Forensic Science International: Genetics 14 (2015) 18–30

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString® barcode probes to specific mRNA targets

Patrick Danaher [a], Robin Lynn White [a], Erin K. Hanson [b], Jack Ballantyne [b,c,*]

[a] NanoString Technologies, 530 Fairview Avenue N, Suite 2000, Seattle, WA 98109, USA
[b] National Center for Forensic Science, P.O. Box 162367, Orlando, FL 32816-2367, USA
[c] Department of Chemistry, University of Central Florida, P.O. Box 162366, Orlando, FL 32816-2366, USA

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form 27 August 2014
Accepted 7 September 2014

Keywords:
Forensic science
Body fluid identification
mRNA profiling
Gene expression
NanoString®
De-convolution of body fluid mixtures

ABSTRACT

A DNA profile from the perpetrator does not reveal, *per se*, the circumstances by which it was transferred. Body fluid identification by mRNA profiling may allow extraction of contextual 'activity level' information from forensic samples. Here we describe the development of a prototype multiplex digital gene expression (DGE) method for forensic body fluid/tissue identification based upon solution hybridization of color-coded NanoString® probes to 23 mRNA targets. The method identifies peripheral blood, semen, saliva, vaginal secretions, menstrual blood and skin. We showed that a simple 5 min room temperature cellular lysis protocol gave equivalent results to standard RNA isolation from the same source material, greatly enhancing the ease-of-use of this method in forensic sample processing.

We first describe a model for gene expression in a sample from a single body fluid and then extend that model to mixtures of body fluids. We then describe calculation of maximum likelihood estimates (MLEs) of body fluid quantities in a sample, and we describe the use of likelihood ratios to test for the presence of each body fluid in a sample. Known single source samples of blood, semen, vaginal secretions, menstrual blood and skin all demonstrated the expected tissue-specific gene expression for at least two of the chosen biomarkers. Saliva samples were more problematic, with their previously identified characteristic genes exhibiting poor specificity. Nonetheless the most specific saliva biomarker, HTN3, was expressed at a higher level in saliva than in any of the other tissues.

Crucially, our algorithm produced zero false positives across this study's 89 unique samples. As a preliminary indication of the ability of the method to discern admixtures of body fluids, five mixtures were prepared. The identities of the component fluids were evident from the gene expression profiles of four of the five mixtures. Further optimization of the biomarker 'CodeSet' will be required before it can be used in casework, particularly with respect to increasing the signal-to-noise ratio of the saliva biomarkers. With suitable modifications, this simplified protocol with minimal hands on requirement should facilitate routine use of mRNA profiling in casework laboratories.

1. Introduction

Genetic identification of the donor of transferred biological traces deposited at the crime scene or on a person using STR analysis is now routine practice worldwide [1]. This represents potentially crucial 'source level' information for investigators [2]. A DNA profile from the perpetrator does not however reveal the circumstances by which it was transferred. This contextual information (sometimes known as the 'activity level' in Cook and Evett's classic 1998 paper [2]) is important for casework investigations because the deposition of the perpetrator's biological material requires some behavioral activity that results in its transfer from the body. The consequences of different modes of transfer of the DNA profile may dramatically affect the investigation

* Corresponding author at: Department of Chemistry, National Center for Forensic Science, University of Central Florida, P.O. Box 162367, Orlando, FL 32816-2367, USA. Tel.: +1 407 823 4041; fax: +1 407 823 4042.
E-mail address: Jack.Ballantyne@ucf.edu (J. Ballantyne).

http://dx.doi.org/10.1016/j.fsigen.2014.09.005
1872-4973/

and prosecution of the crime. For example a DNA profile from a victim that originates from skin versus the same DNA profile that originates from vaginal secretions may support social or sexual contact, respectively. Thus tissue/body fluid sourcing of the DNA profile should be an important concern for, and service from, forensic genetics practitioners who are integral to the investigative team. The problem is that with the routinely used biochemical or immunological testing, it was not possible to definitively identify many of the important body fluids of interest (e.g. vaginal secretions, saliva, and menstrual blood).

In order to overcome the limitations of currently used classical body fluid identification methods, the use of messenger RNA (mRNA) profiling, as described by Juusola and Ballantyne [3], was proposed to supplant conventional methods for body fluid identification. Terminally differentiated cells, whether they comprise blood monocytes or lymphocytes, ejaculated spermatozoa, epithelial cells lining the oral cavity or epidermal cells from the skin become such during a developmentally regulated program in which certain genes are turned off (i.e. transcriptionally silent) and turned on (i.e. are actively transcribed and translated into protein) [4]. Thus, a pattern of gene expression is produced that is unique to each cell type in both the presence and the relative abundance of specific mRNAs [4]. The type and abundance of mRNAs, if determined, would then permit a definitive identification of the body fluid or tissue origin of forensic samples. This is the basis for mRNA profiling for body fluid identification. RNA profiling now offers the ability to identify all forensically relevant biological fluids using methods compatible with the current DNA analysis pipeline [5,6]. Despite the identification of numerous body fluid specific candidates there is some reluctance to utilize RNA profiling assays in the forensic community due to concerns over the perceived instability of RNA in biological samples. However, several studies have been conducted in order to assess the stability of RNA in dried forensic stains [7–10]. These have demonstrated that RNA of sufficient quantity and quality for analysis can be recovered from aged and environmentally compromised forensic samples [7–10]. The effective stability (i.e. 'recoverability') of mRNA in aged and compromised samples is not dissimilar to that of DNA and provides support to the use of mRNA profiling assays in forensic casework (Ballantyne, unpublished observations). The recently published EDNAP collaborative exercises on mRNA profiling for body fluid identification further demonstrate a significant interest in mRNA profiling by the forensic community in Europe and around the world as well as the ease in which this technology can be implemented into forensic casework laboratories [11–15]. Collectively, these studies demonstrate an interest in the use of mRNA profiling in forensic casework and its suitability of use with forensic samples and therefore warrant continued evaluation and development. Other classes of RNA also exist in the cell and one in particular, microRNA (miRNA), has been investigated for potential forensic use since the short size of the molecule (~21–25 bases) makes it an attractive option for analyzing degraded specimens [16–22]. The field of forensic miRNA profiling, although promising, is less mature in terms of there being an international consensus on the identity and specificity of the best body fluid specific miRNA targets. Other non-RNA methods for body fluid identification have been recently investigated including the use of epigenetic [23–29] and proteomic [30–32] biomarkers. Although exhibiting some promise, epigenetic markers have not been identified for all of the important common body fluids and tissues such as vaginal secretions and skin. Proteomic markers suffer from a lack of demonstrated reproducibility studies among different laboratories, and paucity of peer reviewed reports demonstrating their forensic validity.

Gene expression differences are quantitative in nature meaning that a particular biomarker may be expressed in a particular cell type at low, intermediate or high levels. Even when it is not generally regarded as being expressed in a particular cell type it may exhibit basal level (or 'leaky') transcription with a few molecules present per cell. Thus far there have been three main methods developed for mRNA profiling of forensic samples: capillary electrophoresis (CE)-based analysis [5–7,33–41], quantitative RT-PCR (qRT-PCR) [7,42–44] and, more recently, high resolution melt (HRM) analysis [45]. Due to its facile multiplex capabilities and routine use in DNA profiling, CE-based analysis has been the platform of choice for casework mRNA assays [5,6,38,41]. However post-PCR CE peak heights/areas are, at best, semi-quantitative in nature with respect to biomarker expression levels. Similarly, HRM signal amplitude does not appear to correlate precisely with RNA input [45]. Although qRT-PCR permits quantitation of biomarker targets, its low multiplex capability (typically 3–4 targets maximum compared to >20 for CE) appears to have limited its use.

In contrast to the aforementioned, digital gene expression (DGE) methods precisely count the number of individual transcripts in a sample [46] which facilitates the use of advanced statistical methods to better evaluate and interpret the experimental data. This facility would be expected to be of significant benefit when analyzing body fluid mixtures that are commonly encountered in forensic analysis. Deep sequencing of the transcriptome using next generation sequencing (NGS) technologies is capable of directly identifying and quantifying (by counting) all mRNA transcripts in a sample, a DGE technique known as RNA sequencing (RNA-Seq) [47]. RNA-Seq has been spectacularly successful in advancing our knowledge of cell-type-specific gene expression including transcript quantification and elucidation of their sequence diversity [47]. Although NGS heralds a new era of forensic genomics, impediments to its routine implementation in body fluid RNA analysis include its high cost of reagents and time-consuming, complex analysis. In this work we sought an alternative DGE method to NGS that is simpler and requires minimal hands-on experimentation. Here we describe the development of a prototype multiplex DGE method for forensic body fluid identification based upon solution hybridization of color-coded NanoString® probes [48] to 23 tissue/body fluid specific and 10 housekeeping gene mRNA targets present in forensic type samples. We describe calculation of maximum likelihood estimates (MLEs) of body fluid quantities in a sample, and we describe the use of likelihood ratios (LR) to test for the presence of each body fluid in a sample. Concomitantly, to facilitate routine use, we also devised a simple 5 min room temperature cellular lysis protocol as an alternative to standard RNA isolation for forensic sample processing.

2. Methods

2.1. Body fluid samples

Body fluids were collected from volunteers using procedures approved by the University's Institutional Review Board. Informed written consent was obtained from each donor. Blood samples were collected by venipuncture into vacutainers (K3-EDTA preservative) and 50 μl aliquots were placed onto cotton cloth and dried at room temperature. Freshly ejaculated semen was provided in sealed plastic tubes and stored frozen. After thawing, the semen was absorbed onto sterile cotton swabs and allowed to dry. Buccal samples (saliva) were collected from donors using sterile swabs by swabbing the inside of the donor's mouth. Semen-free vaginal secretions and menstrual blood were collected using sterile cotton swabs. Admixed body fluid samples were created by combining ½ of a 50 μl stain or single cotton swab from each body fluid. Environmental samples were prepared by exposing body