ORAL ABSTRACTS

No.	*Presenting Author*	*Title*
1	Thomas Parsons	Forensic genetics and the missing: Striving for certainty and justice
2	Kyoung-Jin Shin	Implementing massively parallel sequencing for forensic DNA analysis using in-house PCR panels
3	Laurence Alicia Elisabeth Devesse	The Tao of MPS: Common novel variants
4	Christopher Phillips	A missing persons identification panel for the MPS era
5	Bruce Budowle	Massively parallel sequencing (MPS) can be considered NGS, i.e., now generation sequencing: Implementation of whole genome mitochondrial DNA sequencing into routine casework
6	Katherine Butler Gettings	STRseq: A resource for sequence-based STR analysis
7	Sascha Willuweit	NOMAUT - NGS STR nomenclature for forensic genetics
8	Sebastian Ganschow	ToaSTR: A web-based forensic tool for the analysis of short tandem repeats in massively parallel sequencing data
9	SallyAnn Harbison	Establishing the limit of detection of massively parallel sequencing using laser micro-dissected cells
10	Chiara Turchi	A microhaplotypes panel for forensic genetics using massive parallel sequencing (MPS)
11	Jennifer Churchill	Massively parallel sequencing-enabled mixture analysis of mitochondrial DNA samples
12	Mark Jobling	Genetic and genomic approaches to estimating ancestry
13	Michael Nothnagel	Revisiting the male genetic landscape of China
14	Torben Tvedebrink	GenoGeographer - A tool for genogeographic inference
15	Martin Bodner	Discerning the "identical": Extended mitogenome diversity behind the most common West Eurasian mtDNA control region haplotype
16	Anu Maria Neuvonen	The effect of a singular population structure on forensic analysis
17	YunGeon Kim	Development of ethnic classification models for Asian group: A data mining approach
18	Runa Daniel	An ancestry-informative assay for the Asia Pacific region
19	Ji hyun Lee	Prediction of Asian ethnic subgroups using HID-Ion AmpliSeq™ Ancestry Panel
20	Arwin Ralf	Increasing the number of Rapidly Mutating (RM) Y-STRs significantly improves male relative differentiation
21	Mehrdad Hajibabaei	Genomics approaches for characterizing biological diversity : From identifying species to monitoring whole ecosystems
22	Adrian Linacre	'ForCyt' DNA Database of wildlife species

23	Sungmin Kim	Non-human forensic identification of complex mixture samples using DNA metabarcoding
24	Burkhard Berger	Canine STR marker get more - Dog breed affiliation with a forensically validated STR panel
25	Sorelle Bowman	Identification of hoax agents, *Bacillus* and *Yersinia* strains by protein profiling using microfluidic capillary electrophoresis with peak detection algorithms
26	Natasha Arora	Microbiome communities analyses to predict body sites: Analyses on published data and on exposed samples in the laboratory
27	Weibo Liang	Microhaplotype: Ability of persanal idenfication and being ancestry informative marker
28	Daniele Podini	Microhaplotype for ancestry prediction
29	Nor Aidora binti Saedon	The applications of Malaysian populations plotted via next generation sequencing
30	Wibhu Kutanan	Complete mitochondrial genomes of the hunter-gatherer populations in Thailand
31	Titia Sijen	DNA transfer and cell type inference to assist activity level reporting
32	Sabrina Ingold	Association of a body fluid with a DNA profile by targeted RNA/DNA deep sequencing
33	Phuvadol Thanakiatkrai	Direct PCR improves STR profiles from substrates of improvised explosive device
34	Lode Sibbens	The development of a forensic clock to determine time of death
35	Alicia Haines	Latent DNA detection using fluorescent DNA binding dyes
36	Bianca Szkuta	Assessment of the prevalence, transfer, persistence and recovery of DNA traces from worn upper garments: Results from an inter-laboratory study
37	Qingzhen Meng	DNA transfer study based on parallel system methods
38	Athina Vidaki	CpGforID: Epigenetic discrimination of identical twins under the forensic scenario
39	Diana H Hall	Application of DIP-STRs to sexual/physical assault investigations: Eight case reports
40	Manfred Kayser	Advances in human appearance genetics and relevance for forensic DNA phenotyping
41	Sae Rom Hong	Epigenetic age signatures in saliva: Age prediction using methylation SNaPshot and massive parallel sequencing
42	Ewelina Pośpiech	Epigenetic prediction of age in different study groups
43	Lin Zhang	Association between BMP4 gene polymorphisms and eyelid traits in Chinese Han population
44	Chie Morimoto	Discriminating between uncle-nephew and grandfather-grandson kinships by analyzing chromosomal sharing patterns

45	James Curran	A new (approximate) method for Y-STR haplotype probability assignment
46	Mikkel Meyer Andersen	How convincing is a matching Y-chromosome profile?
47	Øyvind Bleka	Using EuroForMix to analyse complex SNP mixtures, up to six contributors
48	Catherine Elizabeth McGovern	Establishing a range of foundational validity for complex DNA mixture interpretation using probabilistic genotyping software: A response to PCAST
49	Charles Hallam Brenner	Impossible to specify the number of contributors to a DNA mixture
50	Susan Friedman	Science advancing justice : Lessons learned from DNA exonerations
53	Marta Diepenbroek	First genetic evidence for Jewish holocaust victims in Sobibor, Poland, by phylogeographic analyses
54	Turi King	Who's the daddy? Tracing the break in the Y chromosome lineage of King Richard III
55	Maria Geppert	Wider use of DNA evidence in criminal cases: Ancestry prediction using Y chromosome analysis in Germany
56	Kyleen Elwick	Comparative tolerance of two massively parallel sequencing platforms to common PCR inhibitors for missing person cases
57	Patrick Basset	Lessons from a study of DNA contaminations from police services and forensic laboratories in Switzerland
58	MARTIN EDUARDO MAUTNER	Degraded DNA samples made informative by using superprimers
59	Suni Edson	DNA typing from skeletal remains: A study of inhibitors using mass spectrometry
60	Fabio Macêdo Mendes	Kpop: A Python package for population genetics analysis
61	David John Moore	A large scale study on the characterisation of drop-in within a DNA laboratory
62	Lei Feng	Protein-based forensic identification using hair shaft in east Asian and Eurasian populations
63	Becky Steffen	Beyond the STRs: A comprehensive view of current forensic DNA markers characterized in the PCR-Based DNA Profiling Standard (SRM 2391D)
64	Michelle Peck	Inter-laboratory validation study of the ForenSeq DNA Signature Prep Kit, MiSeq FGx instrument, and ForenSeq universal analysis software for length-based STR analysis

1. FORENSIC GENETICS AND THE MISSING: STRIVING FOR CERTAINTY AND JUSTICE

ISFG Scientific Prize Lecture

T.J. Parsons

International Commission on Missing Persons, Koninginnegracht 12, 2514 AA The Hague, Netherlands
thomas.parsons@icmp.int

In it's magnitude of importance and range of applications, forensic genetics has profoundly influenced society. One area of greatest impact is the identification of missing persons. When loved ones go missing, humans have a fundamental need for knowledge of their fate and the return of mortal remains for commemoration. Society demands effective state and institutional mechanisms for accounting for the missing, including provision for justice. Forensic genetics must therefore accommodate many contexts: DVI from accidents, natural disaster or terrorism; military missing in action; mass graves and war crimes; deaths due to migration; "routine" missing persons; human trafficking; and re-uniting living persons who have been separated from families.
This lecture will review the development and application of DNA methods that have now given rise to unprecedented abilities to address the personal and societal trauma associated with missing persons. Missing persons DNA applications have drawn from developments in other forensic applications, but they have also been a driving force in the development and refinement of methods in many genetic systems. This includes not only DNA marker systems, but informatics and interpretation frameworks, for example in kinship analysis or rigorous multi-disciplinary integration of evidence. Massively Parallel Sequencing is beginning to be implemented, and will offer powerful new capabilities. Our field must ensure that optimal systems are developed for the missing, and to promote as well the establishment of policies and collaborative mechanisms that will permit technical advances to be brought to bear on this area of critical importance to humanity and human rights.

2. IMPLEMENTING MASSIVELY PARALLEL SEQUENCING FOR FORENSIC DNA ANALYSIS USING IN-HOUSE PCR PANELS

Keynote Speech

Kyoung-Jin Shin

Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea
kjshin@yuhs.ac

Massively parallel sequencing (MPS) has been increasingly used for the analysis of various forensic markers to complement the weak points of capillary electrophoresis and to obtain more data while using only one platform. However, there are several obstacles in the current research and practical application of forensic DNA analysis using the MPS method. The first is the lack of feedback for further progress due to limited public information of available commercial MPS panels. The second are the costly and time-consuming procedures used to prepare the MPS library. Finally, the MPS data analysis is somewhat of a hassle. Therefore we have developed in-house PCR panels which can amplify mitochondrial DNA, autosomal/Y chromosomal STRs and microhaplotypes by marker and then investigated sequence variation of these markers using the developed panels and customized MPS protocols. The cost and time spent in lab could be reduced by adopting the PCR based method instead of the adapter ligation method to prepare an MPS library with the amplicons. The generated MPS data

were analyzed using open source software; consequently, a platform independent MPS protocol was able to be established. In addition, autosomal STR genotyping for the same samples was carried out on two different MPS platforms and the results were compared. The speaker hopes that this presentation will be helpful to the researcher and/or practitioner who is using or will soon use MPS in their laboratory by discussing its application for forensic DNA analysis using in-house PCR panels and MPS protocols.

1. Børsting C1, Morling N2. Next generation sequencing and its applications in forensic genetics. Forensic Sci Int Genet. 2015;18:78-89.
2. Kim EH, Lee HY, Kwon SY, Lee EY, Yang WI, Shin KJ. Sequence-based diversity of 23 autosomal STR loci in Koreans investigated using an in-house massively parallel sequencing panel. Forensi Sci. Int. Genet. 2017; Submitted.
3. Kwon SY, Lee HY, Kim EH, Lee EY, Shin KJ. Investigation into the sequence structure of 23 Y chromosomal STR loci using massively parallel sequencing. Forensic Sci Int Genet. 2016;25:132-141.

3. THE TAO OF MPS: COMMON NOVEL VARIANTS

L.A. Devesse[1], D.J. Ballard[1], L.B. Davenport[1], K.B. Gettings[2], L.A. Borsuk[2], P.M. Vallone[2], and D. Syndercombe Court[1]

[1] King's Forensics, King's College London, 150 Stamford Street, London SE1 9NH
[2] National Institute of Standards and Technology, Gaithersburg, MD, USA
Laurence.a.devesse@kcl.ac.uk

The introduction of massively parallel sequencing (MPS) to forensic genetics has led to improvements in multiple aspects of DNA analysis, however additional complexities are concurrently associated with these advances. In relation to STR analysis, the move to assign alleles using sequence rather than length based methodologies has highlighted the extent to which previous allelic variation was masked. In this work, a series of samples (n=1000) from five different population groups (Caucasian, West African, North East African, East Asian and South Asian) were genotyped for 27 forensically validated autosomal STRs. Results were compared to data from the National Institute of Standards and Technology (NIST), with this collaborative project now providing one of the most expansive data sets generated using MPS technology to date. This presentation addresses the benefits and challenges associated with the discovery of novel sequence variants. The large number of these variants characterised at select markers brings into question the strategies for producing representative population data, yet also provides an opportunity to utilize this diversity in unique ways. Results from this collaborative study have demonstrated that the number of samples necessary to capture the breadth of allelic variation is highly dependent on the individual marker and the extent of its sequence variability. As one example, within a single population, all common sequence-based alleles at CSF1PO are captured when genotyping less than 200 individuals, while 400 individuals are insufficient for this purpose at D12S391.

4. A MISSING PERSONS IDENTIFICATION PANEL FOR THE MPS ERA

C. Phillips[1], A. Tillmar[2,3], T.J. Parsons[4], R. Huel[4], K. Kidd[5], M.V. Lareu[1], K. Elliott[6], R. Samara[6], E. Lader[6]

[1] Forensic Genetics Unit, University of Santiago de Compostela, Spain
[2] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping,

Sweden
[3] Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden
[4] International Commission on Missing Persons, Alipasina 45a, 71000 Sarajevo, Bosnia and Herzegovina
[5] Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA
[6] Qiagen Sciences Inc., 6951 Executive Way, Frederick, MD 21703, USA
c.phillips@mac.com

We have developed a large-scale SNP-based forensic identification panel for DNA analysis with massively parallel sequencing (MPS). The panel was specifically designed for the challenges of identifying missing persons; where the DNA can be highly degraded and relationship tests may involve reference samples from across several generations and in a deficient pedigree. For this reason, the panel comprises very short DNA fragments carrying multiple-allele SNPs or small microhaplotypes - both more informative than binary SNPs. We compiled 1411 tri-allelic or tetra-allelic SNPs (1377 autosomal loci, 34 X loci) and 46 microhaplotypes into a single PCR multiplex. In the multiple-allele SNPs, only 1.5% had average Heterozygosities (averaged over five 1000 Genomes population groups) below the binary SNP maximum of 50%. Amongst the 46 microhaplotypes originally compiled by Kiddlab [1], 18 had average Heterozygosities higher than the tri-allelic SNP maximum value of 66.7%. Thirty of the microhaplotypes had their size reduced by an average of 65 nucleotides without significant loss of informativeness due to the exclusion of SNPs at the boundaries of the full haplotype. We report component marker characteristics, including 1000 Genomes population frequency estimates, and simulations of relationship testing scenarios where distant pairwise relationships need to be statistically assessed.

1. Kidd K, et al: Evaluating 130 microhaplotypes across a Global Set of 83 Populations, Forensic Sci. Int. Genet. 2017; accepted for publication.

5. MASSIVELY PARALLEL SEQUENCING (MPS) CAN BE CONSIDERED NGS, I.E., NOW GENERATION SEQUENCING: IMPLEMENTATION OF WHOLE GENOME MITOCHONDRIAL DNA SEQUENCING INTO ROUTINE CASEWORK

Jennifer Churchill[1], Dixie Peters[1], Christina Strobl[2], Walther Parson[2,3], Bruce Budowle[1,4]

[1]Center for Human Identification, University of North Texas Health Science Center, Fort Worth, TX, USA
[2]Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[3]Forensic Science Program, The Pennsylvania State University, PA, USA
[4]Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia

Internal validation and implementation of the Precision ID mtDNA Whole Genome Panel (Thermo Fisher Scientific), has begun in our Missing Persons Unit. The entire genome is enriched by a 162 multiplex reaction (two amplification reactions) resulting in no amplicon greater than 175 bases in length. The mtDNA genome potentially can be sequenced from DNA typically encountered in human remains with greater resolution and no more consumption of sample than by Sanger sequencing. For MPS, library preparation and sequencing are performed on the Ion Chef and S5 (Thermo Fisher Scientific), respectively, reducing the workflow to only three pipetting steps and placing the chip into the sequencer. Since MPS technology is not yet widely used in actual casework analyses, we provide guidance on the criteria and steps to facilitate implementation. Validation entails: concordance studies, comparison of long and short PCR results, accuracy and precision (or reproducibility and repeatability), sensitivity (down to 1 pg of genomic DNA), stochastic effects (amplicon drop out), contamination, because of quantitative nature – mixture studies, degradation and inhibition, mock case samples, species specificity, and different strategies for multiplexing samples. In addition to these typical validation studies, since commercial software is not yet sufficiently robust, a pipeline of freeware and in-house

tools has been developed to process mtDNA results.
We present: internal validation studies, interpretation guidelines, thresholds and heteroplasmy, nomenclature, development of SOPs, training materials and processes for analysts, and novel workflow considerations. Our experiences could assist others considering the challenges of implementing a new platform and chemistry.

6. STRSEQ: A RESOURCE FOR SEQUENCE-BASED STR ANALYSIS

K.B. Gettings, L.A. Borsuk, P.M. Vallone

National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD, USA
katherine.gettings@nist.gov

The STR Sequencing Project was initiated to facilitate the description of sequence-based alleles at the Short Tandem Repeat (STR) loci targeted in human identification assays. This collaborative effort of the international forensic DNA community provides a framework for communication among laboratories. STRSeq data are maintained as GenBank records at the U.S. National Center for Biotechnology Information (NCBI), which participates in a daily data exchange with the DNA DataBank of Japan (DDBJ) and the European Nucleotide Archive (ENA).
Each GenBank record contains: (a) observed sequence of an STR region, (b) annotation of the repeat region ("bracketing" consistent with the guidance of the ISFG [1]) and flanking region polymorphisms, (c) information regarding the sequencing assay and data quality, and (d) backward compatible length-based allelic designation. STRSeq GenBank records are organized within a BioProject at NCBI (The STR Sequencing Project, www.ncbi.nlm.nih.gov/bioproject/380127), which is sub-divided by 1) Commonly used autosomal STR Loci, 2) Alternate autosomal STR Loci, 3) Y-chromosomal STR loci, and 4) X-chromosomal STR loci. Each of these categories is further divided into locus-specific projects. The BioProject hierarchy allows easier access to the GenBank records by browsing, BLAST searching, or ftp download; future plans include interface tools at strseq.nist.gov.
This presentation will inform attendees about the initiative, illustrate ways in which data may be accessed, and provide orientation of this project in relation to other, complementary efforts.

1. Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, Gusmão L, Hares DR, Irwin JA, King JL, Knijff Pd, Morling N, Prinz M, Schneider PM, Neste CV, Willuweit S, and Phillips C: Massively Parallel Sequencing of forensic STRs: Considerations of the DNA Commission of the International Society of Forensic Genetics (ISFG) on minimal nomenclature requirements. Forensic Science International: Genetics. 2016; 22: 54-63.

7. NOMAUT – NGS STR NOMENCLATURE FOR FORENSIC GENETICS

S. Willuweit, on behalf of the DNASeqEx EU project

Charité – Universitätsmedizin Berlin, Institute of Legal Medicine, Department of Legal Medicine
sascha.willuweit@charite.de

In addition to the general adoption of NGS in a forensic lab and implementation effort involved, the exchange of NGS STR results and comparison with legacy STR data is a crucial challenge when it comes to the interpretation of casework samples as part of routine work.
There are some proposals on the nomenclature of NGS STR results. Most of them are driven by the

idea of representing a (genetically) correct variant call against a reference sequence. Others try to reduce the necessary sequence information down to named repeat patterns. And finally, some are just doing in silico CE to call alleles while discarding any sequence information. Only the later does not fail to provide compatibility and comparability to the huge proportion of CE STR results already done and probably still going to be produced. Unfortunately, using this approach, we would not gain any information provided by NGS.
Since the coexistence and transition between NGS and CE is essential to forensic genetics, the NOMAUT NGS STR nomenclature incorporates compatibility and comparability. To archive this, NOMAUT was built as a huge catalogue of acquired sequence variants and the ability to grow in a very convenient but save and robust way. As a catalogue is a centralized structure, it can be imagined as an authoritative oracle answering sequence queries with allele calls - a NOMenclature AUThority. Whereas the underlying nomenclature rules and procedures are rather trivial, the accomplishment of safety and security in a worldwide scale are very challenging. To ensure for reliability and availability NOMAUT was built as a container being easily distributed over web service infrastructures like Amazons AWS.
This presentation will shed some light on the basic idea of NOMAUT as well as the initial implementation of the secure and self-maintaining authority software and the exemplary code samples provided for NGS software producers for easy adoption of NOMAUT in their software.
The intention of this presentation is to motivate scientists and representatives/providers of relevant software products to shape and support NOMAUT as a collaborative effort.

8. TOASTR: A WEB-BASED FORENSIC TOOL FOR THE ANALYSIS OF SHORT TANDEM REPEATS IN MASSIVELY PARALLEL SEQUENCING DATA

Sebastian Ganschow[1], Peter Wiegand[2], Carsten Tiemann[1]

[1]LABCON-OWL GmbH, Bad Salzuflen, Germany
[2]Institute of Forensic Medicine, University Hospital of Ulm, Germany
forensik@labcon-owl.de

In recent studies, massively parallel sequencing (MPS) has demonstrated its potential for the analysis of short tandem repeats (STRs) in challenging forensic casework samples. MPS can detect sequence variants in isoalleles and thus increase resolution over conventional capillary electrophoresis (CE). However, considering currently available software, data analysis turns out to be a cumbersome process especially for laboratories without bioinformatical expertise. We developed the web application toaSTR (toastr.labcon-owl.de) as a user-friendly tool for platform independent STR allele calling in MPS data. toaSTR features an intuitive graphical user interface and transparent configuration options that are continuously optimized based on the user's feedback. It enables highly customizable STR panels and gives the user the ability to select from a wide range of markers. Both commercial and in-house multiplex PCR kits can be analyzed with only minimal requirements on primer design. Due to its stutter modelling algorithm, toaSTR is able to distinguish biological alleles from stutter artefacts and assists mixture interpretation. The result report visualizes called alleles for each STR system and complies with current nomenclature recommendations for sequence alleles. Data can also be exported for further analysis in biostatistical tools. We designed an in-house multiplex STR kit to evaluate toaSTR's performance under defined conditions and to enable comparison with a commercial kit. An initial validation study with ring trial samples and reference DNA confirmed a highly reliable allele calling and full concordance with CE results. Sequence information and stutter labeling proved beneficial for the interpretation of mixed profiles.

9. ESTABLISHING THE LIMIT OF DETECTION OF MASSIVELY PARALLEL SEQUENCING USING LASER MICRO-DISSECTED CELLS

G Nancollis[1], R England[2] and S-A Harbison[2]

[1] *Forensic Science Program, School of Chemical Sciences, University of Auckland, New Zealand*
[2] *Institute of Environmental Science and Research Limited, Auckland, New Zealand*
sallyann.harbison@esr.cri.nz

Massively parallel sequencing is fast emerging as an increasingly useful tool for forensic science. As part of our ongoing validation of this technology we wanted to both ensure compatibility of the technology with laser micro-dissected (LMD) cells and use the LMD to explore the minimal and optimal numbers of cells required to generate informative profiles. To do this we used the ForenSeq™ DNA Signature Prep Kit and a MiSeq FGx™ Sequencer. In this paper we describe the modifications made to our one step DNA extraction protocol to optimize sequencing performance, and address technical issues that require consideration when sequencing DNA from small numbers of cells, below the optimum amounts recommended. We demonstrate that DNA sequence profiles can be obtained from very small numbers of laser micro-dissected epithelial and sperm cells, demonstrating the utility of this approach.

10. A MICROHAPLOTYPES PANEL FOR FORENSIC GENETICS USING MASSIVE PARALLEL SEQUENCING (MPS)

Chiara Turchi, Mauro Pesaresi and Adriano Tagliabracci

Section of Legal Medicine, Department of Biomedical Sciences and Public Health, Polytechnic University of Marche, Ancona, Italy
c.turchi@univpm.it

Microhaplotypes (microhaps) are defined as loci of two or more SNP within the span of a single sequencing run with three or more common allelic combination (haplotypes) of the SNPs [1-2].
Microhaps appear to be useful in forensics for individual identification, ancestry inference, estimating relationships, and deconvoluting mixtures. The most important issue is identifying and characterizing a set of microhaps with the optimum characteristics for specific purposes and developing a suitable genotyping technology.
The MPS technologies are now making microhaplotypes a new type of forensic marker: a single sequence read can cover the expanse of the microhaplotypes and these loci become phase-known (i.e. the allelic combination of multiple SNPs on each chromosome of an individual can be determined).
In the present study we selected a panel of 90 microhaps, from The ALlele FREquency Database (https://alfred.med.yale.edu/), that matched the following criteria: 1- comprised of three, four or five SNPs; 2- comprised of 2-SNPs but with Global Average Effective Number of Alleles (Ae) ranking <= 60; 3- comprised of 2-SNPs but with the Informativeness statistic for ancestry inference (I_n) ranking <= 60.
We amplified the 90 microhaps in a European population sample and performed MPS on the Ion Torrent Personal Genome Machine (PGM).
Our objectives were: to validate a robust sequencing method to separate the allelic phase; to determine, for each microhaps, the effective number and type of alleles, the allele frequencies in the population analyzed; to identify final panels, with appropriate microhaps loci, suitable for the different forensic applications.

1. Kidd KK, Pakstis AJ, Speed WC, Lagacé R, Chang J, Wootton S, Haigh E, Kidd JR: Current

sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics. Forensic Sci Int Genet. 2014;12:215–24
2. Kidd KK, Pakstis AJ, Speed WC , Lagace R, Chang J, Wootton S , Ihuegbu N: Microhaplotype loci are a powerful new type of forensic marker. Forensic Sci Int Genet.Suppl. Ser. 4 2013; e123-e124

11. MASSIVELY PARALLEL SEQUENCING-ENABLED MIXTURE ANALYSIS OF MITOCHONDRIAL DNA SAMPLES

Jennifer D. Churchill[1], Monika Stoljarova[2], Jonathan L. King[1], Bruce Budowle[1,3]

[1]*Center for Human Identification, University of North Texas Health Science Center, Fort Worth, TX, USA*
[2]*Department of Chemistry and Biotechnology, Tallinn University of Technology,Tallinn, Estonia*
[3]*Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia*
Jennifer.Churchill@unthsc.edu

Mitochondrial DNA (mtDNA) has several characteristics that are useful for forensic investigation. Historically, mtDNA has been analyzed with Sanger sequencing methodologies. The time-consuming, labor intensive, and costly nature of this capillary electrophoresis-based methodology has made it difficult to analyze the entire mitochondrial genome. Instead, focus has been on the polymorphic control region. Massively parallel sequencing (MPS) technologies now make it feasible for forensic crime labs to sequence the entire mitochondrial genome, which can provide an increase in discrimination power. Additionally, the quantitative nature of MPS technologies offers opportunities to resolve mixed mitochondrial DNA samples encountered in casework.
A workflow consisting of the Precision ID Whole Mitochondrial Genome Panel, Ion Chef, and Ion S5 sequencer (Thermo Fisher Scientific) was used to evaluate the ability to perform mixture deconvolution. Mixtures in a 1:1, 1:5, 1:10, and 1:20 ratio were prepared and sequenced. Performance metrics, including coverage, strand balance, and noise, were used to evaluate the quality of the sequencing results generated. Single-source reference samples included in the mixed samples study were sequenced on an orthogonal platform to evaluate concordance. The major contributor in each mixed sample was successfully identified illustrating an improvement in the ability to analyze mixed samples. Finally, the bioinformatic pipeline developed and used to analyze the mixed samples in this study will be described in an effort to aid other laboratories interested in the implementation of MPS technologies.

12. GENETIC AND GENOMIC APPROACHES TO ESTIMATING ANCESTRY

Keynote Speech

M.A. Jobling

Department of Genetics & Genome Biology, University of Leicester, Leicester, UK
maj4@le.ac.uk

Ancestry estimation from human DNA sequence variants has applications over a range of fields, including medical genetics, molecular anthropology, genetic genealogy and forensics. Testing has evolved from the use of sets of ancestry-informative markers (AIMs), via genome-wide short-tandem repeats (STRs) and single nucleotide polymorphisms (SNPs) to whole-genome sequences. Markers showing sex-biased inheritance (mitochondrial DNA and the X and Y chromosomes) can reflect past sex-biased processes. Inferences from ancestry testing depend on the range and quality of comparative datasets, and also on how these are subclassified into groups. Classifications vary between fields,

including labels based on language, religion, nationality, geography, and skin-colour. Another key issue is the distinction between population-based and individual-based ancestry testing. The first can rely on the principles of population genetics, and more recent demographic modelling approaches. Inferring individual ancestry is more contentious, and yet this is what genetic genealogists and forensic scientists would like to do. Forensic autosomal STRs are poorly suited to ancestry inference because of their small numbers and high mutation rates. New next-generation sequencing multiplexes promise more power because the number of markers is larger, and because some contain biogeographic ancestry SNPs. The Illumina Forenseq system includes 22 'phenotypic' SNPs, all of which, in fact, are pigmentation SNPs. The fact that some SNPs have shifted between the 'phenotypic' and 'biogeographic' categories as the kit has evolved points to the conflation between ancestry and pigmentation, and also to the ethically concerning issue that DNA-based ancestry testing can act to reinforce the biological definition of race.

13. REVISITING THE MALE GENETIC LANDSCAPE OF CHINA

M. Nothnagel[1], G. Fan[2], F. Guo[3], Y. He[4], Y. Hou[5], S. Hu[6], J. Huang[7], X. Jiang[8], W. Kim[9], K. Kim[10], C. Li[11], H. Li[12], L. Li[12], S. Li[12], Z. Li[13], W. Liang[5], C. Liu[14], D. Lu[15], H. Luo[5], S. Nie[15], M. Shi[16], H. Sun[17], J. Tang[18], L. Wang[19], C. Wang[20], D. Wang[21], S. Wen[12], H. Wu[22], W. Wu[23], J. Xing[24], J. Yan[25], S. Yan[12], H. Yao[26], Y. Ye[5], L. Yun[5], Z. Zeng[27], L. Zha[28], S. Zhang[11], X. Zheng[29], S. Willuweit[30] and L. Roewer[30]

[1]University of Cologne, Germany; [2]Railway Police College, Zhengzhou, P.R. China (PRC); [3]National Police University of China, Shenyang, PRC; [4]Shaanxi Provincial Public Security Bureau, Xi'an, PRC; [5]Sichuan University, Chengdu, PRC; [6]Shantou University Medical College, PRC; [7]Guizhou Medical University, Guiyang, PRC; [8]Liaoning Criminal and Science Technology Research Institute, Shenyang, PRC; [9]Dankook University, Cheonan, Republic of Korea; [10]University of California, San Francisco, USA; [11]Ministry of Justice, Shanghai, PRC; [12]Fudan University, Shanghai, PRC; [13]Hebei Provincial Public Security Bureau, Shijiazhuang City, PRC; [14]Guangzhou Forensic Science Institute, PRC; [15]Kunming Medical University, PRC; [16]Ministry of Education, Beijing, PRC; [17]Sun Yat-sen University, Guangzhou, PRC; [18]Guangdong Medical University, Dongguan, PRC; [19]Police Station of Zhengzhou, Zhengzhou, Henan, PRC; [20]Max Planck Institute for the Science of Human History, Jena, Germany; [21]Jining Medical University, Shandong, PRC; [22]Xinxiang Medical University School of Basic Medical, Henan, PRC; [23]Zhejiang Provincial Public Security Bureau, Hangzhou, PRC; [24]China Medical University, Shenyang, PRC; [25]Chinese Academy of Sciences, Beijing, PRC; [26]Gansu Institute of Political Science and Law, Lanzhou, PRC; [27]Zhengzhou University, Henan, PRC; [28]Central South University, Changsha, PRC; [29]University of Western Ontario, London, Canada; [30]Charité – Universitätsmedizin Berlin, Germany
michael.nothnagel@uni-koeln.de

Numerous studies have investigated China's genetic diversity. Early studies reported a genetic distinction between Northern and Southern Han Chinese, while others showed a picture of more subtle differences. Here, we investigated the distribution of Y chromosome variation across 28 administrative regions as well as 19 recognized Chinese nationalities in continental China to assess the impact of recent demographic processes. To this end, we analyzed 37,994 Y chromosomal 17-marker haplotype profiles from the YRHD database with respect to forensic diversity measures and genetic distance between groups defined by administrative boundaries and ethnic origin, representing the largest genetic study on China to date. We observed high diversity throughout across all investigated Chinese provinces and ethnicities. Kazakhs and Tibetans showed the strongest significant genetic differentiation from the Han and other groups. However, differences between provinces were, except for those located on the Tibetan plateau, less pronounced. This discrepancy is explicable by the sizeable presence of Han speakers, who showed high genetic homogeneity all across China, in nearly all studied provinces. We also observed a subtle genetic North-South gradient in the Han, confirming previous reports of a clinal distribution of Y chromosome variation and being in notable concordance with the previously observed spatial distribution of autosomal variation. Our findings shed light on the demographic changes in China accrued by a fast-growing and increasingly mobile population.

14. GENOGEOGRAPHER - A TOOL FOR GENOGEOGRAPHIC INFERENCE

T. Tvedebrink[1], P.S Eriksen[1], H.S.Mogensen[2] and N. Morling[2]

[1] *Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark*
[2] *Section of Forensic Genetics, Department for Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark*
tvede@math.aau.dk

Ancestry-informative markers (AIMs) are genetic markers that give information about the genoogeographic ancestry of individuals. They are used for predicting the genogeographic origin of the investigated individual in crime and identification cases. In the exploring the genogeographic origin of an AIMs profile, the likelihoods of the AIMs profile in various populations may be calculated. However, there might not be an apt population in the database of reference populations. The fact that the likelihood ratios (LR) of one population compared to other populations are large does not imply that any of the populations are relevant. This is because that even though the populations might be exclusive, they are not exhaustive in the sense that they cover all possible human populations.
To handle this phenomena, we derive a likelihood ratio test (LRT), by which we can judge whether there is at least one population in our reference database that is "sufficiently close" to a profile's "true" genoogeographic population. The LRT is a measure of absolute concordance between a profile and a population, rather than a relative measure of the profile's likelihood in two populations (the LR).
The LRT is similar to a Fisher's exact test. The varying sample sizes of the reference populations in the database is explicitly included in the calculations, and does makes fewer assumptions than for LR calculations. The methodology has been implemented in an free open source interactive platform, GenoGeographer, that enables the forensic geneticist to make explorative analyses, produce various graphical outputs together with evidential weight computations.

15. DISCERNING THE "IDENTICAL": EXTENDED MITOGENOME DIVERSITY BEHIND THE MOST COMMON WEST EURASIAN MTDNA CONTROL REGION HAPLOTYPE

M. Bodner[1], C. Strobl[1], S. Nagl[1], C. Xavier[1], G. Huber[1], I. Cardinali[2], H. Lancioni[2], O. Semino[3], A. Olivieri[3], F. Gandini[3,4], A. Achilli[3], A. Torroni[3] and W. Parson[1,5]

[1] *Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria*
[2] *Dipartimento di Chimica, Biologia e Biotecnologie, University of Perugia, Perugia, Italy*
[3] *Dipartimento di Biologia e Biotecnologie "L. Spallanzani", University of Pavia, Pavia, Italy*
[4] *School of Applied Sciences, University of Huddersfield, Huddersfield, UK*
[5] *Forensic Science Program, The Pennsylvania State University, University Park, PA, USA*
martin.bodner@i-med.ac.at

Mitochondrial (mt)DNA is a vital tool in forensic genetics when nuclear markers do not provide results or maternal relatedness is investigated. The ~1.1 kbp non-coding mtDNA control region (CR) displays highly condensed variation and is therefore routinely typed. In this restricted range, matching haplotypes do not necessarily imply that the entire mitogenomes are identical or even belong to the same phylogenetic lineage. This is especially true for the most frequent West Eurasian mtDNA CR haplotype that occurs at a frequency of 3-4% in many European populations and is observed in numerous clades within haplogroup H ("Helena") and some HV relatives. In a seminal study, we investigated the power of massively parallel complete mitogenome sequencing in 29 Italian samples displaying the most common West Eurasian CR haplotype. This allowed the detection of an unexpected high diversity with 28 distinct haplotypes clustering into 19 clades of haplogroup H and raised the power

of discrimination from 0 to >99% [1]. Here we present novel results from the ongoing investigation of an expanded pan-Italian sample of almost 300 individuals carrying the most common CR haplotype. Even with the currently ~100 samples examined so far, no saturation is reached. This study demonstrates the benefit of complete mitogenome sequencing for forensic applications to enforce maximum discrimination, highest phylogenetic resolution and more comprehensive heteroplasmy detection.

1. Bodner M et al.: Helena, the hidden beauty: Resolving the most common West Eurasian mtDNA control region haplotype by massively parallel sequencing. FSI:Genetics. 2015; 15:21-26

16. THE EFFECT OF A SINGULAR POPULATION STRUCTURE ON FORENSIC ANALYSIS

Anu M. Neuvonen[1]*, Mikko Putkonen[1], Sanni Översti[2,3]Tarja Sundell[3,4] , Päivi Onkamo[3], Antti Sajantila[1], Jukka U. Palo[1]

[1] *Laboratory of Forensic Biology, Department of Forensic Medicine, PO Box 40 (Kytösuontie 11) FI-00014, University of Helsinki, Helsinki, Finland,*
[2] *Department of Biology, Section of Genetics and Physiology, University of Turku, Itäinen Pitkäkatu 4, 20014, Turku, Finland,*
[3] *Department of Biosciences, PO Box 56 (Viikinkaari 5) FI-00014, University of Helsinki, Helsinki, Finland,*
[4] *Department of Philosophy, History, Culture and Art Studies, PO Box 59 (Unioninkatu 38 F) FI-00014, University of Helsinki, Helsinki, Finland.*
anu.neuvonen@helsinki.fi

In Finland, the efficient evaluation of new DNA-markers for forensic purposes has been complicated by issues of low diversity and geographic subdivision, especially in the case of Y-chromosomal profiling. Because of the singular nature of these observed diversity patterns, it is important to further assess their geographical distribution for a more thorough understanding of their origins and history.
In order to refine the ancestry of groups within the population, the geographical patterns of mitochondrial and Y-chromosomal haplogroups of Neolithic and Mesolithic ancestry were assessed in Finnish populations. The distribution of these uniparental markers revealed a northeastern bias for hunter-gatherer haplogroups, while haplogroups associated with the farming lifestyle clustered in the southwest. In addition, a correlation could be observed between more ancient mitochondrial haplogroup age and eastern concentration. These results coupled with prior archeological evidence suggest the genetic northeast/southwest division observed in contemporary Finland represents an ancient vestigial border between Mesolithic and Neolithic populations undetectable in most other regions of Europe.
The characterization of Y-chromosomal and mitochondrial haplogroup patterns contributes to forensic science by clarifying geographical origins and further dissecting the lineages found in a homogenous population. Data gathered on Finnish population diversity can help to understand the processes affecting other small and isolated populations, thus increasing the scope to a global scale.

17. DEVELOPMENT OF ETHNIC CLASSIFICATION MODELS FOR ASIAN GROUP: A DATA MINING APPROACH

Kim. YG[1], Lee. JH[2], Cho. S[3], Lee. SD[2,3], Ahn. JJ[4], Ha. EH[4]

[1]*Department of Applied Statistics, Yonsei University, Seoul, Korea*
[2]*Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Korea*
[3]*Institute of Forensic Science, Seoul National University College of Medicine, Seoul, Korea*
[4]*Department of Information Statistics, Yonsei University, Wonju, Korea*
e-mail : peterpoverty91@yonsei.ac.kr

In the genetic approach of law, genetic testing has been carried out with an emphasis on individual identification, and further studies are being conducted in an effort to obtain meaningful investigation information. Identifying ethnic groups is one of such effort. These studies have been actively carried out mainly in Europe, where there is a large movement between ethnic groups. As the socio-economic level of the Asian region has increased, and the exchange has become active, the need to distinguish the nations in Asia has also increased. In order to do this, it is necessary not only to observe differences among ethnic groups, but also to verify various measures that can be interpreted statistically appropriately. In this study, we developed three-class classification models to categorize Asian ethnic groups using data mining techniques based on Y-chromosomal short tandem repeat(Y-STR) data. In order to design a practical classification model, we considered five different data mining classification models: Decision tree, Random forest, Neural network, Gradient boosting and Ensemble model. The results showed that the classification accuracy was about 80% for each Asian group(Northeast, Southeast and Southwest Asia). It was found that the data mining techniques could play a key role in Ethnic classification problem. Attempts to these data mining approaches can be also applied for autosomes or mitochondria to build classification model. Furthermore, it has been found that the use of various statistical analyzes in the field of genetics of law can be used as a useful tool in investigation information.

1. Jerome H. Friedman: Stochastic Gradient Boosting. Computational Statistics & Data Analysis. 2002; 38: 367-378
2. J. R. Quinlan: Bagging, Boosting, and C4.5. In Proceedings of the Thirteenth National Conference on Artificial Intelligence. 1996; 1: 725-730
3. L. Breiman: Random Forest. Machine Learning. 2001; 45:5-32
4. L. Rokach: Ensemble-based classifiers. Artifical Intelligence Review. 2010; 33:1-39
5. L. Wang and J. Ross L. Rokach: Synchronous neural networks of nonlinear threshold elements with hysteresis. Proceedings of the National Academy of Sciences of the United States of America. 1990; 87:988-992

18. AN ANCESTRY-INFORMATIVE ASSAY FOR THE ASIA PACIFIC REGION

C. Phillips[1], C. Santos[1], D. McNevin[2], M. Eduardoff [3], M. de la Puente[1], A. Heidegger[3], W. Parson[3,4], K. Kidd[5], R. Lagacé[6], S. Wooton[6], D. Power[7], S.M.T. Angustia[8], M. Barash[9], M.C.A. De Ungria[10], S.R. Hong[11], H.Y. Lee[11], C. Oz[12], E. Peters[13], L. Santos[8], N.A. Soliven[10], N. Tuitoga[13], S. Olson[6], M.V.Lareu[1], R. Daniel[14,2]

[1]Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[2]National Centre for Forensic Studies, Faculty of ESTeM, University of Canberra, Canberra, Australia
[3]Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[4]Forensic Science Program, The Pennsylvania State University, PA, USA
[5]Department of Genetics, Yale University School of Medicine, CT, USA
[6]Human Identification Group, Thermo Fisher Scientific, CA, USA
[7]Thermo Fisher Scientific, Victoria, Australia
[8]DNA Laboratory Branch, Philippine National Police Crime Laboratory, Quezon City, Philippines
[9]Centre for Forensic Science, Faculty of Science, University of Technology Sydney, Broadway, NSW, Australia
[10]DNA Analysis Laboratory, Natural Sciences Research Institute, University of the Philippines, Quezon City, Philippines
[11]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea.
[12]Forensic DNA Laboratory, Division of Identification and Forensic Science (DIFS), Israel Police, National H.Q., Jerusalem, Israel.
[13]Forensic Biology & DNA Laboratory, Forensic Science Services, Fiji Police Force, Fiji
[14]Office of the Chief Forensic Scientist, Forensic Services Department, Victoria Police, Australia
runa.daniel@police.vic.gov.au

Forensic DNA Intelligence has rapidly emerged as an important investigative tool. The ability to provide

information relating to the biogeographical ancestry (BGA) of the donor of an evidential sample would enable a more efficient use of valuable police and forensic resources and assist in decision making in the primary stages of an investigation.
Prior to the application of massively parallel sequencing (MPS) in forensic DNA analysis, differentiation of populations within the Asia Pacific region was achieved using a combination of, generally SNaPshot-based, ancestry informative (AIM) assays in multiple analyses [1-3]. As MPS enables the analysis of multiple loci simultaneously, MPS-based assays reduce consumption of limited evidential material and provide greater accuracy through the generation of extensive sequence data.
An AIM panel consisting of 160 markers was constructed with a specific focus on the differentiation of populations from South Asia, East Asia, South East Asia and Oceania in addition to differentiating African, European and indigenous American populations. The panel combines established binary SNPs with two highly informative SNP-based marker sets suited to MPS genotyping; 32 multiallelic SNPs (tri- and tetra-allelic loci) and 21 microhaplotypes, mainly Kiddlab compiled loci [4].
An Asia Pacific trial of this panel was conducted involving laboratories in this region and the Middle East. Participating laboratories contributed population samples of known ancestry. In addition to ancestry predictive performance, this study reports the component marker characteristics, panel optimisation, development of reference population data sets and the adaptation of ancestry-inference analysis tools for the analysis of multiallelic loci.

1. Fondevila, M., et al., Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. Forensic Sci Int Genet, 2013. 7(1): p. 63-74.
2. Phillips, C., et al., Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries. Forensic Sci Int Genet, 2013. 7(3): p. 359-66.
3. Santos, C., et al., Pacifiplex: an ancestry-informative SNP panel centred on Australia and the Pacific region. Forensic Sci Int Genet, 2016. 20: p. 71-80.
4. Kidd K., et al Evaluating 130 Microhaplotypes across a Global Set of 83 Populations, Forensic Sci Int Genet, 2017.

19. PREDICTION OF ASIAN ETHNIC SUBGROUPS USING HID-ION AMPLISEQTM ANCESTRY PANEL

Lee. JH[1], Kim. YK[2], Cho. S[3], Kim. MY[3], Shin. DH[3,4], Ahn. JJ[2], Ha. EH[2], Lee. SD[1,2]

[1]Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Korea
[2]Department of Information Statistics, Yonsei University, Wonju, Korea
[3]Institute of Forensic Science, Seoul National University College of Medicine, Seoul, Korea
[4]Department of Anatomy, Seoul National University, College of Medicine, Seoul, South Korea
e-mail : jh714@snu.ac.kr

Prediction of ethnicity in Asia can be approached in a variety of ways, but the method using Ancestry Informative Markers (AIMs) provides additional information. The HID-Ion AmpliSeq™ Ancestry Panel is a forensic multiplex platform consisted of 165 autosomal markers designed to provide biogeographic ancestry information.
In present study, we have investigated 750 unrelated Asians, from southern China (n=99), Beijing (n=100), Japan (n=101), Korean (n=100), Vietnam (n=100), Nepal (n=100), India (n=51), and Pakistan (n=99). The Torrent Server and the HID SNP genotyper plugin provide the calculated ethnicity probability and likelihood ratio. However, a variety of statistical approaches are needed when considering that Asians are closely related geographically and historically and there is not enough data available. For this reason, several statistical techniques have been tried and compared with the results provided by the plugin. Also, we applied various statistical algorithms for ethnic that classify Northeast Asian, Southeast Asian (Vietnamese) and Southwest Asian using SNP data from panel.
This research is meaningful in terms of the sub-classification of Asian people and applicability without major changes whenever new population is added. Furthermore, if further research is continued from

the viewpoint of usability when considered together with Lineage markers, it is expected that it will provide more comprehensive information.

1. KK Kidd, WC Speed, AJ Pakstis et al: Progress toward an efficient panel of SNPs for ancestry inference. Forensic Science International: Genetics. 2014; 10: 23–32
2. P Larranaga, B Calvo, R Santana et al: Machine learning in bioinformatics. Briefings in bioinformatics. 2006;7(1): 86-112.

20. INCREASING THE NUMBER OF RAPIDLY MUTATING (RM) Y-STRS SIGNIFICANTLY IMPROVES MALE RELATIVE DIFFERENTIATION

A. Ralf, N. Kousouri and M. Kayser

Department of Genetic Identification, Erasmus MC University, Rotterdam, The Netherlands

Y-chromosomal Short Tandem Repeats (Y-STRs) are valuable markers in forensic casework, particularly suitable for solving sexual assault cases, where the excess of female DNA often obscures male perpetrator identification when using standard autosomal DNA markers. Commonly used Y-STR sets, including those available via commercial kits, provide high, but not maximal, differentiation of unrelated men, and typically cannot differentiate between related men. Aiming to overcome this limitation of forensic Y-STR analysis, we previously introduced and further characterized a set of 13 Rapidly Mutating (RM) Y-STRs. In a recent multicenter study, these 13 RM Y-STRs achieved male lineage differentiation of >99% and male relative differentiation of >29% on average, which is superior to any other Y-STR set currently available. Aiming to find new, additional RM Y-STRs for further increasing male relative differentiation, particularly regarding closely related men, we performed in-silico search of the human Y-chromosome sequence for suitable candidate markers. We developed multiplex systems for the most promising candidate markers, and genotyped them in various pedigree and family samples to estimate relative differentiation rates. Our newly identified set of RM Y-STRs performed similarly well as our previously introduced RM Y-STR set. Considering RM Y-STRs from both sets together, we achieved a significant improvement of male relative differentiation, including close relatives, compared to previous findings. We therefore present a big step towards male individualization from DNA by keeping with the forensic advantage of Y-STR analysis to solve cases involving mixed male-female traces.

21. GENOMICS APPROACHES FOR CHARACTERIZING BIOLOGICAL DIVERSITY: FROM IDENTIFYING SPECIES TO MONITORING WHOLE ECOSYSTEMS

Keynote Speech

Mehrdad Hajibabaei

Integrative Biology, University of Guelph, Canada
hajibabaei@gmail.com

Rapid advancements in genome technologies and computational tools have revolutionized bio-medical sciences. The sheer volume of sequence information generated by a single DNA sequencing platform

has significantly increased in less than a decade. Most of this sequence information encompass a small number of organisms with human and their disease models such as mouse being among the most sequenced species of eukaryotes. Aside from assembling a complete (whole) genome sequence from an organism, most investigations focus on specific segments of the genomes. For example, research has shown the usefulness of comparative genomics for identification of units of biological diversity. Similar to DNA fingerprinting approaches that are common in forensic analysis, comparative analysis of a small number of genes can provide a basis for taxonomic identification of all species from microbes to mammals. Over the last decade this "DNA barcoding" approach has gained much momentum and its global sequence database includes over 5M records from 0.5M species. Consequently, the use of sequence information in identification of species has aided challenging problems such as environmental assessment, authentication of food and natural health products, forensic investigations as well as detection of pests, pathogens and species at risk. Additionally, bulk DNA extracted from environmental samples such as soil or seawater have been shown to provide comprehensive biodiversity information at ecosystem level. The widespread availability of genomics technologies and their automation-friendly nature further aid their rapid adoption by various stakeholders from academic researchers to regulatory agencies to any person interested. The integration of genomics in biodiversity analysis will transform our ability to detect and understand biological diversity of any group of organisms at any setting. This will set the stage for a bio-literate society.

22. 'FORCYT' DNA DATABASE OF WILDLIFE SPECIES

Ross McEwing[1], A. Linacre[2], L.M.I. Webster[3], A. Kotze[4] and J. Rovie Ryan Japning[5]

[1] *TRACE Wildlife Forensics Network, ross.mcewing@tracenetwork.org*
[2] *School of Biological Sciences, Flinders University, Adelaide, Australia*
adrian.linacre@flinders.edu.au.
[3] *Wildlife DNA Forensic Unit, SASA, Edinburgh, United Kingdom Lucy.Webster@sasa.gsi.gov.uk*
[4] *National Zoological Gardens, Pretoria, South Africa antoinette@nzg.ac.za*
[5] *National Wildlife Forensics Laboratory, Department of Wildlife and National Parks, Kuala Lumpur, Malaysia*
jeffrine@wildlife.gov.my

The illegal trade in highly endangered species continues unabated. High profile species such as rhinoceros, elephant and tigers are on the brink of extinction in many parts of the world. Robust and reliable forensic science based on DNA sequence comparisons are required to enforce national legislation.

We report on the establishment of ForCyt, a DNA database to aid with prosecutions of illegally traded species. We will present in detail the need for ForCyt and how it will assist in wildlife forensic science.

Current practice compares a questioned DNA sequence to a database. The largest database, GenBank, has a wealth of DNA data but is not regulated and rarely are the DNA sequences from voucher specimens.

ForCyt will be established based on regulated and checked DNA data with supporting information. This information will include: confirmation of correct species identification; up-loaded pictures are provided; any additional means of identification (stud book/microchip); date of collection; the geographic origin of the sample; preservation of sample; integrity of any packaging.

Information of the laboratory practice will be required including: persons performing the analyses; method of any DNA extraction performed; designation of primers used; site of DNA sequencing (in-house or out sourced); and whether the laboratory followed GLP or is accredited to a particular standard. All sequence data will be subject to checking for ambiguities and anomalies.

The outcome will be a repository of validated DNA sequence data to which others may make comparisons. Robust conclusions as to the species origin of samples can be made.

23. NON-HUMAN FORENSIC IDENTIFICATION OF COMPLEX MIXTURE SAMPLES USING DNA METABARCODING

Sungmin Kim and Hye Hyun Oh

Forensic Science Division II, Supreme Prosecutors' Office, Seoul, Korea

DNA identification is a powerful tool and has been widely used in forensic genetic investigations. Recently, the scope of DNA identification has been expanded with the introduction of massive-parallel sequencing (MPS) technology. DNA metabarcoding is a high-throughput approach of quickly and cheaply generating the barcodes of large number of species to be sequenced together in a single reaction. Several universal PCR primers can even be applied to simultaneously detect bacteria, animal, plant, and so on. In order to improve genetic identification of non-human forensic samples, we have developed a forensic database system to perform complex analyses of large datasets and translate the results of these analyses. The ForensicBOL (Forensic Barcode of Life) is an integral database system which analyzes and manages both DNA barcoding and metabarcoding data. Here we report two non-human forensic cases using DNA metabarcoding with universal DNA barcode markers. In the first case of the illegal drugs in a capsule form, we identified a variety of raw materials such as plants and animals (including humans). In the second case, known as the "Fake Back-Su-O" case, we detected a prohibited plant, *Cynanchum auriculatum*, and accurately calculated mixture ratio from the dietary supplements in the form of powders, dried products, and teas. DNA metabarcoding is a promising identification approach to detect prohibited and hazardous species in complex forensic samples. This approach for species identification would widely expand our ability of genetic screening in various non-human forensic cases.

24. CANINE STR MARKER GET MORE – DOG BREED AFFILIATION WITH A FORENSICALLY VALIDATED STR PANEL

B. Berger[1], C. Berger[1], J. Heinrich[1], H. Niederstätter[1], W. Hecht[2], A. Hellmann[3], U. Rohleder[3], U. Schleenbecker[3], N. Morf[4] , A. Freire-Aradas[5], C. Phillips[5], W. Parson[1,6]

[1] *Institute of Legal Medicine, Medical University of Innsbruck, Austria*
[2] *Institute of Veterinary Pathology, Justus-Liebig-University Giessen, Germany*
[3] *Bundeskriminalamt, Kriminaltechnisches Institut, Wiesbaden, Germany*
[4] *Institute of Legal Medicine, University of Zürich, Switzerland*
[5] *Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain*
[6] *Forensic Science Program, The Pennsylvania State University, PA, USA*
burkhard.berger@i-med.ac.at

In many aspects of human social life, dogs (Canis familiaris) play a major role and, therefore, they occupy a significant niche as a source of forensic samples. Dogs can cause accidents or act as perpetrators of attacks, but the transfer of canine DNA evidence that allows victims, suspects, and crime scenes to be linked, can be regarded as even more important. Typically – and in line with human DNA identification approaches – the identification of the "dog of interest" is mainly based on STRs. The Canine DNA Profiling (CaDNAP) Group, a long-standing collaboration between institutes in Austria, Germany and Switzerland, established a panel of 13 forensically validated STR-markers. This panel has repeatedly demonstrated its capability to identify individual dogs. However, in cases where DNA profiles do not match a specific dog, other sources of evidence such as breed membership may gain importance. Due to the genetic structure of dog breeds it is well established that such assignments based on STRs are generally feasible, but a proof of concept in the forensic context is still missing.
We tested the CaDNAP-STR-panel for breed assignments using samples from breeds which are

forensically relevant and common to Austria, Germany and Switzerland. Therefore, a survey of purebred and crossbred dog frequencies was performed in advance, as the analyzed samples should adequately mirror a realistic dog population. Statistical methods applied include Bayesian clustering, population assignments, and discriminant analysis of principal components (DAPC) provided by software programs such as “STRUCTURE” and the “adegenet” R package.

25. IDENTIFICATION OF HOAX AGENTS, *BACILLUS* AND *YERSINIA* STRAINS BY PROTEIN PROFILING USING MICROFLUIDIC CAPILLARY ELECTROPHORESIS WITH PEAK DETECTION ALGORITHMS

Sorelle Bowman[1], Casares de Cal Maria de los Angeles[2], Jose A. Alvarez Dios[2], Antonio Gomez-Tato[2], Paul Roffey[3], Alice Richardson[4], Dennis McNevin[1], Michelle E. Gahan[1]

[1] *National Centre for Forensic Studies, University of Canberra, Bruce ACT 2617, Australia*
[2] *School of Mathematics, University of Santiago de Compostela (Campus Vida), C/Lope Gomez de Marzoa s/n 15782, Santiago de Compostela,Spain*
[3] *Forensics, Specialist Operations, Australian Federal Police, PO Box 401, Canberra ACT 2601, Australia*
[4] *National Centre for Epidemiology & Population Health, Australian National University, Canberra ACT 0200, Australia*
sorelle.bowman@canberra.edu.au

Bacillus anthracis and *Yersinia pestis* are biological agents that pose an increasing concern to national security, if deliberately disseminated. Hoax agents including suspicious white powders and environmental bacterial species can also cause disruption. In either scenario it is of high importance to rapidly and accurately identify any suspicious powder as hazardous or hoax. This can be achieved using protein profiling by microfludic capillary electrophoresis as a field-based automated screening method. Two commonly encountered hoax agents (Dipel containing *B. thuringiensis* and plain flour), three *Bacillus* species (*B. anthracis* Sterne strain, *B. thuringiensis* Kurstaki strain and *B. cereus*), two *Yersinia* species (*Y. enterocolitica B1A S09* and *Y. pseudotuberculosis*) and *E. coli* were profiled using microfluidic capillary electrophoresis on the Experion™ System (Bio-Rad).
Peak detection algorithms allowed for the recognition and identification of protein peaks in raw electropherograms. Boolean logic gates were then employed to model and predict the electrophoretic protein pattern of species based on the presence or absence of indicative protein peaks in specified time windows. Parameters assessed included variation within and between Experion™ Pro260 chips and ability to identify and discriminate between species over time intervals, between operators and between field and laboratory analyses. The results indicate that this technique is rapid (less than 40 minutes), accurate and reproducible for automated front-line identification and differentiation of hoax, biological and environmental agents.

26. MICROBIOME COMMUNITIES ANALYSES TO PREDICT BODY SITES: ANALYSES ON PUBLISHED DATA AND ON EXPOSED SAMPLES IN THE LABORATORY

Janko Tackmann[2], Akos Dobay[1], João Rodrigues[2], Cordula Haas[1], and Natasha Arora[1]

[1] *Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland*
[2] *Institute of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, Zurich, Switzerland*

natasha.arora@irm.uzh.ch

Human microbiome sequencing studies have shown that microbial communities across body habitats are distinct, and therefore potentially useful for the forensic prediction of body sites from fluids or stains. However, studies examining body site classification methods have focused on datasets from single studies, and have generally examined single sources only, while forensic samples are frequently mixtures. We examined the performance of random forest classifiers on a large and heterogeneous dataset comprising publically available 16S rRNA gene sequences from 15,082 saliva, skin, nostril, vagina and feces samples from 57 studies. Additionally, we also produced *in silico* mixtures of sequences for all pairs of body sites, testing the prediction accuracy of the classifiers for different ratios. Our results indicate high prediction performances for the single source sites. We were also able to reliably identify both components in mixtures with fractions as low as 2% for the minor body site. These results are promising, and highlight the potential of microbial 16S rRNA gene sequencing for forensic applications. However, forensic samples are frequently exposed to environmental conditions and therefore subject to degradation. Therefore, we have also initiated a study investigating the differences in the microbial compositions of a total of 36 skin, saliva, menstrual blood, peripheral blood, vaginal fluid and semen samples that have been exposed in the laboratory for 4 weeks. We highlight the main differences across freshly extracted and exposed samples, and the predictive accuracy of our classification methods.

27. MICROHAPLOTYPE: ABILITY OF PERSANAL IDENFICATION AND BEING ANCESTRY INFORMTIVE MARKER

Yan Pu[1], Peng Chen[1], Jing Zhu[1], Youjing Jiang[1], Qingqing Li[2], Tao Feng[2], Weibo Liang[1] and Lin Zhang[1]

[1]*Department of Forensic Genetics, West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, Sichuan, China.*
[2]*Department of Forensic Genetics, Institute of Forensic Science, Chengdu Public Security Bureau, Chengdu 610081, Sichuan, China.*
liangweibo@gmail.com.

Microhaplotype is a new type of genetic marker. Due to the multiallelic characteristic, microhaplotype could have much better personal identification ability compared to the bi-allelic SNPs. Microhaplotype could also have alleles associated with specific populations. With this research, phased data from 1000 genomes project were used to evaluate the ability as personal identification and ancestry informative marker. Totally 32 previously selected microhaplotype loci and 568 unrelated individuals from six region around the globe were enrolled. Including Africans, Europeans, East Asians, Americans. Specifically from African Caribbeans in Barbados (ACB), Americans of African Ancestry in SW USA (ASW), Northern and Western European Ancestry (CEU), Han Chinese in Beijing, China (CHB), Southern Han Chinese (CHS) and Japanese in Tokyo, Japan (JPT). The range of SNPs number per locus among the 32 loci has a range from 3 to 10. The allele number range of the 32 microhaplotpype is 3 to 26. The median heterozygosity for these 32 loci is 0.6717 for the 6 populations studied and ranges from 0.5461 to 0.8693. To assess the power as AIM, we applied standard population analysis approaches of the Nei's genetic distance and STRUCTURE analysis. The genetic distance generated by 32 loci demonstrated that ACB and ASW as well as CHB and CHS were conglomerated together. 6 loci were picked out to run the STRUCTURE analysis for an optimum K:4. Visual inspection of cluster plots suggests a close match between the co-ancestry patterns seen in ACB and ASW as well as in CHB, CHS and JPT.

28. MICROHAPLOTYPE FOR ANCESTRY PREDICTION

D. Podini[1], K. Long[1], K. Maddela[1], S. Cisana[1], M. Schanfield[1], L. Bennett[1], S. Wootton[2], J. Chang[2], R. Lagace[2], K. Kidd[3].

[1] The George Washington University, Washington, DC 20007
[2] Thermo Fisher Scientific, South San Francisco, CA 94080
[3] Yale University, New Haven, CT 06520
podini@gwu.edu

Microhaplotypes (MHs) are loci of two or more SNPs within a short distance from each other (<300 nucleotides) with three or more allelic combinations1. Massively parallel sequencing (MPS) methods, by allowing clonal sequencing of individual strands, can distinguish the parental haplotypes at a locus. MH alleles within a locus all have the same size, do not generate stutter fragments, and have lower mutation rate than STRs. The goal of this project was to evaluate if MHs could provide biogeographic ancestry prediction.
One hundred (100) European-Americans (EAs), 50 African-Americans (AAs), and 92 Southwest Hispanics (SWHs), were analyzed with a 33 MHs panel on Ion S5™ platform. PHASE inferred allele frequencies from 58 populations2 were used to calculate the random match probability (RMP) of each profile in each population. The RMP averaged higher in the populations individuals self-identified as. For example, AAs had the highest RMP in African and African-American populations.
Likelihood ratios (LR) were calculated by dividing the highest RMP obtained from the three US populations by the second highest in that same set. The number obtained represented how much more likely it is to observe that profile if the individual is of the population at the numerator vs the one at the denominator. All individuals were predicted of the correct population except for 38% of SWHs, which were not surprisingly predicted as EAs given it is an admixed population. The 33 MHs were not selected for informativeness, more specifically selected markers will likely improve the prediction's accuracy.

1. Genetic markers for massively parallel sequencing in forensics. Kidd, Kenneth K. et al. Forensic Science International: Genetics Supplement Series , Volume 5 , 677-679
2. Evaluating 130 microhaplotypes across a global set of 83 population. Kidd, Kenneth K. et al. Forensic Science International: Genetics , Volume 29 , 29 - 37

29. THE APPLICATIONS OF MALAYSIAN POPULATIONS PLOTTED VIA NEXT GENERATION SEQUENCING

Nor Aidora Saedon, Sharifah Nur Hafiza Syed A Rahman, Normazlina Zainuddin, Mohd Firdaus Che Amran and Zulhilmi Husni

Forensic DNA Section, Forensic Division, Department of Chemistry, Jalan Sultan, 46661, Petaling Jaya, Selangor, Malaysia
dora@kimia.gov.my

Next Generation Sequencing (NGS) is a brand new approach in forensic DNA analysis where information contained in DNA evidence is materialized such STRs, Phenotypic SNPs such as hair colour, eye colour and skin colour; ancestry and also identity SNPS which is suitable for DVI cases are utilized to identify unknown criminals. This technique generates much more data compared to the current 16 loci adapted in our laboratory. Malaysian population originated from three main ethnicities which is the Malaysian Malay, Malaysian Chinese and Malaysian Indian were analyzed using the MiSeq FGx Next Generation Sequencing.
In May 2015, Malaysia's Wang Kelian villagers were shocked by discovery of 139 mass graves found in jungles of Perlis near the Thailand border, believed they were the victims of human trafficking

syndicates. The remains removed from the jungle were a highly decomposed body that had been left out there for over couple of weeks. A total of 122 mass grave samples which consist of 119 bone samples and 3 teeth samples were received. The results showed that mass grave samples were not clustered in either Malaysian Malay or Malaysian Chinese population but originated from Indian population. However, after analyzing the reference sample of Rohingyas individuals the results indicated that these mass grave samples are likely Rohingyas rather than Malaysian.

30. COMPLETE MITOCHONDRIAL GENOMES OF THE HUNTER-GATHERER POPULATIONS IN THAILAND

W. Kutanan[1], J. Kampuansai[2], D. Kangwanpong[2], M. Stoneking[3]

[1] *Department of Biology, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand*
[1] *Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand*
[3] *Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*
e-mail: wibhu@kku.ac.th

Hunter-gatherers are generally assumed to be maintaining the genetic lineages of anatomically and behaviorally modern humans who dispersed out of Africa at least 50,000 years ago. Part of the rationale of the assumption is the presence of basal mitochondrial (mt) DNA lineages in many hunter-gatherer populations outside Thailand, for example in the Andaman Islands, Malaysia, and the Philippines. There are two hunter-gatherer groups, i.e., the Mlabri and the Sakai, living in northern and southern Thailand, respectively. We sequenced full mitochondrial genomes of 18 and 11 unrelated subjects of Mlabri and Mani. Only 2 haplotypes differing by a single polymorphic site are found in the Mlabri, and all sequences belonged to haplogroup B5a1b1. We found 5 haplotypes in the Mani belonging to 3 basal haplogroups, that is, M17a, M21a and R21. The reduced genetic diversity of these two populations results in increased genetic differentiation when compared to other populations from Thailand. The coalescent age of the two Mlabri haplotypes is 735 years, suggesting that the Mlabri have reverted to a hunter-gather lifestyle, in agreement with a previous study (1). However, the Mani share maternal lineages with other hunter-gatherer populations from outside Thailand, suggesting that they are indigenous hunter-gatherers.

1. Oota H, Pakendorf B, Weiss G, von Haeseler A, Pookajorn S, Settheetham-Ishida W, Tiwawech D and Ishida T and Stoneking M: Recent Origin and Cultural Reversion of a Hunter–Gatherer Group. PLoS Biology. 2005; 3(3): e71.

31. DNA TRANSFER AND CELL TYPE INFERENCE TO ASSIST ACTIVITY LEVEL REPORTING

Keynote Speech

T. Sijen

Netherlands Forensic Institute, The Hague, the Netherlands
t.sijen@nfi.minvenj.nl

Courts are often concerned with interpretation of evidence beyond the proposed identity of an individual, and questions such as ‘what cell material is present and how did it get there?’ need to be addressed [1]. RNA profiling has developed into an more and more accepted methodology to infer cell types present in evidentiary material. Two types of multiplex assays exist: those typing body fluids and

another targeting organ tissues. These assays find their predominant application in sexual assault and violent cases respectively, which will be illustrated by casework examples. Alternatively, DNA methylation may provide useful information to assist cell type inference, markedly when semen is involved. To provide context to observations, subjects such as background, transfer, persistence and recovery require data [2] for which an illustration will be given. Such information will assist interpretation for which commonly hypotheses-based approaches are used.

1. Sijen T: Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. Forensic Science International: Genetics. 2015; 18: 21-32.
2. van den Berge M, Ozcanhan G, Zijlstra S, Lindenbergh A, Sijen T: Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios. Forensic Science International: Genetics. 2016; 21: 81-89.

32. ASSOCIATION OF A BODY FLUID WITH A DNA PROFILE BY TARGETED RNA/DNA DEEP SEQUENCING

S. Ingold[1], C. Haas[1], G. Dørum[1], E. Hanson[2], J. Ballantyne[2, 3]

[1] Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland
[2] National Center for Forensic Science, University of Central Florida, Orlando, USA
[3] Department of Chemistry, University of Central Florida, Orlando, USA
sabrina.ingold@irm.uzh.ch

The identification of body fluids/tissues is an important part of forensic casework and can contribute valuable information about the circumstances of a crime.
Our goal for this study was, in addition to body fluid/tissue identification, to associate tissue specific mRNA transcripts to the donor of a stain. Within the tissue specific sequences we looked for sequence variants (SNPs) that discriminate Caucasian individuals the most. We successfully tested a targeted NGS assay for the Illumina MiSeq platform in 188 European individuals for the detection of the 35 selected SNPs on gDNA level. To assess the usefulness of these 35 markers we estimated how powerful the loci are to discriminate between individuals. The match probabilities for the SNPs within one body fluid and for sets of markers from different body fluids confirmed the discriminatory power of these markers. A separate targeted cSNP panel on RNA level showed good specificity for blood, semen and menstrual blood. For saliva, vaginal secretion and skin, the marker design needs to be optimized with special attention to cross reactivity with DNA contaminants. As a proof-of-principle we could demonstrate that with our targeted DNA and RNA NGS assays we are able to assign a body fluid to a specific individual. This is especially interesting for mixed body fluid samples, where we can find out who contributed which body fluid.

33. DIRECT PCR IMPROVES STR PROFILES FROM SUBSTRATES OF IMPROVISED EXPLOSIVE DEVICE

S. Dangsriwan, T. Kitpipit and P. Thanakiatkrai

Forensic Science Program, Prince of Songkla University, Hat Yai, Thailand
pthanakiatkrai@gmail.com

STR profiling from both pre-blast and post-blast IED evidence are generally unsuccessful due to the low-level of DNA present and high level of degradation. Previous studies have improved the number of alleles recovered by focusing on both the DNA collection process and the downstream processes of

DNA extraction and amplification; however, direct PCR has never been used with touch DNA on IED evidence. In this study, we optimized a direct PCR protocol for touch DNA profiling from various IED substrates. Two swab types with various moistening agents and one type of tapelifting were compared as the collection methods. Different amounts of cuttings from swab heads and tapes were evaluated. Two direct PCR protocols – the direct protocol and the pre-PCR protocol – were performed. As expected, different substrates (i.e. clothes, PVC pipe, circuit board, and electrical tape) required different collection methods and cuttings for optimal DNA recovery and amplification. The direct protocol recovered significantly more alleles than the pre-PCR protocol ($p<0.05$). The optimal direct PCR methods were compared with our standard operating protocol (conventional DNA extraction and STR typing) using mock IED evidence. This improvement could be due to the omission of the inefficient DNA extraction step, which in turn reduces the cost of STR profiling. The result showed that more alleles were recovered using the optimal direct PCR methods ($p<0.05$). The findings of this study suggest that direct STR typing could be used IED evidence, provided that an appropriate, optimal DNA collection method is used.

34. THE DEVELOPMENT OF A FORENSIC CLOCK TO DETERMINE TIME OF DEATH

L. Sibbens[1], E. Cuypers[2], W. Van de Voorde[1,3], B. Bekaert[1,3]

[1] *Forensic Biomedical Sciences, Department of Imaging & Pathology, KU Leuven – University of Leuven, Leuven, Belgium*
[2] *Toxicology and Pharmacology, Department of Pharmaceutical and Pharmacological Sciences, KU Leuven – University of Leuven, Leuven, Belgium*
[3] *Department of Forensic Medicine, Laboratory of Forensic Genetics and Molecular Archaeology, University Hospitals Leuven, Leuven, Belgium*
Lode.sibbens@kuleuven.be

The time and date of a person's death are two major questions police detectives and investigative judges ask a forensic pathologist as it provides a time frame in which the investigators can base their further investigations on. The current time of death (ToD) estimation techniques produce large error ranges since environmental and endogenous factors have to be taken into account. Too large prediction intervals for ToD estimation are of little use to the investigators though at present no tools are available to limit the ranges offered.
Circadian biomarkers (cortisol, melatonin and 3 mRNAs) have already proven their potential use as forensic molecular clock markers to estimate blood-deposition time [1], and intriguingly circadian biomarkers identified from the brain transcriptome were previously shown to be able to estimate ToD with a ± 1.9h error range [2]. The application of circadian biomarkers to estimate ToD would therefore mean a significant improvement in accuracy compared to current techniques, especially after the first 24h after death, as error ranges increase exponentially with increasing postmortem interval.
In order to develop a forensic molecular clock for postmortem blood samples we designed a study to identify circadian markers in three different classes of molecules using postmortem blood samples obtained from deceased individuals with a known ToD (±30 min): circadian transcripts with RNA-seq, circadian metabolites with LC-MS/MS and melatonin and cortisol concentrations using ELISA kits.
During this presentation we will discuss our initial results and demonstrate the applicability of circadian biomarkers for the determination of time of death.

1. Lech K, Liu F, Ackermann K, Revell VL, Lao O, Skene DJ, et al. Evaluation of mRNA markers for estimating blood deposition time: Towards alibi testing from human forensic stains with rhythmic biomarkers. Forensic science international Genetics.2015;21:19-25
2. Li JZ, Bunney BG, Meng F, Hagenauer MH, Walsh DM, Vawter MP, et al. Circadian patterns of gene expression in the human brain and disruption in major depressive disorder. Proc Natl Acad Sci U S A. 2013;110(24):9950-5

35. LATENT DNA DETECTION USING FLUORESCENT DNA BINDING DYES

A.M. Haines[1], H. Kobus[2] and A. Linacre[1]

[1] *School of Biological Sciences, Flinders University, Adelaide, Australia*
[2] *School of Chemical and Physical Sciences, Flinders University, Adelaide, Australia*
Alicia.haines@flinders.edu.au

We report on the use of DNA binding dyes that can target the collection of DNA present on samples. Dyes are available that bind to DNA at high specificity for laboratory-based applications but rarely applied to *in situ* detection.
Common nucleic acid-binding dyes were selected due to their increase in fluorescence when in the presence of DNA (SYBR® Green I, Diamond™ Nucleic Acid Dye, GelGreen™, GelRed™, EvaGreen™ and Redsafe™). The fluorescence from dye/DNA complex was detected using a high intensity light source, the Polilight® (PL500), and emission observed through band pass interference filters with a 40nm band width and central wavelength of 530 nm or 550 nm depending on the dye emission. Some biological samples such as hair and skin were visualised under a fluorescent microscope (Nikon Optiphot) using a B2A filter cube.
Detection of DNA was observed within different biological samples such as saliva, skin, blood and hair which make it possible to select samples that are more likely to produce STR profiles after direct amplification. The use of these dyes as a screening methodology for hairs was conducted where the fluorescent signal was correlated with the quality of STR profile obtained.
The outcome of this work is an innovative means to detect DNA *in situ* within biological samples and on surfaces that make the screening of samples more efficient and successful. The investigation so far has concluded that EvaGreen™ and Diamond™ dye are the optimum dyes for this novel application based on their properties of binding and limited interactions with downstream forensic applications such as DNA extraction, amplification and STR typing.

36. ASSESSMENT OF THE PREVALENCE, TRANSFER, PERSISTENCE AND RECOVERY OF DNA TRACES FROM WORN UPPER GARMENTS: RESULTS FROM AN INTER-LABORATORY STUDY

B. Szkuta[1,2], R. Ansell[3,4], L. Boiso[3], E. Connolly[5], A.D. Kloosterman[6,7], B. Kokshoorn[7], L.G. McKenna[5], K. Steensma[7], R.A.H. van Oorschot[2]

[1] *School of Life and Environmental Sciences, Deakin University, Geelong, Australia*
[2] *Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod, Australia*
[3] *Swedish National Forensic Centre, Linköping, Sweden*
[4] *Department of Physics, Chemistry and Biology (IFM), Linköping University, Linköping, Sweden*
[5] *Forensic Science Ireland, Garda Headquarters, Phoenix Park, Dublin, Ireland*
[6] *Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*
[7] *Division Biological Traces, Netherlands Forensic Institute, The Hague, The Netherlands*
bszk@deakin.edu.au

Clothing items are commonly recovered from crime-scenes and submitted for DNA analysis and subsequent interpretation. Depending on the case circumstances, sampling of the garment may target the wearers DNA or the transferred DNA of one or more offenders. In addition to collecting targeted DNA, traces of background DNA from sources unrelated to the crime event, may also be inadvertently collected. This has the potential to raise questions regarding the relevance of traces. To gain an

understanding of the prevalence, transfer and persistence of DNA on clothing after normal washing and wearing scenarios, samples were collected from several areas of upper garments worn by individuals on working or non-working days, and processed from DNA extraction through to profiling. Activities relating to the garment prior to and during wearing were recorded by the wearer through questionnaires. Reference profiles were obtained from the wearer and their close associates identified within these questionnaires. This was repeated for sixteen individuals across four countries, and the impacts of differences among collaborating laboratories in sample collection through to profiling were considered during the interpretation and analysis of the 400+ profiles generated. Variations in the number of contributors, the composition of the profiles, and inclusion/exclusion of the wearer and their close associates were observed between the collaborating laboratories, participants, garments worn on different occasions, garment types and garment areas sampled. Impacts of activities undertaken during wearing were also observed.

37. DNA TRANSFER STUDY BASED ON PARALLEL SYSTEM METHODS

Qingzhen Meng[1], Chong Wang[1,2], Zheng Tu[1,2] and Mechthild Prinz[3*]

[1] *Institute of Forensic Science, Ministry of Public Security, Beijing, P.R.China*
[2] *Key Laboratory of Forensic Genetics, Ministry of Public Security, Beijing, P.R.China*
[3] *John Jay College of Criminal Justice, New York, United States*
mengqz@cifs.gov.cn

With the ability to detect small amounts of DNA, just a few cells, the occurrence of DNA transfer challenges the probative value of this type of evidence [1]. Several studies have demonstrated factors affecting the occurrence and probability of DNA transfer (for example 2), but these data sets cannot be easily generalized and used to assess transfer in specific court cases. We are proposing the use of complex system modeling to systematically re-examine DNA transfer processes and related research data, establishing relevant parameters. A crime is essentially an open complex system of interactions in an environment, thus the DNA transfer process, which presents spatial-temporal uncertainty and behavioral complexity, can be modeled as such. The ACP approach [3], a combination of artificial societies (A), computational experiments (C), and parallel execution (P), can be used to create a parallel DNA transfer system and model the different factors and interactions. This method has the potential to provide a strong scientific basis and decision support when dealing with DNA transfer issues.

1. Peter Gill (2016) Analysis and implications of the miscarriages of justice of Amanda Knox and Raffaele Sollecito. FSI: Genetics; 3:9-18.
2. Goray et al (2010) Secondary DNA transfer of biological substances under varying test conditions. FSI: Genetics; 4:62-67.
3. Wang Feiyue (2013) Parallel Control: A Method for Data-Driven and Computational Control. Acta Automatica Sinica; 39(4): 293-302.

38. CPGFORID: EPIGENETIC DISCRIMINATION OF IDENTICAL TWINS UNDER THE FORENSIC SCENARIO

A. Vidaki[1], E. Carnero-Montoro[2], R.G. Boers[3], J.B. Boers[3], C. Díez López[1], P. Kalamara[1], T. Spector[2], J. Gribnau[3], J.T. Bell[2], and M. Kayser[1]

[1] *Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, The*

Netherlands
[2] Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom
[3] Department of Developmental Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands
a.vidaki@erasmusmc.nl

Standard DNA profiling is typically non-informative for differentiating between monozygotic (MZ) twins. Whole genome sequencing was recently proposed as a solution; however, both the rarity of such autosomal SNPs and required ultra-high coverage resulting in immense costs, make this approach impractical. Epigenetic variation was identified as suitable source for MZ twin differentiating biomarkers; nevertheless, studies addressing the forensic feasibility are lacking. Here, we performed the first use of DNA methylation for MZ twin differentiation under the forensic scenario, comprising the i) discovery of twin-differentially methylated sites (tDMSs) in reference-type DNA via genome-wide analysis, ii) validation of candidate tDMSs using forensically suitable, targeted methods, and iii) analysis of validated tDMSs in trace-type DNA. Whole blood and buccal cells of 10 MZ twin pairs were analyzed together with forensic-type material including minute bloodstains and cigarette butts. Genome-wide analysis using HumanMethylation450 BeadChips (Illumina) covering 2% of all CpGs, identified hundreds of twin pair-specific tDMSs. Analyzing the top tDMSs per pair using methylation-specific quantitative (q)PCRs in reference-type DNA confirmed many, but not all, markers. Testing the validated tDMSs in trace-type DNA revealed the observed methylation differences for many, but not all, markers. Demonstrating that a proportion of tDMSs selected from reference-type samples eventually becomes non-informative in trace-type samples questions the number of markers needed for successful MZ twin differentiation and identification. Aiming to overcome restrictive marker selection via epigenome screening of both reference and trace-type samples, we applied a novel, highly sensitive enzymatic-based methylation DNA sequencing approach (MeD-seq) covering ~50% of all CpGs.

39. APPLICATION OF DIP-STRS TO SEXUAL/PHYSICAL ASSAULT INVESTIGATIONS: EIGHT CASE REPORTS

D. Hall, F. Oldoni, and V. Castella

Forensic Genetics Unit, University Center of Legal Medicine Lausanne - Geneva, Switzerland
Diana.Hall@chuv.ch

DIP-STRs are compound markers formed by a deletion/insertion polymorphism (DIP) linked to a microsatellite (STR). They enable the deconvolution of unbalanced DNA mixtures from two individuals, up to 1,000-fold excess of the major contributor (1-2). Here, we report on the first use of these new markers in casework to discuss their advantages and limitations in forensic practice.
Trace samples suggestive of containing unbalanced DNA mixtures were selected from eight cases with the permission of the Police and General Prosecutor authorities. DIP-STR results contributed to each case, either by complementing Y-STRs results or by producing novel investigative leads. This was especially the case when considering same sex unbalanced DNA mixtures, female minor/male major unbalanced DNA mixtures or to discriminate two brothers in a sample containing high levels of female DNA. Positive results were obtained at 16,000-fold excess of major DNA using few picograms of input DNA as well as from traces collected several months after an aggression. Likelihood ratios assigned to DIP-STRs' results were modest (10), when based on two markers, and high (106) when determined on six markers. In some cases the detection of extra alleles from additional minor DNA contributors or the observation of stochastic effects with low-template DNA samples, limited the interpretation of the results. In conclusion, the DIP-STRs often provide additional value to the analysis of traces that cannot be exploited by the use of standard methods.

1. Castella V, Gervaix J, Hall D. DIP-STR: highly sensitive markers for the analysis of unbalanced genomic mixtures. Human Mutation 34:644-54 (2013)
2. Oldoni F, Castella V, Hall D. A novel set of DIP-STR markers for improved analysis of challenging

40. ADVANCES IN HUMAN APPEARANCE GENETICS AND RELEVANCE FOR FORENSIC DNA PHENOTYPING

Keynote Speech

M. Kayser

Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands
m.kayser@erasmusmc.nl

Retrieving appearance details of an unknown trace donor from his/her DNA left behind at a crime scene can be useful information to focus police investigation towards finding unknown perpetrators of crime, who typically cannot be identified via standard DNA profiling. Likewise, establishing appearance information from DNA obtained from human remains can be useful for finding potential relatives in missing person cases where no ante-mortem samples or relatives are available. Over recent years, this notion has led to the establishment and further development of Forensic DNA Phenotyping, where phenotypes in the closest sense refer to externally visible characteristics, and in the wider sense additionally include life time age and bio-geographic ancestry also providing investigative information. Currently it is possible, based on genetic knowledge generated and used to develop and forensically validated suitable genotyping and statistical tools, to predict categorical eye, hair, and skin color from crime scene DNA with accuracies deemed suitable for practical applications. Recent activities towards understanding the genetic basis of human appearance comprehensively concern pigmentation traits on more detailed level, various hair traits including head hair structure, head hair loss particularly in men, eye brows, as well as some other externally visible traits. This keynote speech will summarize recent progress in finding genes underlying human appearance and will discuss their relevance in the context of Forensic DNA Phenotyping.

41. EPIGENETIC AGE SIGNATURES IN SALIVA: AGE PREDICTION USING METHYLATION SNAPSHOT AND MASSIVE PARALLEL SEQUENCING

S.R. Hong[1,2], S.E. Jung[1], E.H. Lee[1], K.J. Shin[1,2], W.I. Yang[1], H.Y. Lee[1]

[1] *Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea*
[2] *Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea*
srhong@yuhs.ac

DNA methylation is one of the most promising markers of age prediction. Many DNA methylation-based age predictive models have been developed based on DNA methylation patterns from blood. However, few studies have aimed to predict age from saliva, which can be frequently found at crime scenes. In this study, we generated genome-wide DNA methylation profiles of saliva from 54 males and identified 6 CpG markers on the *SST*, *CGNA3*, *KLF14*, *TSSK6*, *TBR1*, and *SLC12A5* genes that showed a high correlation between methylation and age. With 226 bisulfite-converted saliva DNA samples, we investigated DNA methylation at 6 age-associated CpGs as wll as a cell type-specific CpG from the *PTPN7* gene using the methylation SNaPshot method. Then we constructed an age-predictive model with the age information and the methylation profiles from the 113 training samples. The model showed

a correlation between predicted and chronological age of more than 90%, and a mean absolute deviation from chronological age (MAD) of 3.13 years. Subsequently, the validation set composed of the remaining 113 samples presented a 95.2% correlation between the predicted and chronological age, and a MAD of 3.15 years. Our model based on DNA methylation profiling using multiplex methylation SNaPshot is expected to be easily integrated into the routine forensic laboratory workflow. Also, since the suggested 7 CpG sites could be applied to other platforms through appropriate modeling, we additionally analyzed our proposed CpG sites using Massive Parallel Sequencing and tested age predictability using a newly trained model.

1. Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, Vilain E, Epigenetic predictor of age, PLoS One. 2011; 6: e14821.
2. Eipel M, Mayer F, Arent T, Ferreira MR, Birkhofer C, Gerstenmaier U, Costa IG, Ritz-Timme S, Wagner W, Epigenetic age predictions based on buccal swabs are more precise in combination with cell type-specific DNA methylation signatures, Aging (Albany NY). 2016; 8: pp. 1034–1048.

42. EPIGENETIC PREDICTION OF AGE IN DIFFERENT STUDY GROUPS

M. Spólnicka[1], E. Pośpiech[2,3], B. Pepłońska[4], R. Zbieć-Piekarska[1], Ż. Makowska[1], A. Pięta[1], A. Freire-Aradas[5], C. Phillips[5], M.V. Lareu[5], J. Karłowska-Pik[6], B. Ziemkiewicz[6], M. Wężyk[4], P. Gasperowicz[7], T. Bednarczuk[8], M. Barcikowska[4], C. Żekanowski[4,9], R. Płoski[7], W. Branicki[1,3]

[1]*Central Forensic Laboratory of the Police, Warsaw, Poland*
[2]*Department of Genetics and Evolution, Institute of Zoology, Jagiellonian University, Krakow, Poland*
[3]*Malopolska Centre of Biotechnology of the Jagiellonian University, Krakow, Poland*
[4]*Department of Neurodegenerative Disorders, Mossakowski Medical Research Center, Polish Academy of Sciences, Warszawa, Poland*
[5]*Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain*
[6]*Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Toruń, Poland*
[7]*Department of Medical Genetics, Centre for Biostructure, Medical University of Warsaw, Warsaw, Poland*
[8]*Department of Internal Medicine and Endocrinology, Medical University of Warsaw, Warsaw, Poland*
[9]*Department of Sport's Theory, Jozef Pilsudski University of Physical Education in Warsaw, Warszawa, Poland*
ewelina.pospiech@uj.edu.pl

Insight into the aging rate of individuals may provide a more precise prediction of progressive appearance traits. DNA methylation is considered to be a measure of the cumulative effect of the body's epigenetic maintenance system, suggesting that the aging process may be accelerated by many factors affecting epigenetic stability. Therefore, identification of these factors could be important for the accurate prediction of age, particularly in older individuals, where epigenetic prediction is more prone to error. On the other hand, some studies suggested different sensitivity of DNA methylation loci to environmental factors and showed that, e.g. ELOVL2 methylation increases with cell divisions and to a lesser degree is influenced by other factors. In the present study we address the problem of age prediction accuracy in different individuals by testing DNA methylation measured with pyrosequencing technology using the predictive capacity of five age-related markers (ELOVL2 c7, C1orf132 c1, FHL2 c2, TRIM59 c7, KLF14 c1). Analysis involved several groups including patients of Early Onset Alzheimer's Disease, Late Onset Alzheimer's Disease, patients with Graves' Disease and a group of professional athletes. Our study showed altered DNA methylation patterns and prediction capacity in the tested groups when they were compared to healthy controls in the three markers: TRIM59, KLF14 and FHL2; suggesting their sensitivity to various factors affecting methylation age of an individual. Two important DNA methylation markers: ELOVL2 and C1orf132 were found to correctly predict age in all the tested groups, confirming their power and robustness as chronological age predictors in forensics.

43. ASSOCIATION BETWEEN BMP4 GENE POLYMORPHISMS AND EYELID TRAITS IN CHINESE HAN POPULATION

Qian Wang[a,1], Bo Jin[a,1,2], Xiaoying Luo[1], Tao Feng[3], Yu Xia[3], Weibo Liang[1] and Lin Zhang[1]

[1] *Department of Forensic Genetics, School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, Sichuan, China*
[2] *Department of Forensic Medicine, North Sichuan Medical College, Nanchong 637000, Sichuan, China*
[3] *Department of Forensic Genetics, Institute of Forensic Science, Chengdu Public Security Bureau, Chengdu 610081, Sichuan, China*
[a] *These authors contributed equally to this work*
zhanglin@scu.edu.cn

Forensic DNA phenotyping (FDP) refers to the prediction of unknown sample donors' externally visible characteristics (EVCs) using only genetic information from extracted DNA. And several phenotypic trait estimations have been applied to forensic casework. FDP for single/double eyelids could be quite useful in forensic practice, especially in East Asia. To evaluate the association between polymorphisms of BMP4 gene and eyelid traits, a case-control study including 611 Chinese Han adults was conducted and two candidate SNPs were genotyped using PCR-RFLP. Significant differences were found for rs2761882 CT and CT/TT genotypes in single and double eyelids populations ($p<0.05$); no difference was observed for rs762642 ($p> 0.05$). Single eyelid was further divided into four subgroups: normal eyelid (123 cases), eyelids with thin and soft skin (28 cases), eyelids with thick subcutaneous tissue (28 cases) and eyelids with triangle eyes (27 cases). Typical double eyelids also had four types, including moderate (140 cases), narrow (106 cases), wide (24 cases), a wide and a narrow (42 cases) types. The classification for untypical double eyelids referred to typical double eyelids: moderate (34 cases), narrow (19 cases), wide (22 cases), a wide and a narrow (18 cases) types. Stratification analysis showed significant difference of rs2761882 CT genotype frequency between a wide and a narrow untypical double eyelids and the other subgroups. Significant difference of rs762642 GG genotype frequency also has been found between a wide and a narrow typical double eyelids and the other types. Further analysis is still in process.

44. DISCRIMINATING BETWEEN UNCLE-NEPHEW AND GRANDFATHER-GRANDSON KINSHIPS BY ANALYZING CHROMOSOMAL SHARING PATTERNS

Chie Morimoto[1], Sho Manabe[1], Shuntaro Fujimoto[1], Yuya Hamano[1, 2], Keiji Tamaki[1]

[1] *Department of Forensic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan*
[2] *Forensic Science Laboratory, Kyoto Prefectural Police Headquarters, Kyoto, Japan*
morimoto@fp.med.kyoto-u.ac.jp

We developed a new method for pairwise kinship analysis using the index of chromosomal sharing from high-density single nucleotide polymorphisms (SNPs) that was established in our previous study [1]. Using the previous method, we could determine accurate degrees of kinship up to the third-degree in cases where no predicted relationship existed between two individuals. However, discriminating between kinships of the same degree (e.g., uncle-nephew and grandfather-grandson) was difficult. In the present study, we examined the differences in chromosomal sharing patterns between uncle-nephew and grandfather-grandson kinships.

We computationally generated genotypes of 174,254 autosomal SNPs in 249 uncle-nephew and grandfather-grandson pairs, while considering the effect of linkage disequilibrium among each SNP. We examined the shared chromosomal segments, using our previous method [1], and counted them to investigate the differences in chromosomal sharing patterns. We then classified these chromosome

segments, using a receiver operating characteristic curve analysis. We confirmed the validity by using actual data.
The median values of the number of shared segments in uncle-nephew and grandfather-grandson kinships were 63 and 39, respectively. The area under the curve was 0.99954, indicating that the two relationships were clearly different. Moreover, the actual sample pairs showed similar results. These results suggest that shared chromosomal segments in collateral relationships are greater in number but smaller in length than those in lineal relationships with the same degree of kinship.

1. Morimoto C et al: Pairwise kinship analysis by the index of chromosome sharing using high-density single nucleotide polymorphisms. PLoS One. 2016; 11: e0160287

45. A NEW (APPROXIMATE) METHOD FOR Y-STR HAPLOTYPE PROBABILITY ASSIGNMENT

Duncan A. Taylor[1], James M. Curran[2] and John Buckleton[3]

[1] *Forensic Science South Australia, Adelaide, Australia*
[2] *Department of Statistics, University of Auckland, Auckland, New Zealand*
[3] *ESR, Auckland, New Zealand*
j.curran@auckland.ac.nz

All current statistical models for the interpretation of Y-STR mixtures are hampered by a single point of failure–namely that there is no single agreed model for the assignment of Y haplotype probabilities. There are numerous competing models: the counting method, Brenner's kappa method, the generalized Good method, the coalescent, the discrete Laplace, each with their strengths and weaknesses. In this talk we introduce a new approximate method, that can build probabilities locus by locus. We will discuss the basis for its derivation and show how it may have some grounding in traditional information theory. Finally, we will compare its performance to existing methods.

46. HOW CONVINCING IS A MATCHING Y-CHROMOSOME PROFILE?

Mikkel M Andersen[1,2] and David J Balding[3,4]

[1] *Aalborg University, Denmark*
[2] *University of Copenhagen, Denmark*
[3] *University of Melbourne, Australia*
[4] *University College London, UK*
mikl@math.aau.dk

The introduction of forensic autosomal DNA profiles was accompanied by much controversy over how to evaluate the weight of evidence. The problems have been successfully addressed, and DNA profiling has gone on to revolutionise forensic science. Y-chromosome profiles are valuable when there is a mixture of male-source and female-source DNA, and interest centres on the identity of the male source(s) of the DNA. The problem of evaluating evidential weight is even more challenging for Y profiles than for autosomal profiles. Numerous approaches have been proposed; we show them to be unsatisfactory. Men with matching Y-profiles are related in extended patrilineal clans. The available databases will not represent many of these clans, while others may be over-represented due to sampling biases. Further, the population of forensic relevance is difficult to define, and the number of matching relatives is fixed as population size varies, which means that population-based match

probabilities are of limited value. We propose a very different and surprisingly simple solution to the problem. We provide a simulation model and software to approximate the distribution of the number of males with matching Y profiles. The distribution can be modified to condition on a count from a randomly-sampled database, if available. Our approach allows, for the first time, the value of Y profile evidence to be explained to a court in a way that is scientifically valid, quantitative and easily comprehensible by a finder of fact such as a judge or juror.

47. USING EUROFORMIX TO ANALYSE COMPLEX SNP MIXTURES, UP TO SIX CONTRIBUTORS

Ø. Bleka[1], M. Eduardoff[2], C. Santos[3], C. Phillips[3], W. Parson[2,4] and P. Gill[1,5]

[1] *Department of Forensic Sciences, Division of Laboratory Medicine, Oslo University Hospital, Oslo, Norway*
[2] *Institute of Legal Medicine, Medical University of Innsbruck, Austria*
[3] *Forensic Genetics Unit, Institute of Legal Medicine, University of Santiago de Compostela, Spain*
[4] *Forensic Science Program, The Pennsylvania State University, PA, USA*
[5] *Department of Forensic Medicine, University of Oslo, Oslo, Norway*
oyvble@hotmail.com

A series of two- and three-person mixtures of varying dilutions were prepared and analysed with Life Technologies' HID-Ion AmpliSeq™ Identity Panel v2.2 using the Ion PGM™ massively parallel sequencing system. From this panel we used 134 autosomal SNPs. Using the reference samples of three donors, we evaluated the strength of evidence with likelihood ratio (LR) calculations using the open-source quantitative model EuroForMix and compared the results with a previous study using a qualitative model LRmix. Both models were originally designed for multi-allelic STRs, but can be extended to bi-allelic SNPs. We showed that simple two-person mixtures can be readily analysed with both LRmix and EuroForMix, but the performance of three- or more person mixtures is generally inefficient with LRmix. Taking account of the read number (coverage) sequence information greatly improves the discrimination between true and non-contributors. The higher the mixture proportion (*Mx*) of the person of interest is, the higher the LR. Simulation experiments (up to six contributors) showed that the strength of the evidence is dependent upon *Mx* but relatively insensitive to the number of contributors. Hence estimating the number of contributors is not a limiting factor for the interpretation.

48. ESTABLISHING A RANGE OF FOUNDATIONAL VALIDITY FOR COMPLEX DNA MIXTURE INTERPRETATION USING PROBABILISTIC GENOTYPING SOFTWARE: A RESPONSE TO PCAST

C. McGovern[1], R. Richards[1], H. Kelly[1], J-A. Bright[1], and J. Buckleton[1,2]

[1] *Institute of Environmental Science and Research Ltd, Auckland, New Zealand*
[2]*Department of Biostatistics, University of Washington, Seattle WA, USA*[2]
catherine.mcgovern@esr.cri.nz

The President's Council of Advisors on Science and Technology (PCAST) 2016 report [1] on forensic science in criminal courts included criticism of the interpretation of complex DNA mixtures. The council noted that whilst "probabilistic genotyping software programs clearly represent a major improvement" in DNA interpretation, there is relatively little published in peer reviewed scientific journals about their validity for use with complex DNA mixtures. PCAST defines complex DNA mixtures as having low mixture proportions and high numbers of contributors, and urges the forensic community to collate and

publish validation studies that properly establish the range of reliability for these relatively new and promising approaches.
We present here the findings of a large inter-laboratory validation study, where ground truth known trials were conducted using the STRmix™ software on more than two thousand 3-to-5 person mixtures generated by over 30 laboratories, using a range of mixture proportions, multiplex, and CE platform protocols. The scope and limitations of the probabilistic genotyping software informed by this data are discussed, and evidence is provided that establishes a range of foundational validity for complex mixture interpretation using this method.

1. Executive Office of the President: President's Council of Advisors on Science and Technology. Report to the president Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods (2016).

49. IMPOSSIBLE TO SPECIFY THE NUMBER OF CONTRIBUTORS TO A DNA MIXTURE

Charles Brenner

DNA·VIEW & Sr. Research Fellow UC Berkeley

Traditional advice about DNA mixture analysis says first decide the number of contributors. This advice is not only bad, it's impossible. It dates back twenty years to [1] -- but read the paper and you'll see it really meant "if the mixture is 2-person, here's how; if not, shelve it." Times have changed and now we do seriously consider complicated mixtures. Now is a good time to review and, without disrespect, to reject the old advice.
Sometimes there is no "the" number but rather prosecution and defense must choose different numbers. A mixture example will be presented for which the only reasonable prosecution hypothesis is a four person mixture, while the defense correctly sees it as three person. The resulting LR, for four person versus three, is close to neutral evidence. Calculating instead as four person throughout manufactures the illusion of strong evidence; it frames the suspect. Conversely, calculating uniformly as three person generates the illusion that the suspect is virtually excluded. Correctly employing the unconventional calculation makes a huge difference.
Examples for which the prosecution and defense must differ are moderately frequent in practice and demolish the rule to decide "the" number in advance. The examples do not follow a single pattern. A counterexample to a rule doesn't mean "better look out for this situation." It means the rule is wrong.

1. Clayton et al, Analysis and interpretation of mixed forensic stains using DNA STR profiling, For.Sci.Int 91 (1998) 55–70

50. SCIENCE ADVANCING JUSTICE: LESSONS LEARNED FROM DNA EXONERATIONS

Keynote Speech

S. Friedman

The Innocence Project, New York, United States
sfriedman@innocenceproject.org

The Innocence Project, founded in 1992 by Peter Neufeld and Barry Scheck at Cardozo School of Law,

exonerates the wrongly convicted through DNA testing and reforms the criminal justice system to prevent future injustices. An examination of the 350 DNA exonerations to date has identified five factors that contribute to wrongful convictions: eyewitness misidentification, misapplication of forensic science, false confessions, jailhouse informants, government misconduct, and ineffective defense counsel.
To protect the innocent, we work throughout the United States to ensure access to post-conviction DNA testing, eyewitness identification procedures are based on sound science, interrogations are recorded to reduce false confessions, and forensic disciplines used to identify suspects are valid and reliable.
Still we know the DNA exonerations are the tip of an iceberg since biological evidence is not available in most cases. The National Registry of Exonerations has documented an additional 1,704 non-DNA wrongful convictions. Getting it wrong harms more than an innocent person. There are families, crime victims and the public's confidence in the criminal justice system. Finally, there is a public safety concern as the actual perpetrator remains free to continue to commit crimes. Over the last 25 years, and in the next 25, the Innocence Project will continue to collaborate with a range of stakeholders to pursue reforms that ensure that the evidence used in the criminal justice system is accurate and reliable and thereby minimizes the chances of convicting an innocent person.

1. Brandon L. Garrett, Convicting the Innocent: Where Criminal Prosecutions Go Wrong, 2011; The Innocence Project. https://www.innocenceproject.org/

53. FIRST GENETIC EVIDENCE FOR JEWISH HOLOCAUST VICTIMS IN SOBIBOR, POLAND, BY PHYLOGEOGRAPHIC ANALYSES

Marta Diepenbroek[1], Christina Strobl[2], Harald Niederstätter[2], Bettina Zimmermann[2], Andrzej Ossowski[1], Maria Szargut[1], Grazyna Zielinska[1], Walther Parson[2,3]

[1] *Department of Forensic Genetics, Pomeranian Medical University in Szczecin, Poland*
[2] *Institute of Legal Medicine, Medical University of Innsbruck, Austria*
[3] *Forensic Science Program, The Pennsylvania State University, PA, USA*
marta.diepenbroek@gmail.com

In the course of Europe's largest genocide of the past century victims of the German Nazi regime were systematically killed in extermination camps. One of these was located in Sobibor, in the eastern part of occupied Poland. This camp is known as the deadliest phase of the holocaust in German-occupied Poland with up to 200,000 people being murdered in gas chambers and subsequently cremated. After a revolt in 1943 the Nazis closed and completely destroyed the camp.
In 2014, a team of Polish archaeologists found remains of gas chambers in Sobibor. Surprisingly, also human remains were discovered, of which some showed signs of gunshot traces. According to historical data all Jewish victims of Sobibor were cremated. It was therefore hypothesized that the remains could belong to a group of resistants killed by the Polish communist government and secretly buried there in the 1950s.
The remains were subjected to molecular genetic testing in order to shed light into their possible ethnic background. This study included the analysis of entire mitochondrial genomes obtained from the remains by Massively Parallel Sequencing as well as Y-chromosomal markers, both of which are known to carry phylogeographic signatures. The resulting haplogroups were evaluated in the light of their distribution in present-day populations and were coinciding with lineages that are frequent in modern Ashkenazi Jews. To our knowledge this is the first attempt to use phylogeographic evidence to assign Holocaust victims to their respective source population

54. WHO'S THE DADDY? TRACING THE BREAK IN THE Y CHROMOSOME LINEAGE OF KING RICHARD III

T.King[1] and K. Schürer[2,]

[1] *Department of Genetics and School of Archaeology and Ancient History, University of Leicester, Leicester, England. tek2@le.ac.uk*
[2] *School of English Local History, University of Leicester, Leicester, England.*
k.schurer@le.ac.uk

King Richard III, who was reported as being hunchback, was the last King of England to die in battle and one of the very few whose precise grave location had been lost in the over 500 years since his burial. In September 2012, excavations began to uncover any remnants of the Grey Friars friary, where Richard III was reported to have been buried following his death at the Battle of Bosworth in 1485. The remains of an individual: male, with battle injuries and scoliosis were uncovered during the course of this excavation and subsequent scientific analysis, including forensic, radiocarbon and mitochondrial DNA analysis of the remains, compared with two female-line relatives of the king, indicated that the evidence was overwhelming that these were the remains of Richard III1.
However, analysis of the Y chromosome revealed that there wasn't a DNA match between the skeletal remains and living male-line relatives, indicating that a false paternity/ies must have occurred in the 19 generations separating Richard III and the common ancestor of these living relatives. While unsurprising, we began a project to to narrow down the number of generations, or even pinpoint the exact generation, within which a false-paternity may have taken place.
Dr Turi King will discuss the strategies taken in tracking down where the break in the male line may have occurred, the present results, and next steps in this study which may shed light on the legitimacy on the historical monarchy of England in it's claims to the throne.

1. King, T.E. et al.: Identification of the remains of King Richard III, Nature Communications. 2014; doi: 10.1038/ncomms6631

55. WIDER USE OF DNA EVIDENCE IN CRIMINAL CASES: ANCESTRY PREDICTION USING Y CHROMOSOME ANALYSIS IN GERMANY

Maria Geppert, Lutz Roewer

Department of Forensic Genetics, Charité – Universitätsmedizin Berlin, Germany
maria.geppert@charite.de

In Germany an amendment of legislation concerning wider use of DNA evidence beyond identification is debated. The proposed change to law would allow analyzing the full DNA sequence checking for example hair-, eye color and biological age. Law enforcement, however, is pushing especially for inclusion of biogeographical ancestry prediction. Our laboratory has a 20 years experience in DNA-based prediction of ancestry for example in cases of unidentified persons. Currently, samples are analyzed using SNPs and STRs located on the Y chromosome, sometimes in combination with mtDNA. The eligibility of the Y chromosome for ancestry prediction is based on its unique power to differentiate between migratory events in human history. Descent clusters characterized by high frequencies of specific haplotypes and haplogroups are observed in certain geographical regions and estimates of the ages of such clusters are available due to intense research. The Y chromosome comprises slowly and rapidly mutating sites which can be used to analyze descent clusters in different time windows. Here we demonstrate our strategy for ancestry inference using different Y markers and population

databases. In an actual case we detected the founder mutation M198 (haplogroup R1a) in an unidentified corpse. This mutation occurred in Central Asia and subsequent migrations created descent clusters in Central-South Asia and Eastern Europe. The SNPs Z282 und Z93 differentiate the Asian and the European branch and could thus narrow down the regional origin to Eastern Slavic-speaking Europe. We discuss the prospects of this approach in the light of the current debate.

56. COMPARATIVE TOLERANCE OF TWO MASSIVELY PARALLEL SEQUENCING PLATFORMS TO COMMON PCR INHIBITORS FOR MISSING PERSON CASES

Kyleen Elwick[1], Xiangpei Zeng[2], Jonathan King[2], Bruce Budowle[2,3], Sheree Hughes-Stamm[1]

[1] *Department of Forensic Science, Sam Houston State University, Huntsville, USA*
[2] *Center for Human Identification, University of North Texas, Fort Worth, USA*
[3] *Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia*
kee019@shsu.edu

Human remains are often challenging to identify as they may be highly degraded, fragmented, burnt, decomposed, or contain inhibitory substances. Massively parallel sequencing (MPS) has emerged as an alternative technology to current CE-based genotyping methods. The purpose of this study was to compare the inhibitor tolerance of two MPS-based systems specifically developed for forensic use and human identification (HID) purposes.
DNA (1 ng) was spiked with various concentrations of five inhibitors (humic acid, melanin, hematin, collagen, and calcium). The samples were sequenced in triplicate using the HID-Ion AmpliSeq™ Identity Panel and Ion PGM™ (Thermo Fisher Scientific) in parallel with the ForenSeq™ DNA Signature Prep Kit (Primer Mix A) on the MiSeq FGx (Illumina).
Overall, each MPS system seemed to be more tolerant to some inhibitors than others. Both kits performed well with samples with very low levels of inhibitors. Both systems were tolerant to calcium; however, the AmpliSeq™ panel also performed well with collagen and melanin, but failed when even low levels of hematin were present. The ForenSeq™ kit performed relatively well with four inhibitors, but almost failed to genotype samples spiked with melanin.
This study also examined the relative effectiveness of the most commonly employed DNA extraction protocols used in forensic laboratories to extract high levels of these five inhibitors from blood, hair, muscle, and bone samples prior to genotyping. These results demonstrate the most compatible sample preparation methods for highly inhibited biological samples on the Ion S5 and MiSeq FGx systems.

57. LESSONS FROM A STUDY OF DNA CONTAMINATIONS FROM POLICE SERVICES AND FORENSIC LABORATORIES IN SWITZERLAND

P. Basset and V. Castella

Forensic Genetics Unit, University Center of Legal Medicine, Lausanne-Geneva, Switzerland
Patrick.Basset@chuv.ch

In Switzerland, the DNA profiles of police officers collecting crime scene samples as well as forensic genetic laboratories employees are stored in the staff index of the national DNA database to detect potential contaminations. Our study aimed at making a national inventory of contaminations to better understand their origin and to make recommendations in order to decrease their occurrences. For this

purpose, a retrospective questionnaire was sent to both police services and forensic genetic laboratories for each cases detected.
Between 2011 and 2015, a total of 709 contaminations were detected. This represented a mean of 11.5 (9.6 – 13.4) contaminations per year per 1'000 profiles sent to the Swiss DNA database. Answers to the questionnaire were obtained from the police, the laboratory or both for 552/709 (78%) of the contaminations. About 88% of these contaminations were from police officers whereas only 12% were from genetic laboratories employees. Interestingly, there has been a direct contact between the stain and the contaminant person in only 51% of the laboratory contaminations whereas this number increases to 91% for police collaborators. The high level of indirect DNA transfer in laboratories might be explained by the occurrence of "DNA reservoirs" suggesting that cleaning procedures should be improved in laboratories. At the police level, most contaminations were from the person who collected the sample and likely occurred on the crime scene. This suggests that improving sampling practices (e.g. increasing the frequency of glove changes and systematically wearing masks) could be beneficial.

58. DEGRADED DNA SAMPLES MADE INFORMATIVE BY USING SUPERPRIMERS

M.E. Mautner[1] and D. Corach[2]

[1] *Biodynamics SRL, Buenos Aires, Argentina*
[2] *SHDG-FFYB, Universidad de Buenos Aires, Buenos Aires, Argentina*
martin@biodynamics.com.ar

Available commercial Short Tandem Repeats (STRs) multiplex kits are a reliable and widespread resource for obtaining individual genetic profiles from several types of DNA samples[1,2,3]. However, they are not well suited for fully amplifying the degraded DNA found in most archaeological, aged and forensic samples[4]. This shortfall arises from the need to amplify non-specific sequences[5] in order to obtain large amplicons that are suitable for capillary electrophoresis (CE) size discrimination. That may be overcome by using long ssDNA polynucleotides as primers[6], then greatly improving the detection of genetic markers in degraded DNA. Long primers of up to 200 nucleotides (a.k.a. *superprimers*) allow a closer annealing to the target repeat sequences, thereby reducing the length of the intact DNA required for polymerization, while at the same time rendering amplicons of large size suitable for CE detection in multiplex assays. S*uperprimers* can be designed to anneal adjacently to the target repeat regions, then amplifying fragmented DNA in a similar fashion to the mini-STRs[7]. We have also demonstrated that *superprimer* annealing is sequence-flexible (*i.e.* non-fully complementary), thus allowing for the design of practically any primer sequence. We show hereby that the use of long ssDNA polynucleotides on degraded cadaveric DNA samples can discriminate genotypes otherwise missed by the currently used commercial kits.

1. Caskey CT and Edwards AO: DNA typing with short tandem repeat polymorphisms and identification of polymorphic short tandem repeats. US Patent No. 5,364,759. Houston, 1994.
2. Schumm JW and Sprecher CJ: Multiplex amplification of short tandem repeat loci. US Patent No. 6,479,235 B1. Madison, 2002.
3. Mullis KB: Process for amplifying nucleic acid sequences. US Patent No. 4,683,202. Emeryville, 1987.
4. Butler JM: Forensic DNA Typing, Biology, Technology and Genetics of STR Markers,2nd Ed. Burlington: Elsevier Academic Press; 2005.
5. NIST Short Tandem Repeat DNA Internet Databa: http://www.cstl.nist.gov/strbase/index.htm

59. DNA TYPING FROM SKELETAL REMAINS: A STUDY OF INHIBITORS USING MASS SPECTROMETRY

Suni M. Edson

Flinders University, School of Biological Sciences, Adelaide, South Australia
Armed Forces DNA Identification Laboratory, American Registry of Pathology, Armed Forces Medical Examiner System, 115 Purple Heart Drive, Dover AFB, DE 19902

This project examines the materials co-extracting with DNA from skeletonized remains that have been in the environment for greater than 50 years. A total of 435 samples with known loss locations were collected for this study in the course of ongoing HID processes. During preparation for DNA extraction, a fine powder containing osseous materials and associated environmental detritus was collected from these skeletal elements. Portions of this skeletal residue were extracted with multiple solvents and evaporated to concentrate available materials. Samples were rehydrated in methanol and analyzed using a GC/MS and/or an LC/MS-MS. Additionally, the purified DNA from the associated remains was suspended in methanol for comparison. Initial results indicate that accelerants and other fuels are not completely removed from DNA extracts using an organic extraction protocol. The skeletal materials were a mix of materials present in the environment, by-products of decay (e.g., lipids), and fat-soluble compounds inherent to the remains. Fat-soluble medications (e.g., quinine) were detectable, as were fuels and accelerants. Site-specific biological materials, such as oils from local plants, were also detected. Comparison of skeletal elements from the same site, but not the same individual, showed similar patterns of compounds present with personal variations. Not only is it possible to qualitatively study the presence of DNA inhibitors in real-world situations using MS, but there is the potential to provide an additional metric for individuation or identification of unknown human remains.

60. KPOP: A PYTHON PACKAGE FOR POPULATION GENETICS ANALYSIS

F.M. Mendes[1], C.C. Gontijo[2]

[1] *FGA, University of Brasília, Brasília, Brazil*
[2] *Human Geneticts Laboratory, University of Brasilia, Brasilia, Brazil*
fabiomacedomendes@gmail.com

Kpop is an open source Python package that detects population structure from biallelic data. It implements its own maximum-likelihood routine to estimate admixture coefficients and provides an interface to run Structure [1], ADMIXTURE [2] and Arlequin [3]. Kpop also simulates population dynamics, including a few different models for hybridization and genetic drift. This unified framework makes it convenient to assess how different histories of hybridization and isolation can produce certain patterns of population structure. Kpop integrates Scikit-Learn [4], and Tensorflow [5] packages that provide state-of-the-art implementations of machine learning (ML) algorithms in Python. While some techniques are widely used by the forensic genetics community (e.g., PCA, for dimensionality reduction), other techniques with similar objectives are not so common (e.g., t-distributed Stochastic Neighbor Embedding [6]).
Kpop integrates algorithms for many common Machine learning fields that are useful for processing genetic data: dimensionality reduction, hard and fuzzy clustering, classification, preprocessing of datasets, etc.
This paper introduces Kpop's main functionalities and compares standard methods for dimensionality reduction and clustering of genetic data with some alternatives used by the machine learning community. In many cases, the ML counterparts provide superior performance on difficult to handle

datasets.

1. Jonathan K. Pritchard, Matthew Stephens and Peter Donnelly: Inference of Population Structure Using Multilocus Genotype Data. GENETICS. 2000; 155 (2): 945-959
2. Alexander DH, Novembre J, Lange K: Fast model-based estimation of ancestry in unrelated individuals. Genome Research. 2009;19(9):1655-1664. doi:10.1101/gr.094052.109.
3. Excoffier L, Laval G, Schneider S: Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online. 2005;1:47-50.
4. Pedregosa, F., Varoquaux, G., Gramfort, et. al.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2012;(12):2825–2830. https://doi.org/10.1007/
5. Martín Abadi, Ashish Agarwal, Paul Barham, et. al.: TensorFlow: Large-scale machine learning on heterogeneous systems. 2015. (Software available from tensorflow.org.

61. A LARGE SCALE STUDY ON THE CHARACTERISATION OF DROP-IN WITHIN A DNA LABORATORY

Moore D, Liley J, Thomson J

LGC, Teddington, United Kingdom
david.moore@lgcgroup.com

Forensic DNA analysis has become increasingly sensitive in recent years with the advent of improved PCR chemistries and highly sensitive CE instruments. This has allowed the detection of STR alleles from single copies of input DNA. As a consequence of this increased sensitivity and increases in the number of loci tested; understanding and characterising drop-in events become of greater importance. Drop-in is defined by the DNA commission of the International Society of Forensic Genetics as; one or two alleles that “come from *different* individuals” [1]. Many models have been developed to describe drop-in that can be applied to a continuous probability. However, no large study of drop-in has been published within the literature.
Here we present a review of 13,485 negative control samples processed over a period of 18 months (April 2015 – September 2016). In this review multiple characteristics of drop-in are assessed including drop-in rates, peak heights, locus trends and allele frequencies. Two additional sets of experimental data have been generated for comparison to this data set. These consist of: 1) Experimental ESI 17 Fast data on very low DNA inputs (0.01pg - 20pg); 2) In silico modelling of low-level sampling variation using a Poisson distribution.
These data highlight that the distribution of drop-in alleles is consistent with population allele frequencies and that drop-in rates vary over time, even within the same DNA laboratory. Our experimental data also demonstrate that very low level genomic DNA can result in profiles with just one or two allele peaks which are indistinguishable from drop-in.

1. P. Gill, L. Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, et al., DNA comission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, Forensic Sci. Int. Genet. 6 (2012) 679–688. doi:http://dx.doi.org/10.1016/j.fsigen.2012.06.002.

62.PROTEIN-BASED FORENSIC IDENTIFICATION USING HAIR SHAFT IN EAST ASIAN AND EURASIAN POPULATIONS

Lei Feng, An-Quan Ji, Gui-Qiang Wang, Cai-Xia Li

National Engineering Laboratory for Crime Scene Evidence Investigation and Examination, Institute of Forensic Science, Beijing, China
fengleink@163.com, *Correspondence: Cai-Xia Li (Email: licaixia@tsinghua.org.cn)

The development of forensic DNA methods have advanced rapidly over the past few years, especially short tandem repeats (STRs) genotyping for individual identification and single nucleotide polymorphisms (SNPs) genotyping for biogeographic inference. All these analyses depend on the nuclear DNA and PCR, however, hair shaft is poor source of nuclear DNA template due to keratinocyte apoptosis. One latest research demonstrated that identity discrimination can be obtained from detecting genetically variant peptides (GVP) in hair shaft protein [1].
In total we collected hair shaft samples of 100 Han Chinese (East Asian) and 100 Uyghur (Eurasian) individuals. According to the paper's method, we have established the hair shaft proteomics purification procedure, and analyzed mass spectrometry-based shotgun proteomics in 30 individuals. Using the paper's custom reference protein database, 181 SAPs were found from 212 GVPs in105 genes, and the frequency of 7 SAPs loci is significantly different in Han Chinese and Uyghur subjects. Our ongoing work is to analyze SAPs in all 200 individuals and infer the non-synonymous SNPs, then evaluate the performance of hair shaft in individual identification and biogeographic inference.

1. Parker GJ, Leppert T, Anex DS, Hilmer JK, Matsunami N, Baird L, eRt al. Demonstration of Protein-Based Human Identification Using the Hair Shaft Proteome. PLoS One. 2016;11.

63. BEYOND THE STRS: A COMPREHENSIVE VIEW OF CURRENT FORENSIC DNA MARKERS CHARACTERIZED IN THE PCR-BASED DNA PROFILING STANDARD (SRM 2391D)

C.R. Steffen, K.M. Kiesler, L.A. Borsuk, P.M. Vallone

National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD, USA
becky.steffen@nist.gov

Capillary electrophoresis (CE) of the COmbined DNA Index System (CODIS) core set of 20 Short Tandem Repeat (STR) markers [1] is the primary genotyping method for forensic human identification in the U.S. With the emergence of new technologies beyond CE testing, other categories of forensic DNA markers are now more readily available. The next iteration of the PCR-Based DNA Profiling Standard Reference Material, SRM 2391d, will contain allele calls and sequence information for forensically relevant markers available in commercial kits. The SRM components will be characterized using CE and next generation sequencing (NGS). With NGS technology, more data can be mined from these reference samples than previous generations of this SRM2. This presentation will focus on data from the forensic DNA markers beyond the common autosomal, Y-STR, and X-STR markers to include: insertion/deletions, ancestry single nucleotide polymorphisms (SNPs), identity SNPs, phenotypic SNPs, microhaplotypes and mitochondrial whole genome sequencing. The primary emphasis will be an analysis of the information gained from using NGS vs CE technology with the reference samples going "Beyond the STRs".

1. Hares D: Selection and Implementation of Expanded CODIS Core Loci in the United States. FSI Genet. 2015; 17: 33-34
2. SRM 2391c: PCR-Based DNA Profiling Standard Certificate of Analysis (2015). Available online at https://www-s.nist.gov/srmors/certificates/2391c.pdf. Accessed on March 30, 2017

64. INTER-LABORATORY VALIDATION STUDY OF THE FORENSEQ DNA SIGNATURE PREP KIT, MISEQ FGX INSTRUMENT, AND FORENSEQ UNIVERSAL ANALYSIS SOFTWARE FOR LENGTH-BASED STR ANALYSIS

Michelle A. Peck[1,2], Meghan M. Didier[3], Xiangpei Zeng[4], Maiko Takahashi[4], Jonathan L. King[4], Lindsay Bennett[5], Morgan D. Falk[1,2], Kimberly Sturk-Andreaggi[1,2], Charla Marshall[1,2], Susan Welti[5], Jenifer Smith[5], Timothy P. McMahon[1], Cydne Holt[3], and Bruce Budowle[4,6]

[1]*Armed Forces DNA Identification Laboratory, a division of the Armed Forces Medical Examiner System, 115 Purple Heart Drive, Dover AFB, DE 19902, United States*
[2]*ARP Sciences, LLC, contractor supporting the Armed Forces Medical Examiner System, 9210 Corporate Boulevard, Suite 120, Rockville, MD 20850, United States*
[3]*Illumina, Inc., 5200 Illumina Way, San Diego, CA 92122, United States*
[4]*Center for Human Identification, University of North Texas Health Science Center, 3500 Camp Bowie Blvd. Fort Worth, TX 76107, United States*
[5]*Department of Forensic Sciences, 401 E Street SW, Washington, D.C. 20024, United States*
[6]*Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia*
Michelle.a.peck3.ctr@mail.mil

The ForenSeq DNA Signature Prep kit enables the simultaneous library preparation and massively parallel sequencing of 58 STRs and up to 172 SNPs. The abundance of sequence data generated from a single reaction reduces sample consumption (5 µL volume input) while maximizing genetic marker throughput and power of discrimination. To demonstrate the validity of this kit for the U.S. National DNA Index System, an inter-laboratory validation study was performed at three U.S. laboratories. This collaborative project assessed sensitivity, repeatability and reproducibility, mixtures, non-probative mock samples, and contamination using both DNA primer mixes (DPMA and DPMB). The manufacturer's protocol was followed for processing 32 samples simultaneously, and analysis was performed in the Universal Analysis Software. The testing involved over 1500 samples that were evaluated across 49 sequencing runs on five MiSeq FGx instruments. As a first step to implementing this method, analysis focused strictly on length-based STR calls to maintain backwards compatibility with current databases. Across laboratories and instruments, numerous control DNA and reference sample genotypes were shown to be highly accurate and reproducible. Full STR profiles were observed from DNA inputs down to 63 pg for DPMA and 125 pg for DPMB, and minor allele detection averaged 74% in 1:19 mixture ratios. Non-probative samples resulted in over 99% concordance with CE data. The loci D22S1045 and DYS392 notably underperformed throughout the study. This inter-laboratory validation confirmed the overall robustness of ForenSeq for length-based STR analysis and its utility with current procedures.

The opinions or assertions presented hereafter are the private views of the authors and should not be construed as official or as reflecting the views of the Department of Defense, its branches, the U.S. Army Medical Research and Materiel Command, the Defense Health Agency, or the Armed Forces Medical Examiner System.