AUSTRIAN STR POPULATION DATA USING THE POWERSEQ GY46 SYSTEM AND MASSIVELY PARALLEL SEQUENCING

Petra Müller[1], Burkhard Berger[1], Martin Bodner[1], The DNASEQEX Consortium, Walther Parson[1,2]

[1] *Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria*
[2] *Forensic Science Program, The Pennsylvania State University, PA, USA*

petra.mueller@i-med.ac.at

Background: By using massively parallel sequencing (MPS) technologies to genotype forensic markers, it is possible to characterize sequence variations located in the repeat and flanking regions of short tandem repeats (STRs). To account for sequence based allele frequencies valid for the Austrian population, we performed a population study including 248 unrelated male donors verifiably born in Austria. In accordance with the Austrian law and permission of the Federal Ministry of the Interior, Austria, it is possible to use DNA samples and DNA profiles of the Austrian National DNA Database for population studies, preconditioned the data is made anonymous.

Concordance to capillary electrophoresis (CE) was assessed for all samples using the AmpFlSTR NGM Select PCR Kit analysed on an Applied Biosystems Prism 3500XL Genetic Analyzer with the Gene Mapper ID-X software (all Thermo Fisher Scientific, Waltham, USA). MPS libraries were prepared manually using the PowerSeq GY46 system (Promega, Madison, USA) according to the manufacturer´s recommendations. Sequencing was performed on the MiSeq FGx benchtop sequencer (Illumina, San Diego, USA) and analysed using STRait Razor v2 software [1].

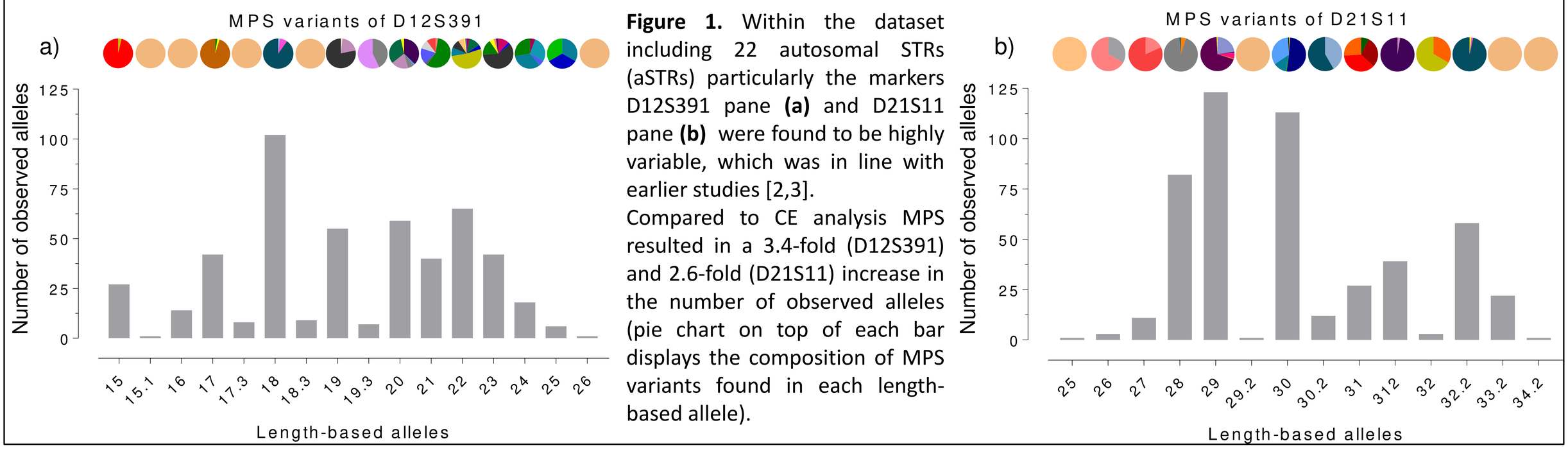


Figure 1. Within the dataset including 22 autosomal STRs (aSTRs) particularly the markers D12S391 pane **(a)** and D21S11 pane **(b)** were found to be highly variable, which was in line with earlier studies [2,3].
Compared to CE analysis MPS resulted in a 3.4-fold (D12S391) and 2.6-fold (D21S11) increase in the number of observed alleles (pie chart on top of each bar displays the composition of MPS variants found in each length-based allele).

Figure 2. We observed a 13 base pair deletion at **Penta D** (Chr21:43636191–43636203, del-GAAAGAAAAAAAA), which results in allele 2.2 using length-based CE technologies.

Penta D exemplifies possible alignment challenges and discordances to CE. Alignment was ambiguous due to the proximity of the deletion to the repeat region.

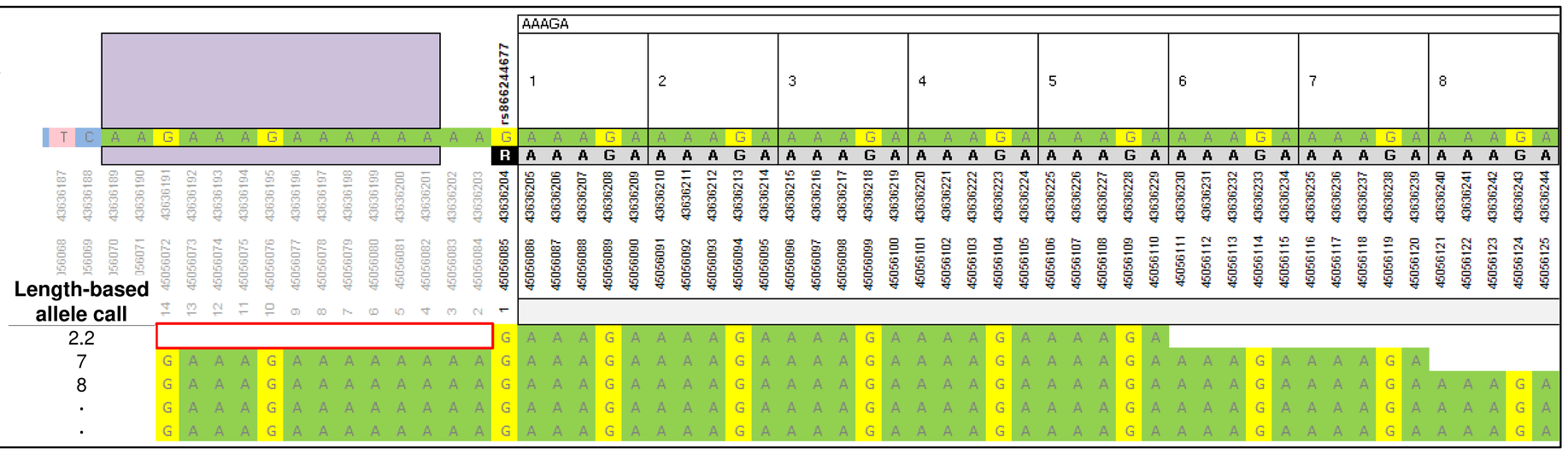
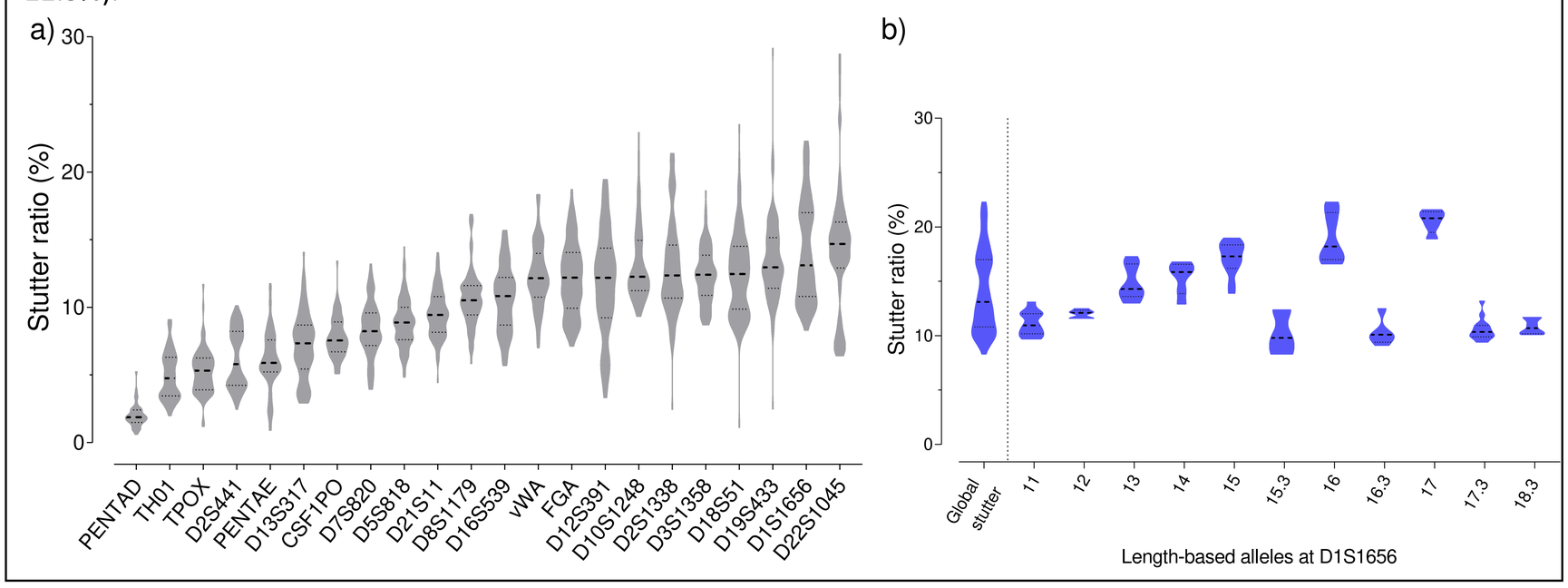


Figure 3. Stutter analysis was performed on a subset of 50 samples (corresponding to 20% of the entire sample set). Samples were selected according to the total numbers of reads (selection criteria: ≤63,300 reads or ≥199,000 reads).
a) Violin plot of relative stutter ratios for 22 aSTRs included in the PowerSeq GY46 panel. Median stutter ratios ranged from 1.9% (PentaD) to 14.7% (D22S1045). **b)** Locus specific violin plot for D1S1656 showing distinct differences in stutter ratios between integer alleles and micro-variants (global median: 13.1%, median stutter ratios ranged from 8.3% to 22.3%).



Conclusions:
In general, the here presented preliminary results underline the significance of sequence based studies. Our results were 99.99% concordant to known genotypes (derived from CE-analysis) including 15 aSTRs (AmpFlSTR NGM Select PCR Kit). Analysing 248 samples, MPS was able to successfully type 7439 out of 7440 alleles. The reason for discordance was due to allelic drop-out of the longer allele in D2S1338 (typed alleles: CE: 18, 28; MPS: 18).
The Austrian dataset is expected to provide sequence based allele frequencies to support the implementation of MPS into forensic routine casework.

References: **[1]** King JL, Wendt FR, Sun J, Budowle B. STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet. 2017;29:21-8. **[2]** Devesse L, Ballard D, Davenport L, Riethorst I, Mason-Buck G, Syndercombe Court D. Concordance of the ForenSeq system and characterisation of sequence-specific autosomal STR alleles across two major population groups. Forensic Sci Int Genet. 2017;34:57-61. **[3]** Gettings KB, Borsuk LA, Steffen CR, Kiesler KM, Vallone PM. Sequence-based U.S. population data for 27 autosomal STR loci. Forensic Sci Int Genet. 2018;37:106-15.

Acknowledgements: This project has been funded with support from the European Commission (grant HOME/2014/ISFP/AG/LAWX/4000007135 under the Internal Security Funding Police program of the European Commission-Directorate General Justice and Home Affairs). This project has been funded with support from European Union 779485-STEFA-ISFP-2016-AG-IBA-ENFSI Steps Towards a European Forensic Science Area; WP7; Empowering Forensic Genetic DNA Databases for the Interpretation of Next Generation Sequencing Profiles (dna.bases). This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use which may be made of the information contained herein. The authors are grateful to Reinhard Schmid and the database staff at the Federal Ministry of the Interior, Vienna, Austria for the excellent cooperation. The authors would like to thank D. Niederwieser, A. Kaindl-Lindinger, L. Schnaller (all Institute of Legal Medicine Innsbruck), J.L. King (UNT Health Science Center, USA) for technical support as well as H. Niederstätter (Institute of Legal Medicine Innsbruck) for helpful discussions. The authors thank Spencer Hermanson and Doug Storts (both Promega, USA) for technical support.