Forensic Science International: Genetics 36 (2018) 95–103

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Systematic evaluation of the early access applied biosystems precision ID Globalfiler mixture ID and Globalfiler NGS STR panels for the ion S5 system

Petra Müller[a], Antonio Alonso[b], Pedro A. Barrio[b], Burkhard Berger[a], Martin Bodner[a], Pablo Martin[b], Walther Parson[a,c,*], The DNASEQEX Consortium

[a] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[b] National Institute of Toxicology and Forensic Sciences, Madrid Department, Las Rozas de Madrid, Spain
[c] Forensic Science Program, The Pennsylvania State University, PA, USA

ARTICLE INFO

Keywords:
Short tandem repeats
Massively parallel sequencing
Globalfiler NGS STR ID panel
Globalfiler mixture ID panel
NGS

1. Introduction

The current gold standard for analysing forensic genetic variation focuses on human identity testing using multiplex autosomal short tandem repeat (STR) genotyping using polymerase chain reaction (PCR)- and capillary electrophoresis (CE)-based approaches [1,2]. In recent years, massively parallel sequencing (MPS) has been under continuous development for forensic genetics and, while CE-based STR genotyping is generally sufficiently discriminatory for routine forensic applications, shown some benefits over this technology [3]. For instance, in CE-based PCR kits (containing up to 30 STR markers) amplicons need to be designed to avoid overlapping fragment lengths. Consequently, some PCR products are longer than would be necessary to reflect the information of the polymorphic repeat block only. Target amplicons between 80 and 500 bp are common [4–8]. In contrast, MPS offers the ability to target dozens to hundreds of amplicons regardless of their size in one assay. Short amplicons are beneficial for the analysis of degraded DNA [9,10]. Due to the nature of degraded DNA, samples analysed using conventional DNA typing methods often result either in partial profiles with lower discriminatory power or in the total lack of retrievable genetic information.

Further, many studies demonstrated that sequencing of STRs provides new insights into sequence variation of established STR markers, elucidating the structure of micro-variants and improving the detection of mixtures [11–22]. However, before MPS can be implemented into routine forensic casework both the new technology as well as resulting data have to be rigorously validated with respect to robustness, performance and backward compatibility to the large body of CE-based STR data stored in national DNA intelligence databases. For instance, in 2017 Alonso et al. [23] reported the need for a sequencing platform independent analysis software based on the considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) [24].

We evaluated the current state of MPS technology and chemistry to facilitate STR sequencing in the framework of the EU funded project DNASEQEX (DNA-STR Massive Sequencing & International Information Exchange (HOME/2014/ISFP/AG/LAWX/4000007135)) including commercially available products from Illumina (San Diego, USA), pre-release products from Thermo Fisher Scientific (TFS; Waltham, USA) and analysis software at their respective developmental stages. The consortium consisted of the following partners: The Biology Service of the National Institute of Toxicology and Forensic Science, Madrid (Spain) acting as coordinator, the Department of Forensic Genetics of the Institute of Legal Medicine and Forensics Science, Berlin (Germany) and the Institute of Legal Medicine, Medical University of Innsbruck (Austria) as beneficiaries and the Institute of Applied Genetics, Department of Molecular and Medical Genetics, University of North Texas Health Science Center, USA acting as consultant to the consortium.

Here, we present the evaluation of two MPS-prototypes, Early Access Applied Biosystems Precision ID Globalfiler Mixture ID and Early Access Applied Biosystems Precision ID Globalfiler NGS STR

* Corresponding author at: Institute of Legal Medicine, Medical University of Innsbruck, Müllerstraße 44, 6020 Innsbruck, Austria.
E-mail address: walther.parson@i-med.ac.at (W. Parson).

https://doi.org/10.1016/j.fsigen.2018.06.016
Received 8 May 2018; Received in revised form 20 June 2018; Accepted 25 June 2018
Available online 30 June 2018

Panels (both Thermo Fisher Scientific (TFS), Waltham, USA), including experiments to assess concordance, sensitivity, mock casework (single source and mixtures) as well as degraded DNA samples. We note that only the STR genotype calls were considered for this study in agreement with the goals defined in the DNASEQEX project. Inter-laboratory concordance was assessed by performing all experiments in a uniform manner at the Institute of Legal Medicine, Medical University of Innsbruck, Austria (GMI) and the Biology Service of the National Institute of Toxicology and Forensic Science, Madrid, Spain (INTCF).

2. Materials and methods

This inter-laboratory evaluation study combines data produced at the GMI and the INTCF. A set of forensically relevant samples was agreed upon in order to assess the inter-laboratory performance of the Early Access Applied Biosystems Precision ID Globalfiler Mixture ID (GF-Mix) and the Globalfiler NGS STR Panels (GF-STR) for the Ion S5 System (TFS, see Section 2.1). Time estimates for a typical automated workflow for both panels amounted to 24.5 h for eight samples. This includes 7.3 h (0.3 h hands-on time, 7 h instrument time) for library preparation, 1.2 h (0.5 h hands-on time, 0.7 h instrument time) for library quantitation, 13 h (0.3 h hands-on time, 12.7 h instrument time) for template preparation and 2–3 h for sequencing. The GF-Mix panel included 113 markers (29 autosomal STRs (aSTRs), one Y-STR, 42 SNPs, two Y-SNPs, 36 micro-haplotypes, Amelogenin and Y-InDel (rs2032678)) whereas with the GF-STR panel 33 targets could be analysed simultaneously (29 aSTRs, one Y-STR, Amelogenin and Y-InDel (rs2032678)).

2.1. Sample selection

2.1.1. Concordance

Both laboratories performed a concordance study using the same single donor standard reference material (SRM) purchased from the National Institute of Standards and Technology (NIST; component 2391c A–C [25]; Gaithersburg, USA) and the forensic standard control DNA 9947A (TFS). The final DNA input for the concordance study was 1 ng according to the manufacturer's recommendation. Concordance was determined using 29 aSTRs and one Y-STR for the GF-Mix and 28 aSTRs and one Y-STR for the GF-STR panel for 2391c A–C, known from reference [25], and 20 aSTRs for 9947A and both panels known from CE results.

2.1.2. Sensitivity

At GMI serial dilutions for both Globalfiler trials consisted of forensic standard control DNA 9947A and NIST SRM 2372A [26]. To examine the impact of manual library preparation on sensitivity testing, additional libraries using 2372A for the GF-STR and control DNA 2800 M (Promega, Madison, WI, USA) for the GF-Mix panel (DNA input/assay: 1 ng to 31 pg, single approach) were manually prepared at GMI following [27]. At INTCF serial dilutions were prepared using DNA control 9947A and 2372A for the GF-Mix panel as well as 2372A for the GF-STR panel. Dilutions were prepared enabling DNA input of 250 pg and 125 pg, following the manufacturer´s workbooks. In addition DNA input (2372A) of 500 pg, 250 pg, 125 pg, 62 pg were examined in duplicate at INTCF according to the manufacturer's recommendations. Moreover, to improve the total number of sequencing reads INTCF tested adapted PCR cycling conditions (28 instead of 24 PCR cycles) using 2372A at the following genomic DNA input amounts in duplicate: 125 pg, 62 pg, 31 pg and 15 pg. Sensitivity was determined using 20 aSTRs and one Y-STR known from CE experiments.

2.1.3. Casework

2.1.3.1. Globalfiler mixture ID panel. The casework study sample set at GMI consisted of 16 single source and six mixture stains (2–3 persons) from past GEDNAP (German DNA Profiling; http://www.gednap.org) proficiency tests listed in Table S1. The final DNA input for each sample was 1 ng according to the manufacturer´s recommendation.

The casework study sample set at INTCF consisted of 12 single source and 13 mixture stains (2–3 persons) from past GEDNAP proficiency tests listed in Table S1. The final DNA input was adjusted to 1 ng for all single source samples, while mixtures were analysed using 2–3 ng final DNA input (see Tables S1–S6_INTCF) according to the manufacturer´s recommendation.

2.1.3.2. Globalfiler NGS STR panel for the S5 system. The casework study sample set utilized at GMI consisted of 17 single source and six mixture stains (2–3 persons) from past GEDNAP proficiency tests plus SRM 2391c E [28] (Table S1). The final DNA input for each sample was 1 ng according to the manufacturer's recommendation.

The casework study sample set at INTCF consisted of 10 single source and 13 mixture stains (2–3 persons) from past GEDNAP proficiency tests, three single source samples with known profile from volunteers, one single source and two mixture samples (2–3 persons) from past GHEP-ISFG (Spanish and Portuguese-Speaking Group of the ISFG; https://ghep-isfg.org/) proficiency tests and four challenging DNA samples (one tooth, one neonatal femur, one neonatal *pars petrosa* bone and one FFPE (formalin fixed and paraffin embedded) sample) (Table S1).

2.2. DNA quantification

At both laboratories the amount of genomic DNA was determined using the Quantifiler Trio DNA Quantification Kit (TFS) [29]. For calibration the Quantifiler THP DNA Standard was diluted in Quantifiler THP DNA Dilution Buffer to 50 ng μL^{-1}, 5 ng μL^{-1}, 0.500 ng μL^{-1}, 0.050 ng μL^{-1} and 0.005 ng μL^{-1}. The final volume of each reaction was 20 µL consisting of 10 µL Quantifiler THP PCR Reaction Mix, 8 µL Quantifiler Trio Primer Mix and 2 µL sample of unknown quantity. Samples were run in duplicate using an Applied Biosystems 7500 Fast Real-Time PCR Instrument (TFS) using the HID Real-Time PCR Software v 2.3. Thermal cycler protocol was performed according to [29].

2.3. Capillary electrophoresis (CE) concordance data

GMI: STR typing was conducted using the Globalfiler PCR Amplification Kit (GF-PCR) (TFS). DNA samples were diluted in UV-irradiated Molecular Biology Grade water according to the manufacturer´s protocol [4] to the final DNA input amount of 1 ng. All DNA extracts were amplified with 29 PCR cycles except of GEDNAP-47_S1 and GEDNAP-43_S3, where 30 PCR cycles were applied. Amplification was performed on an Applied Biosystems GeneAmp 9700 thermal cycler (TFS) following the manufacturer´s protocol [4]. CE analysis of PCR products was performed on an Applied Biosystems Prism 3500XL Genetic Analyzer (TFS). The analysis of STR fragments utilized the GeneMapper ID-X software, version 1.2 (TFS) by applying the in-house validated colour channel specific analytical thresholds: blue – 50 relative fluorescence units (RFU), green – 80 RFU, yellow – 100 RFU, red – 100 RFU, purple – 175 RFU.

INTCF: Standardized STR typing was conducted using the GF-PCR kit according to Martin et al. [30], PowerPlex Fusion 6C System (Promega, Madison, WI, USA) according to Ensenberger et al. [31], Minifiler, Identifiler Plus, NGM SElect NGM Detect (all four TFS) as well as PowerPlex ESX 17, PowerPlex ESI 17 and PowerPlex CS6 (all three Promega). Amplification was performed according to previous studies [30,31] and the manufacturer's protocols [32–39] on an Applied Biosystems GeneAmp 9700 thermal cycler (TFS). CE analysis of PCR products was performed on an Applied Biosystems Prism 3500 Genetic Analyzer (TFS). The analysis of STR fragments utilized the GeneMapper ID-X software, version 1.4 (TFS) by applying 100 RFU as an analytical threshold for all channels (blue, green, yellow, red and purple).