Biomarkers, September–October 2006; 11(5): 472–479

informa
healthcare

Quantitation of DNA in buccal cell samples collected in epidemiological studies

M. GARCÍA-CLOSAS[1], L. E. MOORE[1], C. S. RABKIN[1], T. FRANKLIN[2], J. STRUEWING[3], D. GINZINGER[4], J. ALGUACIL[1], & N. ROTHMAN[1]

[1]*Division of Cancer Epidemiology and Genetics, and* [3]*Laboratory of Population Genetics, National Cancer Institute, Bethesda, MD, USA,* [2]*ATCC, Cell Biology, Manassas, VA, USA and* [4]*University of California at San Francisco, San Francisco, CA, USA*

Abstract

Buccal cell samples are increasingly used in epidemiological studies as a source of genomic DNA. The accurate and precise quantitation of human DNA is critical for the optimal use of these samples. However, it is complicated by the presence of bacterial DNA and wide inter-individual variation in DNA concentration from buccal cell collections. The paper evaluated the use of ultraviolet light (UV) spectroscopy, Höechst (H33258) and PicoGreen™ as measures of total DNA, and real-time quantitative polymerase chain reaction (PCR) as a measure of human amplifiable DNA in buccal samples. Using serially diluted white blood cell DNA samples (at a concentration range of 300 to 0.5 ng μl^{-1}), UV spectroscopy showed the largest bias, followed by Höechst, especially for low concentrations. PicoGreen and real-time PCR provided the most accurate and precise estimates across the range of concentrations evaluated, although an increase in bias with decreasing concentrations was observed. The ratio of real-time PCR to PicoGreen provided a reasonable estimate of the percentage of human DNA in samples containing known mixtures of human and bacterial DNA. Quantification of buccal DNA from samples collected in a breast cancer case-control study by PicoGreen and real-time PCR indicated that cytobrush and mouthwash DNA samples contain similar percentages of human amplifiable DNA. Real-time PCR is recommended for the quantification of buccal cell DNA in epidemiological studies since it provides precise estimates of human amplifiable DNA across the wide range of DNA concentrations commonly observed in buccal cell DNA samples.

Keywords: *Buccal cells, DNA, epidemiology, quantitation.*

(Received 15 November 2005; accepted 28 March 2006)

Introduction

Epidemiological studies are increasingly using buccal epithelial cell samples as a source of genomic DNA since blood samples are often difficult to obtain due to logistical or cost issues. The two main methods currently used to collect buccal cells are cytobrush and mouthwash collections (Lum & Le Marchand 1998, García-Closas et al. 2001, Heath et al. 2001, Le Marchand et al. 2001, Moore et al. 2001, Feigelson

Correspondence: M. García-Closas, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892, USA. Tel: 1-301-435-3981. Fax: 1-301-402-0916. Email: garciacm@exchange.nih.gov

ISSN 1354-750X print/ISSN 1366-5804 online
DOI: 10.1080/13547500600733820

et al. 2002). Mouthwash samples provide larger amounts of high molecular weight DNA than cytobrush samples. However, the amount of DNA recovered per subject by either of these methods varies widely between individuals, from a few nanograms to a few hundred micrograms per sample (García-Closas et al. 2001, Moore et al. 2001).

Accurate quantitation of human DNA concentrations from buccal cell samples is critical for the optimal use of DNA, especially for samples with low yields, and for the successful use of techniques that require narrow ranges of DNA concentrations such as the determination of microsatellite repeats or highly multiplex genotyping platforms. Quantitation of DNA in buccal cell samples is complicated by the fact that the concentration varies widely between individuals and can contain non-human sources of DNA, degraded DNA and contaminants that could interfere with the quantification technique (Moore et al. 2001). Thus, different quantification techniques can provide very different estimates of DNA concentration and yield for a given sample, depending on the type of DNA they measure (e.g. total single- or double-stranded DNA, total human DNA, polymerase chain reaction (PCR)-amplifiable human DNA), their accuracy across the concentration ranges typically found in these samples and the interference of contaminants.

Measuring absorbance at 260 nm by ultraviolet light (UV) spectroscopy is the most commonly used technique for determining nucleic acid concentrations because it is quick and inexpensive. Höechst (H33258) and PicoGreen™ are two fluorescent stains selective for double-strand DNA, and PicoGreen is the most sensitive of these methods (Singer et al. 1997). Real-time quantitative PCR can measure small amounts of human amplifiable DNA within a mixture of DNA species (Walker 2002). A major advantage of real-time PCR over PicoGreen to quantify DNA extracts from buccal cell samples is its ability to quantitate only amplifiable human DNA. The present study evaluates the accuracy and precision of these different methods to measure total DNA (i.e. UV spectroscopy, Höechst (H33258), PicoGreen) and human amplifiable DNA (i.e. real-time PCR) for a range of concentrations typically found in buccal cell DNA extracts. It also evaluates the accuracy of the ratio of real-time PCR to PicoGreen in measuring the percentage of human DNA using samples with known mixtures of human and bacterial DNA.

Materials and methods

Study samples

Serial dilutions (1:2, 1:5, 1:25, 1:600) of white blood cell DNA obtained from one volunteer were prepared using a standard phenol–chlorophorm extraction protocol. The starting concentration of the reference sample was targeted to 300 ng μl^{-1} using UV spectroscopy. Four aliquots of each dilution were sent to different laboratories to perform the following quantitation assays: laboratory 1 (UV spectroscopy, PicoGreen and real-time quantitative PCR); laboratory 2 (Höechst); and laboratory 3 (PicoGreen and real-time quantitative PCR).

Samples were prepared with known percentages of human and bacterial DNA by pooling DNA extracted from white blood cell samples and DNA extracted from *Streptococcus mutans*, a bacterium commonly found in the human oral cavity. Two sets of samples were prepared containing targeted percentages of human DNA of 100, 80, 60 and 40%. The final human DNA concentrations in the mixtures were 12 and

60 ng μl^{-1} for each set of samples, and the total DNA concentrations were 15–60 ng μl^{-1} for the first set and 75–360 ng μl^{-1} for the second set.

DNA from cytobrush and mouthwash DNA samples collected by mail in a case-control study of breast cancer in the USA were re-quantified (García-Closas et al. 2001) using PicoGreen and real-time PCR. Quantitation by these methods was compared with earlier estimates by UV spectroscopy and hybridization with a human DNA probe, respectively. The study population and methods of collection for the breast cancer study are described in detail in García-Closas et al. (2001). In brief, cytobrush samples were collected by a self-collection kit that contained two cytobrush samples and a return envelope. Women were asked to brush their teeth, after 10–15 min 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 them to the laboratory. Upon arrival to the laboratory, the brushes were separated from the handle, placed in a cryovial and frozen at –70/80°C until DNA extraction. The mouthwash samples were collected by self-collection kit that used Scope® mouthwash. Women were asked to swish 10 ml of Scope mouthwash for 30–45 s, expectorate into a collection cup and mail the sample to the laboratory. Upon arrival at the laboratory, mouthwash sample cell pellets were separated from the supernatant, resuspended in 3 ml of TE buffer solution — Tris ethylenediamine tetra-acetic acid (EDTA) (100 × concentration, pH 8.0) in diethylpyrocarbonate (DEPC) water, 1:100 solution—aliquoted, and stored at –70/80°C until DNA extraction.

DNA quantification assays

DNA concentrations were determined by UV spectroscopy using a Beckman DU-640 Spectrophotometer (Beckman Scientific Instruments, Fullerton, CA, USA), by flurometery Höechst (H33258) and the PicoGreen™ dsDNA Quantitation Kit (Molecular Probes, Inc., Eugene, OR, USA) according to the manufacturer's instructions, and real-time quantitative PCR using an amplification target located within the human *BRCA1* locus and the TaqMan detection system. DNA yields were calculated by multiplying the DNA concentration by the total volume of DNA solution after DNA extraction.

Statistical analysis

Mean values and coefficients of variation were calculated for each set of four replicate samples of white blood cell DNA: reference sample, 1:2 dilution, 1:5 dilution, 1:25 dilution and 1:600 dilution. The expected concentration for each dilution factor was calculated by multiplying the concentration of the reference sample times the corresponding dilution factor, for each quantification technique. The percentage bias was calculated as: (observed mean concentration – expected concentration)/expected concentration. The percentage of human DNA was estimated as the ratio between the real-time PCR:PicoGreen concentration. The adjusted percentage of human DNA was calculated by dividing each mean percentage by the mean percentage for the sample containing only white blood cell DNA. The percentage bias was calculated as: (expected percentage – observed percentage)/expected percentage.

Differences in median DNA yields between DNA quantitation methods were tested with a Wilcoxon matched-pair signed-rank test for paired observations.

Results and discussion

Table I shows the observed mean concentrations and coefficients of variation for four replicates of the reference and four diluted samples, by UV spectroscopy, Höechst, PicoGreen and real-time quantitative PCR. The percentage bias for each dilution was calculated with respect to the reference sample. UV spectroscopy overestimated the DNA concentrations and the bias increased dramatically as the sample concentration decreased. Höechst fluorometry overestimated DNA concentration to a lesser extent than UV spectrometry and the bias also increased as the DNA concentration decreased. Estimates of DNA concentrations by PicoGreen and real-time PCR were very similar and highly accurate (coefficient of variation ranging from 1.5–3.2 and 1.3–8.9%, respectively). However, both methods tended to overestimate low DNA concentrations.

The percentage of human DNA present in buccal samples can be quantified by the ratio of measures of the total to the human DNA concentration. We evaluated the bias and precision of the ratio of real-time PCR to PicoGreen DNA concentrations as a measure of the percentage of human DNA in samples with mixtures of human and

Table I. Bias and coefficient of variation for DNA concentrations from serial dilutions of a white blood cell DNA sample using different DNA quantification techniques.

Quantification technique Dilution factors	N[1]	Observed (ng μl^{-1})	Expected[2] (ng μl^{-1})	Bias (%)[3]	Coefficient of variation (%)
Ultraviolet light spectroscopy					
Reference sample	4	368.7			3.6
1:2	4	194.1	184.3	5.3	4.5
1:5	4	88.5	73.7	20.0	3.2
1:25	4	43.2	14.7	192.7	15.5
1:600	4	44.7	0.6	7172.2	8.0
Hoechst					
Reference sample	4	321.1			9.8
1:2	4	165.6	160.6	3.2	33.9
1:5	4	103.0	64.2	60.4	48.3
1:25	4	17.8	12.8	38.2	13.9
1:600	4	10.0	0.5	1768.4	15.8
PicoGreen™					
Reference sample	4	305.9			2.0
1:2	4	157.3	152.9	2.9	2.3
1:5	4	61.1	61.2	−0.2	1.5
1:25	4	12.9	12.2	5.3	3.2
1:600	4	0.8	0.5	64.8	1.7
Real-time polymerase chain reaction (*PCR*) (BRCA1 *probe*)					
Reference sample	4	305.5			1.8
1:2	4	162.0	152.8	6.0	3.3
1:5	4	70.1	61.1	14.7	8.9
1:25	4	14.5	12.2	18.7	4.8
1:600	4	0.7	0.5	40.9	1.3

[1]Number of replicate samples.
[2]Expected DNA concentrations are based on the concentration of the neat sample divided by each dilution factor.
[3]Bias = (observed − expected)/expected.

bacterial DNA at known proportions (Table II). The ratio of these two measurements provided a reasonably accurate (bias ranging from 12.6 to 43.3%) and precise (coefficient of variation ranging from 1.2 to 14.3%) estimates of the percentage of human DNA. When estimates where adjusted by setting the percentage of human DNA from the samples containing only human DNA to 100%, bias was substantially reduced (bias ranging from 4.6 to 27.3%). Therefore, it is recommended that the ratio of real-time PCR to PicoGreen in buccal DNA samples be adjusted by the ratio in samples known to contain only human DNA. Because between-laboratory and between-batch differences in quantification are possible, human DNA samples should preferably be quantified at the same laboratory and within the same batches as the buccal DNA samples.

We have previously reported that total and human DNA yields obtained by one 10-ml sample of alcohol-containing mouthwash are higher than those obtained by two buccal cytobrush samples, based on samples collected from women participating in a breast cancer case-control study in the USA (García-Closas et al. 2001). UV spectroscopy and hybridization with a human DNA probe using the ACES 2.0+ DNA Quantification System (Life Technologies, Inc., Grand Island, NY, USA) were used in that report to estimate total and human DNA concentrations, respectively. Because we have shown in the present paper that UV overestimates DNA concentrations typically found in buccal specimens, and the ACES 2.0+ DNA Quantification System is no longer available, we re-quantified the samples from the previous report (García-Closas et al. 2001) by PicoGreen and real-time quantitative PCR (Table III). Compared with PicoGreen, UV spectroscopy overestimated the total DNA yield from cytobrush samples (median yield of 13.5 and 6.8 μg for UV spectroscopy and PicoGreen, respectively; $p < 0.001$). Compared with real-time quantitative PCR, hybridization underestimated human DNA yields from cytobrush samples (median

Table II. Bias in the estimation of the percentage of human DNA by TaqMan/PicoGreen.

				Observed values			Adjusted values[4]	
Human DNA concentration (ng μl^{-1})	Total DNA concentration (ng μl^{-1})	Expected[1] human DNA (%)	N[2]	Mean (%)	Coefficient of variation (%)	Bias (%)[3]	Mean (%)	Bias (%)[3]
12	12	100.0	4	114.8	8.3	14.8	100.0	0.0
12	15	80.0	4	96.1	4.4	20.2	83.7	4.6
12	20	60.0	4	77.6	5.6	29.3	67.5	12.6
12	30	40.0	4	48.8	9.0	21.9	42.5	6.2
12	60	20.0	4	26.6	4.5	33.2	23.2	16.0
60	60	100.0	4	112.6	6.0	12.6	100.0	0.0
60	75	80.0	4	97.9	5.5	22.4	87.0	8.7
60	100	60.0	4	71.0	9.6	18.3	63.1	5.1
60	150	40.0	4	57.2	1.2	42.9	50.8	26.9
60	360	16.7	4	28.7	14.3	43.3	25.5	27.3

[1]Based on mixtures of white blood cell DNA and bacterial DNA at known percentages.
[2]Number of replicate samples.
[3]Bias is defined as: (observed – expected)/expected.
[4]Adjusted values are calculated by setting the percentage of human DNA from the first sample containing only human DNA to 100%.

yield of 1.7 and 3.7 μg for hybridization and real-time PCR respectively; $p < 0.001$). On the other hand, differences in yield for mouthwash samples were not statistically significant for either total or human DNA yield (Table III).

The bias in UV spectroscopy measurement increases with decreasing 'true' concentrations. Thus, the differences in total DNA yields observed for cytobrush but not for mouthwash DNA samples are likely to be explained by lower DNA concentrations found in cytobrush compared with mouthwash samples (median (range) total DNA concentration by PicoGreen was 91.6 (1.8–373) versus 192.7 (406–3323) ng μl^{-1}, respectively). Differences in the yield between hybridization and real-time PCR are likely to be explained by a decreased sensitivity of the hybridization method for low DNA concentrations, although we could not confirm this because of the unavailability of this quantification system.

These findings (Table I) indicate that both PicoGreen and real-time PCR assays can provide accurate and precise estimates of DNA concentrations across a wide range of concentrations (300 to 0.5 ng μl^{-1}) typically found in buccal DNA extracts. Since real-time PCR assays provide human-specific measures, they would be preferable to quantify buccal DNA samples. Given the low accuracy and precision of UV spectroscopy to measure low DNA concentrations, this method is not recommended to quantify DNA in buccal samples. Additional limitations of UV include large volume requirements (approximately 7–14 μl), an inability to distinguish between different types of nucleic acids (RNA and DNA of human and non-human origin), and potential interferences caused by contaminants found in DNA extracts.

The bias in total and human concentrations in cytobrush DNA samples measured by UV spectroscopy and hybridization, respectively, resulted in an apparently lower percentage of human DNA in cytobrush than mouthwash DNA samples (10.1 versus 47.0%; Table III) (García-Closas et al. 2001). However, the ratio of real-time PCR to PicoGreen indicated that both types of DNA samples contain a similar percentage of human amplifiable DNA (57.9 versus 58.2%; Table III). The range in the ratio of human DNA concentration measured by real-time PCR to total DNA concentration measured by PicoGreen was highly variable between subjects, ranging from 11.2 to 88.9% for cytobrush samples and from less than 0.01 to more than 100% for mouthwash samples (Table III). This indicates that by PicoGreen estimates of total DNA would not be good surrogates for human DNA present in these buccal DNA samples. The estimated ratios were not adjusted by measures in human DNA samples; therefore, they might not reflect the true magnitude of the ratio of human to bacterial DNA, as indicated by the data shown in Table II.

In conclusion, real-time quantitative PCR provides accurate estimates of amplifiable human DNA present in buccal cell samples, allowing for the most efficient and optimal use of these samples. Quantification of total rather than human DNA in buccal cell samples, even when using highly sensitive techniques such as PicoGreen, is not recommended for assays that require narrow ranges of amplifiable human DNA concentrations because of the wide variation in the ratio of real-time PCR to PicoGreen concentrations found in buccal cell samples. Finally, the ratio of real-time quantitative PCR to PicoGreen concentration indicated that although the amount of DNA is larger for mouthwash than cytobrush samples, as previously reported, both types of samples contain similar percentages of human DNA.

Table III. Distribution of DNA yields estimated by different quantification techniques using samples in García-Closas et al. (2001) (μg/mouthwash sample or μg/two cytobrush samples).

Method of collection	N	Spectrophotometer[1], median (range)	Hybridization with human DNA probe[2], median (range)	Ratio[1,2] (%), median (range)	PicoGreen, median (range)	TaqMan real-time polymerase chain reaction (PCR) (*BRCA1* probe), median (range)	Ratio[3] (%), median (range)
Cytobrush	28	13.5 (2.0, 38.2)	1.7 (0.1, 7.7)	10.1 (2.0–73.3)	6.8 (0.2, 37.3)	3.7 (0.1, 25.6)	57.9 (11.2–88.9)
Mouthwash	25	38.7 (11.9, 124.2)	16.6 (1.6, 140.3)	47.0 (3.6–187.2)	35.5 (6.6, 81.3)	17.2 (<0.01, 225.4)	58.2 (0.0–304.0)

[1]Data were previously published in García-Closas et al. (2001).
[2]Defined as: (hybridization with a human DNA probe/ultraviolet light spectroscopy)*100.
[3]Defined as: (TaqMan/PicoGreen)*100.

Acknowledgements

This work was supported by funds from the intramural program of the National Cancer Institute, National Institutes of Health, USA. None of the authors of this manuscript declares competing interests.

References

Feigelson H, Rodriguez C, Robertson A, Jacobs E, Calle E, Reid Y, Thun M. 2002. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. Cancer Epidemiology and Biomarkers Prevention 10:1005–1008.

García-Closas M, Egan KM, Abruzzo J, Newcomb PA, Titus-Ernstoff L, Franklin T, Bender PK, Beck JC, Le Marchand L, Lum A, Alavanja M, Hayes RB, Rutter J, Buetow K, Brinton LA, Rothman N. 2001. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. Cancer Epidemiology and Biomarkers Prevention 10:687–696.

Heath EM, Morken NW, Campbell KA, Tkach D, Boyd EA, Strom DA. 2001. Use of buccal cells collected in mouthwash as a source of DNA for clinical testing. Archives in Pathology and Laboratory Medicine 125:127–133.

Le Marchand L, Lum-Jones A, Saltzman B, Visaya V, Nomura AM, Kolonel LN. 2001. Feasibility of collecting buccal cell DNA by mail in a cohort study. Cancer Epidemiology and Biomarkers Prevention 10:701–703.

Lum A, Le Marchand L. 1998. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. Cancer Epidemiology and Biomarkers Prevention 7:719–724.

Moore L, Wiencke JK, Eng C, Zheng S, Smith A. 2001. Evaluation of buccal cell collection protocols for genetic susceptibility studies. Biomarkers 6:448–454.

Singer VL, Jones LJ, Yue ST, Haugland RP. 1997. Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. Annals of Biochemistry 249:228–238.

Walker N. 2002. A technique whose time has come. Science 296:557–559.