained from the 2 abstinent controls.

A typical gel pattern, shown in Figure 2, demonstrates the disappearance of the Y–PCR product band over the 3-week study period, with nearly full extinction at 15 days. This pattern was replicated in all of the gel patterns. We arbitrarily chose a 2-week cutoff because of the system’s limited resolution beyond 15 days and for the wide confidence limits at low concentrations.
example, Gaydos found that chlamydial DNA persists in the female genital tract up to 15 to 21 days posttreatment\textsuperscript{35} and is the basis for the current recommendations that nucleic acid amplification tests not be used to demonstrate treatment efficacy. Our assay has potential use in field settings, because PCR-based assays can be collected on dry swabs without need for refrigeration or other transport requirements.

Y chromosome sequence detection would produce false-negative erroneous results in only a small number of situations. For example, from the male, vasectomized males or those with azoospermia would produce false-negative results. Cross-sectional population-based studies have found the prevalence of azoospermia to be 0.4\% to 4\%, in which men of all ages were considered.\textsuperscript{36,37} In most field settings evaluating STD risk, these conditions would be relatively rare.

Our data are subject to limitations. Our sample size in this pilot study was small, and consisted of volunteers who were students and employees at our medical center. Use of different feminine hygiene products such as tampons or pads may impact on Yc clearance, but our sample size was not large enough to detect differences, if they existed. None of the subjects reported douching, a common practice in many populations at risk for STDs, which may affect Yc clearance. The assay sensitivity would be substantially reduced in women who practice coitus interruptus, in which the number of spermatozoa deposited in the vagina is substantially reduced. Our assay can only be used in women, because they do not have native Yc sequences. Therefore, this assay would not be useful to determine either heterosexual or homosexual exposure in men.

We have demonstrated that the Yc assay is sensitive and that detectable Yc components are present for 2 weeks postcoitus in our subjects. This assay has the potential of being a nondisease biomarker of unprotected sexual behavior in women.

References


