ELSEVIER

Cancer Detection and Prevention 31 (2007) 480–488

Cancer Detection and Prevention

www.elsevier.com/locate/cdp

Original article

Whole genome amplification of DNA extracted from hair samples: Potential for use in molecular epidemiologic studies

Suzanne M. Leanza BS[a,*], Robert D. Burk MD[b,c,d], Thomas E. Rohan MD, PhD[d]

[a] *Marion Bessin Liver Research Center, Albert Einstein College of Medicine, Ullmann Building, Room 503, 1300 Morris Park Avenue, Bronx, NY 10461, United States*

[b] *Department of Pediatrics, Albert Einstein College of Medicine, Bronx, NY, United States*

[c] *Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY, United States*

[d] *Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, United States*

Accepted 29 October 2007

Abstract

Background: Because of concerns regarding the quality and quantity of DNA isolated from human hair, such samples are often overlooked as a source of DNA for molecular epidemiological studies. Nevertheless, there are many potential benefits to using hair: it is easily self-collected; it does not require costly collection kits; it can be mailed for a nominal fee; and the hair specimens can be stored at room temperature. However, the amount of DNA that can be extracted from hair samples is somewhat limited. Therefore, we assessed the feasibility of using whole genome amplification (WGA) on genomic DNA extracted from archived human hairs (stored for 7 to 11 years) to increase the quantity of DNA available for genotyping analysis. *Methods*: We evaluated two methods of WGA, multiple displacement amplification and the Genomeplex® method. Both WGA methods were performed on each of 44 DNA samples isolated from archived human hair specimens. The resulting WGA products where then screened for the presence of three single nucleotide polymorphisms. The genotyping results were compared to genotyping data obtained from DNA isolated from mouthwash samples collected from the same individuals. *Results*: When we focused on DNA extracted from the hair root, we observed excellent agreement between the genotypes determined from both the hair (pre-WGA) samples and Genomeplex® WGA when compared to their corresponding mouthwash DNA samples (kappa = 0.83–0.91 and 0.79–0.92, respectively); whereas the agreement between the MDA samples and mouthwash DNA samples was poor (kappa = 0.27–0.51). *Conclusions*: Our data suggest that, when combined with Genomeplex® WGA, hair specimens containing the root portion can serve as a reliable and renewable source of DNA.

Keywords: Genomic DNA; Hair follicle; Specimen collection; Specimen handling; Amplification techniques; Nucleic acid; Phage phi 29; Genotype; Polymorphism; Single nucleotide; Epidemiology; Molecular biology

1. Introduction

Many molecular epidemiological studies are currently employing large SNP panels to identify disease-related genes [1]. In order for these studies to be successful, it is usually important to obtain high quality genomic DNA (gDNA) from a large number of individuals. The high cost of such studies requires that the collection and storage of the DNA samples be cost-effective. Although blood and buccal (brush or mouthwash) DNA have been used extensively in epidemiological studies because they are excellent sources of gDNA [2], the collection, processing, and storage of these types of samples are relatively costly [3]. Given these considerations, samples that can be collected inexpensively, non-invasively, and easily, represent an appealing option for large-scale epidemiological studies.

Cephalic hair DNA is frequently used in forensics, archeology, paternity testing, and wildlife studies, but seldom, if at all, in molecular epidemiologic studies. In general, the quantity and quality of DNA acquired from hair is limited, especially when stored for long periods of time [4,5].

* Corresponding author.
E-mail address: semerlin@aecom.yu.edu (S.M. Leanza).

0361-090X/$30.00
doi:10.1016/j.cdp.2007.10.009

However, when hair is plucked from the scalp, the probability of collecting a sample with remnants of the root portion intact is relatively high because at any given time it is estimated that 75–95% of hair follicles are in the anagen or active growth phase [6]. It has been reported that a single, freshly plucked hair root can yield anywhere from 50 to 1500 ng of DNA [7]. When DNA extraction is combined with whole genome amplification (WGA), it is possible to generate enough gDNA to perform virtually unlimited genotyping analysis per sample.

Currently, there are several WGA methods, all of which generate substantial quantities of gDNA from a minute amount of starting DNA template. Although some of the early methods of WGA, such as primer extension PCR (PEP) and degenerate oligonucleotide-primed PCR (DOP-PCR), were reported to result in preferential amplification, incomplete genome coverage, and the generation of relatively short fragments [8], two relatively new methods of WGA, multiple displacement amplification (MDA) and the Genomeplex® technique developed by Rubicon Genomics, have been shown to address some of the aforementioned concerns [9–13].

The MDA method utilizes ϕ29 DNA polymerase and a mix of random hexamer primers to isothermally replicate the genome. The Genomeplex® method begins with random fragmentation of the genome into short overlapping fragments followed by the ligation of adaptors to each fragment, and finally PCR amplification using primers complementary to the ligated adaptors. Recently, both methods have been reported to successfully amplify high quality genomic DNA with high fidelity and ample genome coverage [9–13]. However, the utility of these methods with respect to DNA sources of less than optimal quality has received little attention.

As part of an ongoing cohort study of diet, lifestyle, molecular markers and cancer risk, hair specimens were collected from the study participants. Given concerns about the quality and quantity of DNA that might be available from the stored hair samples, we assessed the feasibility of using WGA on gDNA extracted from the archived human hair to increase the quantity of gDNA available for molecular epidemiological studies. To this end, we evaluated two WGA methods, Genomiphi®, an MDA based method, and the Genomeplex® method. One major concern when using WGA on low copy number or highly degraded DNA samples is the preferential amplification of one allele leading to a false homozygous genotype assignment in samples that are known to be heterozygous (amplification bias or allele bias) [14]. Therefore, to analyze the fidelity of the two WGA methods evaluated, we specifically selected three single nucleotide polymorphisms (SNPs) with heterozygosities of ∼0.40.

2. Materials and methods

2.1. Study subjects

The feasibility study was conducted using hair samples provided by participants in the Canadian Study of Diet, Lifestyle, and Health, a prospective study of 73,909 individuals (34,291 males and 39,618 females) who were recruited mostly between 1992 and 1998 [15]. At recruitment, participants completed diet and lifestyle questionnaires, provided clippings from each toenail, and provided a hair sample. For the study described here, we retrieved from storage the hair samples of 44 randomly selected subjects from whom buccal cell samples had been collected in an earlier pilot study [13].

2.2. Sample collection and DNA extraction

At the time of their recruitment into the cohort study, the participants were given detailed instructions on how to collect the hair sample. Specifically, they were asked to pull out about six strands of hair, including the hair roots, either from the back of the scalp or from other parts of the body. The samples[1] were returned to the study-coordinating center (together with the completed questionnaires) and stored in paper coin envelopes at room temperature in the dark. Collection of the mouthwash samples was described previously [13].

For the purpose of DNA extraction from the hair samples, approximately 1 cm of the root end of each of the four to six hairs was cut with a new sterile blade and placed inside a 1.7 mL microcentrifuge tube containing 500 μL of a lysis buffer (cell lysis buffer: 10 mM Tris/HCl pH 8.0, 10 mM EDTA pH 8.0, 0.1 M NaCl, 2% SDS) and 20 μL (20 mg/mL) Proteinase K (Roche, Indianapolis, IN), incubated for 2 h at 55 °C and followed by phenol/chloroform extraction according to the Phase lock Gel Tube protocol. For the remaining hair samples that appeared to have been cut with scissors rather than plucked, the QIAamp DNA Micro Kit (Qiagen, Valencia, CA) was used following the manufacturer's protocol for hair shaft samples using approximately 20 cm of hair shaft cut into 1cm sections.

For some of the buccal cell samples, DNA had already been extracted, as described previously by Thompson et al. [13]. For the remainder, the mouthwash samples were subjected to proteinase K digestion followed by phenol/chloroform extraction according to the Phase lock Gel Tube protocol (Eppendorf, Westbury, NY). The DNA content of all of the samples were quantified using the PicoGreen assay (Molecular Probes, Eugene, OR) and by UV Spectrophotometry (Nanodrop, Wilmington, DE).

2.3. DNA integrity

To assess the integrity of the mouthwash and archived hair DNA, 2 μL of each sample was electrophoresed on an

[1] The sample set chosen should ideally contain three to five samples of hair each of four to five hairs of each of (i) various shades, and (ii) various types, i.e. plain, wavy, curly, frizzy, etc., as well as samples from various racial background. Although, the samples were not chosen to be representative of different hair types, many of the aforementioned textures and a variety of shades were present in the sampling.