Institut Pasteur
Microbes and Infection 18 (2016) 399–405
www.elsevier.com/locate/micinf

Original article

The salivary microbiome for differentiating individuals: proof of principle

Sarah L. Leake [a], Marco Pagni [b], Laurent Falquet [b,c], Franco Taroni [a,1], Gilbert Greub [d,*,1]

[a] School of Criminal Justice, University of Lausanne, Lausanne, Switzerland
[b] Swiss Institute of Bioinformatics, Vital-IT Group, Lausanne, Switzerland
[c] Department of Biology, University of Fribourg, Fribourg, Switzerland
[d] Institute of Microbiology, Lausanne, Switzerland

Received 17 March 2016; accepted 25 March 2016
Available online 8 April 2016

Abstract

Human identification has played a prominent role in forensic science for the past two decades. Identification based on unique genetic traits is driving the field. However, this may have limitations, for instance, for twins. Moreover, high-throughput sequencing techniques are now available and may provide a high amount of data likely useful in forensic science.

This study investigates the potential for bacteria found in the salivary microbiome to be used to differentiate individuals. Two different targets (16S rRNA and *rpoB*) were chosen to maximise coverage of the salivary microbiome and when combined, they increase the power of differentiation (identification). Paired-end Illumina high-throughput sequencing was used to analyse the bacterial composition of saliva from two different people at four different time points (t = 0 and t = 28 days and then one year later at t = 0 and t = 28 days). Five major phyla dominate the samples: Firmicutes, Proteobacteria, Actinobacteria, Bacteroidetes and Fusobacteria. *Streptococcus*, a Firmicutes, is one of the most abundant aerobic genera found in saliva and targeting *Streptococcus rpoB* has enabled a deeper characterisation of the different streptococci species, which cannot be differentiated using 16S rRNA alone. We have observed that samples from the same person group together regardless of time of sampling. The results indicate that it is possible to distinguish two people using the bacterial microbiota present in their saliva.

Keywords: Saliva microbiota; Forensic science; Illumina sequencing; Human identification

1. Introduction

Current methods of human identification in forensic science rely heavily upon the analysis of human DNA. However, there are limitations to the use of human DNA namely its degradation and low quantity. For example, in sexual assault cases, the DNA from the perpetrator is often masked by the DNA of the victim making identification difficult. In such cases saliva is commonly found due to it being transferred through, amongst others, biting, kissing and licking. To overcome the current unsatisfactory situation, the potential of other targets, for example bacteria, needs to be investigated. Why is bacterial DNA interesting in this context? Firstly, bacterial DNA is better protected than human DNA and more resistant to degradation. Therefore, bacterial DNA will persist better once deposited on a surface. Secondly, it may be possible to distinguish twins using bacterial DNA [1], a feat impossible with current human DNA based methods.

It has been estimated that 99% of bacteria found in the environment cannot be cultured [2]. However, with the arrival of next generation sequencing (NGS) the analysis of bacterial community composition has reached depths previously unachievable. There is now potential to exploit bacteria for forensic purposes. Fierer et al. demonstrated that the analysis of the skin microbiome could be used to link an individual to an object they touched and that the bacterial community found

* Corresponding author. Institute of Microbiology, University of Lausanne, Bugnon 48, 1011 Lausanne, Switzerland. Tel.: +41 21 314 49 79; fax: +41 21 314 40 60.
E-mail address: gilbert.greub@chuv.ch (G. Greub).
[1] Both authors contributed equally to this work.

http://dx.doi.org/10.1016/j.micinf.2016.03.011

on the object was more similar to the community on the owners hand than to 270 other hands, indicating the potential of this technique for forensic identification [3]. This study extends the idea presented by Fierer et al. by demonstrating the potential of NGS analysis of the salivary microbiota for forensic identification.

A number of studies showing saliva bacterial community composition using NGS have been published [1,4–9]. To date the main gene targeted is 16S rRNA because it is ubiquitous and essential for bacterial life [10,11]. However, there are limitations to targeting 16S rRNA namely, intra-genomic heterogeneity, mosaicism and the lack of a universal threshold sequence identity value [12]. Therefore, in order to have a more complete picture of a microbiome, analysing a second (single-copy) target is essential. In this study the second gene targeted was *rpoB* which, encodes the beta-subunit of RNA polymerase, a very important enzyme that is highly conserved throughout bacteria. It has been shown that like the 16S rRNA gene the *rpoB* gene contains alternating variable and conserved regions [13]. The hypervariable regions of rpoB have shown promise for bacterial identification down to the species and subspecies levels [14–16]. Specifically studies have shown that humans have many different strains of the same \textit{Streptococcus} species, the most prevalent genus in saliva, with many strains being unique to individuals [17,18]. Using 16S rRNA alone these strains would not be detected and therefore an important part of the salivary microbiome would be missed out. By combining *rpoB* with 16S rRNA a deeper level of identification is possible.

Saliva unlike sperm and blood, the other main biological fluids found in criminal cases, is not sterile. Indeed, saliva contains, as many as 500 million bacterial cells per millilitre (ml) and at least 700 different bacterial species [19]. The average composition of the salivary microbiome being known [1,8], we wondered whether there is enough variation to differentiate salivary microbiomes of two different people. To date, studies have shown that differences in salivary microbial communities between individuals are present [5,20], however whether these differences are great enough to differentiate individuals has yet to be explored. Additionally, the salivary microbiome has been shown to be stable over a couple of months [5,8] but no longer, however studies on gut microbiota show stability over a few years [21,22], further work is required to see if this pattern is observed in saliva microbiota. Thus, this study investigates the intra and inter-individual variation of the salivary microbiome of two healthy subjects to investigate the potential of saliva microbiota in forensic science.

2. Materials and methods

2.1. Sampling and DNA extraction

This study was approved by the Ethics Committee of the Canton of Vaud, Switzerland (protocol 357/11). Saliva samples were obtained from two healthy adult individuals at four time points; t = 0 and t = 30 days and one year later at t = 0 and t = 30, with informed consent. Volunteers were asked to brush their teeth in the morning and not eat or drink 1 h before sampling. The saliva was collected by spitting into a sterile tube and then stored at −20 °C until processing. DNA extraction was performed using the automated MagNA Pure 96 DNA and Viral Nucleic Acid small volume kit (Roche) following the Pathogen Universal 200 v2.0 protocol [23]. Samples were then stored at −20 °C.

2.2. PCR and sequencing

In order to maximise coverage of the salivary microbiome, two different targets were chosen; 16S rRNA and *rpoB*. Practically two different pairs of primers targeting *rpoB* were used to investigate the biodiversity of streptococci (*rpoB*1) and other bacteria (*rpoB*2). For 16S rRNA, primers were designed to amplify the V5 region and for *rpoB*, two sets of primers covered the V1 region. Primers were designed using general target species then checked against species known to be found in saliva (see table Table 1 for final primer sequences). Each target was amplified separately in a reaction containing 5 µl of DNA extract, 0.5 µM of both forward and reverse primer, 1× Phusion® HF buffer, 200 µM each dNTP, 0.02 U/µl Phusion® Hot Start II DNA polymerase, 3% DMSO and 1 mM $MgCl_2$ in a total volume of 50 µl. The following thermal cycling parameters were used: initial denaturation at 98 °C for 30 s, 35 cycles of denaturation at 98 °C for 5 s, primer dependant annealing temperature (see Table 1 for annealing temperatures) for 15 s and extension at 72 °C for 10 s with a final extension of 5 min at 72 °C.

All amplified targets from the same sample were pooled together and the pooled sample barcoded. To pool samples equal molar amounts of each sample are necessary, in this case approximately 10 pmol of each were used. The samples were then purified using Agencourt AMPure XP PCR purification (Beckman Coulter). The purified products were then separated on an agarose gel and the band corresponding to the target size (120 bp) excised. Finally, the sequencing libraries were prepared using the TruSeq DNA sample preparation kit (Illumina) [24]. Then, 100 cycles of paired-end sequencing were performed on a HiSeq 2000 (Illumina).

2.3. Sequence analysis

Base-calling was performed by HCS 2.0.12/RTA 1.17.21.3 and quality control by the CASAVA 1.8.2 pipeline using

Table 1
Primers designed for each gene target. Primer name for 16S rRNA and *rpoB*2 corresponds to the *Escherichia coli* positions and for *rpoB*1 to the *Streptococcus bovis* positions.

Gene	Primer name	Primer sequence (5′-3′)	Tm (°C)
16S rRNA	792 F	AGGATTAGATACCCTGGTAG	56
	891R	CGTACTCCCCAGGCGG	
*rpoB*1	130F	GGACCTGGTGGTTTGAC	64
	220R	CGATGTTAGGTCCTTCAGG	
rpoB2	340F	GGACCAGAACAACCCG	60
	434R	GGGTGTCCGTCTCGAAC	