Behavior Genetics, Vol. 33, No. 1, January 2003 (© 2003)

Brief Communication

DNA from Buccal Swabs Recruited by Mail: Evaluation of Storage Effects on Long-term Stability and Suitability for Multiplex Polymerase Chain Reaction Genotyping

B. Freeman,[1] N. Smith,[1] C. Curtis,[1] L. Huckett,[1] J. Mill,[1] and I. W. Craig[1,2]

Received 30 July 2002—Final 27 Aug. 2002

We provide details of an inexpensive and rapid method for extraction of DNA from buccal swabs (including samples received through the mail) and from a range of other tissue samples. The procedure we have developed provides amounts of DNA adequate for several thousand polymerase chain reactions (PCRs), and we have validated its potential for long-term storage. Samples stored for >4 years are of comparable concentration and provide as robust PCR templates as those tested immediately after extraction. The availability of this technology is of considerable significance in planning DNA banks from population collections and cohorts.

KEY WORDS: DNA extraction; DNA stability; multiplex polymerase chain reaction; sample storage.

INTRODUCTION

New technologies have brought the possibility of genotyping and sequencing of DNA from very large numbers of individuals within the reach of most laboratories and have allowed both linkage and association approaches to be introduced to the routine analysis of multifactorial traits, including behavior (e.g., see Craig and McClay, 2002). Apart from sample collection itself, the most fundamental and potentially rate-limiting step in such population-based molecular studies is that of DNA extraction. We have previously published a simple method for the collection of material for DNA extraction by mail (Freeman *et al.*, 1997). This route lends itself particularly well to studies in behavioral genetics, where the individuals to be recruited are home-based rather than accessible in hospitals and clinics. Furthermore, in some cases, they may also be reluctant to provide blood samples for analysis. The procedure now in routine use in such circumstances is based on the collection of buccal cavity cells obtained by rubbing the inside of the mouth with cotton swabs (such as Q-Tips). Up to 10 of these swabs are placed in storage buffer in a single tube and sent by mail to the laboratory. To complement this convenient sample collection service, we have developed a simple, reliable, and inexpensive method for obtaining DNA, normally sufficient for several thousand PCRs from this material. This procedure that we have developed obviates the need for access to robotic extraction procedures. It comprises a two-step process that can be employed directly on the tissue samples received by mail and that, as will be shown, may be stored for long periods (up to several months) before DNA extraction. Apart from the production of high-purity DNA, suitable for the most demanding contemporary molecular analysis and compatibility with long-term storage, the other criteria employed to evaluate the suitability of the procedures adopted were those of simplicity and cost, such that the process would be compatible with large-scale population screening.

[1] Institute of Psychiatry, SGDP Research Centre, Box P. O. 82, De Crespigny Park, Denmark Hill, London SE5 8AF, United Kingdom.
[2] To whom all correspondence should be addressed at Tel: 0207-848-0017. Fax: 0207-848-0407. e-mail: i.craig@iop.kcl.ac.uk

0001-8244/03/0100-0067/0

MATERIALS AND METHODS

Buccal mucosa cells were collected according to the protocol previously described (Freeman *et al.*, 1997) with minor modifications. In brief, cotton swabs on plastic (previously wooden) sticks are used to remove cells by scraping the inside of the mouth, and the impregnated ends of the swabs are then placed in 15-ml plastic tubes (Sarstedt Ltd.) containing 2.5 mls of collection buffer. The buffer is composed of 100 mM NaCl; 10 mM Tris-HCl, pH 8.0; 10 mM ethylene diaminetetraacetic acid (EDTA) pH 8.0; with 0.2 mg/ml proteinease potassium and 0.5% w/v sodium dodecyl sulphate (Meulenbelt *et al.,* 1995). Tubes are labeled with an adhesive bar-coded waterproof strip, which acts as a unique identifier for the samples throughout the remainder of the procedures. When a suitable number of samples have been collected, the tubes containing the swabs are spun for 4 min at 300 *g* av. to collect all the liquid at the base. The tubes are then incubated in a water bath at 65°C for 2 hours to activate the proteinase potassium, and the tubes are subsequently centrifuged again at 300 *g* av. for 4 min to recover any condensation. The caps are removed from the 15-ml tubes, which are then inverted into larger (50-ml) tubes (Sarstedt Ltd.) and spun at 300 *g* av. for 4 min. The cotton buds remaining in place are then removed and the buds discarded. The original 15-ml tubes are removed and the supernatant decanted into them. Aliquots of 300 μl of an organic deproteinization reagent (ODPR; see Results and Discussion) are added to each tube, which is then capped and shaken vigorously by hand for about 30 s. The denatured debris and remaining organic mix are then compacted by centrifugation at 5000 *g* av for 25 min. Then, with one swift movement, the supernatant from the tube is tipped into a fresh, appropriately labeled, 15-ml tube (Sarstedt Ltd.) and a further 300 μl ODPR added. Mixing and centrifugation is then repeated and the supernatant collected in a fresh tube. Isopropyl alcohol, 1.2 ml, at room temperature is then added; after mixing gently for 1 min, the DNA is collected by centrifugation at 5000 *g* av. for 25 min. The isopropyl alcohol mixture is then decanted, and the pellets gently resuspended in 2 ml cold (4°C) ethanol 70% (v/v) for 10 min before final pelleting at 5000 *g* av. for 10 min. The ethanol wash is discarded, tubes inverted, and pellets left to dry at room temperature for 30 min. The DNA is re-suspended in 400 μl Tris Edta (TE) buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) by mixing overnight (Techne Ltd., UK).

The DNA is then quantified by ultraviolet absorption, using standard settings on a Genequant II spectrophotometer. In addition, a proportion of samples are further examined by removal of 10- to 20-μl aliquots, which are electrophoresed in 1.2% w/v agarose gels, stained with ethidium bromide (1 μg/ml). The fluorescent bands are compared with standards of known concentration using a gel document system (UVP GDS 8000) and a digital camera, which enables the comparison of areas under the peaks recorded for the various bands.

DNA samples are bar coded and stored at −80°C in TE and periodically thawed for dispensing aliquots, dilution, and analysis.

Multiplex PCR Analysis of Simple Sequence Repeat (SSR) Loci

A multiplex assay was developed enabling analysis of a series of unlinked, highly polymorphic microsatellite loci to be genotyped in a single PCR, which includes a combination of locus-specific primers. Initially, this assay was composed of five such markers; however, the number of loci included was increased incrementally to a current methodology including 12 such markers. (See Fig. 1 for a description of loci.) DNA, 75 mg, were amplified in a total reaction volume of 10 μl containing mixed primers (Fig. 1) and deoxyribonucleoside triphosphates (dNTPs). The PCR conditions were 30 cycles (94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1 min) following an initial denaturation at 94°C for 3 min. The reaction was completed with a final extension at 72°C for 10 min. Analysis was carried out employing ABI 310 or 3100 sequencers, with an injection time of 5 s.

Real-Time PCR (TaqMan)

The concentration of available DNA templates was assessed using "real-time" PCR. Quantitative PCR was performed in triplicate for each sample on an ABI Prism 7900HT with TaqMan universal PCR master mix and standard conditions (Applied Biosystems, Foster City, USA) using a predeveloped primer/probe set for ribosomal 18S RNA genes. The data produced were analyzed and converted into threshold cycle values (Ct values) using the computer program "SDS 2.0" (Applied Biosystems, Foster City, USA). Relative template levels in each sample were determined by comparing Ct values. A unitary increase in Ct value represents a requirement for an additional PCR cycle to reach the same threshold fluorescence output and a fold difference in original template concentration.

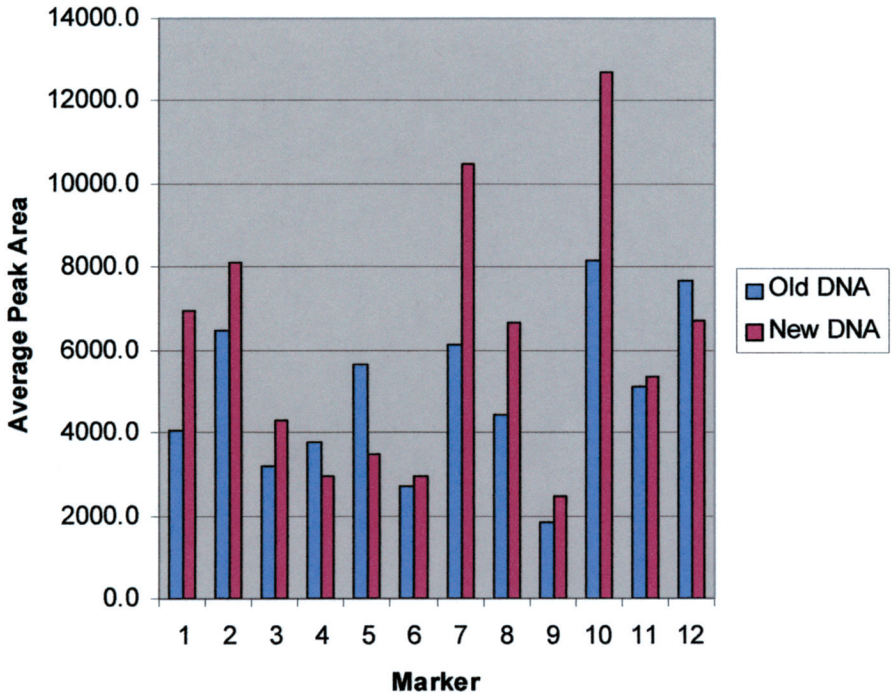

Fig. 1. Ten DNA samples isolated from cheek swabs in 1998 and 10 samples isolated similarly in 2002 were subjected to multiplex polymerase chain reaction amplification employing 12 primer sets in a single tube reaction. Fluorescence outputs for the 12 markers were analyzed (using "Genotyper" software and an ABI 310 instrument). Average peak areas for the alleles displayed at each locus were estimated and the samples compared. The markers employed (1–12) were microsatellites at the following loci, respectively: *CYAR, D14S74, D16S519, D17S798, D18S51, D1S255, D22S264, D3S1300, FABP, FRP2, PLA2A,* and *TH.* All primer sets were modified by the addition of pigtails. [Further information on these loci can be found at http://gdbwww.gdb.org/ and in Kimpton *et al.* (1993).]

RESULTS AND DISCUSSION

Protocol

The procedure described in the Materials and Methods section represents the protocol finally adopted for DNA extraction. After several trials, we selected a commercial product for deproteinization (Yeast reagent 3; Autogen, Bioclear, Ltd., UK) diluted with an equal volume of absolute ethanol. It was found that, apart from the widely ranging quality of starting material inherent in the collection process, the greatest variable affecting recovery was the volume of aqueous extract obtained following the deproteinization steps. With the plastic stemmed buds now routinely adopted (Medical Wire and Equipment Co. Ltd., UK) this ranges from 1.5 to 2.5 ml, with a median value of 2.2 ml. Because the purity and recovery of DNA is dependent on the relative proportions of isopropyl alcohol employed to precipitate DNA from the aqueous mix (the optimal ratio being 0.54:1.0 v/v isopropyl alcohol:mix), a standard amount of isopropyl alcohol was chosen, which was sufficient to ensure precipitation from the entire range of volumes likely to be recovered. This approach was adopted to obviate the need to measure the volume of deproteinized extract recovered for each sample. The volume of isopropyl alcohol originally employed was 0.8 ml, but was later modified to 1.4 ml. Use of the larger volume of isopropyl alcohol appeared not to influence the purity of DNA isolated, as judged by spectrophotometric evaluation. Following the development of the protocol, we evaluated its efficacy with regard to efficiency, the effects of long-term storage on stability and quality of the product, and its applicability to other tissue sources.

Efficiency of Extraction

The overall efficiency of the extraction process was examined by taking known quantities of previously purified DNA through the protocol. Duplicate samples

representing a range of DNA contents (5, 10, 20, 40, and 80 μg) in 2.5-ml collection buffer were subjected to the purification procedure described. The overall recovery from these was 80% [standard deviation (SD) = ±4%], and there was no evidence of significant variation in recovery with starting concentration of DNA. Alternative recovery protocols can result in significant loss of material through loss at aqueous–organic interfaces or through failure to elute from resin columns.

Effects of Long-term Storage on Collected Samples

Here, two aspects are relevant: first, the length of time between collection and extraction, and second, the effects of long-term storage on extracted DNA. First, we evaluated the performance of the procedure on samples collected by mail in the manner previously described (Freeman *et al.*, 1997) and which had been stored for long periods before extraction. Twenty-three samples were selected that had been stored at room temperature for over a period of one year before processing by the standard protocol. The purity of the samples gauged by the optical density ratio of 260 nm/280 nm was 98% to 100%, and the average yield was 53.3 μg (range, 18.5–158). The overall yield exceeded that reported by us previously (Freeman *et al.*, 1997) for samples that had been extracted within one week of receipt by mail and the DNA extracted by a robotically assisted process (AutoGen 740). They are only slightly below the mean value of all samples extracted since the procedure was introduced into the SGDP Research Laboratory (23,141 mouth swab samples extracted with a mean yield of 76 μg).

In addition to ultraviolet analysis, the quality of the DNA was evaluated by examination of its performance in a multiplex PCR. The latter simultaneously amplifies a series of microsatellite loci located on different chromosomes, the products being analyzed employing ABI 310 or 3100 sequencers. At the beginning of the project, the assay contained fewer markers, the total having been increased stepwise during the period of development from five to the current total of 12. The markers were chosen such that, where appropriate, tests of zygosity could be carried out that provide monozygosity ascertainments with predicted confidence ranges of >99.9998 to be achieved when data for all 12 markers are available (manuscript in preparation). Successful analysis was achieved, with 95% of samples giving peak heights in the range of 100–5000 arbitrary units in "Genotyper" displays. The few that failed to amplify were those with very low yields of DNA, suggesting an error in the original collection of buccal swabs undertaken by the individual concerned.

Effects of Long-term Storage on DNA Stability and Quality

We have employed both "real-time" PCR and the zygosity evaluation procedure based on multiplex analysis of 12 SSR loci to establish the template activity of the DNA after long-term storage following extraction. Samples extracted 4 years ago (February 1998) have been compared with those prepared at the time of the experiment (February 2002). Based on the original spectrophotometrically determined concentrations, samples were diluted to 25 ng/μl and template availability estimated by "real-time" PCR on aliquots using primers and probe specific for the *18S RNA* locus employing the ABI 9700HT instrument. Nine samples for each condition were analyzed in triplicate. The average Ct value for the long-term stored samples was 19.53 (SD = 0.96), and for the samples extracted and analyzed contemporaneously it was 20.75 (SD = 1.01). These data suggest that the range of concentrations were very similar between the two groups. Given the difficulties in estimating DNA and diluting samples to standard concentrations, it appears that the long-term stored samples were at least as robust in providing PCR templates as the short-term stored samples. Template availability depends on purity and fragment length, among other factors. It is possible that some degradation has reduced the overall fragment size over long-term storage—possibly enhancing the accessibility of particular genomic regions. However, this has not been detrimental to its usefulness for PCR-based assays.

The same samples were also examined by the multiplex microsatellite PCR, described previously. All gave reliable amplification of all 12 markers, with peak areas of similar magnitude for several markers. Some, however, including *CYAR*, *D22S264*, and *FRP2*, gave better signals for the recently purified samples; *D18S51*, *D1S255* performed better with the older samples (see Fig. 1).

Comparison with Commercial Resin-Based Extraction Kits

In our experience, the yields and purity obtained by the method described here have consistently exceeded those employing commercial kits based on resin-extraction protocols. In a direct comparison, the

liquid from 10 buccal swabs was combined and re-aliquotted. Five were extracted employing our standard protocol and five employing a commercial kit system and the protocol suggested by the manufacturer (Nucleon). The yields were 84.7 μg (SD = 9.7) and 80.6 μg (SD = 15.8), respectively, with a similar range of purity indicated by ultraviolet analysis and agarose gel evaluation (results supplied by Teprel Life Sciences PLC, Wythershave, U.K.).

Application to Other Sample Sources

Although developed for buccal swabs, the method has also proved to be inexpensive and effective for crude lymphocyte preparations from blood. In a comparative test, eight aliquots of 6 ml of blood were taken from a single individual. Four were processed by a resin-kit–based method according to the manufacturer's instructions; four were extracted employing the standard method, following harvesting of the lymphocytes by spinning at 700 *g* and resuspension in 2.5 ml of collection buffer. DNA was dissolved in 1.2 ml of TE. Average yield for the commercial kits was 247 μg (SD = 20.4) and for the standard protocol was 322 μg (SD = 23.8). Further comparisons on blood samples from 10 additional individuals employing ultraviolet analysis and examination on agarose gels indicated that the standard procedure routinely gave similar or better of DNA yield (results not shown). We have also examined DNA prepared from blood by both procedures after long-term storage (3–4 years) at −80°C by multiplex PCR. Again, with very few exceptions, similar average peak areas for the marker alleles were observed for aliquots prepared by either procedure (Fig. 2).

CONCLUSION

We have described an inexpensive, rapid method for extraction of DNA from buccal swabs and from a range of other sources. We have further demonstrated that despite anecdotal concerns about the quantity and stability of DNA prepared from this resource, the procedure we have developed provides amounts of DNA adequate for several thousand PCRs and has long-term storage potential. This is of considerable significance in planning DNA banks from population collections

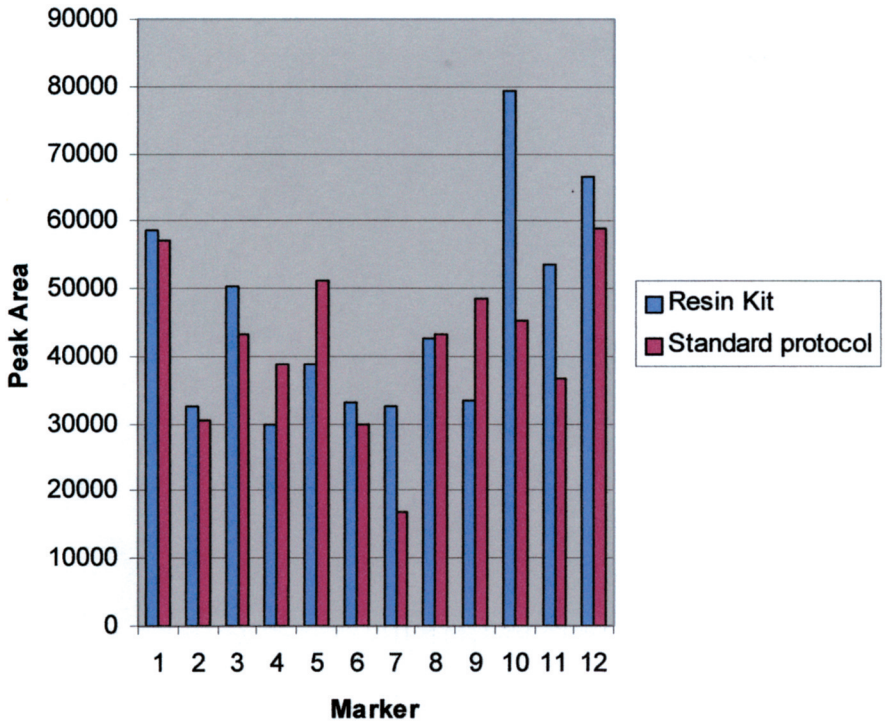

Fig. 2. Ten DNA samples prepared from blood by the standard procedure, or by a commercial resin-based purification kit, were examined after long-term storage (3–4 years) at −80 degrees C, employing multiplex polymerase chain reaction, as described in Fig. 1. Fluorescence outputs for the 12 markers were analyzed (using "Genotyper" software and an ABI 310 instrument, as described in the legend to Fig. 1).

and cohorts. This procedure has been adopted as one of the approaches for DNA extraction from the U. K. 1946 birth cohort. (Swallow and Wadsworth, personal communication).

REFERENCES

Craig, I. W., and McClay, J. (2002). The role of molecular genetics in the postgenomic era. In: Plomin, R., DeFries, J., Craig, I., and McGuffin, P. (eds.), *Behavioral Genetics in the Postgenomic Era*, American Psychological Association, Washington, D.C.

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

Kimpton, C. P., Gill, P., Walton, A., Urquart, A., Millican, E. S., and Adams, M. (1993). Automated DNA profiling employing multiplex amplification of short tandem repeats. In: *PCR Methods and Applications.* Cold Spring Harbour Laboratory Press, New York, USA, pp. 13–12.

Meulenbelt, I., Droog, S., Trommelen, G. J., Boomsma, D. I., and Slagboom, P. E. (1995). High yield non-invasive genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am. J. Hum. Genet.* **57**:1252–1254.

Edited by John Hewitt and Andrew Smollen