

Quality and Quantity of Saliva DNA Obtained from the Self-administrated Oragene Method—A Pilot Study on the Cohort of Swedish Men

Tove Rylander-Rudqvist,¹ Niclas Håkansson,¹ Gunnel Tybring,² and Alicja Wolk¹

¹Division of Nutritional Epidemiology, Institute of Environmental Medicine and ²KI Biobank, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Abstract

Self-collection of saliva has the potential to provide molecular epidemiologic studies with DNA in a user-friendly way. We evaluated the new Oragene saliva collection method and requested saliva samples by mail from 611 men (ages 53-87 years). We obtained a response rate of, on average, 80% [varying from 89% (ages 67-71 years) to 71% (ages 77-87 years)]. DNA was extracted from 90 randomly selected samples, and its usefulness was evaluated with respect to quality, quantity, and whole-genome amplification (WGA). Visual inspection of DNA on agarose gels showed high molecular weight DNA (>23 kb) and no degradation. Total DNA yield measured with PicoGreen ranged from 1.2 to 169.7 µg, with a mean of 40.3 µg (SD, 36.5 µg) and a median of 29.4 µg. Human DNA yield was estimated by real-time PCR of the human prothrombin gene to account for 68% (SD, 20%)

of total DNA. We did WGA on 81 saliva DNA samples by using the GenomiPhi DNA kit and genotyped both saliva DNA and WGA DNA for 10 single-nucleotide polymorphisms randomly selected from the human genome. Overall genotyping success rate was 96% for saliva DNA and 95% for WGA DNA; 79% of saliva DNA samples and 79% of WGA DNA samples were successfully genotyped for all 10 single-nucleotide polymorphisms. For the 10 specific assays, the success rates ranged between 88% and 100%. Almost complete genotypic concordance (99.7%) was observed between saliva DNA and WGA DNA. In conclusion, Oragene saliva DNA in this study collected from men is of high quality and can be used as an alternative to blood DNA in molecular epidemiologic studies. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1742-5)

Introduction

For the establishment of DNA biobanks from large cohorts, it is of great importance to collect from as many subjects as possible to get a representative sample and to increase the power in association studies on genetic variants and disease. DNA quality is essential for the success rates in genotype analyses, and the amount of collected genomic DNA is of importance for future studies. A peripheral blood sample is the preferred source of DNA with respect to both quality and quantity. However, the need for the donors to go to a primary health care facility to have the sample drawn as well as the invasive character of this method may reduce response rate. A simple, self-administrated sample collection protocol may increase participation rates especially among elderly participants. Saliva and buccal cells are alternative DNA sources that can be self-collected through rinses, brushes, or swabs (1-12). Disadvantages with these methods are that the donors need to either swish and spit an alcohol-containing mouthwash solution (1, 3-5, 8-10), which have been reported to give a burning sensation after swishing (3), or follow sometimes complicated instructions on how to collect buccal cells with brushes or swabs (1, 3, 12). Furthermore, mouthwash solu-

tions with relatively high alcohol content (e.g., Scope) are not available in all countries. A new method for saliva self-collection, Oragene, was recently introduced on the market in which the donors simply just spit into a vial. When the vial is capped, a solution containing antibacterial and DNA preserving chemicals mixes with the saliva, resulting in immediate conservation of the sample. In a pilot study encompassing 611 men ages 53 to 87 years, we evaluated response rate, DNA quality and quantity, and whole-genome amplification (WGA) after using this novel collection method.

Materials and Methods

Participants. From the population-based prospective cohort of Swedish men consisting of 48,645 subjects, born 1918 to 1952 and living in Västmanland and Örebro counties (13), 625 men, equally distributed in five age groups, were randomly selected. After exclusion of men who died or moved from the study area between baseline and September 2004, 611 men were eligible for the study.

Sample Collection. All men received a preliminary notification, with short information about the study and that a saliva collection kit will be arriving within a few days. Three days later, they received the Oragene kit (DNA Genotek, Inc., Ottawa, Ontario, Canada), a longer information letter, detailed instructions on how to deliver the saliva sample, an informed consent form for signature, and a prepaid return envelope. Nonresponders received two reminders by mail; the first after 2 weeks and the second after a further 2 weeks. Participants were requested to spit ~2 mL saliva. Four hundred ninety sent a saliva sample and a signed informed consent, 43 sent an answer saying that they did not want to participate, and 78 did

Received 9/8/05; revised 6/7/06; accepted 7/6/06.

Grant support: Swedish Research Council (longitudinal studies), Karolinska Institutet, and the Swedish National Biobank programme within Wallenberg Consortium North. The DNA extraction facility in Malmö, Sweden is financially supported by SWEGENE and the Wallenberg Consortium North.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Tove Rylander-Rudqvist, Institute of Environmental Medicine, Karolinska Institutet, Nobels väg 13, Box 210, SE-171 77 Stockholm, Sweden. Phone: 46-8-5248-72-56; Fax: 46-8-30-45-71. E-mail: tove.rylander-rudqvist@ki.se

Copyright © 2006 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-05-0706

not respond. Saliva samples from 90 respondents were randomly chosen for DNA extraction. Extraction was done 1 to 2 months after saliva collection and storage in room temperature.

DNA Quantity and Quality. DNA was robotically extracted by the Autopure LS system using the Puregene DNA purification kit (Gentra Systems, Minneapolis, MN), and the yield and A260/A280 ratio were determined with UV at 260 and 280 nm. Quantitation of DNA yield was also done at the DNA extraction facility in Malmö, Sweden by using PicoGreen and a FLUOstar Optima device (BMG Labtech, Offenburg, Germany), and 100 μ L of PicoGreen reagent (diluted 1:200 in Tris-EDTA buffer) were dispensed into 100 μ L of sample (5 μ L DNA plus 95 μ L Tris-EDTA buffer). Samples were incubated in darkness for 5 minutes before fluorescence reading at the excitation and emission wavelengths of 485 nm and 520 nm, respectively. DNA (5 μ L) was visualized on agarose gels together with 125 ng λ DNA.

To get an estimate of the relative fraction of human and bacterial DNA, real-time PCR of the human prothrombin gene with 5' exonuclease (Taqman) probes was done on an ABI Prism 7900HT system (Applied Biosystems, Foster City, CA). Primer and probe sequences are available on request. The PCR of 20 μ L contained 20 to 200 ng DNA, 1.25 mmol/L MgCl₂, 0.1 μ L Custom Assay mix, including primers and probe, and 5 μ L Taqman Universal Master Mix (Applied Biosystems). Denaturation was done at 50°C for 2 minutes and 95°C for 10 minutes followed by 50 cycles each consisting of 95°C for 15 seconds and 60°C for 1 minute. The proportion of human DNA was calculated by dividing received number of prothrombin PCR copies with estimated number of total DNA copies added in the PCR. Number of total DNA copies was calculated from the DNA concentration quantified with PicoGreen by assuming that 1 ng of genomic DNA contains ~270 copies. This figure is based on the approximation that one cell (equivalent to two DNA copies) contains ~7 pg DNA.

The saliva DNA samples were genotyped for 10 single-nucleotide polymorphisms (SNP; Table 1) randomly selected from the human genome by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Sequenom, Inc., San Diego, CA; ref. 14). PCR assays and associated extension reactions were designed using the SpectroDESIGNER software (Sequenom). Primer sequences are available on request (Metabion GmbH, Planegg-Martinsried, Germany). All amplification reactions were run in the same conditions in a total volume of 5 μ L with 2.5 ng genomic DNA, 1 pmol of each amplification primer, 0.2 mmol/L deoxynucleotide triphosphate, 2.5 mmol/L MgCl₂, and 0.2 unit HotStarTaq DNA polymerase (Qiagen, Inc., Valencia, CA). Reactions were heated at 95°C for 15 minutes and subjected to 45 cycles of amplification (20 seconds at 94°C, 30 seconds at 60°C, 30 seconds at 72°C) before a final extension of 7 minutes at 72°C. Extension reactions were conducted in a total volume of 9 μ L using 5 pmol of allele-specific extension primer and the MassEXTEND Reagent kits before being cleaned using SpectroCLEANER (Sequenom) on a Multimek 96 automated 96-channel robot (Beckman Coulter, Fullerton, CA). Clean primer extension products were loaded onto a 384-element chip with a nanoliter pipetting system (Sequenom) and analyzed by a MassARRAY mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). The resulting mass spectra were analyzed for peak identification using the SpectroTYPER RT 2.0 software (Sequenom). For each SNP, two independent scorers confirmed all genotypes. Hardy-Weinberg calculations were done to ensure that each marker was within allelic population equilibrium (>0.01; ref. 15) in our sample set.

WGA (16) of saliva DNA samples was done by using the GenomiPhi DNA kit (Amersham Biosciences, Uppsala,

Sweden). After measurement of DNA yield, gel electrophoreses, and SNP analysis of the saliva DNA samples, there was enough material left for WGA on 81 of 89 samples. One to 5 ng of template DNA were used per reaction. The amplified DNA (referred to hereafter as WGA DNA) was visualized on 0.8% agarose gels. The WGA DNA was genotyped for the same 10 SNPs (Table 1) as the saliva DNA samples by using the protocol described above.

Statistical Analysis. Student's *t* tests were used to evaluate differences of response rate between age groups. Spearman rank correlation (r_s) was used to test correlations between DNA yields measured with PicoGreen, UV absorbance, and real-time PCR as well as between DNA yield and age.

Results

We obtained a total response rate of 80%, 7% did not want to participate in the study, and 13% did not respond after two reminders. Of the total number of 490 samples, 52% were received during the 1st week, 24% during the 2nd week, and 17% during the 3rd week of the sample collection. We observed statistically significant differences of response rates between age groups. The highest rates were observed in age groups 67 to 71 (89%) and 62 to 66 (85%) compared with age group 53 to 61 (73%; $P_s < 0.01$ and 0.02, respectively). The lowest response was observed in the oldest age group 77 to 87 (71%).

By visual inspection of the received samples, the saliva volume was estimated to vary between 1.5 and 2.5 mL; phlegm or other contaminants were not visible. We extracted DNA from 90 randomly selected saliva samples and roughly estimated total DNA yield by UV absorption. The mean total yield was 135.9 μ g (SD, 118.2), and the mean A260/A280 ratio, used as an estimate of DNA purity, was 1.76 (SD, 0.12). We also measured the total DNA yield by using the PicoGreen method. In one sample, the DNA was not measurable. The total DNA yield ranged from 1.2 to 169.7 μ g, with a mean of 40.3 μ g (SD, 36.5 μ g) and a median of 29.4 μ g. By real-time PCR of the human prothrombin gene, we estimated that the yield of human DNA ranged from 11% to 100% of total DNA, with a median of 68% (SD, 20%). The human DNA yield ranged from 0.8 to 85.6 μ g, with a mean of 25.4 μ g (SD, 22.4) and a median of 19.2 μ g. Electrophoretic analysis of the extracted DNA showed detectable levels of high molecular weight genomic DNA (>23 kb) in all 89 samples. No degradation of DNA was observed. Total DNA (bacterial and human), as measured with PicoGreen, was highly correlated with both total DNA as measured with UV absorbance ($r_s = 0.90$) and human DNA as measured with real-time PCR ($r_s = 0.92$; Fig. 1). We found no correlation between age and DNA yield [$r_s = 0.02$ for total DNA (PicoGreen), $r_s = 0.07$ for total DNA (UV absorbance), and $r_s = 0.005$ for human DNA (real-time PCR)].

The DNA amount, as measured with PicoGreen, obtained after WGA of 1 to 5 ng of saliva DNA was ~4 μ g. Quality assessment of both saliva DNA and WGA DNA was done by genotype analyses of 10 SNPs. All genotype frequencies were in Hardy-Weinberg equilibrium (Table 1). The success rates for the 10 specific SNP assays ranged between 91% and 100% for the saliva DNA samples and between 88% and 99% for the WGA DNA samples (Table 1). We obtained all 10 genotypes for 70 of 89 (79%) saliva DNA samples and 64 of 81 (79%) WGA DNA samples. Of the samples that were not completely genotyped for all 10 SNPs, none missed more than five genotypes. Eleven percent of the saliva DNA and 29% of the WGA DNA samples missed four or five genotypes. The rest missed one, two, or three genotypes. Almost complete genotypic concordance (99.7%) was observed between saliva DNA and WGA DNA samples, and 736 of 738 genotypes were identical.

Table 1. Information on 10 SNPs genotyped on 89 saliva DNA samples and on 81 WGA DNA samples; success rates, *P*s for Hardy-Weinberg equilibrium for each assay, and genotypic concordance

SNP identification no.*	Chromosome	Gene/Protein	Saliva DNA		WGA DNA		Concordance
			Success rate of genotype analyses (%)	<i>P</i> for Hardy-Weinberg equilibrium	Success rate of genotype analyses (%)	<i>P</i> for Hardy-Weinberg equilibrium	WGA DNA vs saliva DNA (%)
Specific SNPs							
rs746320	1	No gene	94	0.28	99	0.49	100
rs2364731	2	Hypothetical protein DKFZp451M2119	94	0.08	91	0.55	100
rs4306954	4	No gene	98	0.75	91	0.35	100
rs1500320	9	Protein tyrosine phosphatase, receptor type, D	100	0.79	94	0.64	100
rs3781058	10	BRCA2 and CDKN1A interacting protein	94	0.42	96	0.66	100
rs7994365	13	A kinase (PRKA) anchor protein 11	94	0.996	95	0.29	100
rs1465471	16	Sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingomyelinase II)	91	0.52	88	0.52	99
rs8110400	19	Zinc finger protein 543	99	0.08	99	0.49	100
rs744533	20	Chromosome 20 open reading frame 151	91	0.43	96	0.75	100
rs139842	22	Eukaryotic translation initiation factor 3, subunit 6 interacting protein	99	0.83	99	0.85	99
All 10 SNPs			96	NA	95	NA	99.7

*National Center for Biotechnology Information Reference SNP Cluster Identifier.

Discussion

In our work with planning a DNA biobank from subjects in the cohort of Swedish men (13), we conducted this pilot study to evaluate collection of saliva through the self-administrated Oragene method, which to our knowledge has not been previously evaluated in a scientific report. We contacted 611 men and asked them to spit roughly 2 mL saliva into a collection vial, which they returned by mail to our laboratory. The method turned out to be user friendly, and 80% responded with a saliva sample. Of the received samples, 93% were mailed to us within 3 weeks from the send out date. Extraction yielded high molecular weight DNA, which was of sufficiently high quality for genotype analyses and for WGA.

Other methods for collection of saliva or buccal cells include mouthwash, buccal swabs/brushes, and treated cards

(1-12, 17). The human DNA yield from the Oragene method (19 µg) is in the same range as the human yield from the mouthwash method (3, 4) but substantially higher compared with buccal swabs/brushes (0.2-2.7 µg per swab/brush; refs. 2, 3, 7, 12). The foremost advantage with the Oragene protocol compared with these methods is its simplicity; neither do the donors need to put something in their mouth before collecting the sample nor do they need to follow a certain collection protocol, such as rubbing their cheeks against their teeth to prepare the buccal mucosa before mouthwash (1) or tooth brushing (11). Moreover, when closing the collection vial directly after spitting, the saliva is mixed with DNA preserving and purifying chemicals, which prevents bacterial growth and degradation of human DNA. Feigelson et al. (4) have shown that handling of mouthwash samples 10 to 30 days after collection gave a statistically significant reduction of human DNA yield, which indicates that the time between collection and processing of the samples when samples are held at room temperature may be important for the human DNA yield. According to the manufacturers of the Oragene kit, the saliva samples can be stored at room temperature for up to at least 1 year without DNA degradation.³

One obstacle with saliva DNA is the high variability of bacterial DNA. In DNA extracted from buccal swabs, human DNA accounted for only 11% (3). Our data showed that 68% of the total DNA was of human origin, which is higher compared with Garcia-Closas study (3) on the mouthwash method in which 49% was of human origin. However, the higher proportion of human DNA from the Oragene samples compared with the mouthwash samples may be due to the different methods used to measure human DNA. We also found that the DNA yield varied substantially between saliva samples, the amount of human DNA ranged between 1 and 86 µg, and ~20% of the samples had a yield <5 µg. Considering genotype analyses using 2 to 5 ng DNA per PCR, biobank DNA samples in this lower yield range would not be sufficient for future use. We therefore evaluated WGA as a backup approach and

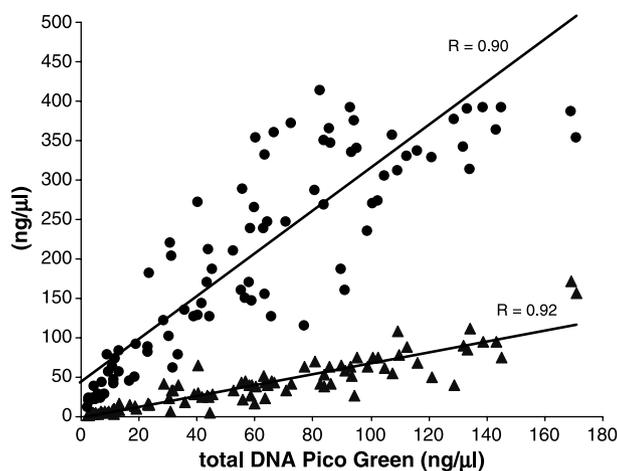


Figure 1. Spearman rank correlation of total DNA concentration as measured by PicoGreen with human DNA concentration as measured by real-time PCR (▲) as well as total DNA concentration as measured by UV absorbance (●).

³ Personal communication.

showed that this method can be used to generate additional DNA. Genotyping of the WGA DNA worked satisfactorily; however, to evaluate a potential risk for allelic dropout derived from WGA, >10 SNPs need to be tested in the future.

In conclusion, the Oragene collection method, which in this study was evaluated in men, is a suitable way to collect DNA from large groups of subjects where self-administration is an advantage. The DNA quality allows for genotyping with high success rate, and the yield is sufficient for a large number of analyses. Furthermore, the DNA is amplifiable, which insure near-unlimited biobank material for future studies.

Acknowledgments

We thank Cecilia Lindgren (Mutation Analysis Facility, Karolinska Institutet, Stockholm, Sweden) and Joyce Carlson (DNA Extraction Facility, Malmö, Sweden) for providing invaluable methodologic knowledge and Camilla Lagerberg (KI Biobank, Stockholm, Sweden) for excellent technical assistance.

References

- King IB, Satia-Abouta J, Thornquist MD, et al. Buccal cell DNA yield, quality, and collection costs: comparison of methods for large-scale studies. *Cancer Epidemiol Biomarkers Prev* 2002;11:1130–3.
- Walker AH, Najarian D, White DL, Jaffe JF, Kanetsky PA, Rebbeck TR. Collection of genomic DNA by buccal swabs for polymerase chain reaction-based biomarker assays. *Environ Health Perspect* 1999;107:517–20.
- Garcia-Closas M, Egan KM, Abruozzo J, et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;10:687–96.
- Feigelson HS, Rodriguez C, Robertson AS, et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;10:1005–8.
- Le Marchand L, Lum-Jones A, Saltzman B, Visaya V, Nomura AM, Kolonel LN. Feasibility of collecting buccal cell DNA by mail in a cohort study. *Cancer Epidemiol Biomarkers Prev* 2001;10:701–3.
- Freeman B, Powell J, Ball D, Hill L, Craig I, Plomin R. DNA by mail: an inexpensive and noninvasive method for collecting DNA samples from widely dispersed populations. *Behav Genet* 1997;27:251–7.
- Meulenbelt I, Droog S, Trommelen GJ, Boomsma DI, Slagboom PE. High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 1995;57:1252–4.
- Lum A, Le Marchand L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. *Cancer Epidemiol Biomarkers Prev* 1998;7:719–24.
- Heath EM, Morken NW, Campbell KA, Tkach D, Boyd EA, Strom DA. Use of buccal cells collected in mouthwash as a source of DNA for clinical testing. *Arch Pathol Lab Med* 2001;125:127–33.
- Mulot C, Stucker I, Clavel J, Beaune P, Lorient MA. Collection of human genomic DNA from buccal cells for genetics studies: comparison between cytobrush, mouthwash, and treated card. *J Biomed Biotechnol* 2005;2005:291–6.
- London SJ, Xia J, Lehman TA, et al. Collection of buccal cell DNA in seventh-grade children using water and a toothbrush. *Cancer Epidemiol Biomarkers Prev* 2001;10:1227–30.
- Harty LC, Garcia-Closas M, Rothman N, Reid YA, Tucker MA, Hartge P. Collection of buccal cell DNA using treated cards. *Cancer Epidemiol Biomarkers Prev* 2000;9:501–6.
- Norman A, Bellocco R, Vaida F, Wolk A. Total physical activity in relation to age, body mass, health, and other factors in a cohort of Swedish men. *Int J Obes Relat Metab Disord* 2002;26:670–5.
- Jurinke C, van den Boom D, Cantor CR, Koster H. Automated genotyping using the DNA MassArray technology. *Methods Mol Biol* 2002;187:179–92.
- Phillips MS, Lawrence R, Sachidanandam R, et al. Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nat Genet* 2003;33:382–7.
- Holbrook JF, Stabley D, Sol-Church K. Exploring whole genome amplification as a DNA recovery tool for molecular genetic studies. *J Biomol Tech* 2005;16:125–33.
- Steinberg K, Beck J, Nickerson D, et al. DNA banking for epidemiologic studies: a review of current practices. *Epidemiology* 2002;13:246–54.

Quality and Quantity of Saliva DNA Obtained from the Self-administrated Oragene Method—A Pilot Study on the Cohort of Swedish Men

Tove Rylander-Rudqvist, Niclas Håkansson, Gunnel Tybring, et al.

Cancer Epidemiol Biomarkers Prev 2006;15:1742-1745.

Updated version Access the most recent version of this article at:
<http://cebp.aacrjournals.org/content/15/9/1742>

Cited articles This article cites 17 articles, 7 of which you can access for free at:
<http://cebp.aacrjournals.org/content/15/9/1742.full#ref-list-1>

Citing articles This article has been cited by 12 HighWire-hosted articles. Access the articles at:
<http://cebp.aacrjournals.org/content/15/9/1742.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cebp.aacrjournals.org/content/15/9/1742>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.