Short Communication

Whole Genome Amplification Increases the Efficiency and Validity of Buccal Cell Genotyping in Pediatric Populations[1]

Shichun Zheng, Xiaomei Ma, Patricia A. Buffler, Martyn T. Smith, and John K. Wiencke[2]

Laboratory for Molecular Epidemiology, Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California 94143-0560 [S. Z., J. K. W.], and Divisions of Public Health Biology and Epidemiology [X. M., P. A. B.] and Environmental Health Sciences [M. T. S.], University of California Berkeley, School of Public Health, Berkeley, California 94720-7360

Abstract

The collection of buccal cells provides a noninvasive method for obtaining DNA for genetic studies. Here we report the results on buccal cell genotyping from our ongoing study of childhood leukemia in Northern California. We have collected buccal samples from children ranging in age from 4 months to 15 years using an interviewer- or nurse-administered protocol using a cytology brush. Initial results of the genotyping, including the *glutathione S-transferase* μ, *glutathione S-transferase* θ, *NAD(P)H:quinone oxidoreductase*, and methylenetetrahydrofolate reductase polymorphisms, were disappointing because many specimens contained little DNA, failed repeated attempts at PCR amplification, and produced unreliable results. Here we evaluate a solution to the problem that involves whole genome amplification using the improved primer extension preamplification methodology. Sixty cases of pediatric acute leukemia were studied; five PCR-based genotypes were attempted using buccal cell DNA and whole genome amplified (WGA) buccal DNA. Results were compared with genotyping results using DNA isolated from peripheral whole blood or bone marrow for each child. The standard buccal protocol failed to yield successful PCR reactions in 30–57% of specimens, whereas WGA-buccal was markedly more efficient (2–5% failed PCR). A success rate of 100% was achieved with one repeat test of the failed WGA-PCR reactions. Misclassification of genotype was common for the glutathione *S*-transferase θ marker using the standard buccal procedure. The WGA-buccal protocol, however, produced genotyping results fully concordant with the referent blood or bone marrow DNA results for all five loci. DNA yields were increased

by WGA to allow for ~900 PCR reactions/brush. WGA is very useful for improving the efficiency and validity of PCR-based genotyping in pediatric populations.

Introduction

A number of protocols have been developed to obtain DNA for genotyping from cheek swabs (buccal cells), mostly in studies of adult populations (1–5). Several studies have evaluated methods for buccal cell collection in younger subjects, including babies (6, 7). Because of their age, infants and children pose an additional challenge for the collection of DNA specimens. In our population-based case-control study of childhood acute leukemia, we have field-tested a cytobrush protocol widely used in adults and have found that the yields of successful buccal specimens were less than that expected from our experience with adults. To address this problem, we have evaluated a method to amplify the entire genome from limited buccal specimens using a method called WGA.[3]

Materials and Methods

Study Population. Our larger population is comprised of incident cases of childhood leukemia who were under 15 years of age at the time of diagnosis. The subset included here were enrolled from the beginning of our study (1996) to the present time and were representative of the larger case population. Peripheral blood and/or bone marrow slides for each case were obtained from the clinical center that first diagnosed the case. In addition, buccal cell specimens were collected by a research nurse at the clinical center (31 cases) or by an interviewer at the time of the in-home personal interview (29 cases) shortly after a case consented to participate. Specimens were usually obtained after therapy. Because controls do not provide blood or bone marrow samples, we did not include controls in our comparison of buccal cell methodologies. The biological samples were sent to the laboratories for processing via courier or express mail. To compare the buccal cell specimens between children and adults, data on DNA concentration and total DNA yield from 116 adult lung cancer patients were obtained from the University of California San Francisco Laboratory for Molecular Epidemiology (San Francisco, CA). The same protocol and brushes were used for adults and children, except that adults self-administered the buccal collection and mailed in their specimens.

Buccal Cell Collection and Processing. Buccal cell collection with the cytobrush was performed as described previously by Richards *et al.* (1). Briefly, cells were collected on a sterile cytology brush by twirling the brush (Medical Packaging Corp,

Camarillo, CA) on the inner cheek for 30 s. The brush was then transferred to a tube containing 50 mM NaOH and transported to the laboratory at ambient temperature. In processing, samples were boiled and centrifuged, and the supernatant volume was measured and neutralized with Tris-HCl to pH 7.8. Solutions were then stored at −70°C. DNA was measured in all samples by Hoechst 33258 fluorometry (TKO100; Hoefer Scientific Instruments, San Francisco, CA) and then stored at −70°C.

WGA. We used a version of WGA known as PEP-PCR, using cycling modifications as described by Dietmaier *et al.* (8) and degenerate PCR primers 15 nucleotides in length as described by Zhang *et al.* (9). Primers (random 15-mers) were obtained from Operon Technologies (Alameda, CA). Improved PEP-PCR was set up on the basis of instructions in the Expand High Fidelity PCR System (Boehringer Mannheim) as follows. First, begin with 20 μl of mix 1 (final concentration: 0.05 mg/ml gelatin, 16 μmol/liter degenerate 15-mer primers, and 0.1 mM dNTPs) to 10 μl of buccal cell solution (10–100 ng DNA). Second, add 30 μl of mix-proprietary part of a kit 2 (final concentration: 3.5 units Expand High Fidelity polymerase and proofreading Pwo polymerase, 2.5 mM MgCl$_2$ in 1× PCR buffer No. 3) to the buccal cell DNA primer mixture. PCR cycling conditions were as follows. A single 2-min predenaturation step at 94°C was performed and then 50 cycles of step 1: denaturation at 94°C for 40 s; step 2: 2 min at 37°C; step 3: ramping up to (0.1°C/s) and holding 55°C for 4 min; and step 4: 30 s at 68°C. During the last cycle, a final elongation step of 68°C for 7 min was performed. Total PCR cycling time was approximately 10.5 h.

Genotyping. PCR was performed with a Gene Amp PCR 9600 Thermal Cycler (Perkin-Elmer) in 50 μl reaction volumes. Unless otherwise stated, each PCR reaction contained 1.5 mM MgCl$_2$, 0.2 mM deoxynucleotide triphosphate, 2.5 units Taq polymerase in 1× PCR Buffer (Perkin-Elmer), and 0.4 μM each of forward and reverse primers. Amplifications were performed with 10 μl of buccal cell solution, 4 μl of WGA product, or 20–50 ng of whole blood or bone marrow genomic DNA.

***NQO1* Polymorphism.** In this PCR analysis, the sense primer (5′-TCCTCAGAG TGGCATTCTGC-3′) and antisense primer (5′-TCTCCTCATCCTGTACCTCT-3′) amplify a 230-bp region, including the last 7 bases of intron 5 and the first 223 bases of exon 6 (Ref. 10; GenBank accession no. J05348). Reaction mixtures were preincubated for 2 min at 94°C. The reactions were heated to 94°C for 30 s, cooled to 60°C for 30 s, and heated to 72°C for 30 s, for 35 cycles, with a final extension of 7 min at 72°C. Fifteen μl of the PCR products were then digested with 2 μl of *Hin*fI (NE Biolabs) for 4 h at 37°C. *Hin*fI recognizes GA*N*TC, and so the transition from wild-type GAA CCT to GAA TCT in the variant creates a new *Hin*fI restriction site. Fragments were separated by electrophoresis through 3% agarose gels containing ethidium bromide (0.001 mg/ml). Controls for the PCR method consisted of known *NQO1* heterozygous DNA, which was included in each assay batch.

***GSTT1*, *GSTM1*, and *MTHFR677* Polymorphisms.** A multiplex PCR reaction was used to characterize these polymorphisms. Primers for *GSTT1* (GenBank accession no. 000853) were 5′CTTACTGGTCCTCACATCTC-3′ (sense), 5′-CAGG-GCATCAGCTTCTGCTTT-3′ (antisense), which amplify a 434-bp fragment of exons 4 and 5 spanning intron 4 (204 bp). These were combined with previously published primers for *GSTM1* (11) and *MTHFR*677 (12). In the multiplex reaction, the final concentrations of primers were as follows: *GSTT1* (0.4 μM); *GSTM1* (0.4 μM); and *MTHFR*677 (0.2 μM). Cycling conditions were as follows: preheat at 94°C for 2 min; 40 cycles

of 94°C for 30 s; 62°C for 30 s; and 72°C for 1 min; with a final extension at 72°C for 7 min.

***MTHFR*1298 Polymorphism.** Primers were designed to amplify a 144-bp fragment of exon 7; sense 5′-GCAAG-TCCCCCAAGGAGG-3′, antisense(intronic) 5′-GGTCCCC-ACTTCCAGCATC-3′ (GenBank accession no. AF105983). Cycling conditions were the same as described above for *NQO1*. Fifteen μl of the PCR product was digested with *Mbo*II (New England Biolabs, Beverly, MA).

Statistical Analysis. All statistical analyses were carried out using SAS (version 6; Ref. 13). The distributions of outcome variables were checked for normality. Student's *t* test was used to compare the mean DNA concentration and total DNA yield of the buccal cell specimens between the 60 children diagnosed with leukemia and the 116 adults diagnosed with lung cancer.

Univariate and multivariate linear regression models were fitted to explore whether DNA concentration or total DNA yield is correlated with one or more of the following variables among the 60 children: duration of storage, *i.e.,* the interval between initial specimen processing and WGA; age of the child; gender of the child; and duration of shipping *i.e.,* the interval between sample collection and initial processing. ANOVA and logistic regression were used to evaluate factors affecting DNA yield and successful PCR reactions. Using the genotyping results from the peripheral blood or bone marrow slides as the gold standard, validity of the buccal cell genotyping with and without WGA was evaluated.

Results

Buccal Swabs from Children Produce Lower DNA Yields than Adults. The patient group studied here is typical of our larger population of childhood leukemia cases, with a median age of 5.1 years [Quartile (Q1): 3.15 years, Q3: 9.1 years] and ranging from 4 months to 15 years. Approximately half of the children were between 2 and 5 years of age. The mean storage time before analysis was 30 months (range, 6–44 months), the median shipping time was 3 days (range, 0–34 days). There were 33 male and 27 female subjects. Forty-two subjects were non-Hispanic white, 11 were Hispanic, and 7 were other race/ethnicities. Buccal specimens were shipped at ambient temperature and stored at −20°C until extraction. Our uniform success with a mail-in protocol for self-administered buccal cell collection in adult lung cancer patients was contrasted with many failed PCR attempts using the children's interviewer-collected samples. The mean DNA concentrations of the buccal cell specimens collected from children ($n = 60$) and adults ($n = 116$) are 5.9 ng/μl (SD, 3.1; median, 6) and 11.6 ng/μl (SD, 10.8; median, 9), respectively. The difference is statistically significant ($P < 0.01$). The mean DNA yield of the children's specimens (2.0 μg; SD, 1.1 μg; median 1.7 μg/brush) is also significantly lower than that of the specimens from adults (mean 4.4 μg; SD, 4.6 μg; median 3.2 μg). Neither duration of storage, duration of shipping, age of the child, gender of the child, nor collection method (interviewer or nurse) was statistically significantly associated with DNA yield from children's samples.

WGA Increases the Number of Successful PCR Reactions from Buccal Cells. The proportion of successful PCR reactions yielding genotypic results using the buccal cell specimens from children with and without the WGA are shown in Table 1. These results reflect the yields of interpretable PCR results when the specimens are run a single time for each of the five polymorphic loci. The PCR results were poorest with the GST-deletion polymorphisms (43% success rate) and somewhat bet-

Table 1 Percentage of PCR reactions yielding genotyping results using children's buccal cell specimens with and without WGA

Gene loci	% amplified without WGA (genotypes)	% amplified with WGA (genotypes)
GSTT1	43.3 (26/60)[a]	95.0 (57/60)[a]
	(deletion, 23; non-deletion, 3)	(deletion, 11; non-deletion, 46)
GSTM1	43.3 (26/60)[a]	95.0 (57/60)[a]
	(deletion, 13; non-deletion, 13)	(deletion, 29; non-deletion, 28)
MTHFR677	63.3 (38/60)[a]	95.0 (57/60)[a]
	(w/w: 12; w/v: 22; v/v: 4)[b]	(w/w: 24; w/v: 29; v/v: 4)[b]
MTHFR1298	70.0 (42/60)	98.3 (59/60)
	(w/w: 21; w/v: 19; v/v: 2)[b]	(w/w: 32; w/v: 23; v/v: 4)[b]
NQO1	63.3 (38/60)	98.3 (59/60)
	(w/w: 23; w/v: 14; v/v: 1)[b]	(w/w: 32; w/v: 24; v/v: 3)[b]

[a] Numbers of successful PCR reactions (numerator)/total number of individual samples tested (denominator).
[b] Distribution of genotypes are shown in parentheses. w/w, wild type/wild type; w/v, wild type/variant; v/v, variant/variant. Variant: *MTHFR* C677T, *MTHFR* A1298C, or *NQO1* C609T.

Table 2 Validity measures for buccal cell genotyping of 5 polymorphisms with and without WGA

	GSTT1	*GSTM1*	*MTHFR677*	*MTHFR1298*	*NQO1*
Without WGA					
Sensitivity	1.00	1.00	0.96	1.00	1.00
	(6/6)[a]	(13/13)	(26/27)	(20/20)	(15/15)
Specificity	0.15	1.00	1.00	0.95	1.00
	(3/20)	(13/13)	(11/11)	(21/22)	(23/23)
PPV	0.26	1.00	1.00	0.95	1.00
	(6/23)	(13/13)	(26/26)	(20/21)	(15/15)
NPV	1.00	1.00	0.92	1.00	1.00
	(3/3)	(13/13)	(11/12)	(21/21)	(23/23)
With WGA					
Sensitivity	1.00	1.00	1.00	1.00	1.00
	(11/11)	(29/29)	(24/24)	(26/26)	(27/27)
Specificity	1.00	1.00	1.00	1.00	1.00
	(46/46)	(28/28)	(23/23)	(33/33)	(32/32)
PPV	1.00	1.00	1.00	1.00	1.00
	(11/11)	(29/29)	(24/24)	(26/26)	(27/27)
NPV	1.00	1.00	1.00	1.00	1.00
	(46/46)	(28/28)	(23/23)	(33/33)	(32/32)

[a] Numbers in parentheses are the numbers used to calculate the corresponding measures. For example, 6/6 for *GSTT1* without WGA means that among the six children who actually have the characteristic (deletion of *GSTT1*), six tested positive.

ter for the bp substitutions (63–70% success rate). Success rates were greater for WGA samples (95–98%); only 1–3 failed PCRs occurred at each locus in the 60 experiments. We repeated these failed samples and were successful in all cases on the second attempt. In contrast, many repeated attempts of the original buccal samples yielded only an ~80% success rate.

Validation of WGA Genotyping Compared with DNA from Whole Blood. Validity was evaluated by comparing the buccal cell genotyping results for children (single PCR attempt) with and without WGA with those results obtained from peripheral blood or bone marrow slides. Individual data from loci that failed to amplify PCR were excluded from this subanalysis. Parameters calculated included sensitivity, specificity, positive predictive value, and negative predictive value. As shown in Table 2, a test was considered positive with the *GSTT1* and *GSTM1* markers when the PCR indicated a "null genotype," the presumptive high risk form of the polymorphism. As for *NQO1* and *MTHFR* markers, any variants (*NQO1* C609T, *MTHFR* C677T, or *MTHFR* A1298C), either homozygous or heterozygous, were considered positive. As indicated in Table 2, there was a systematic misclassification in genotyping without WGA for some loci; very often a positive *GSTT1* test (-positive for deletion) was found nondeleted using blood DNA. Interestingly, the largest PCR fragment amplified in all these experiments (*GSTT1*; 434 bp) was associated with the high false-positive null genotype result. In no instance was a nondeletion found for either *GSTT1* or *GSTM1* using the buccal cells that was later shown to be a deletion in blood or bone marrow analysis.

We observed high positive predictive values when the WGA protocol was applied to buccal cells. The genotyping results obtained using the buccal cells with WGA were concordant in all experiments with results obtained from blood or bone marrow. Therefore, all of the four validity measures equaled unity with WGA.

Discussion

Of the several protocols available for collecting buccal cells, we reasoned that an interviewer-collected cytobrush method would be superior to a mouth rinse for our study of childhood leukemia. Mouthwash typically contains alcohol, and the possibility of swallowing or aspirating these solutions poses a hazard to young subjects and is contraindicated. The limited success of

the cytobrush method in the current pediatric studies often necessitated repeated PCR runs, and a substantial number of brushes (about 20%) never yielded complete genotypic data for all of the five polymorphisms, even with additional laboratory work. We have found that buccal brushes from children consistently yield less DNA compared with brushes taken from adults. To obtain sufficient DNA, an earlier protocol (6) proposed collecting 20 cotton swabs/child and achieved a recovery of 2 μg of DNA/swab, which is identical to our DNA recovery using cytobrushes. Insufficient DNA may be one reason for the higher failure rate of PCR analyses from children.

To address this problem, we have used WGA of the buccal solution, which led to nearly complete PCR amplifications with one PCR run. Interestingly, the failure of PCR using the direct method does not appear random, but was greatest for the *GSTT1* deletion polymorphism. In this PCR the end point is the detection of a 434-bp fragment, which is the largest product of the five genes under study. Failure to detect the *GSTT1* product is evidence of homozygous deletion of the gene. It is likely that the buccal cell DNA is degraded by the action of cellular enzymes and the NaOH preservative, which leads to greater difficulty in amplifying larger PCR products. As an internal control, the *GSTT1* assay uses another product (*i.e., MTHFR*); however, this product is smaller (*i.e.,* 198 bp) and may be amplified relatively easily even with poor template. The result is a differential misclassification of *GSTT1* genotype; amplification failure is scored as deletion leading to a poor predictive value of a negative test. We attempted to find variables associated with poor DNA recovery from buccal brushes, which may be associated with these failures, but were unable to see any obvious relationships. Whereas the WGA procedure increased the efficiency and validity of the genotyping procedure using pediatric buccal cells, it should be cautioned that other types of polymorphisms (*e.g.,* nucleotide repeats) may require additional validation.

A concern in applying WGA is that the resultant population of DNA molecules after WGA may not reflect the allelic distribution of the original cell population. At very low genome copy numbers, it has been shown that preferential amplification

of some alleles can occur (8), leading to misclassification of genotype. This is not a limitation for the buccal protocol we have described here. In their validation study of improved PEP, Dietmaier *et al.* (8) showed that no preferential amplification takes place if the initial WGA reaction includes greater than approximately 10 cell equivalents. We measured the DNA concentration of each of our specimens with a sensitive fluorescent method and determined that we added about 10–60 ng of genomic DNA to the initial WGA reaction; this amount of DNA is equivalent to about 1,600–10,000 cells. Furthermore, we have found that the Hoechst 33258 fluorescence measurement of DNA collected with buccal brushes is equivalent to the genome copy number measured by quantitative PCR.[4] This indicates that, whereas degradation of DNA template is likely to occur during buccal cell processing, the input DNA measured by fluorescence is a reliable indicator of genome copy number. Hence, quality control of the WGA protocol can be maintained by measuring the buccal brush DNA concentration before WGA and insuring that a minimum number of genome copies are added to maintain the distribution of alleles during WGA. The WGA protocol tested here involves considerably longer PCR amplification times compared with conventional PCR and additional costs for reagents. The high-fidelity polymerase is the most costly reagent involved. The potential expansion of PCR reactions made possible with WGA, however, is considerable. If the entire volume of the buccal solution were carried through the WGA protocol, we estimate about 900 PCR reactions could be run on a single cytobrush.

Acknowledgments

References

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

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

3. 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.

4. Watson, J. B., and Grunenwald, H. Comparison between two methods for human buccal cell DNA extraction: a high throughput method *versus* one that generates ultra-pure DNA. Am. J. Hum. Genet., *65* (Suppl.)*:* A412, 1999.

5. Garcia-Closas, M., Egan, K., Alavanja, M., Newcomb, P., Titus-Ernstoff, L., LeMarchand, L., Hayes, R., Brinton, L., and Rothman, N. Collection of genomic DNA from buccal cells by mouthwash and buccal cytobrush in molecular epidemiology studies. San Francisco, CA: American Association for Cancer Research 91st Annual Meeting, 2000.

6. Meulenbelt, I., Droog, S., Trommelen, G. J. M., 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:* 1254–1255, 1995.

7. Hagerman, R. J., Wilson, P., Staley, L. W., Lang, K. A., Fan, T., Uhlhorn, C., Jewell-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.

8. Dietmaier, W., Harmann, A., Wallinger, S., Heinmoller, E., Kerner, T., Endl, E., Jauch, K. W., Gofstadter, F., and Ruschoff, J. Multiple mutation analyses in single tumor cells with improved whole genome amplification. Am. J. Pathol., *154:* 83–95, 1999.

9. Zhang, L., Cui, X., Schnitt, K., Hubert, R., Navidi, W., and Amheim, N. Whole genome amplification from a single cell: implications for genetic analysis. Proc. Natl. Acad. Sci. USA, *89:* 5847–5851, 1992.

10. Wiencke, J. K., Spitz, M. R., McMillan, A., and Kelsey, K. T. Lung cancer in Mexican-Americans and African-Americans is associated with the wild-type genotype of the NAD(P)H:quinone oxidoreductase polymorphism. Cancer Epidemiol. Biomarkers Prev., *6:* 87–92, 1997.

11. Smith, M. C., Kelsey, K. T., Wiencke, J. K., Leyden, K., Levin, S., and Christiani, D. C. Inherited glutathione *S*-transferase deficiency is a risk factor for pulmonary asbestosis. Cancer Epidemiol. Biomarkers Prev., *3:* 471–477, 1994.

12. Frosst, P., Blom, H. J., Milos, R., Goyette, P., Sheppard, C. A., Matthews, R. G., Boers, G. J. H., den Heijer, M., Kluijtmans, L. A. J., van den Heuvel, L. P., and Rozen, R. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat. Genet., *10:* 111–113, 1995.

13. SAS, SAS/STAT User's Guide, version 6, Ed. 4, Vol. 2. Cary, NC: SAS Institute, Inc., 1990.

---

[4] L. Moore and C. Eng, personal communication.

Cancer Epidemiology, Biomarkers & Prevention

Whole Genome Amplification Increases the Efficiency and Validity of Buccal Cell Genotyping in Pediatric Populations

Shichun Zheng, Xiaomei Ma, Patricia A. Buffler, et al.

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/10/6/697

Cited articles	This article cites 11 articles, 4 of which you can access for free at: http://cebp.aacrjournals.org/content/10/6/697.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/10/6/697.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/10/6/697. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.