Generalised Osteoarthritis: from Mendelian Disorder to Complex Disease

Min, J.L.

CHAPTER

2

High microsatellite and SNP genotyping successrates established in a large number of genomic DNA samples extracted from mouth swabs and genotypes

Josine L. Min, Nico Lakenberg, Margreet Bakker-Verweij, Eka Suchiman, Dorret I. Boomsma, P. Eline Slagboom and Ingrid Meulenbelt

Twin Research and Human Genetics 2006; 9(4):501-506

Abstract

In this paper, we present the genomic DNA yield and the microsatellite and single nucleotide polymorphism (SNP) genotyping success rates of genomic DNA extracted from a large number of mouth swab samples. In total, the median yield and quality was determined in 714 individuals and the success rates in 378,480 genotypings of 915 individuals. The median yield of genomic DNA per mouth swab was 4.1 µg (range 0.1 - 42.2 µg) and was not reduced when mouth swabs were stored for at least 21 months prior to extraction. A maximum of 20 mouth swabs is collected per participant. Mouth swab samples showed in, respectively, 89% for 390 microsatellites and 99% for 24 SNPs a genotyping success rate higher than 75%. A very low success rate of genotyping (0 - 10%) was obtained for 3.2% of the 915 mouth swabs samples using microsatellite markers. Only 0.005% of the mouth swab samples showed a genotyping success rate lower than 75% (range 58 - 71%) using SNPs. Our results show that mouth swabs can be easily collected, stored by our conditions for months prior to DNA extraction and result in high yield and high quality DNA appropriate for genotyping with high success rate including whole genome-wide scans using microsatellites or SNPs.

Introduction

Genomic DNA is commonly extracted from peripheral blood samples for genetic studies of families and populations. Blood sampling, however, is expensive and an invasive procedure to which, for ethical reasons, objections may be raised. Several non-invasive, DNA sampling methods using buccal cells were reported including mouth swabs, cytobrushes or rinses[1-11]. In contrast to blood sampling, involving clinically trained personnel, these self-administered procedures are fast, less expensive and suitable especially for large-scale studies involved in geographically scattered subjects. Hesitations to use genomic DNA extracted by such non-invasive procedures remain among researchers and laboratories, since a large-scale genotyping success rate of genomic DNA extracted from mouth swabs is lacking. In this paper, extensive data is presented on the genotyping success rate of genomic DNA extracted from mouth swabs samples using cotton buds[1]. Moreover, several options to accommodate the collection and storage of these mouth swabs are described based on six years of experience.

Material and Methods

Study populations

For the current study, mouth swab samples from two populations were used. All subjects participating in the study signed informed consents approved by institutional review boards. For the quantity and quality of genomic DNA extracted from mouth swabs, DNA samples collected from additional family members from the Genetics osteoARthritis and Progression (GARP) were used[12]. The GARP study consists of 191 Caucasian sibling pairs of Dutch ancestry with predominantly symptomatic osteoarthritis at multiple sites, and is aimed at identifying determinants of osteoarthritis susceptibility and progression. In addition, family members from each sibling pair were asked to participate by completing a questionnaire and sampling of 20 mouth swabs sent by mail.

The genotyping success rates were calculated for 915 mouth swab samples from subjects participating in the anxious and depression (NETSAD) study. This study comprises a longitudinal study of Dutch adolescent and young adult twins, their parents and their siblings of which questionnaire data were collected on depression, anxiety and correlated personality traits[13]. The data on anxiety and depression, collected over a 10-year period, have been used to select families with sibling pairs who are most informative for linkage and association studies and DNA samples have been collected in these families[13].

DNA collection

Each participant was mailed a sample collection kit containing a tube with cotton buds, a tube with collection buffer, a sampling protocol, informed consent and a prepaid envelope. Subjects took mouth swabs themselves following the sampling protocol. Participants were asked to take series of mouth swabs before a meal, preferably in a mouth without food remains, by rubbing a maximum of five consecutive cotton buds (Antonides C.V.), with a time interval of at least four hours, along the inside of the mouth. After rubbing, the mouth swab was placed in a Falcon tube, containing 0.5 ml of STE buffer (100 mM NaCl, 10 mM Tris and 10 mM EDTA) with proteinase K (0.1 mg/ml) and sodium dodecyl sulphate (0.5%) per mouth swab. After taking the mouth swab sample, participants sent these tubes to the research centre by mail. In the GARP and NETSAD study, 20 and 12 mouth swabs from each individual were collected, respectively.

DNA processing and DNA yield

On arrival at the laboratory, the tube with the mouth swabs was stored at room temperature until genomic DNA extraction. High molecular genomic DNA from mouth swab samples was extracted using a high salt (KAc) precipitation procedure followed by a standard DNA extraction method as previously described[1;14]. All DNA samples were stored at 4 °C in TE (10 mM Tris and 0.1 mM EDTA) buffer. After extraction the genomic DNA, a DNAse activity test was performed by incubating 1 µl of the genomic DNA sample in 10 µl 10 mM Tris pH 8.0 at 37 °C for 16 hours. By electrophoresis on 0.8% agarose gels and visualization with ethidium bromide, the genomic DNA samples were tested for DNA degradation and high molecular weight (> 23 kb). The concentration of the samples was assessed on 0.8% agarose gels by comparison with λDNA quantity standards using Geldoc 2000 (Bio-Rad Laboratories) and Quantity One® software for image analysis (Bio-Rad Laboratories). The total yield of genomic DNA of the participant was calculated by multiplying the DNA concentration with the volume of the extract. The genomic DNA yield per mouth swab was calculated by dividing the total yield of the participant by the total amount of mouth swabs.

For a comparison of extraction of DNA after a long-term and short-term storage period, a random subset of 546 mouth swab samples from the GARP study were selected. Long-term storage of the sample before genomic DNA extraction was tested by comparison of the DNA yield of 259 mouth swab samples extracted after a maximum of 259 days (range 6 - 259 days) with 287 mouth swab samples extracted after a minimum of 303 days (range 303 - 672 days).

Bacterial Content

The percentage of bacterial DNA was estimated in a subset of 30 mouth swab samples from the GARP study and in five blood samples (from laboratory personnel) as negative controls. The total DNA yield (human and nonhuman) was measured with the Picogreen® dsDNA Quantitation Kit (Molecular Probes, Inc). Bacterial DNA yield present in the mouth swab sample was estimated by a semi-quantitative method using real-time polymerase chain reaction (PCR). A region of the 16S rRNA gene, which is known to be conserved across a wide variety of micro organisms and is not found in humans[15] was amplified using a Roche LightCycler™ (Roche) and SYBR Green I dye (Molecular Probes, Inc.). The nucleotide sequences of the primers are as follows: primer 1, 5'-CTACGGGAGGCAGCAG-3'; primer 2, 5'-ATTACCGCGGCTGCTGG-3'[15]. The estimated bacterial DNA content of the mouth swab samples divided by the measured total DNA yield represents the estimated proportion of bacterial DNA in the mouth swab sample.

Genotyping Success Rate

The genotyping success rate of mouth swab samples from the NETSAD study was calculated using 396 microsatellite markers from Screening Set 10, which were genotyped by the NHLBI Mammalian Genotyping Service[16]. A very low genotyping success rate was obtained for six microsatellite markers (< 40%), which were omitted in the further analyses for calculation of the genotyping success rate of mouth swab samples. The remaining 390 microsatellite markers showed a genotyping success rate of at least 82%.

The genotyping success rate of 915 mouth swab DNA samples from the NETSAD study was also calculated for 24 single nucleotide polymorphisms (SNPs). These SNPs were genotyped in five multiplexes by mass spectrometry (the homogeneous Mass ARRAY system; Sequenom, San Diego, CA) using standard conditions. Genotypes were analysed using Genotyper version 3.0 software (Sequenom). For the calculation of the genotyping success rate of the mouth swab samples, 330 'Bad Spectrum' calls, which are due to technical events, were eliminated. Genotyped SNPs showed a success rate of at least 97%. The genotyping success rates of mouth swab samples represent the proportion of successful genotypings of the total amount of genotypings and were calculated without recurrent measurements of mouth swab samples.

Statistical analysis

Differences between short and long-term storage were tested using the Mann-Whitney test. Correlations were calculated with the Spearman's rank correlation coefficient. All analyses were performed with SPSS version 11 software (SPSS, Chicago, IL).

Results

From August 2000 through December 2003, 1063 additional family members from the GARP sibling pair study received a mouth swab collection kit and a questionnaire: 852 subjects agreed to participate in this study, while 211 refused. After two mailings and a follow-up phone call, the total participation rate was 80.2%.

Using the standard extraction protocol, genomic DNA of 714 individuals from the GARP study was extracted by using the mouth swab procedure. High molecular genomic DNA (> 23 kb) was successfully extracted without DNAse activity. The median of the yield of genomic DNA per mouth swab was 4.1 µg (range 0.1 - 42.2 µg). The distribution of the yield per mouth swab of these samples is shown in Figure 1. The median yield per participant (20 mouth swabs) was 78.1 µg (range 5.0 - 843 µg).

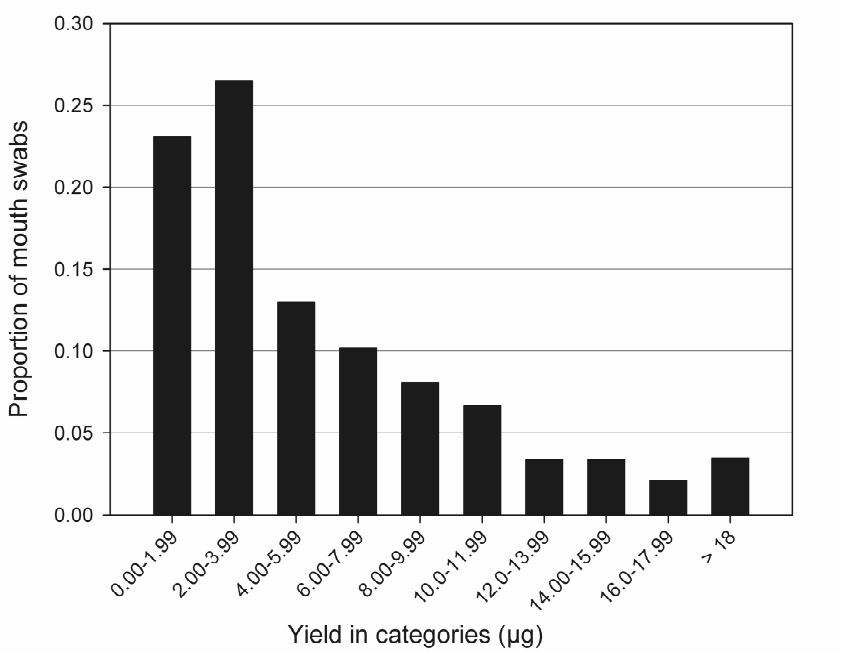

Figure 1 Genomic DNA yield per mouth swab collected from 714 participants from the GARP study.

For laboratories, it might be desirable to collect all samples first and extract genomic DNA from the mouth swabs later without an effect on the DNA yield. Return and extraction date was available of 546 mouth swab samples from the GARP study, which were extracted over two time periods. The median DNA yield of 259 mouth swab samples (median yield per mouth swab = 4.8 µg; range 0.25 - 30.7 µg) extracted after a short-term period (median = 160 days; range 6 - 259 days) provided higher median yields than 287 mouth swab samples (median yield per mouth swab = 3.5 µg; range = 0.10 - 23.9 µg) extracted after a long-term period (median = 570 days; range 303 - 672 days, P = .003). The overall correlation, however, between yield and days was low (r = -.11, P = .01) in all mouth swab samples and both groups provided high-quality DNA without any signs of degradation.

The bacterial DNA yield was estimated for 30 mouth swab samples from the GARP study and for five blood samples as negative controls. As illustrated in Figure 2, a median bacterial DNA yield percentage of 52% (mean ± standard deviation (SD) = 48.1% ± 23.4%) was observed in the mouth swab samples and a median bacterial DNA yield of 0% in the blood samples.

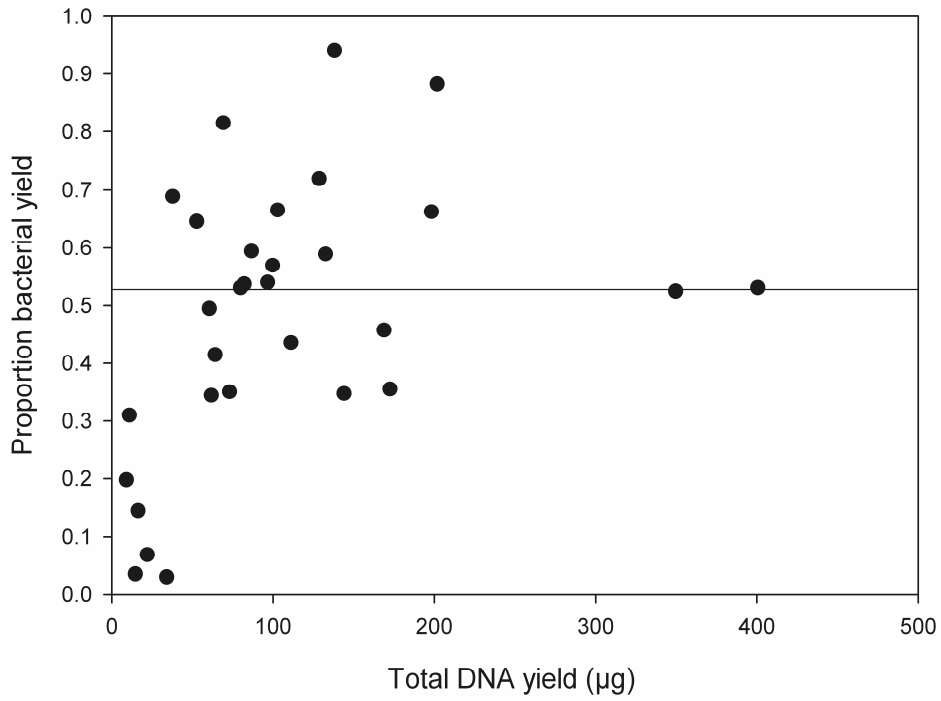

Figure 2 Proportion bacterial yield per sample (20 mouth swabs) from 30 participants from the GARP study. The black line represents median bacterial yield proportion.

Genotyping success rates of mouth swab samples, comprising microsatellite or SNP genotypings, were calculated for DNA samples collected in the NETSAD study. These mouth swab samples were collected, stored and extracted using identical methods as the mouth swab samples from the GARP study in the same laboratory. The genotyping success rate of 915 mouth swab samples, measured for 390 microsatellite markers, was calculated for 356,850 genotypings. The overall genotyping success rate of these mouth swab samples was high with a median success rate of 98.7% (mean ± SD = 90.8% ± 0.21%). A high genotyping success rate (90 - 100%) was observed for a large percentage of mouth swab samples (84.4%). Mouth swab samples showed in 88.9% a genotyping success rate higher than 75%. In contrast, a very low success rate of genotypings (0 - 10%) was observed for 3.2% of the mouth swab samples.

The genotyping success rate of 915 mouth swab samples, comprising 24 SNPs, was computed for 21,630 SNP genotypings. A median success rate of 100% (mean ± SD = 98.8% ± 0.04%) was observed. Only five mouth swab samples (0.005%) showed a success rate lower than 75% (range 58 - 71%). In contrast, the genotyping success rate of microsatellites for these five mouth swab samples was high.

Discussion

Our study demonstrates that mouth swab sampling is a convenient method for collecting high-quality and quantity DNA in large-scale studies and results in high genotyping success rates using microsatellites and SNPs. The median yield of genomic DNA of high molecular weight per mouth swab was 4.1 µg (range = 0.1 - 42.2 µg) and was not substantially reduced in quantity or quality when mouth swabs were stored for up to 21 months prior to extraction. The median yield per participant was 78.1 µg (range 5.0 - 843 µg) which is sufficient to perform genome-wide scans consisting of 225 PCR reactions (400 microsatellite markers). A large variation in DNA yield within our study was observed and may be primarily caused by the difference in pressure exerted during the mouth swab sampling.

The overall genotyping success rate of our mouth swab samples extracted from (n = 378,480 genotypings) was very high for both microsatellites (98.7%) and SNPs (100%). A percentage of 3.2 of the mouth swab samples had a range of success between 0 - 10% for the microsatellites, indicating that once the 'quality' of a mouth swab sample is diminishing the success rate drops easily. There was no correlation between the genotyping success rate and the DNA concentration or DNA yield of the mouth swab samples (data not shown). Genotyping failure of mouth swabs may be caused by protein contamination of the mouth swab sample. Since mouth swab

extraction is cheap and easy to collect it may be worthwhile to recollect or purify DNA samples that fail at early stages in the project. Previously, in three other studies a genotyping success rate of mouth swabs or cytobrushes was determined which were similar to the success rate of our mouth swabs[2;3;17]. However, direct comparison is difficult since these were small-scale studies using only an optimised genotyping assay (n = 60 - 276 genotypings)[2;3;17].

Previous studies have reported that genomic DNA extracted from buccal cells contains a substantial amount of DNA from bacterial origin[3;8]. The percentage of human DNA yields range from 11.5% of the total yield for cytobrushes to 49.5% for mouth washes[3]. In our study, we estimated that mouth swabs samples contained 52% bacterial DNA, which is significantly lower than previously observed in cytobrush samples but consistent with mouth wash samples. Storage in collection buffer (e.g., alcohol-containing mouthwash or proteinase K (0.1 mg/ml) and sodium dodecyl sulphate (0.5%) immediately after taking the sample may, therefore, result in a reduction of the percentage of bacterial yield and inhibit bacterial growth. The percentage of bacterial DNA may, therefore, not be influenced by growth after mouth swab sampling, however, is likely to be a result of bacteria that reside in the mouth which are swabbed together with the buccal cells. In contrast, cytobrushes that are stored dry after wiping, that is, without collection buffer in a tube decrease in the human yield of genomic DNA to 11.5% of the total yield[3]. Moreover, DNAse activity and a reduction in the success rate of the PCR were also observed for our mouth swabs stored dry without a tube for three days[1].

Because a high percentage of bacterial DNA is present in samples collected from buccal cells, measurements of concentrations and yields of human DNA in mouth swab samples might be more accurate using human-specific techniques. In addition, inappropriate quantification due to high bacterial yields might result into genotyping failure. However, the high genotyping success rates in our study obtained with standard quantification measurements suitable for large-scale studies indicate that it is not a major determinant of genotyping failure. Moreover, these methods to determine human yields are relatively laborious, expensive and semi-quantitative.

The mouth swab procedure presented here provides long-term easy storage of the mouth swabs for at least 652 days at room temperature before genomic DNA extraction without substantial loss of quality (high molecular weight and DNA degradation) or quantity of the genomic DNA which were consistent with findings reported by Freeman et al.[2] using a similar mouth swab procedure.

Substantially higher yields are obtained from one mouth wash as compared to one mouth swab sample[3;11;16;18]. In our mouth swab procedure we, therefore, recommend multiple sampling with a minimal (but no maximal) interval of four hours to obtain similar yields as mouth washes. Possibly, such repeated sampling may result in a

decline of the participation rate, which was not observed in our study (participation rate = 80%) suggesting a high level of acceptance. Sample collection was sometimes simplified for our participants by taking fewer samples a day using large cotton buds (two times larger diameter) which have resulted in similar yields (data not shown). In contrast, mouth wash techniques are expensive and inconvenient for infants and elderly since a large amount of specimen has to be produced[11;19].

Apart from cheap and easy sample collection of our mouth swab samples, DNA extraction method is currently a rate-limiting step in the DNA collection process. In the study conducted by Freeman et al.[2], DNA extraction from mouth swabs was performed by an automatic extraction procedure without any signs of a reduction in DNA quality or quantity, which may be a final prerequisite for large-scale population- based DNA collection methods.

Our optimised mouth swab procedure is suitable for large population-based genetic studies in which DNA collection is required of a large number of geographically scattered subjects. Samples can be easily collected by participants themselves, stored for months prior to DNA extraction and results in high human yield and high quality DNA appropriate for genotyping including genome-wide scans using microsatellites or SNPs.

Acknowledgments

We thank all participants of the GARP and NETSAD study for providing mouth swab samples. We further thank Joanne Chartier from DNA Genotek Inc. for estimating bacterial contents in our mouth swab samples.

References

1. 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(5):1252-1254.
2. Freeman B, Smith N, Curtis C, Huckett L, Mill J, Craig IW. DNA from buccal swabs recruited by mail: evaluation of storage effects on long-term stability and suitability for multiplex polymerase chain reaction genotyping. Behav Genet 2003; 33(1):67-72.
3. Garcia-Closas M, Egan KM, Abruzzo J, Newcomb PA, Titus-Ernstoff L, Franklin T et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. Cancer Epidemiol Biomarkers Prev 2001; 10(6):687-696.
4. Lum A, Le Marchand L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. Cancer Epidemiol Biomarkers Prev 1998; 7(8):719-724.

5. 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(5):501-506.
6. Richards B, Skoletsky J, Shuber AP, Balfour R, Stern RC, Dorkin HL et al. Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. Hum Mol Genet 1993; 2(2):159-163.
7. Lench N, Stanier P, Williamson R. Simple non-invasive method to obtain DNA for gene analysis. Lancet 1988; 1(8599):1356-1358.
8. Feigelson HS, Rodriguez C, Robertson AS, Jacobs EJ, Calle EE, Reid YA et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. Cancer Epidemiol Biomarkers Prev 2001; 10(9):1005-1008.
9. Burger MF, Song EY, Schumm JW. Buccal DNA samples for DNA typing: new collection and processing methods. Biotechniques 2005; 39(2):257-261.
10. Steinberg K, Beck J, Nickerson D, Garcia-Closas M, Gallagher M, Caggana M et al. DNA banking for epidemiologic studies: a review of current practices. Epidemiology 2002; 13(3):246-254.
11. King IB, Satia-Abouta J, Thornquist MD, Bigler J, Patterson RE, Kristal AR et al. Buccal cell DNA yield, quality, and collection costs: comparison of methods for large-scale studies. Cancer Epidemiol Biomarkers Prev 2002; 11(10 Pt 1):1130-1133.
12. Riyazi N, Meulenbelt I, Kroon HM, Ronday KH, Hellio le Graverand MP, Rosendaal FR et al. Evidence for familial aggregation of hand, hip, and spine osteoarthritis (OA), but not knee OA in siblings with OA at multiple sites. The GARP study. Ann Rheum Dis 2005; 64(3):438-443.
13. Boomsma DI, Beem AL, van den BM, Dolan CV, Koopmans JR, Vink JM et al. Netherlands twin family study of anxious depression (NETSAD). Twin Res 2000; 3(4):323-334.
14. Beekman M, Lakenberg N, Cherny SS, de Knijff P, Kluft CC, van Ommen GJ et al. A powerful and rapid approach to human genome scanning using small quantities of genomic DNA. Genet Res 2001; 77(2):129-134.
15. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 1993; 59(3):695-700.
16. Yuan B, Vaske D, Weber JL, Beck J, Sheffield VC. Improved set of short-tandem-repeat polymorphisms for screening the human genome. Am J Hum Genet 1997; 60(2):459-460.
17. 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(7):517-520.
18. Cozier YC, Palmer JR, Rosenberg L. Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study. Ann Epidemiol 2004; 14(2):117-122.
19. Saftlas AF, Waldschmidt M, Logsden-Sackett N, Triche E, Field E. Optimizing buccal cell DNA yields in mothers and infants for human leukocyte antigen genotyping. Am J Epidemiol 2004; 160(1):77-84.