Collection of Genomic DNA from Adults in Epidemiological Studies by Buccal Cytobrush and Mouthwash[1]

Montserrat García-Closas,[2] Kathleen M. Egan,
Jeannine Abruzzo, Polly A. Newcomb,
Linda Titus-Ernstoff, Tracie Franklin, Patrick K. Bender,
Jeanne C. Beck, Loïc Le Marchand, Annette Lum,
Michael Alavanja, Richard B. Hayes, Joni Rutter,
Kenneth Buetow, Louise A. Brinton, and
Nathaniel Rothman

Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland 20852 [M. G-C., M. A., R. B. H., J. R., K. B., L. A. B., N. R.]; Harvard School of Public Health, Boston, Massachusetts 02115 [K. M. E., J. A.]; Fred Hutchinson Cancer Research Center, Seattle, Washington 98104 [P. A. N.]; University of Wisconsin, Madison, Wisconsin 53706 [P. A. N.]; Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire 03756 [L. T-E.]; American Type Culture Collection, Cell Biology, Manassas, Virginia 20110 [T. F.]; Coriell Institute, Camden, New Jersey 08103 [P. K. B., J. B.]; and Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii 96813 [L. L. M., A. L.]

Abstract

Blood samples are an excellent source of large amounts of genomic DNA. However, alternative sources are often needed in epidemiological studies because of difficulties in obtaining blood samples. This report evaluates the buccal cytobrush and alcohol-containing mouthwash protocols for collecting DNA by mail. Several DNA extraction techniques are also evaluated. The study was conducted in two phases. In phase 1, we compared cytobrush and mouthwash samples collected by mail in two different epidemiological studies: (a) cytobrush samples ($n = 120$) from a United States case-control study of breast cancer; and (b) mouthwash samples ($n = 40$) from a prospective cohort of male United States farmers. Findings from phase 1 were confirmed in phase 2, where we randomized cytobrush ($n = 28$) and mouthwash ($n = 25$) samples among participants in the breast cancer study to directly compare both collection methods. The median human DNA yield determined by hybridization with a human DNA probe from phenol-chloroform extracts was 1.0 and 1.6 μg/2 brushes for phases 1 and 2, respectively, and 27.5 and 16.6 μg/mouthwash sample for phases 1 and 2, respectively. Most (94–100%) mouthwash extracts contained high molecular weight DNA (>23 kb), in contrast to 55–61% of the brush extracts. PCR success rates for amplification of β-globin gene fragments (268, 536, and 989 bp) were similar for cytobrush and mouthwash phenol-chloroform extracts (range, 94.4–100%). Also, we obtained high success rates in determining the number of CAG repeats in the androgen receptor gene, characterizing tetranucleotide microsatellites in six gene loci, and screening for mutations in the BRCA1/2 genes in a subset of phenol-chloroform DNA extracts. Relative to DNA extracted by phenol-chloroform from cytobrush samples, DNA extracted by NaOH had lower molecular weight, decreased PCR success rates for most assays performed, and unreliably high spectrophotometer readings for DNA yields. In conclusion, although DNA isolated from either mouthwash or cytobrush samples collected by mail from adults is adequate for a wide range of PCR-based assays, a single mouthwash sample provides substantially larger amounts and higher molecular weight DNA than two cytobrush samples.

Introduction

Obtaining high quality genomic DNA is critical for epidemiological studies that aim to evaluate the role of genetic factors in human disease susceptibility. Blood samples are an excellent source of large amounts of genomic DNA. However, epidemiological studies often need alternative sources when study subjects are reluctant to provide a blood sample, when only a self-administered collection protocol is logistically or economically feasible, or as a back-up source of DNA in studies that collect blood samples. Exfoliated buccal epithelial cells and other cells found in saliva are a very promising alternative source of DNA because they can be obtained using self-administered, noninvasive, and relatively inexpensive techniques (1–12). Buccal swabs and mouthwash protocols are the most commonly used protocols for buccal cell collection. Studies using different types of buccal swabs, i.e., cotton swabs or cytobrushes, have obtained similar DNA yields and PCR success rates (1–5). The first studies using mouthwash to collect DNA for PCR-based assays used saline rinses that were processed or frozen immediately after collection (6–8). Hayney et al. (8) evaluated the stability of saline mouthwash samples stored for 7 days at temperatures to which samples are likely to be exposed if collected by mail. This study indicated that samples stored at 25°C and 37°C tended to have higher amounts of high molecular weight DNA than samples stored at lower temperatures (−20°C and 4°C), suggesting the presence of DNA of bacterial origin. Similarly, a study conducted by Walsh et al. (5) suggested the presence of DNA of predominantly bacterial origin on cotton swabs stored for 4 days at 37°C. Lum and Le Marchand (9) proposed the use of an alcohol-containing mouthwash of a familiar brand that would be more appropriate for self-collection of samples by mail in epidemiological studies because the alcohol content is likely to reduce bacterial growth during mailing. Indeed, these authors found that storage
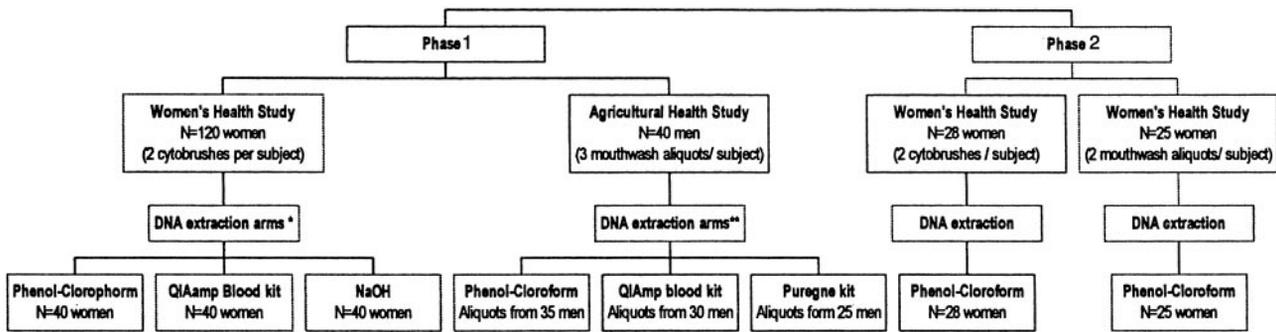
Fig. 1.   Subjects included in different study arms from phase 1 and 2 of the study. Each DNA extraction arm in phase 1 includes different women from the WHS (*), whereas each arm in phase 2 includes one of three aliquots of equal volume made from single mouthwash samples collected from 40 men in the AHS (**).

of alcohol-containing mouthwash at room temperature or at 37°C for 7 days did not affect the DNA yields or the ability to PCR amplify the samples when compared with samples stored at −20°C. This mouthwash protocol was successfully used in a community-based cohort study in Hawaii that included subjects of Japanese, Caucasian, and Hawaiian origin, with participation rates ranging from 59–76% (12).

The above-mentioned studies (1–12) have shown that DNA isolated from buccal cells under different protocols can be successfully used to perform PCR-based assays. In addition, two studies reported complete agreement in results from PCR-based assays when DNA was isolated from either buccal or blood cells from the same individual (13, 14). However, important questions remain with regard to the yield and quality of human DNA that can be obtained from different buccal cell collection protocols and different DNA extraction methods.

This report compares two protocols for self-collection of buccal cells by mail, cytobrush and alcohol-containing mouthwash. The study was conducted in two phases (Fig. 1). In phase 1 of the study, we had the opportunity to compare mailed cytobrush samples from cases participating in the WHS,[3] a United States case-control study of breast cancer, with mailed mouthwash samples from male United States farmers in the prospective AHS (15). In phase 2 of the study, we randomized cytobrush and mouthwash samples among cases participating in the WHS to confirm findings from phase 1 when both collection methods are directly compared in the same population. In addition to the comparison of collection protocols, this report compares phenol-chloroform DNA extraction (9) to the following alternative methods: (a) QIAamp DNA Blood Mini Kit (Qiagen Inc., Chatsworth, CA); (b) Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN); and (c) NaOH extraction (2).

The two collection methods and the DNA extraction methods are compared in terms of (a) total and human DNA yield, (b) DNA integrity, and (c) PCR success in amplifying β-actin gene fragments. This report also includes results from assays performed in a subset of study samples that evaluate the feasibility of using buccal DNA to perform other PCR-based assays.

Materials and Methods

Study Population

As indicated in the introduction, the current study was conducted in two phases (Fig. 1). All subjects participating in the study provided signed informed consents approved by institutional review boards.

Phase 1. Cytobrush samples in phase 1 of the study were collected by mail from 120 women with breast cancer who had granted a telephone interview in the WHS between June and July 1998. Women had a median age of 55 years (range, 29–74 years); most were Caucasian (99.2%) and resided in Massachusetts or New Hampshire. Mouthwash samples were collected by mail from 40 men participating in a pilot study within the AHS between August and October 1998. This pilot study was aimed at evaluating different strategies for collecting mailed mouthwash samples after a telephone interview. Participants were males working in the farming industry, with a median age of 49 years (range, 29–77 years), of Caucasian origin, and residing in Iowa or North Carolina.

Phase 2. In September 1999, we randomly mailed 28 cytobrush and 25 mouthwash collection kits to women with breast cancer who had been interviewed for the WHS. The median age for women who provided cytobrush samples was 53 years (range, 32–74 years), and the median age for women who provided a mouthwash sample was 57 years (range, 38–71 years). All cytobrush samples and all but one mouthwash sample came from Caucasian women residing in Massachusetts or New Hampshire.

Methods of Sample Collection

Phase 1. Each woman from the WHS participating in phase 1 was mailed a sample collection kit containing two sterile cytobrushes in sealed plastic tubes (Cyto-Pack Cytosoft Brush; Medical Packing Corp., Camarillo, CA), instructions for collection, and a prepaid return envelope. Women were asked to brush their teeth and 10–15 min later brush the inside of the right and left cheeks for at least 30 s each using one cytobrush per cheek, place the cytobrushes back into the plastic tubes, and mail to the laboratory. On arrival at the laboratory, the brush was separated from the handle, placed in a cryovial, and frozen at −70°/−80°C until DNA extraction in May 1999 (i.e., 10–11 months of storage before extraction). From June through July 1998, we received a total of 192 buccal cell kits with a median shipping time of 3 days (range, 0–19 days). To minimize the effects of shipping time on our assays, we included a random

sample of specimens that had been in the mail 4 days or less ($n = 120$). The median shipping time for the 120 included kits was 2 days (range, 1–4 days).

Each man from the AHS participating in the buccal cell pilot study was mailed a sample collection kit containing a sealed trial size bottle of Scope mouthwash with a 14.3 wt% alcohol content (Procter & Gamble, Cincinnati, OH), a collection cup with a 10-ml fill line, instructions for collection, and a prepaid return envelope. Participants were asked not to eat or drink the hour before the sample collection, fill the cup with mouthwash to the fill line, swish the mouthwash throughout the mouth for 45 s, expectorate into the cup, and mail the sample to the laboratory. This protocol was based on the protocol proposed by Lum and Le Marchand (9) with the following modifications. First, we did not instruct subjects to brush their teeth before sample collection because we had no evidence that this would improve DNA quality, and teeth brushing could result in a decrease in DNA yields by rinsing out exfoliated buccal cells. Second, we used Scope (14.3 wt% alcohol) rather than FreshBurst Listerine (21.6% alcohol) mouthwash because of reports of a burning sensation after swishing Listerine for 60 s and no evidence of differences in DNA yields and PCR success rates between these two brands (12). Comparability in DNA yields and PCR success rates (*NAT2*, *GSTM1,* and *β-globin* genes) for Scope and Listerine mouthwash samples stored under different conditions (frozen immediately and held for 3 days at room temperature or at 37°C before freezing) was confirmed in a small pilot study conducted by our group.[4] Third, we asked subjects to swish the mouthwash for 45 s rather than the 60 s specified in the original protocol because data from a small pilot study indicated that DNA yields increase with increasing swishing time, with a plateau at about 30–45 s.[5] Samples arrived at the laboratory within a few days after collection. On arrival at the laboratory, each mouthwash sample was transferred to a 15-ml conical tube for centrifugation at $1500 \times g$ for 15 min. The supernatant was decanted, and the cell pellet was resuspended in 3 ml of TE buffer solution [Tris-EDTA ($100\times$ concentration; pH 8.0) in DEPC water; 1:100 solution], aliquoted in three cryovials of equal volume, and stored at $-70°/-80°$C until DNA extraction in May 1999 (*i.e.*, after 7–9 months of storage). That was modified from the original protocol, in which entire mouthwash samples were frozen at $-20°$C to minimize storage space.

Phase 2. Cytobrush samples in phase 2 of the study were collected following the same protocol used in phase 1. For the mouthwash collection, women from the WHS were asked to brush their teeth 10–15 min before the collection of the sample to make the protocol strictly comparable with the cytobrush protocol. Then women were asked to swish 10 ml of Scope mouthwash for approximately 30 s. The median shipping time for cytobrush and mouthwash samples was 2 days (range, 2–3 days) and 3 days (range, 2–5 days), respectively. On arrival at the laboratory, mouthwash samples were processed as described for phase 1, except that two aliquots were made, rather than three. In October 1999, DNA was extracted from 28 cytobrush and 25 mouthwash samples received from breast cancer cases in September 1999 (*i.e.*, after approximately 1 month of storage).

DNA Extraction Methods

Phase 1. To compare different DNA extraction methods, cytobrush and mouthwash samples in phase 1 of the study were randomly assigned to different extraction arms (Fig. 1). Cytobrush samples from the WHS were extracted by three different methods (see "Appendix" for detailed protocols): (*a*) phenol-chloroform extraction (9); (*b*) QIAamp DNA Blood Mini Kit (Qiagen Inc.); and (*c*) NaOH extraction used previously on cytobrush samples (2). Each extraction arm included cytobrush samples from 40 women. DNA was extracted from one of the two brushes provided by 35 of the 40 women and from both cytobrushes provided by the other 5 women. The two cytobrushes from the same women were used to assess differences in DNA yields from right and left cheeks.

Mouthwash aliquots from each of the 40 men in the AHS were assigned to three extraction arms (see "Appendix" for detailed protocols): (*a*) phenol-chloroform extraction (9); (*b*) QIAamp DNA Blood Mini Kit (Qiagen Inc.); and (*c*) Puregene DNA Isolation Kit for mouthwash samples (Gentra Systems). The main modification in the phenol-chloroform extraction method from that of Lum and Le Marchand (9) was to extend the DNA precipitation time from 2 h to overnight because this resulted in improved DNA yields.[6] The phenol-chloroform arm included samples from 35 subjects (one aliquot from 30 subjects and all three aliquots from 5 subjects). The 5 subjects with all three aliquots included in this arm were used to assess differences in DNA yields among aliquots. The QIAamp arm included samples from 30 subjects (one aliquot from 25 subjects and all three aliquots from 5 subjects). The 5 subjects with all three aliquots included in this arm were used to evaluate the effect of modifications to the manufacturer's extraction protocol on DNA yields. Finally, the Puregene arm included samples from 20 subjects (one aliquot per subject). Thus, of the 30 subjects with one aliquot included in the phenol-chloroform arm, 25 had a second aliquot in the QIAamp arm (*i.e.*, 25 paired samples), and 20 had a third aliquot in the Puregene arm (*i.e.*, 20 paired samples).

DNA extracts from five randomly selected subjects in each extraction arm were split into two aliquots to assess the reproducibility of our assays.

Phase 2. DNA from cytobrush and mouthwash samples in phase 2 of the study was extracted using the phenol-chloroform protocols described in the "Appendix."

DNA Quantification Assays

Total DNA concentration was determined by spectrophotometry at 260 and 280 nm using a Beckman DU-640 Spectrophotometer (Beckman Scientific Instruments, Fullerton, CA). Human DNA concentration was determined by hybridization with human DNA probe D17Z1 using the ACES 2.0+ DNA Quantification System (Life Technologies, Inc., Grand Island, NY). One to two serial dilutions of samples were prepared with Tris-low EDTA, and the DNA concentration was quantified by comparison with the human DNA standards using image analysis. The DNA yield per cytobrush or mouthwash aliquot was calculated by multiplying the DNA concentration by the final volume of DNA extract.

When DNA was extracted from all specimens collected per subject (two cytobrushes or two or three mouthwash aliquots), DNA yield per subject was estimated by adding the

---

yield from each sample. We found no significant differences in DNA yields from each cytobrush or mouthwash aliquot collected from the same individual. Therefore, when DNA was extracted from only one specimen, total DNA yield per subject was estimated by multiplying the DNA yield in the measured specimen by the total number of specimens. DNA yields for subjects that had the DNA extracts split into two aliquots were estimated as the average of the DNA yield for the two split samples.

The ratio of human DNA concentration measured by hybridization:total DNA concentration measured by spectrophotometry was used as an estimate of the percentage of human DNA in a given sample.

DNA Integrity

The integrity of genomic DNA was assessed by resolving DNA extracts on a 0.4% agarose gel by electrophoresis, followed by visualization with ethidium bromide staining. Each DNA sample was graded according to the electrophoretic migration of sample DNA in comparison to a known molecular weight marker (Ready-Load λ DNA/*Hin*dIII Fragments; Life Technologies, Inc.). Grade 1 was assigned when the gel showed no visible DNA, grade 2 was assigned when DNA was ≤564 bp, grade 3 was assigned when DNA was <6,557 bp, grade 4 was assigned when DNA was <23,130 bp (diffuse head between approximately 10,000 and 23,000 bp with a smeared tail), and grade 5 was assigned when DNA was ≥23,130 bp. When two or more samples from the same subject differed in DNA grade, the average grade was assigned.

PCR-based Assays

The adequacy of the buccal DNA extracts for PCR-based assays was assessed by amplifying three different-sized fragments of the β-globin gene (268, 536, and 989 bp) as described by Greer *et al.* (Ref. 16; see "Appendix"). Amplicons were separated in either 4% agarose gels for 268-bp fragments or 3% agarose gels for 536- and 989-bp fragments by electrophoresis, followed by visualization with ethidium bromide staining. When no PCR product was visible on the agarose gel, the PCR product was rerun in a modified polyacrylamide matrix (GeneAmp Solution; Perkin-Elmer, Foster City, CA), followed by visualization with silver staining. A second amplification attempt was made for samples with no visible PCR product on either the agarose or the polyacrylamide gel. Subjects that had more than one sample included in the study were classified as having a successful amplification when at least one sample was amplified.

A subset of 10 buccal DNA samples from each of the DNA extraction arms in phase 1 of the study (except for the mouthwash Puregene arm; *n* = 50) was sent to Coriell Institute (Camden, NJ) to evaluate the feasibility of using buccal DNA to perform different types of genetic analyses, including analysis of the number of CAG repeats in the *AR* gene and tetranucleotide microsatellite analysis in six gene loci. Six WBC DNA samples were included as quality control. Here we present results from the CAG repeats in the *AR* and microsatellite analyses. PCR for amplification of the *AR* gene fragment encompassing the CAG repeat was performed using a seminested PCR procedure modified from those reported by Hakimi *et al.* (17) and Irvine *et al.* (Ref. 18; see "Appendix"). The size of the amplified fragments ranged from 178–220 bp. Genotyping by microsatellites was based on the method of Edwards *et al.* (Ref. 19; see "Appendix"). The tetranucleotide microsatel-
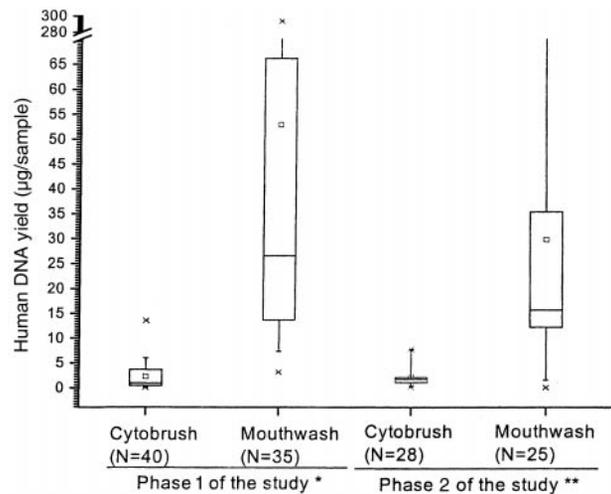

Fig. 2. Distribution of human DNA yield for two cytobrush samples or a 10-ml mouthwash sample extracted by phenol-chloroform in phase 1 and 2 of the study. Phase 1 includes cytobrush samples collected from women participating in the WHS, and mouthwash samples were collected from men participating in the AHS in 1998 (∗). Phase 2 includes cytobrush and mouthwash samples collected from women participating in the WHS in 1999 (∗∗). Box range, 25 and 75 percentiles; whisker range, 5 and 95 percentiles; □, mean.

lite assay was considered to be successful if alleles could be scored for five of the six loci.

Four randomly selected mouthwash DNA samples extracted by phenol-chloroform in phase 1 of the study were screened for mutations in portions of the *BRCA1* and *BRCA2* genes using dHPLC (20). Specifically, 10 ng of human DNA were used as template to amplify 8 amplicons of the 84 total amplicons required for complete analysis of the *BRCA1/BRCA2* genes. Sequence differences in the 8 amplicons were confirmed by DNA sequencing.

Statistical Analysis

Differences in median DNA yields between methods of collection or DNA extraction were tested with Wilcoxon's rank-sum test for independent observations or the Wilcoxon matched-pair signed-rank test for paired observations. Differences between collection or extraction methods in the percentage of samples with high molecular weight DNA or successful PCR reaction were tested with Fisher's exact test for independent observations or McNemar's exact test for paired observations.

Quality Control Procedures

Each batch of DNA samples (*n* = 33 batches) included one aliquot from a single solution of WBC DNA extracted previously by phenol-chloroform for quality control. The mean ± SD DNA concentration for the WBC DNA samples was 81.0 ± 7.8 μg/μl (CV = 9.6%) for spectrophotometry readings and 80.1 ± 15.2 μg/μl (CV = 19.0%) for hybridization readings. The mean ± SD percentage of human DNA (estimated by the ratio of hybridization:spectrophotometry readings) was 100 ± 24% (CV = 24.0%). All WBC DNA aliquots had a DNA grade from the agarose gels of 4, and the three different β-globin fragments were successfully amplified on all aliquots.

The reproducibility of the assays for DNA quantification and DNA quality were assessed using data from DNA extracts that were split into two aliquots (five split samples per DNA

Table 1 Total and human DNA yields (μg/two cytobrushes or μg/one 10-ml mouthwash sample) from phase 1 of the study by method of collection and DNA extraction

Method of collection DNA extraction	N	Spectrophotometry				Hybridization	
		Total DNA yield		$A_{260\,nm}$:$A_{280\,nm}$		Human DNA yield	
		Median	(range)	Median	(range)	Median	(range)
Cytobrushes[a]							
Phenol-chloroform	40	13.6[b,c]	(2.2–43.5)	1.9	(1.6–2.1)	1.0[d,e,f]	(0.006–13.5)
QIAamp kit	40	8.8[c]	(1.5–28.6)	1.7	(1.4–1.9)	0.4[e]	(0.006–19.0)
NaOH	40	126.0	(7.6–396.7)	1.3	(1.2–2.2)	1.3[f]	(0.006–20.5)
Mouthwash[g]							
Phenol-chloroform	35	57.3[b,h,i]	(10.8–286.9)	1.9	(1.7–2.0)	27.5[d,i,k]	(3.1–293.5)
QIAamp kit	30	35.2[h]	(10.5–90.6)	1.8	(1.7–2.1)	10.6[j]	(1.5–79.9)
Puregene kit	20	52.5[i]	(10.6–199.2)	1.9	(1.7–1.9)	28.5[k]	(6.2–243.3)

[a] Each DNA extraction arm includes 40 cytobrush samples from a total of 120 women participating in the WHS (*i.e.*, independent observations in each arm).

[b] $P < 0.001$, cytobrush phenol-chloroform *versus* mouthwash phenol-chloroform, Wilcoxon's rank-sum test.

[c] $P = 0.008$, cytobrush phenol-chloroform *versus* cytobrush QIAamp kit, Wilcoxon's rank-sum test.

[d] $P < 0.001$, cytobrush phenol-chloroform *versus* mouthwash phenol-chloroform, Wilcoxon's rank-sum test.

[e] $P = 0.005$, cytobrush phenol-chloroform *versus* cytobrush QIAamp kit, Wilcoxon's rank-sum test.

[f] $P = 0.97$, cytobrush phenol-chloroform *versus* cytobrush NaOH, Wilcoxon's rank-sum test.

[g] Each DNA extraction arm includes one of three aliquots of equal volume made from single mouthwash samples collected from a total of 40 men participating in the AHS (*i.e.*, paired observations in each arm).

[h] $P < 0.001$, mouthwash phenol-chloroform *versus* mouthwash QIAamp kit ($N = 25$ pairs), Wilcoxon's sign-rank tests for matched pairs.

[i] $P = 0.22$, mouthwash phenol-chloroform *versus* mouthwash Puregene kit ($N = 20$ pairs), Wilcoxon's sign-rank tests for matched pairs.

[j] $P < 0.001$, mouthwash phenol-chloroform *versus* mouthwash QIAamp kit ($N = 25$ pairs), Wilcoxon's sign-rank tests for matched pairs.

[k] $P = 0.83$, mouthwash phenol-chloroform *versus* mouthwash Puregene kit ($N = 20$ pairs), Wilcoxon's sign-rank tests for matched pairs.

Table 2 Total and human DNA yields (μg/two cytobrushes or μg/one 10-ml mouthwash sample) from phase 2[a] of the study by method of collection

Method of collection	N	Spectrophotometry		Hybridization
		Total DNA yield	$A_{260\,nm}$:$A_{280\,nm}$	Human DNA yield
		Median (range)	Median (range)	Median (range)
Cytobrushes	28	13.5[b] (2.0–38.2)	1.8 (1.6–2.0)	1.6[c] (0.11–7.7)
Mouthwash	25	38.7[b] (11.9–124.2)	1.8 (1.7–1.9)	16.6[c] (1.55–140.3)

[a] Samples were collected from women participating in the WHS and extracted by phenol-chloroform.

[b] $P < 0.001$, cytobrush *versus* mouthwash, Wilcoxon's rank-sum test.

[c] $P < 0.001$, cytobrush *versus* mouthwash, Wilcoxon's rank-sum test.

extraction arm in phase 1 of the study). We found no significant differences in DNA yields from split samples for any of the DNA extraction arms (data not shown). Three of 15 (20%) cytobrush pairs of DNA extracts had different DNA grades (one pair of NaOH extracts had grades of 3 and 4, and two pairs of phenol-chloroform extracts had grades of 3 and 4 and 3 and 2). PCR success for the three β-globin gene fragments agreed in all cytobrush DNA pairs. DNA grades from all mouthwash DNA pairs were in complete agreement, and all samples were successfully amplified for the three β-globin gene fragments.

Results

Phenol-chloroform DNA Extracts

Total and Human DNA Yields. The distribution of human DNA yields for two cytobrush samples and a single mouthwash sample extracted by phenol-chloroform is shown in Fig. 2. This figure shows that in both phases of the study, human DNA yields from cytobrush samples were substantially lower than yields from mouthwash samples. In phase 1 of the study, median (mean $\pm$ SD) total and human DNA yields for cytobrushes were 13.6 μg/two cytobrushes (15.8 $\pm$ 9.8 μg/two cytobrushes) and 1.0 μg/two cytobrushes (2.4 $\pm$ 3.0 μg/two cytobrushes), respectively, whereas they were 57.3 μg/mouth-

wash sample (88.4 $\pm$ 73.4 μg/mouthwash sample) and 27.5 μg/mouthwash sample (56.7 $\pm$ 64.2 μg/mouthwash sample), respectively (Table 1). In phase 2 of the study, median (mean $\pm$ SD) total and human DNA yields were 13.5 μg/two cytobrushes (16.0 $\pm$ 8.7 μg/two cytobrushes) and 1.6 μg/two cytobrushes (2.2 $\pm$ 0.9 μg/two cytobrushes), respectively, and 38.7 μg/mouthwash sample (46.1 $\pm$ 28.4 μg/mouthwash sample) and 16.6 μg/mouthwash sample (31.0 $\pm$ 33.8 μg/mouthwash sample), respectively (Table 2). There were no significant differences in DNA yields from cytobrush samples collected from women in the WHS during phases 1 and 2 of the study, despite differences in storage time before DNA extraction (9–11 months in phase 1 and ~1 month in phase 2). Similarly, we found no significant differences in DNA yields from mouthwash samples collected in phases 1 and 2, although mouthwash samples from the two phases came from different study populations (*i.e.*, WHS and AHS) of differing gender and age distribution.

The percentage of human DNA present in cytobrush and mouthwash samples was estimated after combining data from samples in phase 1 and 2 that were extracted by phenol-chloroform. The median percentage of human DNA in cytobrush samples [$n = 68$ (40 samples from phase 1 and 28 samples from phase 2)] was 11.5%, and in mouthwash samples [$n = 60$ (35 samples from phase 1 and 25 samples from phase 2)] the median percentage of human DNA was 49.5%. For comparison, the median percentage found in the WBC DNA aliquots included as quality control in the study ($n = 33$) was 94.4%. Given that the WBC DNA aliquots contained only human DNA, in expectation, the mean value of the percentage of human DNA should be 100%. This is what we observed in our data (mean $\pm$ SD = 100 $\pm$ 24%). The fact the median was 94.4% rather than 100% could indicate random variation or that the distribution is not centered around 100%.

DNA Integrity. In both phase 1 and 2 of the study, the percentage of phenol-chloroform extracts with high molecular weight DNA (grade 5 or >23 kb) was significantly lower in cytobrush samples (55.0% in phase 1 and 60.7% in phase 2)

Table 3 Percentage of samples with high molecular weight DNA (>23 kb) and PCR success rates to amplify fragments of the *β-globin* gene from phase 1 of the study by method of collection and DNA extraction

Method of collection DNA extraction	N	% samples with high molecular weight DNA[a]	PCR success rate		
			268 bp	536 bp	989 bp[b]
Cytobrushes[c]					
Phenol-chloroform	40	55%[d,e,f]	100%[g]	97.5%[h]	94.4%[i]
QIAamp kit	40	2.5%[d,j]	100%	92.5%	89.5%
NaOH	40	0.0%[e]	40.0%[g]	20.0%[h]	15.0%[i]
Mouthwash[k]					
Phenol-chloroform	35	94.3%[f,l,m]	100%	100%	100%
QIAamp kit	30	13.3%[j,l]	100%	100%	100%
Puregene kit	20	100%[m]	100%	100%	100%

[a] High molecular weight DNA was defined as grade 5 on the agarose gel (*i.e.*, DNA fragments >23 kb).
[b] Respectively, amplification of the 989-bp fragment was not attempted in 4, 21, and 26 cytobrush samples extracted by phenol-chloroform, QIAamp kit, and NaOH due to low sample concentration.
[c] Each DNA extraction arm includes 40 cytobrush samples from a total of 120 women participating in the WHS (*i.e.*, independent observations in each arm).
[d] P < 0.001, cytobrush phenol-chloroform *versus* cytobrush QIAamp kit, Fisher's exact test.
[e] P < 0.001, cytobrush phenol-chloroform *versus* cytobrush NaOH, Fisher's exact test.
[f] P < 0.001, cytobrush phenol-chloroform *versus* mouthwash phenol-chloroform, Fisher's exact test.
[g] P < 0.001 cytobrush phenol-chloroform *versus* cytobrush NaOH, Fisher's exact test.
[h] P < 0.001 cytobrush phenol-chloroform *versus* cytobrush NaOH, Fisher's exact test.
[i] P < 0.001 cytobrush phenol-chloroform *versus* cytobrush NaOH, Fisher's exact test.
[j] P = 0.16, cytobrush QIAamp *versus* mouthwash QIAamp kit, Fisher's exact test.
[k] Each DNA extraction arm includes one of three aliquots of equal volume made from single mouthwash samples collected from a total of 40 men participating in the AHS (*i.e.*, paired observations in each arm).
[l] P < 0.001, mouthwash phenol-chloroform *versus* mouthwash QIAamp kit (N = 25 pairs), McNemar's exact test for matched pairs.
[m] Mouthwash phenol-chloroform *versus* mouthwash Puregene: all pairs had high molecular weight DNA (N = 20 pairs).

Table 4 Percentage of samples with high molecular weight DNA and PCR success rates to amplify fragments of the *β-globin* gene from phase 2[a] of the study by method of collection

Method of collection	N	% samples with high molecular weight DNA[b]	PCR success rate		
			268 bp	536 bp	989 bp
Cytobrushes	28	60.7%[c]	100%	100%	96.4%
Mouthwash	25	100.0%[c]	100%	100%	96.0%

[a] Samples were collected from women participating in the WHS and extracted by phenol-chloroform.
[b] High molecular weight DNA was defined as grade 5 on the agarose gel (*i.e.*, DNA fragments >23 kb).
[c] P < 0.001, cytobrush *versus* mouthwash Fisher's exact test.

than in mouthwash samples (94% in phase 1 and 100% in phase 2; Tables 3 and 4).

PCR Amplification of *β-Globin* Gene Fragments. Most cytobrush and all mouthwash phenol-chloroform extracts in phase 1 (Table 3) and all cytobrush and mouthwash samples in phase 2 (Table 4) were successfully amplified for the three fragments of *β*-globin. However, in both phases of the study, a larger percentage of cytobrush samples than mouthwash DNA samples required a second PCR reaction to obtain a successful amplification, especially for longer PCR fragments. For instance, in phase 1 of the study, 11.8% of cytobrush DNA samples and 2.9% of mouthwash DNA samples (P = 0.20) with a successful amplification of the 989-bp fragment required a second attempt. Similarly, in phase 2 of the study, 22.2% of cytobrush DNA samples and 4.2% of mouthwash DNA samples (P = 0.10) with a successful amplification of the 989-bp fragment required a second attempt.

Other DNA Extraction Techniques
In this section, several DNA extraction techniques are compared with phenol-chloroform extraction in terms of DNA

yields, DNA integrity, and PCR amplification of *β*-globin gene fragments, using samples collected in phase 1 of the study. It should be noted that the three DNA extraction arms for cytobrush samples include different women participating in the WHS (*i.e.*, independent observations), whereas each of the three extraction arms for mouthwash samples includes different aliquots from the same men participating in the AHS (*i.e.*, paired observations).

Median DNA yields obtained from the QIAamp blood DNA extraction kit were less than half of the phenol-chloroform yields for both cytobrush and mouthwash samples (Table 1). In addition, the percentage of cytobrush or mouthwash QIAamp DNA extracts with high molecular weight DNA (grade 5 or >23 kb) was also lower than that for phenol-chloroform extracts (Table 3). However, the success rate for PCR amplification of *β*-globin fragments (Table 3) and the percentage of samples that required a second attempt for a successful amplification were similar for both extraction methods.

Total DNA yields estimated for the NaOH extracted cytobrushes using a spectrophotometer were unreliably high (as compared with the other extraction methods), with a low $A_{260\ nm}:A_{280\ nm}$ ratio (median, 1.3; range, 1.2–2.2) indicating the presence of substantial contamination of the DNA extract (Table 1). Whereas the human DNA yield was similar to that of the phenol-chloroform extracts, the quality of the DNA was substantially lower, as indicated by the absence of high molecular weight DNA (>23 kb) in NaOH extracts (DNA grade ranged from 1–4) and significantly lower success rates for PCR amplification of *β*-globin fragments (Table 3). In addition, a higher percentage of NaOH extracts than phenol-chloroform extracts required two attempts for a successful PCR reaction, especially for longer PCR fragments.

Total and human DNA yields as well as DNA grade, PCR success rates, and number of attempts for a successful amplification were very similar for mouthwash samples extracted by Puregene or by phenol-chloroform (Tables 1 and 3).

Other PCR-based Assays Performed in a Subset of Samples

Subsets of DNA extracts from phase 1 of the study were sent to other laboratories to evaluate the feasibility of using buccal DNA to perform other PCR-based assays.

(*a*) Ten buccal DNA samples from each extraction arm in phase 1 of the study (except for the mouthwash Puregene arm; *n* = 50) and six WBC DNA aliquots were sent to Coriell Institute to determine the number of CAG repeats in the *AR* gene and for tetranucleotide microsatellite analysis. The number of CAG repeats in the AR gene was determined successfully in all types of DNA extracts. The success rates for the tetranucleotide microsatellite assay in six loci (*TH-01, D10S526, D5S592, D22S417, FES/FPS,* and *vWA31*) were 85.7%, 100%, and 62.5% for the phenol-chloroform, QIAamp, and NaOH cytobrush extraction arms, respectively. The success rate for the phenol-chloroform and QIAamp mouthwash extraction arms was 100%. This assay was successful in all but one (five of six) of the WBC DNA aliquots included as quality control. None of the differences in success rates across extraction arms were statistically significant.

(*b*) Four randomly selected phenol-chloroform mouthwash DNA samples were screened successfully for mutations in portions of the *BRCA1* and *BRCA2* genes using dHPLC (20). Using 10 ng of human DNA as a template, 8 amplicons of the 84 total amplicons required for complete analysis gave strong signals on the dHPLC assays. Both polymorphic and wild-type samples were identified by dHPLC, and they were subsequently confirmed by sequence analysis. Therefore, using current technology, complete *BRCA1/BRCA2* analyses could be completed on as little as 1 μg of human DNA from mouthwash samples.

Discussion

We have shown that a single self-collected 10-ml Scope mouthwash sample (14.3 wt% alcohol) sent by mail at room temperature provides substantially larger amounts of human genomic DNA than two self-collected cytobrush samples sent by mail at room temperature without a transport medium. Although most (>93%) of the mouthwash phenol-chloroform DNA extracts and only 60% of the cytobrush phenol-chloroform DNA extracts contained high molecular weight DNA (>23 kb), all extracts by either method were adequate for a wide range of PCR-based assays.

Cytobrush and Mouthwash DNA Extracted by Phenol-chloroform

DNA Yield. Previous studies have reported totals rather than human DNA yields. Reported total DNA yields from buccal swabs vary from an average of 1.3–3.8 μg/cotton swab (3, 4) and an average of 2.7 μg/cytobrush (1). The total DNA yields by phenol-chloroform extraction found in our study (respectively, a median and mean of ~13 and 16 μg/two cytobrushes) are higher than previously reported yields. Because our cytobrushes were processed after 1–4 days in the mail, as opposed to immediately after collection in previous studies, the differences in total DNA yields could reflect a larger amount of nonhuman DNA present in our samples due to bacterial growth during shipping.

Total mouthwash DNA yields from our study [median (mean ± SD) of 57.3 μg/mouthwash sample (88.4 ± 73.4 μg/mouthwash sample) and 38.7 μg/mouthwash sample (46.1 ± 28.4 μg/mouthwash sample) in phase 1 and 2, respectively] are similar to total yields reported by Lum and Le Marchand (9) from 10-ml rinses with an alcohol-containing mouthwash (FreshBurst Listerine) for 60 s followed by DNA extraction within 1 week (mean ± SD of 49.7 ± 31.7 μg/mouthwash sample). Harty *et al.* (11) reported somewhat lower total yields in samples of 10-ml sterile water rinses for 10 s, which were mixed with a transport medium immediately after collection and stored at −70°C for 3–36 months before DNA extraction [median (range) of 25.9 μg (2.0–204.5 μg)]. Finally, Tobal *et al.* (7) reported a range of 10–240 μg/10-ml normal saline rinse for 30 s and immediate DNA extraction by phenol-chloroform. Mouthwash DNA yields obtained per subject can be substantially increased by asking subjects to provide more than one sample at a time without reducing the participation rates in the study (12). Other modifications to the collection protocol such as brushing the inside of the cheeks before mouthwash collection may also impact the DNA yield.

Our study indicated that cytobrush and mouthwash samples contain a mixture of DNA of human and nonhuman origin and that only a median of about 11% and 49% of the total DNA found in cytobrush and mouthwash samples, respectively, is of human origin. Thus, to obtain accurate measures of the amount of human DNA present in these types of samples, it is important to use human-specific techniques such as the hybridization method used in our study. A limitation of this method is that is relatively expensive and labor intensive. Thus, we are currently evaluating other techniques to quantify human DNA concentration that would be better suited for DNA quantification of large numbers of samples.

DNA Integrity. Most mouthwash phenol-chloroform extracts (>94%) contained high molecular weight DNA (>23 kb), as compared with only 55–60% of the cytobrush DNA extracts. This suggests that more DNA degradation occurs in cytobrush samples than in mouthwash samples collected by mail at room temperature. The DNA fragment size in mouthwash phenol-chloroform extracts from our study was consistent with that found by investigators at Coriell Institute in similarly collected mouthwash DNA samples that were analyzed by pulse-field gel electrophoresis to assess the integrity of DNA.[7] Specifically, Coriell Institute assessed the integrity of phenol-chloroform DNA extracts from buccal cells collected in Scope and stored for 3 days at room temperature in the dark. The majority of the recovered DNA was between 35 and 63 kb in length (range, >15–97 kb). For comparison, the bulk of the DNA prepared from tissue cultured cells by a similar DNA extraction method had sizes ranging from 45 kb to >200 kb. There was also evidence for DNA damage detected by the Comet assay (21) in the Scope DNA samples.[7] However, this did not appear to affect the ability to carry out the PCR-based assays presented here.

PCR Amplification of β-*Globin* Gene Fragments and Other PCR-based Assays. Phenol-chloroform-extracted DNA from both mouthwash and cytobrush samples was successfully used to amplify β-*globin* gene fragments of 268, 536, and 989 bp in length; *AR* CAG repeats; and tetranucleotide microsatellites. In addition, phenol-chloroform-extracted DNA from mouthwash samples was used successfully to screen for mutations in portions of the *BRCA1* and *BRCA2* genes using dHPLC and to sequence PCR products to confirm differences in sequences. These results, together with previous reports that used similar DNA extracts to determine genetic polymorphisms in a wide range of genes (3, 4, 6, 8, 9, 11, 22), support the appropriateness

----

[7] Jeanne C. Beck and Patrick K. Bender, Coriell Institute, unpublished data.

of using phenol-chloroform DNA extracts from mouthwash or buccal swabs to perform a wide range of PCR-based assays.

Other DNA Extraction Techniques

Total and human DNA yields from cytobrush or mouthwash samples extracted by the QIAamp blood kit (Qiagen Inc.) were lower than the yields obtained by a phenol-chloroform extraction method modified from the method of Lum and Le Marchand (9). The lower DNA yields obtained from this column-based extraction method may be explained by difficulties in choosing a column size adequate for the wide range of number of cells recovered using the cytobrush or mouthwash protocols. QIAamp DNA extracts were of lower molecular weight than phenol-chloroform extracts. This is supported by experiments at Coriell Institute that showed that DNA prepared by binding to silica-based matrices such as Qiagen DNA extraction kits tended to be of smaller size than phenol-chloroform extracts.[7]

The NaOH-extracted DNA was of low molecular weight and presented difficulties for PCR amplification of *β-globin* gene fragments. In contrast, Walker *et al.* (1) reported high success rates (range, 75.5–79.6%) for similar PCR-based genotype assays on NaOH DNA extracts from mailed cytobrushes. Richards *et al.* (2) also reported a high success rate (99%) for a multiplex PCR amplification assay from the *CFTR* gene using NaOH-extracted DNA from mailed cytobrushes. We made no special efforts to optimize the PCR conditions for NaOH DNA extracts because PCR reactions were performed blindly for different types of DNA extracts. Thus, differences in efforts to optimize PCR conditions could explain the lower PCR success rates found in our study, as compared with previous reports.

Conclusion

In conclusion, our data indicate that the mouthwash protocol originally proposed by Lum and LeMarchand (9) should have wide application for epidemiological studies in adult populations that require collection of genomic DNA. Additional research is necessary to evaluate modifications to the collection protocols that could further improve DNA yields and/or quality, to evaluate techniques for quantification of human DNA concentration, and to evaluate the suitability of buccal DNA extracts for new genetic technology (*e.g.,* whole genome amplification and whole genome scans). As more experience is gained in the collection of this type of sample, information will accumulate on human DNA yields obtained from populations of different ethnic origin and participation rates accomplished in different study populations and under different study designs (12).

Appendix

DNA Extraction Protocols

Phenol-chloroform DNA Extraction from Cytobrushes [Modified from the Method of Lum and Le Marchand (9)]. Place the cytobrush in 700 μl of lysis buffer, vortex, and incubate for 10 min at room temperature. Add 20 μl of RNase A (10 mg/ml), mix the suspension, and digest at 37°C for 30 min. Add 35 μl of proteinase K, mix the suspension, and digest at 58°C for 2 h. Remove the cytobrush, add an equal volume of phenol-chloroform (1:1), vortex for 10 s, centrifuge at 15,800 $\times$ *g* for 2 min, and transfer the aqueous layer to a new tube. Add an equal volume of chloroform, vortex for 10 s, centrifuge at 15,800 $\times$ *g* for 2 min, and transfer the aqueous layer to a new tube. Add 0.1$\times$ volume of 3 M NaOAc (pH 6.0) and 2$\times$ volume of 100% ethanol and store overnight at -80°C.

Pellet the DNA at 15,800 $\times$ *g* for 10 min, wash the DNA with 70% ethanol, and dry for 15 min. Resuspend the DNA pellet in 50 μl of TE. Place at 4°C for short-term storage.

QIAamp DNA Blood Mini Kit (Qiagen Inc.) for Cyto-brushes (Modified Manufacturer's Instructions). Place the brush in a 2-ml centrifuge tube, add 600 μl of PBS to the sample, and incubate for 5 min at room temperature. Add 20 μl of RNase A (10 mg/ml), incubate for 30 min at 37°C, and add 20 μl of proteinase K (20 mg/ml) and 600 μl of Buffer AL. Vortex for 15 s, incubate at 56°C for 2 h, and remove the cytobrush. Add 20 μl of proteinase K and incubate overnight at 56°C. Add 600 μl of 100% ethanol, vortex, and apply 700 μl of the mixture to QIAamp spin column in a 2-ml collection tube. Centrifuge the collection tube at 6000 $\times$ *g* for 1 min and add the remaining mixture to the column in a clean 2-ml collection tube. Centrifuge the collection tube at 6000 $\times$ *g* for 1 min and add 500 μl of Buffer AW1 to QIAamp spin column in a clean 2-ml collection tube. Centrifuge the collection tube at full speed for 3 min and add 500 μl of Buffer AW2 to QIAamp spin column in a clean 2-ml collection tube. Centrifuge the collection tube at full speed for 3 min. Centrifuge the QIAamp spin column in a clean 2-ml collection tube at full speed for 1 min. Elute DNA with 150 μl of Buffer AE into a clean 1.5-ml microfuge tube and incubate at room temperature for 1 min. Centrifuge at 6000 $\times$ *g* for 1 min and place at 4°C for short-term storage.

NaOH Extraction from Cytobrushes (2). Immerse the brush in 600 μl of 50 mM NaOH contained in a polypropylene tube and vortex. Heat the tube with the brush at 95°C for 5 min and remove the brush from the tube. Neutralize the DNA solution with 60 μl of 1 M Tris (pH 8.0), vortex, and store at 4°C.

QIAamp DNA Blood Mini Kit (Qiagen Inc.) from Mouth-wash Cell Pellets (Modified Manufacturer's Instructions). Three variations of the QIAamp protocol were evaluated using all three buccal cell aliquots collected from five subjects. The first aliquot was extracted using the QIAamp DNA Blood Mini Kit according to the manufacturer's instructions, the second aliquot was extracted using the QIAamp DNA Blood Mini Kit with a double volume of lysis buffer (360 μl instead of 180 μl) added to the cell pellet before incubation with RNase, and the third aliquot was extracted using the QIAamp DNA Blood Midi Kit according to manufacturer's instructions. The QIAamp DNA Blood Mini Kit with a double volume of lysis buffer yielded larger amounts of DNA (median, 15.4 μg human DNA/aliquot; range, 2.3–18.2 μg human DNA/aliquot) than either the QIAamp DNA Blood Mini Kit (median, 4.4 μg human DNA/aliquot; range, 1.3–9.8 μg human DNA/aliquot; Sign Test $P = 0.06$) or the QIAamp DNA Blood Midi Kit (median, 3.7 μg human DNA/aliquot; range, 0.5–5.6 μg human DNA/aliquot; Sign Test $P = 0.06$). Therefore, this was the protocol used in the rest of the buccal cell aliquots in this extraction arm. Details of the protocol are described below.

Thaw the sample, add 1.0 ml of TE to the tube, vortex, and centrifuge for 15 min at 1500 $\times$ *g*. Remove the supernatant, add 360 μl of Buffer ATL to the pellet and 20 μl of proteinase K (20 mg/ml), and incubate at 55°C for 2 h. Add 20 μl of proteinase K (20 mg/ml) and incubate at 55°C overnight. Add 20 μl of RNase A (10 mg/ml), vortex for 15 s, and incubate for 30 min at 37°C. Add 400 μl of Buffer AL, mix, and incubate at 70°C for 10 min. Centrifuge to remove drops from the lid, add 400 μl of ethanol, and vortex. Apply the mixture to QIAamp spin column in a 2-ml collection tube and centrifuge at 6000 $\times$ *g* for 1 min. Repeat for the remaining solution. Add 500 μl of Buffer AW1 to QIAamp spin column in a clean 2-ml

collection tube and centrifuge at 6000 × *g* for 1 min. Add 500 μl of Buffer AW2 to QIAamp spin column in a clean 2-ml collection tube and centrifuge at full speed for 3 min. Centrifuge QIAamp spin column in a clean 2-ml collection tube at full speed for 1 min. Elute DNA with 150 μl of Buffer AE into a clean 1.5-ml microfuge tube and incubate at room temperature for 1 min. Centrifuge at 6000 × *g* for 1 min and repeat the elution process with the same 150 μl of Buffer AE. Place at 4°C for short-term storage.

Phenol-chloroform DNA Extraction from Mouthwash Cell Pellets [Modified from the Method of Lum and Le Marchand (9)]. Thaw sample, add 1.0 ml of TE to the tube, vortex, and centrifuge for 15 min at 1,500 × *g*. Remove the supernatant, resuspend cell pellets in 700 μl of lysis buffer, add 20 μl of RNase A (10 mg/ml), mix the suspension, and digest at 37°C for 30 min. Add 35 μl of proteinase K (20 mg/ml), mix the suspension, and digest at 58°C for 2 h. Add an equal volume of phenol-chloroform (1:1), vortex for 10 s, and centrifuge at 15,800 × *g* for 2 min. Transfer the aqueous layer to a new tube, add an equal volume of chloroform, vortex for 10 s, and centrifuge at 15,800 × *g* for 2 min. Transfer the aqueous layer to a new tube, add an equal volume of chloroform, vortex for 10 s, and centrifuge at 15,800 × *g* for 2 min. Transfer the aqueous layer to a new tube, and add 0.1× volume of 3 M NaOAc (pH 6.0) and 2× volume of 100% ethanol. Store overnight at −20°C. Pellet the DNA at 15,800 × *g* for 10 min, wash the DNA with 70% ethanol, dry, and resuspend the DNA pellet in 100 μl of TE buffer. Place at 4°C for short-term storage.

Puregene DNA Isolation Kit (Gentra Systems) for Mouthwash Cell Pellets (Manufacturer's Instructions). Thaw samples on ice and transfer the sample to a 15-ml conical tube. Rinse the original tube with 1 ml of TE, transfer TE to the 15-ml conical tube with the sample. Centrifuge at 2000 × *g* for 10 min to concentrate the cells. Pour off the supernatant, leaving behind 100 μl of residual liquid. Vortex vigorously to resuspend the cells in the residual supernatant. Add 3 ml of Cell Lysis Solution to the resuspended cells and vortex for 5 s at medium speed. Add 20 μl of proteinase K (20 mg/ml) and incubate the lysate at 55°C for 1 h. Add 20 μl of proteinase K (20 mg/ml) and incubate the lysate at 55°C overnight. Add 20 μl of RNase A Solution (4 mg/ml) to the cell lysate. Mix the sample by inverting the tube 25 times and incubate at 37°C for 15 min. Cool the sample to room temperature. Add 1 ml of Protein Precipitation Solution to the lysate. Vortex samples at high speed for 20 s to mix the Protein Precipitation Solution uniformly with the lysate. Place the sample into an ice bath for 10 min. Centrifuge at 2000 × *g* for 10 min. The precipitated proteins should form a tight, green pellet. Pour the supernatant containing the DNA (leaving behind the precipitated protein pellet) into a clean 50-ml tube containing 3 ml of 100% Isopropanol containing 5 μl of Glycogen Solution (20 mg/ml). Mix the sample by inverting gently 50 times and keep the tube at room temperature for at least 5 min. Centrifuge at 2000 × *g* for 10 min. The DNA may or may not be visible as a small white pellet, depending on the yield. Pour off the supernatant and drain tube briefly on clean absorbent paper. Add 3 ml of 70% ethanol and invert the tube several times to wash the DNA pellet. Centrifuge at 2000 × *g* for 3 min. Carefully pour off the ethanol. Invert and drain the tube on clean absorbent paper and allow to air dry 10–15 min. Add 100 μl of DNA Hydration Solution. Allow DNA to rehydrate overnight at 37°C. Tap the tube periodically to aid in dispersing the DNA. Vortex briefly and pulse spin before use. Place at 4°C for short-term storage.

PCR-based Assays

Amplification of β-*Globin* Gene Fragments (16). In separate PCR reactions, primers GH20 (5′-GAAGAGCCAAGGACAG-GTAC-3′) and PC04 (5′-CAACTTCATCCACGTTCACC-3′) were used to generate a 268-bp fragment, primers RS42 (5′-GCTCACTCAGTGTGGCAAAG-3′) and KM29 (5′-GGTTG-GCCAATCTACTCCCAGG-3′) were used to generate a 536-bp fragment, and primers RS80 (5′-TGGTAGCTGGAT-TGTAGCTG-3′) and RS40 (5′-ATTTTCCCACCCTTAGG-CTG-3′) were used to generate a 989-bp fragment. The PCR reaction contained 10 ng of DNA sample, 200 μM each primer, 200 μM deoxynucleotide triphosphate mix, 2.5 units of Ampli-Taq Gold DNA polymerase (Perkin-Elmer), 1× PCR buffer, and 4 mM of 25 mM MgCl. Samples underwent 39 amplification cycles of 95°C for 1 min, 55°C for 1 min, and 72°C for 2 min followed by a final extension of 5 min at 72°C.

Amplification of the *AR* Gene Fragment Encompassing the CAG Repeat [Modified from Hakimi *et al.* (17) and Irvine *et al.* (18)]. The forward primer sequence for the first PCR was CGTGCGCGAAGTGATCCAG, and the reverse primer sequence was GCTGTGAAGGTTGCTGTTCCTCAT. For the seminested PCR, the forward primer sequence was the same as that in the first PCR, but the primer was synthesized with a 6-carboxy-fluorescein label on its 5′ end. This label permits the detection of the fragment during size analysis on the model 377 Fluorescent Sequencer (PE Biosystems, Foster City, CA). The seminested reverse primer sequence was AGAACCATCCT-CACCCTGCTG. The buffer for both the first and seminested PCR contained 20 mM Tris (pH 8.3), 2.5 mM MgCl₂, 50 mM KCl, 0.2 mM each deoxynucleotide triphosphate, 2.5 units of AmpliTaq DNA polymerase (PE Biosystems), and 0.16 μM each primer. In the first PCR, 300 ng of template DNA were used in a 25-μl reaction, whereas in the seminested PCR, 2.5 μl of the first PCR products were used in a 25-μl reaction. Conditions for both PCRs are an initial soak at 95°C for 5 min followed by 28 cycles of 95°C for 2 min, 61°C for 1.5 min, and 72°C for 1.5 min, and a final soak at 72°C for 5 min before storage at 4°C. The size of the PCR-generated fragment(s) was determined by fragment size analysis on an ABI model 377 fluorescent sequencer using GeneScan-400HD (TAMARA) molecular weight standards (PE Biosystems). The number of CAG repeats was calculated from the size of the fragment using the sequence for the *AR* gene reported in GenBank (accession number NM 000044) as a reference. The size of the amplified fragments ranged from 178–220 bp.

Genotyping for Microsatellites [Modified from Edwards *et al.* (19)]. Six microsatellites were selected for genotyping. Three of these are commonly used in forensics. These are *TH-01* (The Genome Database, accession number 212652), *vWA31* (GenBank accession number M258581; Ref. 23), and *FES/FPS* (GenBank Accession number AC003004; Ref. 24). The other three were selected because of their high heterozygosity. These are *D5S592* (The Genome Database, accession number 198533), *D10S526* (The Genome Database, accession number 198813), and *D22S417* (The Genome Database, accession number 198522). Amplification of each microsatellite marker was performed by PCR using the primers reported in the Genome Database or the cited references. One primer in each primer pair was labeled with either 6-carboxy-fluorescein or 4,7,2′,4′,7′-hexachloro-6-carboxylfluorescein fluorescent reporters for detection on the ABI fluorescent sequencer (PE Biosystems). Conditions for PCR were for the AmpliTaq Gold DNA polymerase using the manufacturer's supplied buffer (PE Biosystems), 2.5 mM MgCl₂, and primer concentrations of 0.1

μM for *D10S526*, 0.6 μM for *D5S592*, 0.7 μM for *D22S417*, and 0.2 μM for the other primers. Scoring the alleles was performed with the aid of the Genscan and Genotyper software (PE Biosystems) after electrophoresis on an ABI model 377 fluorescent sequencer. The tetranucleotide microsatellite assay was considered to be successful if alleles could be scored for five of the six loci.

References

1. Walker, A. H., Najarian, D., Rebbeck, T. R., Jaffe, J. M., Kanetsky, P. A., and Rebbeck, T. R. Collection of genomic DNA by buccal swabs for polymerase chain reaction-based biomarker assays. Environ. Health Perspect., *107:* 517–520, 1999.

2. Richards, B., Skoletsky, J., Shuber, A. P., Balfour, R., Stern, R., Dorkin, H. L., Parad, R. B., Witt, D., and Klinger, K. W. Multiplex PCR amplification from CFTR gene using DNA prepared from buccal brushes/swabs. Hum. Mol. Genet., *2:* 159–163, 1993.

3. Freeman, B., Powell, J., Ball, D., Hill, L., Craig, I., and Plomin, R. DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. Behav. Genet., *27:* 251–257, 1997.

4. 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. Am. J. Hum. Genet., *57:* 1252–1254, 1995.

5. Walsh, D. J., Corey, A. C., Cotton, R. W., Forman, L., Herrin, G. L., Jr., Word, C. J., and Garner, D. D. Isolation of deoxyribonucleic acid (DNA) from saliva and forensic science samples containing saliva. J. Forensic Sci., *37:* 387–395, 1992.

6. Lench, N., Stanier, P., and Williamson, R. Simple non-invasive method to obtain DNA for gene analysis. Lancet, *1:* 1356–1358, 1988.

7. Tobal, K., Layton, D. M., and Mufti, G. J. Non-invasive isolation of constitutional DNA for genetic analysis. Lancet, *2:* 1281–1282, 1989.

8. Hayney, M. S., Dimanlig, P., Lipsky, J. J., and Poland, G. A. Utility of a "swish and spit" technique for the collection of buccal cells for TAP haplotype determination. Mayo Clin. Proc., *70:* 951–954, 1995.

9. Lum, A., and Le Marchand, L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. Cancer Epidemiol. Biomark. Prev, *7:* 719–724, 1998.

10. Harty, L. C., Garcia-Closas, M., Rothman, N., Reid, Y. A., Tucker, M. A., and Hartge, P. Collection of buccal cell DNA using treated cards. Cancer Epidemiol. Biomark. Prev, *9:* 501–506, 2000.

11. Harty, L. C., Shields, P. G., Winn, D. M., Caporaso, N. E., and Hayes, R. B. Self-collection of oral epithelial cell DNA under instruction from epidemiologic interviewers. Am. J. Epidemiol., *151:* 199–205, 2000.

12. Le Marchand, L., Lum-Jones, A., Saltzman, B., Visaya, V., Nomura, A. M. Y., and Kolonel, L. N. Feasibility of collecting buccal DNA by mail in a cohort study. Cancer Epidemiol. Biomark. Prev., *10:* 701–703, 2001.

13. de Vries, H. G., Collee, J. M., van Veldhuizen, M. H., Kate, L. P., Achterhof, L., Smit Sibinga, C. T., Scheffer, H., and Buys, C. H. Validation of the determination of ΔF508 mutations of the cystic fibrosis gene in over 11,000 mouthwashes. Hum. Genet., *97:* 334–336, 1996.

14. Hagerman, R. J., Wilson, P., Staley, L. W., Lang, K. A., Fan, T., Uhlhorn, C., Jewel-Smart, S., Hull, C., Drisko, J., Flom, K., and Taylor, A. K. Evaluation of school children at high risk for fragile X syndrome utilizing buccal cell FMR-1 testing. Am. J. Med. Genet., *51:* 474–481, 1994.

15. Alavanja, M. C., Sandler, D. P., McMaster, S. B., Zahm, S. H., McDonnell, C. J., Lynch, C. F., Pennybacker, M., Rothman, N., Dosemeci, M., Bond, A., and Blair, A. The Agricultural Health Study. Environ. Health Perspect., *104:* 362–369, 1996.

16. Greer, C. E., Wheeler, C. M., and Manos, M. M. Sample preparation and PCR amplification from paraffin-embedded tissues. PCR Methods Applications, *3:* S113–S122, 1994.

17. Hakimi, J. M., Schoenberg, M. P., Rondinelli, R. H., Piantadosi, S., and Barrack, E. R. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. Clin. Cancer Res., *3:* 1599–1608, 1997.

18. Irvine, R. A., Yu, M. C., Ross, R. K., and Coetzee, G. A. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. Cancer Res., *55:* 1937–1940, 1995.

19. Edwards, A., Civitello, A., Hammond, H. A., and Caskey, C. T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am. J. Hum. Genet., *49:* 746–756, 1991.

20. Wagner, Y., Steoppa-Lyonnet, D., Fleischmann, E., Muhr, D., Pages, S., Sandberg, T., Caux, V., Moeslinger, R., Langbauer, G., Borg, A., and Oefner, P. Denaturing high-performance liquid chromatography detects reliably BRCA1 and BRCA2 mutations. Genomics, *62:* 369–376, 1999.

21. Hellman, B., Vaghef, H., and Bostrom, B. The concepts of tail moment and tail inertia in the single cell gel electrophoresis assay. Mutat. Res., *336:* 123–131, 1995.

22. Hayney, M. S., Poland, G. A., and Lipsky, J. J. A noninvasive "swish and spit" method for collecting nucleated cells for HLA typing by PCR in population studies. Hum. Hered., *46:* 108–111, 1996.

23. Kimpton, C., Walton, A., and Gill, P. A further tetranucleotide repeat polymorphism in the vWF gene. Hum. Mol. Genet., *1:* 287, 1992.

24. Polymeropoulos, M. H., Rath, D. S., Xiao, H., and Merril, C. R. Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene (FES). Nucleic Acids Res., *19:* 4018, 1991.