Forensic Science International: Genetics 31 (2017) 67–80

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Epigenetic discrimination of identical twins from blood under the forensic scenario

Athina Vidaki[a], Celia Díez López[a], Elena Carnero-Montoro[b], Arwin Ralf[a], Kirsten Ward[b], Timothy Spector[b], Jordana T. Bell[b,1], Manfred Kayser[a,*,1]

[a] Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands
[b] Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

ARTICLE INFO

Article history:
Received 26 April 2017
Received in revised form 15 July 2017
Accepted 28 July 2017
Available online 18 August 2017

Keywords:
Monozygotic twins
Forensic epigenetics
DNA methylation
Illumina 450K array
Quantitative PCR
Whole blood

ABSTRACT

Monozygotic (MZ) twins share the same STR profile, demonstrating a practical problem in forensic casework. DNA methylation has provided a suitable resource for MZ twin differentiation; however, studies addressing the forensic feasibility are lacking. Here, we investigated epigenetic MZ twin differentiation from blood under the forensic scenario comprising i) the discovery of candidate markers in reference-type blood DNA via genome-wide analysis, ii) the technical validation of candidate markers in reference-type blood DNA using a suitable targeted method, and iii) the analysis of the validated markers in trace-type DNA. Genome-wide methylation analysis in blood DNA from 10 MZ twin pairs resulted in 19–111 twin-differentially methylated sites (tDMSs) per pair with >0.3 twin-to-twin differences. Considering all top three candidate tDMSs across all pairs in the technical validation based on methylation-specific qPCR, 67.85% generated >0.1 twin-to-twin differences. Of the validated tDMSs, 68.4% showed >0.1 twin-to-twin differences with qPCR in trace-type DNA across 8 pairs. Using an updated marker selection strategy, 8 additional candidate tDMSs were obtained for an example MZ pair, of which 7 showed >0.1 twin-to-twin differences in both reference- and trace-type DNA. Lastly, we introduce a high-resolution melting curve analysis of the entire fragment that can complement the proposed approach. Overall, our study demonstrates the general feasibility of epigenetic twin differentiation in the forensic context and highlights that the number of informative tDMSs in the final trace DNA analysis is crucial, as some candidate markers identified in reference DNA were shown not informative in the trace DNA due to various, including technical, reasons. Future studies will need to address the optimal number of epigenetic markers required for reliable identification of MZ twin individuals including statistical considerations.

1. Introduction

Individual identification of monozygotic (MZ) twins in either criminal or paternity casework has always been challenging. Since MZ twins are derived from the same zygote, they share a literally identical genomic DNA sequence. Except for very rare cases [1], conventional DNA profiling techniques using short tandem repeat (STR) genetic markers typically fail to distinguish between individuals within the same MZ twin pair. This inevitably leads to a dead-end in police investigation and may result in court decisions of prime suspects being set free, which highlights the need for a suitable approach to be applied in such cases.

There is increasing evidence that MZ twins can demonstrate a very small degree of genetic differences, in terms of single nucleotide polymorphisms (SNPs) [2], copy number variations (CNVs) [3–5], Y-chromosome satellite DNA [6], and bone-marrow-derived memory B lymphocytes DNA sequences [7]. However, these sequence differences are extremely rare and difficult to locate. Furthermore, their detection in forensic-type samples might be impractical in case that they are seen only in a small portion of cells (mosaicism). Proposed genome-wide screening methods, such as ultra-deep whole genome sequencing [2], require DNA of high quality and quantity and lead to high costs.

Contrary to the stable genomic DNA sequences, epigenetic patterns, especially DNA methylation, are more dynamic as they

* Corresponding author.
E-mail address: m.kayser@erasmusmc.nl (M. Kayser).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.fsigen.2017.07.014
1872-4973/

are influenced by genetic, environmental and stochastic factors, the latter two varying throughout an individual's life [8,9]. DNA methylation differences between both phenotypically concordant and discordant MZ twins have been studied previously, not only to investigate naturally occurring variation in DNA methylation and epigenetic heritability [10–14], but also to unravel the impact of epigenetics in development and ageing [15,16] as well as disease [17–20]. The observed epigenetic drift within MZ twin pairs can be seen at a genome-wide level [12] and also at specific loci [21,22], which are most likely linked with gene regulation involved in specific phenotypes.

Although the extent of these differences, and to what degree they are shared by different MZ twin pairs, is yet not understood, the potential of using DNA methylation to distinguish between MZ twins for forensic purposes has already been recognized [23–28], albeit via limited studies. In an initial study, Li et al. tested blood DNA samples of 22 adult MZ twin pairs by using the Illumina Infinium Human Methylation 27K BeadChip array that allows for the co-analysis of 27,578 CpG sites, which revealed significant DNA methylation differences (377 CpG sites, methylation difference >0.17) [23]. Applying the criterion of how frequent the observed methylation differences were amongst MZ twin pairs, authors sub-selected a total of 92 CpGs, that were differentially methylated in all 22 pairs [24]. Furthermore, in another candidate-finding approach, Du et al. [25] analysed four MZ twin pairs with methylated DNA immunoprecipitation (MeDIP) and identified a set of 38 differential methylation regions demonstrating pair-specific methylation differences, all of which are located within CpG islands (<500 bp long, >55% GC content, [29]). The majority of these markers are associated with cell differentiation, proliferation and development, but future validation is necessary to reveal whether they can be used for MZ twin differentiation in routine forensic practice.

Using a more targeted approach, forensic researchers have also explored the possibility of using differential methylation patterns within satellite DNA, and more specifically within interspersed repeats such as LINE-1 [26] and Alu sequences [27]. LINE-1 methylation was investigated in both blood and buccal cells from a total of 119 MZ pairs using bisulfite pyrosequencing. Authors detected statistically significant tissue differences and were able to distinguish only a subset (12.61%) of MZ pairs (n = 15 pairs) using this specific region (3 CpG sites) in blood [26]. Also, the detected CpG methylation was significantly correlated with gender in blood and with age in buccal cells ($p = 0.001$), highlighting the complexity of epigenetic patterns. Lastly, in the study by Stewart et al. [27], all 5 MZ twins tested were distinguished by analysing Alu methylation in buccal cells (2 fragments, 19 CpG sites). The authors used high-resolution melting curve analysis, which can differentiate between fragments having differential methylation levels on the basis of different melting temperatures. Limitations of such approach include the low resolution (investigation of an entire fragment rather than specific CpG sites), the required large sample volume and the questioned applicability in other tissues (such as blood or saliva) [27].

In the present study, our aim was to execute epigenetic differentiation of MZ twins under the forensic scenario, comprising i) the discovery of candidate markers in reference DNA via genome-wide screening, ii) validating selected candidate markers in reference DNA using a method suitable for forensic trace analysis, and iii) applying such suitable method for analysing the validated markers in trace DNA. To the best of our knowledge, this is the first study addressing the differentiation of MZ twins in reference-type samples as well as in forensic-type samples, using the combination of genome-wide screening and methylation-specific qPCR for targeted analysis.

2. Materials and methods

2.1. Sample collection

This study was approved by the St Thomas' Hospital Local Ethics Committee, and all participants provided signed informed consent prior to sample collection. In total, 10 pairs of female MZ twins from the TwinsUK cohort [30], aged 52–62 years and of European ancestry were included in the study. Whole blood was collected in EDTA-treated tubes as reference sample, while a small amount of blood (a small drop) was also used to make up small bloodstains on cotton material (~0.4–0.5 cm^2 stains when dried). Samples were stored at −80 °C before proceeding to DNA extraction and analysis. The monozygosity of the twins used in the study was confirmed by genotyping 15 highly polymorphic STR loci using the AmpFLSTR™ Identifiler™ PCR Amplification kit (Thermo Fisher Scientific, USA), which resulted in the same STR profile per MZ pair respectively.

2.2. DNA sample preparation

Total DNA from whole blood, representing reference-type DNA samples, was extracted using the DNeasy Blood & Tissue kit (QIAGEN, Germany) according to the manufacturer's instructions. Genomic DNA from the dried bloodstains on cotton, representing trace-type DNA samples, was isolated using the QIAamp® DNA Investigator kit (QIAGEN) following the recommended protocol for 'Isolation of total DNA from body fluid stains'. The whole blood stain was used for DNA extraction and DNA samples were eluted in 20 μl. To maximise DNA yield, samples were lysed at 56 °C for at least 2 h and QIAshredder spin columns (QIAGEN) were used to harvest lysate remaining in the cotton material. To access the accuracy and linearity of methylation quantification, 7 DNA standards of known methylation levels (0, 0.1, 0.25, 0.5, 0.75, 0.9, 1) (EpigenDx, USA) were also analysed. All DNA samples were quantified using the Quantifiler® Human DNA Quantification kit (Applied Biosystems, USA), according to the manufacturer's instructions. For genome-wide analysis, 750 ng of extracted DNA was bisulfite-converted using the EZ DNA Methylation™ Kit (ZymoResearch, USA). For qPCR analysis, depending on the available DNA quality and/or quantity, 200 ng (whole blood), 20 ng (bloodstains) and 250 ng (DNA standards) of DNA were treated using the MethylEdge™ Bisulfite Conversion system (Promega, USA). Considering 80% DNA recovery after conversion as suggested by the manufacturer, bisulfite DNA samples were diluted down to 1 ng/μl.

2.3. Genome-wide DNA methylation profiling

Genome-wide DNA methylation profiles used in this study were generated using the Illumina Infinium Human Methylation 450K BeadChip array (Illumina, USA). Intensity images were captured by GenomeStudio (2010.3) Methylation module (1.8.5) software. The Illumina 450K BeadChip assay allows for quantification of DNA methylation levels at 485,512 CpG dinucleotides. For each CpG site, a beta value was estimated, ranging from 0, representing completely unmethylated, to 1, representing completely methylated sites. The beta value is interpreted as average methylation in a particular site taking into account the collection of cells that form each sample. The Human Methylation 450K BeadChip array contains two types of beads associated with the two different chemical assays, Infinium I and II, which can potentially cause bias in probe design [31]. DNA methylation probes that mapped to multiple locations to the reference sequence (with exact sequence match and within up to two base pair mismatches) and probes that contain a non-rare polymorphism in the CG site, minor allele frequency (MAF) >0.1 in European population from 1000 genomes,