Detection of Y Chromosome DNA as Evidence of Semen in Cervicovaginal Secretions of Sexually Active Women

NICOLAS CHOMONT,1 GÉRARD GRESENGUET,2 MICHEL LÉVY,3 HAKIM HOCINI,1 PIERRE BECUUART,1 MATHIEU MATTA,4 JULIETTE TRANCHOT-DIALLO,1 LAURENT ANDREOLETTI,4 MARIE-PAULE CARRENO,1 MICHEL D. KAZATCHKINE,1 AND LAURENT BÉLEC1,4*

INSERM U430, Hôpital Broussais, 1 Département de Statistiques, Université René Descartes,2 and Laboratoire de Virologie, Hôpital Européen Georges Pompidou, Paris, France, and Centre National de Référence des Maladies Sexuellement Transmissibles et du SIDA, Bangui, Central African Republic.2

Received 26 February 2001/Returned for modification 1 May 2001/Accepted 29 May 2001

The detection of traces of semen in cervicovaginal secretions (CVS) from sexually active women practicing unprotected sex is a prerequisite for the accurate study of cervicovaginal immunity. Two semen markers, the prostatic-specific antigen (PSA) and the Y chromosome, were detected in parallel in CVS obtained by a standardized vaginal washing of consecutive women attending the principal medical center for sexually transmitted diseases of Bangui, Central African Republic. PSA was detected by immunoenzymatic capture assay in the cell-free fraction of CVS, and the Y chromosome was detected by a single PCR assay of DNA extracted by silica from the cell fraction (Y PCR). Fifty (19%) cell-free fractions of the 264 β-globin-positive CVS samples were positive for PSA, and 100 (38%) cell fractions of the CVS samples were positive for the Y chromosome. All the 50 (19%) PSA-containing CVS samples were also positive for the Y chromosome. Fifty (19%) CVS samples were positive only for the Y chromosome, with no detectable PSA. The remaining 164 (62%) CVS samples were both PSA and Y chromosome negative. These findings demonstrate that CVS from sexually active women may contain cell-associated semen residues unrecognized by conventional immunoenzymatic assays used to detect semen components. The detection of cell-associated male DNA with a highly sensitive and specific procedure such as Y PCR constitutes a method of choice to detect semen traces in female genital secretions.

Mucosal immunity of the female genital tract has recently gained special attention as an important factor that could modulate the transmission of many sexually transmitted infections (STIs), including human immunodeficiency virus (HIV) infection. Furthermore, current concepts of designing vaccines against viral infections acquired through sexual portals focus on the potential interest in inducing specific mucosal immunity at the sites of sexual exposure in association with systemic and cellular immune responses.

Mucosal immunity is investigated by collecting cervicovaginal secretions (CVS), either by vaginal washing (1) or by a vaginal or cervical swab further treated with collecting buffer (2). One potential methodological pitfall when sampling CVS of sexually active women is the presence of contaminating semen in the vaginal fluid that will bias the immunological characterization of the collected genital fluid. Female participants in clinical studies are generally asked to avoid sexual intercourse and intravaginal medications for 3 (10, 11) to 5 (4) days before sampling of CVS. However, semen residues may be detected in the lower female genital tract up to 5 days after sexual intercourse (12), and CVS collected from women at high risk for sexually transmitted diseases have frequently been found to contain traces of semen (13). Thus, ensuring that vaginal fluid is free of semen is essential to avoid misinterpretation of the data and accurately assess the immune response in the female genital tract. Similar precautionary measures should be undertaken when analyzing genital shedding of HIV in infected women. Finally, sensitive methods to detect traces of semen may be required in forensic medicine.

The presence of semen in CVS is assessed by microscopic observation of motile spermatozoids (3), determination of acid phosphatase activity in CVS (9), and the detection of semen components, including prostatic acid phosphatase, prostatic-specific antigen (PSA) (6), and seminal vesicle-specific antigen (5). The latter methods, including those based on the immunochemical detection of semen-derived molecules by immunocapture assays, may lack specificity and sensitivity. The present study was undertaken to assess the validity of using a highly sensitive PCR assay for the Y chromosome (designated Y PCR) in the cellular fraction of CVS for detecting contaminating semen in female genital fluids.

MATERIALS AND METHODS

Study population. Two hundred seventy-four unselected women attending the National Reference Center for Sexually Transmissible Diseases and AIDS in Bangui, Central African Republic, participated in the study. The Center offers multipurpose reproductive health services, including STI services, and operates as the main voluntary HIV testing and counselling site in Bangui. We followed the ethical recommendations of the Ministry of Health of the Central African Republic, and verbal informed consent was obtained from all participants.

Women entering the study underwent genital and pelvic examinations, during which CVS were collected as described below. A 7-day follow-up appointment was arranged for all women, and appropriate treatment was provided free of charge for any treatable STI syndrome or genital pathogen diagnosed.

Cervicovaginal sampling. CVS were collected by a standardized nontraumatic 60-s vaginal washing with 3.0 ml of phosphate-buffered saline, as previously

* Corresponding author. Mailing address: Laboratoire de Virologie, Hôpital Européen Georges Pompidou, 20 rue Leblanc, 75 908 Paris Cedex 15, France. Phone: 331 56 09 39 59. Fax: 331 56 09 24 47. E-mail: laurent.belec@cgp.ap-hop-paris.fr.
described (1). The cellular fraction and the cell-free fraction of CVS were separated by centrifugation at 1,000 × g for 10 min and kept frozen at −80°C until processing. The ratio of dilution of native CVS introduced by the washing procedure has previously been calculated as being 1:10 (1). Menstruating women and those with genital bleeding were excluded from the study.

**Detection of semen traces.** Detection of PSA and PCR amplification of DNA of the Y chromosome were performed in parallel in all collected CVS samples.

The detection and quantitation of PSA were performed with 150 μl of the acellular fraction of CVS by an immunoenzymatic assay with a threshold of positivity of 0.1 ng/ml (PSA IMX System; Abbott Laboratories, Chicago, IL). The cutoff for the presence of PSA antigen in cervicovaginal fluid was 0.4 ng/ml, determined as the mean plus 2 standard deviations of the values obtained with this assay with 150-μl samples of CVS obtained from 30 healthy childbearing-aged HIV-seronegative Caucasian women claiming to be non-sexually active at the time of sampling and recruited as controls. Positive and negative controls for enzyme-linked immunosorbent assay (ELISA) were those proposed by the manufacturer.

For PCR of the Y chromosome, DNA was extracted from the cellular pellet of CVS using the QIAamp DNA kit, according to the manufacturer’s recommendations (Qiagen AG, Basel, Switzerland). One microgram of extracted DNA was processed for Y chromosome DNA amplification by means of a single PCR assay with the primer set SRY3F (5′-CCG ATT CAT CGT GTG GTC TCG-3′) and SRY3R (5′-ATT CTT CGG CAG CAT CGT CTC-3′); specific for a 229-bp region in the sex-determining region (SRY), a gene located on the short arm of the Y chromosome (7). The PCR consisted of an initial denaturation at 94°C for 4 min, followed by 38 cycles of amplification (94°C, 60 s; 66°C, 60 s; and 72°C, 120 s) and a final elongation for 10 min at 72°C. The final PCR products were visualized under UV transillumination by means of ethidium bromide staining after electrophoresis with a 1.5% agarose gel. The positive control for Y PCR was a 1:1000 dilution (in distilled water) of DNA extracted from 100,000 peripheral blood mononuclear cells from a male donor; the negative control was undiluted DNA extracted from 100,000 peripheral blood mononuclear cells from a female donor. This Y PCR is able to detect five copies of Y DNA chromosome per μg of total extracted DNA (7). Furthermore, we checked that all CVS (cellular fraction) obtained from the 30 not-at-risk, HIV-seronegative control women claiming to be non-sexually active at time of sampling were Y DNA chromosome negative. To control the quality of extracted DNA and the lack of PCR inhibitors, the ubiquitous β-globin gene was amplified by PCR, as previously described (8).

**RESULTS**

**Population characteristics and sample processing.** Two hundred seventy-four women (mean age, 27 years; range, 15 to 48 years) were eligible for enrollment. None refused to participate in the study. The median age of first sexual intercourse was 16 years, with a median of two (range, one to eight) reported lifetime partners. Seventy-one women (26%) were found to be seropositive for HIV-1. DNA extracted from a cellular pellet of CVS tested positive by PCR for the β-globin gene in 264 samples (96%). The 10 cervicovaginal samples testing negative for the β-globin gene, suggesting poor conservation or a low amount of DNA in these samples, were excluded from the analysis.

**Detection of PSA and Y chromosome in CVS.** When tested for the presence of the PSA antigen, 50 of the 264 β-globin-positive CVS samples (19%) showed an optical density above the cutoff of positivity (Fig. 1). The mean concentration of PSA antigen ± standard deviation was 18.7 ± 19.9 ng/ml, with important differences among CVS samples. Thus, the concentrations of PSA ranged from 0.4 to 2 ng/ml in 10 samples, from 2.1 to 10 ng/ml in 17 samples, and from 10.1 to 50 ng/ml in 12 samples and were greater than 50.1 ng/ml in 11 samples (interquartile range, 2.8 to 39.3).

The cellular fractions of the 264 β-globin-positive CVS samples were tested further for the SRY gene. One hundred samples (38%) gave an amplicon as a unique and clearly distinguishable band of 129 bp and were considered positive for the Y chromosome. All PSA-containing CVS samples (n = 50; 19% of all CVS samples) were also positive for SRY DNA. Fifty CVS samples (19%) were positive only for the presence of the Y chromosome, with no detectable PSA. The number of semen-containing CVS samples detected by Y PCR (n = 100) was significantly higher than the number of semen-containing CVS samples detected by PSA assay (n = 50). The remaining 164 cervicovaginal samples (62%) were negative for both PSA and the Y chromosome.

**Choice of the best strategy.** Taken together, our experimental observations indicate that the detection of semen by means of an immunocapture assay for PSA is easy (automatic) and relatively inexpensive, but it lacks sensitivity. The detection of Y DNA in the cellular fraction of CVS by PCR is both very specific (no cross-amplification is possible in the female genital tract) and sensitive (high sensitivity of the gene amplification procedure used), but it is time consuming and relatively expensive.

A mathematical approach for the choice of the best strategy to be used is proposed in the Appendix. One can suppose that all PSA-containing CVS samples are TP_{PSA} (where TP_{PSA} is the number of true-positive samples), since all PSA-positive samples were confirmed to be Y PCR positive; then, 1 − α_{1} = 50/(50 + 10), and α_{1} = 0.50; one can also assume that β_{1} is very

![FIG. 1. Distribution of PSA concentrations among the 50 PSA-positive CVS samples. The box comprises the first to third quartiles; the horizontal line shows the median.](http://cvi.asm.org/)
FIG. 2. Estimated costs of strategy A (detection of PSA in all cervicovaginal samples and use of Y PCR for PSA-negative samples) (C_A) and strategy B (Y PCR in all samples), C_B, as a function of the expected prevalence of semen residue in the study population (P_e), according to equations A2 and A3, assuming that factor β₁ is very close to 0 (β₁ ≈ 0) and that the cost of Y PCR is q-fold that of PSA immunochemical detection (q > 1). When the P_e reaches the critical value P_c in the study population (~0.66, according to equation A8, where α₁ = 0.50 and q = 3), C_A and C_B are equivalent. Strategy A is preferable if P_e is >66% in the study population; below P_c, strategy B is less expensive and should be preferred.

close to 0 (β₁ ≈ 0), since all PSA-positive samples were defined according to a threshold of positivity used in the PSA assay established with cervicovaginal samples from a control group of women who were not sexually active at the time of sampling; q is approximately equal to 3, since an ELISA is approximately three times less expensive than a simple PCR (14). Thus, equation A8 (see Appendix) may be written as follows: P_e = 1/1.50 ≈ 0.66. Strategy A (i.e., the use of PSA for all cervicovaginal samples and Y PCR with PSA-negative samples) should be considered if the value of P_e reaches or exceeds 66%; below the critical value P_c, strategy B (Y PCR in all samples) appears less expensive and should be preferred (Fig. 2).

If one supposes that the initial screening ELISA (SELISA) used in strategy A has a sensitivity, Se, such that SeSELISA = 1 − α₃ and a specificity Sp, such that SpSELISA = 1 − β₃ that is very close to 1 (β₃ ≈ 0), P_e may be considered a function of α₃, as follows:

\[ P_e(\alpha_3) = 1/q (1 - \alpha_3) \]  \hspace{1cm} (1)

or

\[ P_e(SeSELISA) = 1/q SeSELISA \]  \hspace{1cm} (2)

Under these experimental conditions, P_e is inversely proportional to the sensitivity of the screening immunoenzymatic assay. In other terms, the higher the sensitivity of the ELISA (SeSELISA), the lower is the critical value P_c and the greater are the chances that strategy A is preferred for a population of sexually active women in which the P_e is high.

**DISCUSSION**

In the present study, we demonstrate that the PSA immunocapture assay, one of the most sensitive, specific, and commonly used immunoenzymatic assays available to detect semen in CVS collected from women practicing unprotected sexual intercourse, did not show 50 (50%) of 100 Y PCR-positive CVS. Our findings show that CVS from sexually active women may contain semen unrecognized by conventional immunoenzymatic assays used to detect semen components. The detection of semen components in the female genital secretions after penovaginal intercourse depends on the clearance of the semen components and on the sensitivity of methods used. Although the rate of clearance of semen-associated DNA deposited in the vagina is unknown, it is likely that the DNA protected in the nucleus of spermatozoids or male gamete precursors is relatively stable in the female lower genital tract. One may hypothesize that the rate of clearance of semen-associated DNA in the female genital tract is lower than that of soluble molecules such as PSA, in agreement with our observation that the Y chromosome was always amplified when PSA was detected in the CVS of sexually active women. The data suggest that the detection of male DNA by a highly sensitive and specific procedure such as Y PCR constitutes a method of choice to detect semen traces in female genital secretions.

Because PCR is time consuming and relatively expensive, the possibility of screening CVS for soluble semen components by means of an immunoenzymatic assay and reserving Y PCR for confirmation of the negative samples is interesting. In the study population of women living in Central Africa, the prevalence of semen retention in cervicovaginal fluid was as high as 38%. Using a mathematical approach taking into account the expected prevalence of semen-containing CVS in a given population, the strategy of screening CVS samples for PSA and to further confirming only the PSA-negative samples by Y PCR was estimated to be more expensive than systematically searching for Y DNA material in the cellular fraction of each genital secretion. In studies requiring genital secretion samples from
voluntary participants claiming to have avoided sexual intercourse prior to genital sampling, as well as in pathophysiological studies including women having unprotected sex, the prevalence of women with semen-containing CVS may be expected to be low or at least below the critical value of $P_c$. Under these conditions, our model shows that the best strategy consists of directly detecting the Y chromosome without prescreening by the detection of soluble semen components.

The PCR detection of Y DNA in the cellular fraction of CVS appears more sensitive and is likely to be more specific than the immunochemical detection of soluble semen components, such as PSA, in establishing the presence of semen. Obviously, the use of Y PCR might not be recommended as the best tool to assess recent sexual intercourse, for example, in forensic medicine. We conclude that the detection of semen in cervicovaginal fluid is an essential prerequisite to accurately assess HIV-specific cervicovaginal immunity or genital shedding of HIV in HIV-infected women, as well as in HIV-negative women exposed to the virus through sexual activity.

**APPENDIX**

Two strategies may be considered in assessing the presence of semen in CVS. In strategy A, all samples are tested for the presence of PSA, and then PSA-negative samples are tested by PCR. In strategy B, all CVS samples are screened for the SRY gene. We have developed a mathematical approach that takes into account the expected prevalence of semen residues in the study population ($P_s$) and the relative cost of each procedure used to detect semen residues ($C$).

The PSA immunoassay is characterized by its sensitivity ($Se_{PSA}$) and specificity ($Sp_{PSA}$), where $TP_{PSA}$ represents the number of true-positive samples in a population of $n$ women, $TN_{PSA}$ is the number of true-negative samples, $FP_{PSA}$ is the number of false-positive samples, and $FN_{PSA}$ is the number of false-negative samples, as follows:

$$S_{PSA} = \frac{TP_{PSA}}{TP_{PSA} + TN_{PSA} + FN_{PSA}} = 1 - \alpha_1$$

where $\alpha_1$ is a constant and $S_{PSA}$ corresponds to the percentage of semen-containing cervical samples that have been correctly detected by PSA detection. In the following equation

$$Sp_{PSA} = \frac{TN_{PSA}}{TN_{PSA} + FP_{PSA}} = 1 - \beta_1$$

$\beta_1$ is a constant and $Sp_{PSA}$ corresponds to the percentage of semen-free cervicovaginal samples that have been found to be truly negative for the presence of PSA.

Similarly, Y PCR is characterized by its sensitivity ($Se_Y$) and specificity ($Sp_Y$), as follows, where $TP_Y$ represents the number of true-positive samples in the tested population of women, $TN_Y$ is the number of true-negative samples, $FP_Y$ is the number of false-positive samples, and $FN_Y$ is the number of false-negative samples:

$$S_{PSA} = \frac{TP_Y}{TP_Y + TN_Y + FN_Y} = 1 - \alpha_2$$

and

$$Sp_{PSA} = \frac{TN_Y}{TN_Y + FP_Y} = 1 - \beta_2$$

where $\alpha_2$ and $\beta_2$ are constants. According to the high specificity and sensitivity of PCR procedure, one can assume that the specificity of the Y PCR is close to 1, giving $\alpha_2 = 0$, and that its sensitivity is also very close to 1, giving $\beta_2 = 0$.

The total cost of the strategy A, $C_A$, may be expressed as follows:

$$C_A = n(C_{PSA}) + \left[ n P_s \cdot \alpha_1 + n(1 - P_s) \cdot (1 - \beta_1) \right] C_Y$$

where $C_{PSA}$ represents the cost per sample of the PSA immunoassay assay, and $C_Y$ is the cost per sample of the Y PCR. If one considers that $C_Y$ is $q$-fold $C_{PSA}$ ($q > 1$), equation A1 becomes:

$$C_A = n(C_{PSA}) + \left[ n P_s \cdot \alpha_1 + n(1 - P_s) \cdot (1 - \beta_1) \right] q C_{PSA}$$

$$C_A = n(C_{PSA}) + \left[ n P_s \cdot \alpha_1 + n(1 - P_s) \cdot (1 - \beta_1) \right] q C_{PSA}$$ (A2)

$$C_A = n(C_{PSA}) + \left[ n P_s \cdot \alpha_1 + n(1 - P_s) \cdot (1 - \beta_1) \right] q C_{PSA}$$ (A3)

The total cost of strategy B, $C_B$, may be expressed as follows:

$$C_B = n(C_Y)$$

$$C_B = n(C_Y)$$ (A4)

$$C_B = n(C_Y)$$ (A5)

The costs of strategies A and B are equivalent when $P_s$ reaches the critical value, $P_c$, deduced from equation A6 and A7:

$$C_A = C_B$$

$$C_A = C_B$$ (A6)

$$1 + q + q(\alpha_1 + \beta_1) - P_c - q\beta_1 = q$$

$$P_c = 1 - q (1 - \alpha_1 - \beta_1)$$

(A7)

(A8)

**ACKNOWLEDGMENTS**

This study was supported by the Institut National de la Santé et de la Recherche Médicale, the Université Pierre et Marie-Curie (Paris VI), and the Agence Nationale de Recherches sur le SIDA. N.C. is recipient of a scholarship from the Ministère de l’Education Nationale, de la Recherche et de la Technologie, Paris, France.

**REFERENCES**


