Tuesday 3 September, 2013

Opening Address

A. Jeffreys[1]

[1]*University of Liecester, United Kingdom*

It is now 27 years since DNA analysis was first implemented in forensic investigations. This talk will be a personal retrospective on how DNA fingerprinting was accidentally discovered and how minisatellite profiling underwent the transition from an academic curiosity to being speedily implemented in casework ranging from immigration disputes to paternity cases and the first murder investigation solved by DNA. The second revolution in DNA identification came in the late 1980s with the development of rapid, sensitive PCR-based typing systems. I will discuss how STR typing became the preferred platform and how it led to the creation of major national criminal intelligence DNA databases, with the UK database established in April 1995 being the first of its kind. I will discuss the current usage of such databases, plus the potentially worrisome directions in which some databases are now evolving. More recent DNA marker systems allowing ethnogeographic classification or the recovery of phenotypic information from crime-scene DNA will be discussed, along with concerns about genetic privacy. Future directions such as database expansion and high-speed DNA typing will also be considered.

Wednesday 4 September, 2013

Hunting the Molecular Past

E. Willerslev[1]

[1]*University of Copenhagen, Denmark*

In the past two decades, ancient DNA research has progressed from the retrieval of small fragments of mitochondrial DNA from a few specimens to large-scale studies of ancient populations, reconstructions of past environments, and whole genome sequencing. Increasingly, ancient genetic information is providing a unique means to directly test theories in archaeology, palaeontology, ecology, and evolutionary biology. Initial results have changed the way we look at long debated topics such as the massive extinction of ice age mammals, early peopling of the Americas and early spread of modern humans outside Africa.

Next Generation mtDNA sequencing of forensic samples with the Ion Torrent PGM

W. Parson[1,2], G. Huber[1], H. Niederstätter[1], M. Mayr-Eduardoff[1], C. Strobl[1], A.W. Röck[1], M. Bodner[1] and the EUROFORGEN-NoE Consortium

[1]*Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria*

[2]*Penn State Eberly College of Science, University Park, PA, USA*

In forensic genetics, the mitochondrial control region (CR, ~1.1 kbp) is typically analyzed using single target or small multiplex amplification and Sanger-type sequencing (STS). The coding region (15.4 kbp) is usually not sequenced, as this approach is labor intensive and known to require more DNA than usually available in limited forensic samples. It has been demonstrated that Next Generation Sequencing (NGS) techniques produce much larger amounts of sequence data from even more challenged samples as various amplification methods can be applied. If valid in forensics, those techniques would significantly increase the discrimination power of an mtDNA assay.

We have previously shown that NGS of full mitochondrial genomes is feasible with the Ion Torrent PGM, by sequencing two overlapping 8.5 kbp PCR amplicons with a high degree of concordance to STS (> 99.98%). However, this approach is not suitable for degraded DNA. We have therefore explored the coding region for discriminatory sites using 15,512 full mitochondrial genomes and revealed 74 discriminatory SNPs that significantly increase the discriminatory power compared to the CR and roughly 50 haplogroup-specific SNPs for each of the four major phylogenies, African, East Asian, Native American and West Eurasian. Amplification primers were designed to specifically target coding region sites in midi-scale format (300-400 bp), which is applicable to degraded or limited forensic samples. We evaluated the performance of the PCR multiplex assays with subsequent NGS on the PGM.

We acknowledge the European Union Seventh Framework Programme (FP7/2007-2013, grant agreement n° 285487, EuroForGen-NoE) for funding.

Multiplex DNA amplification and barcoding in a single reaction for 454 Roche sequencing: a comprehensive study on the control region of the mitochondrial genome.

B. Bekaert[1,5], C. Massoli[2], A. Anandarajah[3], W. Van de Voorde[4,5], R. Decorte[1,4,5]

[1]*Laboratory of Forensic Genetics and Molecular Archaeology, UZ Leuven, Belgium*

[2]*Faculty of Medicine, University of Perugia, Italy*

[3]*Faculty of Health and Medical Sciences, University of Surrey, UK*

[4]*Department of Imaging and Pathology, KU Leuven, Belgium*

[5]*Forensic Medicine, UZ Leuven, Belgium*

The largest bottleneck of the next generation sequencing (NGS) amplicon assay workflow is the construction of a DNA library in which amplicons are barcoded to allow pooling of samples on a single flow cell or lane either through sequence-specific MID-containing primers in a single PCR or by adding MID primers in a second PCR. While the first approach requires unique primers for each sample, a two-step PCR protocol increases cost, time, risk of contamination and sample-mix up.

In order to circumvent these disadvantages we developed an efficient protocol to multiplex amplicons with separate MID primers within a single PCR. We demonstrate its effectiveness by sequencing two multiplex reactions of two amplicons each covering the complete hypervariable regions 1 and 2 of the mitochondrial control region from 58 reference and 30 casework samples on a 454 Roche DNA sequencer with GS FLX chemistry. The experiment was designed to (1) investigate the possibility of multiplexing sequence-specific primers with MID primers in a single PCR reaction, (2) make a comparison between results produced by individual or pooled samples after qPCR quantification, (3) compare coverage of the amplicons after quantification by qPCR or PicoGreen and (4) to compare results produced between the NGS instrument and Sanger sequencing.

Successful amplification of MID-containing amplicons was obtained for all samples as shown by qPCR and NGS coverage. Samples pooled after qPCR quantification produced similar coverage while reducing purification costs. Finally, NGS data correlated 100% (except for the C-tracks) with Sanger data confirming the effectiveness of the protocol.

Assessment of rapid DNA reference sample prototype instruments

T.F. Callaghan[1], L. Moreno[1], B. O'Brien[2], K. Olson[2], R.E. Tontarski[2], M.J. Salyards[2], C.A. Miles[3], E.L.R. Butts[4], P.M. Vallone[4]

[1]FBI Laboratory, Quantico, Virginia, USA

[2]Defense Forensic Science Center, Fort Gillem, Georgia, USA

[3]Department of Homeland Security, Washington, DC. USA

[4]National Institute of Standards and Technology, Biomolecular Measurement Division, Gaithersburg, Maryland, USA

The first generation of Rapid DNA (R-DNA) prototype instruments are becoming available to the forensic and biometric communities. These platforms are capable of automated typing and allele calling for the 13 CODIS core STR loci. Results are produced in approximately 90 minutes with minimal operator intervention. A high-level assessment of the performance of R-DNA prototypes will include STR profile accuracy, reproducibility, and reliability. This presentation will discuss the experimental design and data analysis for the first assessments of R-DNA instrumentation conducted in collaboration by the United States National Institute of Standards and Technology, the Defense Forensic Science Center and the FBI. The information gathered by this initial performance testing will help determine the design of future studies as well as drive equipment and expert software improvements. This initial assessment provides a starting point toward the validation and acceptance of laboratory-generated R-DNA reference samples for CODIS. The final operational goal of the R-DNA initiative is commercial instrumentation producing CODIS-compatible profiles in one hour for effective integration into existing CODIS structure to register and search reference samples during the booking process.

Next generation sequencing provides comprehensive multiplex capabilities

B. Budowle[1,2], D.H. Warshauer[1,2], S.B. Seo[1], J.L. King[1], C. Davis[1,2], B. LaRue[1,2]

[1] Institute of Applied Genetics, [2] Dept of Forensic and Investigative Genetics, UNTHSC, Ft Worth, Texas, USA

Next generation sequencing (NGS) provides data with unprecedented capacity and speed at a reduced cost. With barcoding many different samples can be typed simultaneously. This study describes a multiplex of >40 autosomal, 28 Y, >30 X STRs, and >375 SNPs. Whole genome mtDNA initially is sequenced separately. The findings that impact robustness and reliability are presented. For all systems library preparation and depth of coverage are critical. Library preparation has been tested using a PCR-based approach, TruSeq (Illumina), Nextera XT (Illumina), and Haloplex (Agilent). Advantages and limitations of sample preparation include amount of template required, impact on coverage, effective read length, and labor. Based on sequencing results each marker system has criteria (beyond population genetics) to consider for developing a robust system. For STRs read length spanning the repeat region is the highest priority. To support STR analysis STRait Razor, a Linux-based (free) software tool, was developed to detect forensically-relevant STR alleles in FASTQ sequence data, based on allele length. For SNPs chemistry compatibility is important (rs1029047 is a prime example). For mtDNA depth of coverage, homopolymer regions and alignment software, and baseline read errors are considered data from HL-60, several buccal, hair, blood, bone samples define the needs). The outcome is that with NGS can produce reliable results and the findings support that it is feasible to include a more comprehensive set of markers for reference samples, which will allow the needs of casework to drive marker selection for analysis, foster investigations, and yet maintain compatibility with legacy data.

Microhaplotypes Are a Powerful New Type of Forensic Marker

K. K. Kidd[1], W. C. Speed[1], A. J. Pakstis[1], E. Haigh[1], J. R. Kidd[1]

[1]Yale University School of Medicine, New Haven, CT, USA

In mass disaster and missing person cases, scientists often need to link unknowns to a family. We are proposing that haplotypes of SNPs can become the markers of choice for such tasks. Minihaplotype loci comprised of two or more SNPs in close proximity (<10kb) can be highly informative both for ancestry inference and for lineage/clan/family inference (Pakstis et al., 2012; Eur J Hum Gen 20:1148). Our recent results validate minihaplotypes.

Now we are turning to microhaplotypes: loci of two or more SNPs within a span of ~200bp that define at least three haplotypes (alleles) with no evidence of recurrent recombination. Such microhaplotypes can be genotyped by next generation sequencing since existing high throughput desktop machines (e.g., the LifeTechnologies PGM or the Illumina MySeq) allow read lengths of at least 200bp so that the phase will be unambiguous from the individual sequence reads. Large numbers of microhaplotype loci can be multiplexed at affordable costs allowing high statistical power.

We have now documented the population genetics globally of 34 microhaplotype loci in over 2500 individuals from 56 different populations. All have multiple alleles (haplotypes) and most have average heterozygosities >50%. Several show sufficient population differences in haplotype frequencies that ancestry inference is feasible. We have easily identified in HapMap and 1000 Genomes an additional three dozen such loci now being studied on our population samples. Microhaplotype loci constitute a statistically powerful new type of genetic marker ready for forensic applications using existing sequencing methods.

The development and implementation of direct PCR in casework

R. Ottens[1], J. Templeton[1], D. Taylor[2], D. Abarno[2], O. Handt[2], A. Linacre[1]

[1] School of Biological Sciences, Flinders University, Adelaide, South Australia, Australia

[2] Forensic Science SA, Adelaide, South Australia, Australia

DNA isolation and purification procedures are standard practice in the initial steps of DNA profiling when examining swabs, adhesive tapes and sections of fabric. Significant loss of DNA occurs during this process resulting in no DNA profile generated. Direct PCR circumvents the extraction process such that a DNA profile may be generated directly from the substrate. This potentially saves time, increases the sensitivity, reduces tube changes, and minimises steps open to error or contamination in the laboratory. We report on the generation of DNA profiles from a range of substrates such as hairs, fibres, and swabs taken from touch substrates. Sections of hair shafts were placed in the reaction solution with no prior treatment. Fibres from clothing were treated likewise. Individual fibres from swabs used to remove latent DNA on plastics and metals were removed and placed directly in the reaction tube. The number of amplification cycles remained as recommended by the supplier. The only alteration required to generate DNA profiles that can be uploaded to the Australian National Criminal Investigation DNA Database (NCIDD) was to increase the amount of DNA polymerase. The result is that DNA profiles can be generated from single hair shafts, single fibres and substrates touched for 5 seconds.

Funding for Renée Ottens, Jennifer Templeton and Adrian Linacre was provided by the Department of Justice South Australia.

Blood and semen identification from human DNA using copy number variation markers

D. Zubakov[1], J. Jakubowska[2], I. Kokmeijer[1], A.W. Langerak[3], P. Martínez[4], A. Maciejewska[5], R. Pawłowski[5], M. Kayser[1]

[1] Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[2] Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Gdańsk, Poland

[3] Department of Immunology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[4] Biology Service, National Institute of Toxicology & Forensic Sciences, Madrid, Spain

[5] Institute of Forensic Medicine, Medical University of Gdańsk, Gdańsk, Poland

Forensic tissue identification (FTI) aims to attribute biological traces to their donating persons. FTI is usually achieved via non-DNA evidence, while individual identification is performed by DNA analysis. This scenario leaves opportunities for false matching of traces to persons, which could be avoided if FTI was based on human DNA. Here we introduce a new FTI approach employing tissue-specific copy number variations (CNVs) in human DNA. We performed NimbleGen whole-genome comparative genomic hybridization microarray screening of samples from peripheral and menstrual blood, saliva, semen, vaginal secretion, and skin. We found most striking CNV differences for blood and semen compared to all other tissues tested. The most promising blood- and semen-specific CNV markers were selected and validated using quantitative PCR in a large set of individual samples. We found that ranges of normal CNV variation in the target tissues were not overlapping with non-target tissues, independent from the donors' age, sex, and disease status. Sensitivity of the developed qPCR assays allowed successful analysis of picogram amounts of DNA, DNA degradation caused no genotyping problems, and testing a panel of animal DNA revealed human specificity. We further demonstrated that CNV-based FTI is not destructive to DNA samples; full STR profiles could be obtained by re-using DNA after qPCR for tissue-specific CNV markers. Overall, we introduce a new concept of forensic tissue identification based on DNA copy number variation and practically demonstrate its suitability for blood and semen while in the future this approach may be extended to additional forensically relevant tissue types.

Duplex-specific nuclease (DSN) normalization: A method for the rehabilitation of low copy number DNA profiles

J. Creecy[1], N. Sambol[1]

[1] University of Central Oklahoma, Edmond, Oklahoma, USA

The greatest challenge to the science of forensic DNA analysis remains the analysis of evidence containing low-copy number (LCN) DNA. It is well established that DNA profiles obtained from low quality DNA samples are often of limited value due to artifacts associated with polymerase chain reaction (PCR). While it is theoretically possible to obtain a complete DNA profile from a single cell, in reality, profiles obtained from suboptimal amounts of DNA are difficult to interpret and often inconsistent when replicated. Here we introduce a method for the rehabilitation of evidentiary samples containing low quality and/or low quantity DNA using a novel double strand DNA specific exonuclease called duplex-specific nuclease (DSN).

The use of duplex-specific nuclease (DSN) during PCR amplification is a novel approach to improving the quality of a LCN DNA profile. DSN is a highly thermostable enzyme found in the Kamchatka crab, and is capable of cleaving double-stranded DNA over a wide pH range. DSN has been used successfully in a number of well-established research protocols. Experimentation has demonstrated that DSN can reduce the amount of the highly conserved sequences without altering the representation of the unique sequences. It is our hypothesis that the addition of DSN treatment to existing DNA analysis protocols will alleviate the preferential amplification of small molecular weight alleles observed in LCN DNA sample, thereby allowing equal representation of all of the alleles present in a given DNA sample.

Programmable DNA hybridisation as a new tool for forensic DNA genotyping.

D. Khodakov[1], A. Khodakova[2], A. Linacre[2] and A. Ellis[1]

[1] Flinders Centre for Nanoscale Science and Technology, Flinders University, Adelaide, South Australia, Australia

[2] School of Biological Sciences, Flinders University, Adelaide, South Australia, Australia

Current DNA profiling methods using PCR based amplification of STR loci are the standard means of human identification but the application of methods of genetic testing developed for diagnostic purposes has the potential to result in quicker means of human identification and with no loss of precision and reproducibility.

Here we present for the first time the implementation of a toehold-mediated non-enzymatic DNA strand displacement reaction for forensic related DNA genotyping. Toehold DNA strand displacement processes have recently gained acceptance and wide usage in the fields of DNA nanotechnology and bio-sensing. A "zipper" mode, otherwise known as a consecutive "base-by-base" uni-directional hybridization mechanism, allows for the highly efficient and selective discrimination of the reacting DNA molecules. This separation permits single nucleotide polymorphisms to be distinguished and identified rapidly and accurately. The amelogenin gene, the STR locus D16S539 and SNPs with the mtDNA hypervariable region were all targeted to demonstrate the use of this method across the range of loci encountered in forensic practice. The result was that the gender, STR genotype and SNP mitochondrial variant was determined accurately and reproducibly from all the samples tested. The potential is to apply this methodology to high-throughput microarray technologies with all the allied benefits.

Funding for Anastasia Khodakova and Adrian Linacre was provided by the Department of Justice South Australia.

Assessment of HaloPlex amplification for Sequence Capture and Massive Parallel Sequencing of Sudden Cardiac Death / Arrhythmogenic Right Ventricular Cardiomyopathy genes

A. Gréen[1], H. Gréen[2,3], M. Larsson[1], C. Gunnarsson[1], J. Jonasson[1]

[1] Department of Clinical Genetics, Linköping University Hospital, Linköping, Sweden.

[2] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden.

[3] Division of Drug Research, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden.

Introduction: Arrhythmogenic right ventricular cardiomyopathy (ARVC) is mainly a genetically determined autosomal dominant form of cardiomyopathy with reduced penetrance. Mutations in genes encoding the cardiac desmosome have been implicated as pathogenic. Next generation sequencing (NGS) has a tremendous capacity for sequencing gene panels.

Aims: In this study we designed and validated a NGS method for sequencing of a gene panel, exons and surrounding intronic

sequences, of ten ARVC genes.

Materials and methods: We used SureDesign to design a HaloPlex target enrichment system (Agilent) for the genes DES, DSC2, DSG2, DSP, JUP, PKP2, RYR2, TGTB3, TMEM43 and TTN. Twelve samples from sudden cardiac death victims or ARVC-patients were sequenced using Haloplex and MiSeq-sequencer (2x150 bp, Illumina), multiplexing 6-8 samples/run. Cut-Adapt, Burrows-Wheeeler aligner, Samtools, GATK and IGV were used for alignment and variantcalling. For validation, all exons and splice-sites of PKP2, DSP, DSG2, DSC2 and JUP were Sanger-sequenced in the samples.

Results and discussion: The targeted region covered 163025 bp (99.7 % of the exons). All samples were successfully sequenced using NGS with >200x average coverage for most samples. The variants found by Sanger in the five genes were all found using NGS. Some additional mutations were found using NGS and revising the Sanger data, these were found in low quality electropherograms regions or due to allele specific amplification in the Sanger-method, and thereby false negatives. Redesigning the Sanger-method confirmed all NGS found mutations. Our validation of Haloplex and MiSeq sequencing showed equal or better sensitivity compared to Sanger sequencing.

The Changing Face of Body Fluid Identification

SA. Harbison[1]

[1]*The Institute of Environmental Science and Research Limited, New Zealand*

For many years the identification of body fluids has relied upon the use of presumptive (bio) chemical tests. These tests were generally limited in specificity or lacking in sensitivity or incompatible with current DNA profiling methodology and/or consumed potentially limited material. Recent advances in this area have resulted in the development of immunochemical tests and methods that rely upon the detection of the expression of body fluid specific mRNA molecules. These later tests can be customized and multiplexed to identify a wider range of tissue types and are being considered by forensic scientists for use in casework, as their detection can have a significant bearing on the outcome of a forensic case. In this presentation, the changing face of body fluid identification will be described as the results of research and evaluation of these RNA tools is contrasted with traditional methods demonstrating the potential applications to forensic casework. Advantages, limitations and challenges of adopting this technology will also be discussed alongside casework examples, providing a peek into the future.

On RNA profiling, RNA data interpretation and an inter-laboratory EUROFORGEN RNA exercise.

T. Sijen[1,2], M. van den Berge[1], A. Lindenbergh, J. Harteveld[1], B. Bhoelai[1]

[1]*Netherlands Forensic Institute, The Hague, The Netherlands.*

[2]*On behalf of the EUROFORGEN Network of Excellence: http://www.euroforgen.eu/.*

RNA profiling has emerged as an alternative strategy to obtain information about the biological origin of an evidentiary trace. At the Netherlands Forensic Institute, we developed two mRNA-based multiplexes: (1) a cell typing multiplex identifying blood, saliva, semen, vaginal mucosa, menstrual secretion and skin and (2) an organ typing multiplex marking brain, lung, heart, kidney, liver, skeletal muscle and skin. RNA-based cell or organ typing occurs in combination with DNA profiling so that information on 'who is the donor?' and 'what cells are present?' derive from the exact same sample. However, from mixtures of two body fluids provided by two different donors it was established that it is inadvisable to associate donors and cell types unless assumptions can be derived from gender or sampling location. A general interpretation strategy was developed that builds on replicate RNA analyses and interpretation rules that categorize the results per each cell type. This approach accommodates several features we found to accompany mRNA profiling such as spurious marker expression in non-target tissues, marker dis-balance, variation between replicates and co-expression of body fluids (eg blood in menstrual secretion). For the cell typing multiplex, an extensive exercise was performed between nine laboratories of the EUROFORGEN Network of Excellence (http://www.euroforgen.eu/). The exercise had a step-wise approach and went from analysis of prepared multiplex PCR products to full investigation (id est extraction, DNA profiling, RNA analysis and data interpretation) of samples resembling compromised casework specimens. Outcomes and experiences of the partner laboratories will be discussed.

The development of a method for FISH identification of forensically relevant body fluids

E. Williams[1,2], M. Lin[2], S. Harbison[2], R. Fleming[2]

[1] *School of Chemical Sciences, University of Auckland, New Zealand*

[2] *The Institute of Environmental Science and Research Limited, Mount Albert Science Centre, Auckland, New Zealand*

Identifying the tissue origin of cells is increasingly important in forensic casework, and many methods, presumptive and confirmatory, have been developed to aid forensic scientists in this endeavour. Recent research has shown mRNA profiling to be a useful confirmatory test for body fluid identification, but there are limitations to this method including the level of sensitivity and the inability to link body fluids with the corresponding DNA profile in mixed samples. We have been investigating the use of fluorescent in situ hybridization (RNA-FISH) to label different cell types based on their gene expression and the separation of these labelled cells. We have developed a method for RNA-FISH suitable for forensic samples using an adapted suspension-FISH protocol with locked nucleic acid probes. Vaginal and buccal epithelial cells, sperm, seminal round cells and white blood cells have been targeted using this method. We have successfully visualised and separated these cells using laser microdissection and have obtained full DNA profiles from the separated cell populations. The development of this method shows much promise for the use of this technique in forensic casework in the future.

LINE-1 DNA methylation: a marker for discriminating monozygotic twins

J. Xu[1], G. Fu[1], L. Yan[1], F. Yu[1], S. Li[1], B. Cong[1]

[1]*Department of Forensic Medicine, Hebei Medical University, Hebei Key Laboratory of Forensic Medicine, Shijiazhuang 050017, PR China*

Aims: To evaluate the possibility of long interspersed nuclear element-1 (LINE-1) DNA methylation for discriminating two individuals of monozygotic twins (MZ).

Methods: We collected blood samples and buccal samples from 119 pairs of MZ and 57 pairs of dizygotic twins (DZ). Genome DNA was extracted and performed bisulphite conversion. LINE-1 target sequence was amplified. PCR products were sequenced by pyrosequencing and the methylation level of 3CpG sites was obtained.

Results and discussion: The mean methylation level of the three CpG sites in the blood sample among the 176 unrelated individuals was 76.60%. However it was 70.08% in buccal cell samples, which was significantly lower than that in blood samples, indicating the tissue

specificity of LINE-1 DNA methylation. Among 119 pairs of MZ, 15 pairs could be discriminated according to the difference of CpG methylation level between them, accounted for 12.61% of total number of MZ. As for DZ, 10 pairs had significant difference between two individuals, accounted for 17.54% of the total 56 pairs of DZ. The methylation level in the blood samples was associated with gender and profession. However it was correlated with age and profession in buccal samples. There was an obvious positive relationship between age and LINE-1 DNA methylation in buccal samples. In conclusion, LINE-1 DNA methylation might be a marker for discriminating MZ. However, more CpG sites should be studied and added to improve the cumulative discrimination power, and the tissue specificity and other affect factors must be considered.

Corresponding author: Shujin Li, Email: shujinli@163.com

Acknowledgements: This study was supported by National Natural Science Foundation of China (30973364) and Hebei Province Outstanding Youth Science Fund (H2012206103).

Forensic Epigenetic Age Signature - Biological Age Determination using Pyrosequencing Analysis of Methylation Markers.

M. Allen[1], M. Nilsson[2]

[1]Uppsala University, Dept. of Immunology, Genetics and Pathology, Uppsala, Sweden

[2]Stockholm County Police, Forensic Unit, Stockholm, Sweden

Introduction and aims: Knowledge of the biological age of a person leaving a biological trace at a crime scene can be an important investigative lead in some cases. A wide range of age-related changes are seen in cells as DNA damage, DNA rearrangements, deletions, shortening of telomeres and altered levels of epigenetic modifications such as methylation at CpG sites. The aim of this study is to use and determine age-specific DNA methylation patterns for chronological age of donors of forensic samples.

Materials and methods: Several markers with age-specific methylation patterns have been evaluated by Pyrosequencing analysis. Markers with 3 to 8 CpG sites were designed using the software methprimer. DNA from donors of different age was bisulfite-converted using the EpiTect Bisulfate kit from Quiagen. PCR amplification was performed using the Quiagen PyroMark PCR Kit and pyrosequencing was carried out using the PyroMark Q24 System. The percentage of methylation at each site was determined with the PyroMark software 2.0.

Results and discussion: Several markers, both previously associated with and novel to age associated methylation, were tested. Some markers showed significant correlation in methylation levels at CpG positions with age (R^2>0.6), whereas some markers did not show any statistically significant difference. Comparison of predicted and observed age for markers and CpG sites with ageing showed that the method can be used to predict age, although additional markers and samples need to be tested. Overall, our preliminary data is promising for a creation of a novel Forensic Epigenetic Age Signature test.

New RNA methods for the identification of body fluids and cell types.

R. Fleming[1], E. Williams[1,2], MH. Lin[1], SA. Harbison[1]

[1] Institute of Environmental Science and Research Limited (ESR), Auckland, New Zealand

[2] Forensic Science Program, Department of Chemistry, University of Auckland, Auckland, New Zealand

The identification of body fluids and tissues using messenger RNA (mRNA) markers and RT-PCR has proved to be a useful technique in forensic casework. We can now identify the origin of single cells by combining the use of next generation sequencing and RNA fluorescent in situ hybridization to improve the sensitivity of using RNA for body fluid and cell type origin. We can also use FISH to help identify cells from different individuals. Comparing different methods for separating the RNA-FISH labelled cells (laser microdissection and flow cytometry/cell sorting), we can separate different cell types from different individuals before DNA profiling, avoiding difficult, mixed DNA profiles. We will discuss how next generation sequencing coupled with RNA-FISH can be used for both individual and cell type identification. The casework type samples that these methods are suitable for and the advantages and disadvantages of each of these techniques will be discussed.

Body Fluid Identification by Simultaneous Analysis of DNA Methylation and Body Fluid-Specific Bacteria

A. Choi[1], K.J. Shin[1], W.I. Yang[1], H.Y. Lee[1]

[1]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea

Identification of body fluids found at crime scenes provides important clues that can support a link between sample donors and actual criminal act. Previous studies reported that DNA methylation analysis at several tDMRs (tissue-specific differentially methylated regions) enables successful identification of semen and detection of certain bacterial DNA can allow identification of saliva and vaginal fluid. In the present study, the method to detect bacterial DNA was integrated into the previously reported multiplex MSRE-PCR (methylation-sensitive restriction enzyme polymerase chain reaction). The developed multiplex PCR was improved by the addition of a new semen-specific marker and by including amplicons for 16S rRNA gene of saliva and vaginal fluid-specific bacteria. Using the developed multiplex system, semen was distinguishable by unmethylation at the USP49, DACT1, PFN3 tDMRs and hypermethylation at L81528, saliva could be identified by detection of bacteria, V. atypica and/or S. salivarius, and vaginal fluid and menstrual blood were differentiated from other body fluids by hypomethylation at the PFN3 tDMR and the presence of bacteria, L. crispatus and/or L. gasseri. Because the multiplex system uses DNA and simultaneously analyzes various body fluids in one PCR reaction, the method will facilitate efficient body fluid identification in forensic casework.

This study was supported by the Basic Science Research Program through the National Research Foundation of Korea (2010-0005208).

Selection of miRNA reference genes for accurate normalization of gene expression data for body fluid identification

I. Gomes[1], W. Liang[1,2], M. Sirker[1], L. Zhang[2], P.M. Schneider[1] and the EUROFORGEN-NoE Consortium

[1] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Cologne, Germany

[2] Institute of Forensic Medicine, Sichuan University, Chengdu, Sichuan, China

Body fluid/tissue identification has become a field of wide interest in forensic genetics in the past few years. The potential use of molecular genetic markers for identification of the biological source of a crime scene stain has been demonstrated in many studies. miRNA is of particular interest in forensics due to the small size molecules when compared to mRNA, which can provide an additional advantage for testing degraded samples. The method of choice for miRNA gene expression analysis is real-time quantitative PCR due to its high accuracy and specificity. The accuracy of this method is highly dependent on a previous selection and validation of appropriate

endogenous control genes for normalization in a given set of samples. In this work, the small non-coding genes miR-26b, miR-92, miR-423, miR-374, miR-484 and miR-191, and the small-nucleolar RNAs RNU24, RNU48, RNU44 and RNU47 were selected for an endogenous control validation study for normalization of miRNA gene expression data. The gene expression patterns of the ten candidate genes were analysed using the SYBR green based fluorescence detection system and the targeted body fluid types included representative samples from semen, saliva and peripheral blood. To identify the most suitable genes the global mean normalization method included in the reference gene validation software, geNorm (Vandesompele et al., 2002) was used. Results demonstrated that miR-92, miR-374 and miR-191 showed the highest expression stabilities and are the best candidates for endogenous controls for normalising miRNA qPCR gene expression data in the studied samples sets.

Identification of blood, saliva and semen using second generation sequencing of total miRNA

C. H. Petersen[1], B. B. Hjort[1], L. Kielpinski[2], J. Vinther[2], N. Morling[1]

[1]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

[2]Department of Biology, Faculty of Science, University of Copenhagen, Denmark

In addition to the DNA profile from a sample collected at a crime scene, the knowledge of the origin of the sample can be of great use in uncovering crime scene events or sexual assaults.

Some of the current methods of bodyfluid and tissue identification are laborious and differ in specificity and sensitivity. Thus, alternative methods using bodyfluid specific RNA has therefore been investigated. miRNA are small (18-22 nukleotides) non-coding RNA molecules that regulate the expression of mRNA and some are stably expressed in a bodyfluid specific manner.

Here, we describe a new method for bodyfluid identification using custom miRNA libraries and Next-Generation Sequencing.

Small RNA was extracted from blood, saliva and semen.

cDNA libraries were created by reverse transcriptase using a poly-dT primer on polyadenylated RNA. Following a 3´end linker ligation the libraries were amplified with multiplexed PCR primers and sequenced on an Illumina Hiseq 2000.

The sequencing reads from two sequencing runs were subjected to quality and adapter trimming and annotated to miRBase. In order to distinguish the expression patterns of miRNA between the three tissues the pattern recognition algorithm Random Forest was used to create a prediction model from the data.

Based on their impact in separating the tissues, variables were selected by the model and when used on data not included in the training set, combinations of only two or three miRNA were found to predict the origin of the tissues investigated.

Two-Factor Logistic Regression Models Enable miRNA Profiling to Provide Accurate Identification of Forensically Relevant Body Fluids and Tissues

E.K. Hanson[1], K. Rekab[2], J.Ballantyne[1,3]

[1]National Center for Forensic Science, University of Central Florida, Orlando, FL USA

[2]University of Missouri – Kansas City, Kansas City, MO USA

[3]University of Central Florida, Department of Chemistry, Orlando, FL USA

Numerous studies have demonstrated the ability to identify the body fluid of origin of forensic biological stains using messenger (mRNA) profiling. However, the size of the amplification product used in these assays (100-400 bases) may not be ideal for use with environmentally degraded samples. MiRNA profiling represents a potential alternative to mRNA profiling, since the small size of the miRNAs (~22 bases) might still permit their detection in degraded stains. Previously, we reported the first study involving the forensic use of microRNA (miRNA) profiling, which required screening of 452 candidates. Since our initial screening, hundreds of novel miRNAs have been identified. We have therefore evaluated additional miRNA candidates to further improve the sensitivity and specificity of the body fluid assays. Consequently we have expanded our body fluid identification panel to include 18 miRNAs (comprising 4 original and 14 novel miRNAs). This panel permits the identification of all forensically relevant body fluids and, uniquely, includes miRNAs for the identification of skin.

Using normalized miRNA expression data, we constructed body fluid specific two-factor (i.e. two distinct miRNA species) logistic regression models to permit an accurate identification of the body fluid of interest. Using the developed models, we have obtained 100% accuracy in predicting the body fluid of interest.

Thursday 5 September, 2013

Y-chromosomal variation from whole-genome sequences

C. Tyler-Smith[1]

[1]Wellcome Trust, Cambridgeshire, United Kingdom

Advances in technology have made sequencing whole human genomes routine and are providing a flood of data on Y-chromosomal variation. With careful quality control and validation, reliable sequence information from about 10 Mb of the Y chromosome can be extracted. This allows a robust high-resolution phylogeny to be constructed, and the resulting tree recapitulates and refines the known phylogeny, allowing almost every Y chromosome to be distinguished. In addition, mutation rates can be measured in families and the phylogeny can be calibrated directly. This dated phylogeny provides many insights into male history and the most novel and exciting ones available in September 2013 will be presented.

Expanding the knowledge of genetic variation in indigenous populations of South America by resequencing the Y chromosome

M. Geppert[1], Q. Ayub[2], Y. Xue[2], L. Gusmao[3,4], J. Gozi[3], S. Santos[5], Â. Ribeiro-dos-Santos[5], D. Corach[6], C. Tyler-Smith[2], L. Roewer[1]

[1]Institute of Legal Medicine and Forensic Sciences, Department of Forensic Genetics, Charité - Universitätsmedizin Berlin, 13353 Berlin, Germany

[2]The Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire CB10 1SA, United Kingdom

[3]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), 20550-900 Rio de Janeiro - RJ, Brazil

[4]Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), 4200-465 Porto, Portugal

[5]Universidade Federal do Pará, Laboratório de Genética Humana e Médica, 8615 Belém, Pará, Brazil

[6]Universidad de Buenos Aires, Facultad de Farmacia y Bioquimica, Servicio de Huellas Digitales Geneticas, 1113 Buenos Aires, Argentina

The peopling of the Americas and especially of South America is a fascinating field of research involving numerous scientific disciplines, namely anthropology, archeology, linguistics and human evolutionary genetics. For decades, scientists have tried to reconstruct the migrations of the first Americans after their initial arrival from Siberia between 13 and 20 KYA. Analyses of uniparentally inherited markers on the Y chromosome have significantly contributed to this understanding because indigenous populations carry similar haplogroups in Siberian and New World populations. All indigenous Y chromosomes of the Americas belong to the broad Asian-derived haplogroups C and Q, with subsequent founder mutations occurring during the colonization process (notably C-P39 and Q-M3). The haplogroup Q chromosomes in South America exhibit limited variation, with almost all of them designated to subgroup Q-M3. However, taking into account that the last common ancestor of all these males lived a minimum of 13 KYA, and the descendants mostly evolved in small, isolated groups, a fraction of variation could have escaped detection so far. Therefore, we initiated a study to find new phylogenetic Y-SNP variation within this large group of Q-M3 chromosomes. Using two different approaches based on next-generation sequencing, we identified about 90 promising new SNPs in two indigenous male individuals from Ecuador, both belonging to haplogroup Q-M3. Here, we present the process of SNP validation in order to create a new minisequencing approach to rapidly analyze Native American population samples with the aim of detecting novel genetic substructures which are informative about past demographic events.

Rapidly mutating Y-chromosomal STRs – a multi-center assessment of global male lineage and relative differentiation

K.N. Ballantyne[1,2], A.F. Ralf[1] and M. Kayser[1] on behalf of the International RM Y-STR Study Group*

1 Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

2 Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod, Victoria, Australia

The previously identified [1,2] set of 13 Rapidly Mutating (RM) Y-STRs is attracting significant attention through providing improved forensic Y-chromosome analysis. To allow enhanced quantification of the improvement effect of male lineage and relative differentiation available with RM Y-STRs, we have conducted a multicenter study involving 56 international laboratories. Quality-controlled RM Y-STR haplotypes from over 12,100 unrelated males of 111 worldwide populations were generated, 7,800 of which additionally had Yfiler data, and over 1000 father-son pairs were genotyped. These data have provided further evidence of the exceptional performance of the RM Y-STR set in male lineage and relative differentiation. The global RM Y-STR haplotype diversity was 0.999997, with 12,029 (99%) of the 12,142 worldwide males tested individualized. Haplotype matches, although rarely observed, were all between males from the same source population, with no matching between populations. In comparison to Yfiler, the RM set increased haplotype diversity in 94% of populations tested. The rate of male relative differentiation across 2528 relative pairs, including new and previous data [1,2], was 26.9% for fathers/sons, 56.3% for brothers and 80.0% for cousins. An average of 78% of pairs related by 1-20 generations were separated with RM Y-STRs - a substantial increase from the Yfiler average of 31%. Our data and results not only demonstrate the significant benefits of the RM Y-STR set to identify and separate unrelated and related males for forensic purposes, but also provides a large set of worldwide population data that will be made available for forensic and other applications.

References:

1. Ballantyne et al. (2010) The American Journal of Human Genetics 87, 341–353

2. Ballantyne et al. (2012) Forensic Science International: Genetics 6, 208–218

Disentangling close and distant relatives in Tyrolean mountain villages by rapidly mutating Y-STRs

B. Berger[1], H. Niederstätter[1], D. Erhart[1], C. Gassner[2,3], H. Schennach[2], W. Parson[1,4]

[1] Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

[2] Central Institute for Blood Transfusion & Immunological Department, Innsbruck, Austria

[3] Blood Transfusion Service Zürich, SRC, Schlieren, Switzerland

[4] Penn State Eberly College of Science, University Park, PA, USA

The Y-STRs currently used in forensic genetics have an adequate resolution to differentiate between unrelated men but usually fail to separate between close relatives. The simplistic dichotomy between "related" and "un-related" is sufficiently accurate for large, panmictic populations. However, in smaller endogamous populations most members are embedded in a network of close and distant relatives. This strongly impacts haplotype diversity and allele frequencies and significantly weakens the Y-STR evidence, if this issue is not properly considered. The recently introduced rapidly mutating (RM) Y-STRs provide the opportunity to examine this often neglected aspect of small isolated populations with high accuracy.

We present a local-scale Y-chromosomal study among ~550 men from five municipalities located in Tyrol (Austria). Whereas two sites had an urban character the remaining three mountain villages were remote and geographically isolated. The samples were screened for 17 standard Y-STRs and 19 Y-SNPs. The three isolated villages showed high patrilocality, deviating haplogroup compositions, and decreased levels of surname and haplotype diversity. We further

refined the haplotypes using 13 RM Y-STRs and tested their potential to differentiate between closely related individuals and to disentangle the complex patterns of cryptic distant relatives. Particularly in populations with low Y chromosome diversity these features would clearly improve forensic Y-STR analysis.

Regarding marker selection, analysis methods and nomenclature we refer to the International RM Y-STR Study Group (unpubl.). The study was approved by the ethics commission of the Innsbruck Medical University and it was financially supported by the Austrian Academy of Sciences.

The global PPY23 databasing project to support the implementation of high-resolution Y chromosome diagnostics in the forensic workflow

J.Purps[1], S.Siegert[2], S. Willuweit[1], M. Nagy[1], M.Nothnagel[2], L.Roewer[1] and the PPY23 Network[3]

[1]Institute of Legal Medicine and Forensic Sciences, Department of Forensic Genetics,

Charité - Universitätsmedizin Berlin, 13353 Berlin, Germany

[2]Cologne Center for Genomics, University of Cologne, 50931Cologne, Germany

[3]Currently consisting of 60 institutions from 32 countries

In 2012 we have started a large collaborative network to collect, analyze, database and publish population data of the 23 Y-STR loci included in the PowerPlex® Y23 system (Promega Corp.) which has been released in August 2012. At the time of this submission about 6,500 haplotypes from more than 50 populations have already been collected (Europe 24, Americas 16, Asia 9, Africa 3). More than 60 institutions from 32 countries participated in an ad hoc quality assessment organized by the YHRD which includes blind testing of DNA samples. Allele and haplotype frequencies were estimated by counting. Gene diversities and haplotype diversities were calculated per country, ethnic group and continent. Genetic relationships between different population groups were quantified by means of RST, thereby taking the evolutionary distance between individual Y-STR haplotypes into account. In order to graphically visualize the Y-STR genetic landscape of different continents, principal coordinates were identified by subjecting the pair-wise RST estimates between all samples to a Multidimensional scaling (MDS) analysis. Plots of the first two MDS components capturing most of the variation were generated and superimposed on maps using the GRASS (Geographical Resources Analysis Support System) software. The discrimination indices were compared for 17 loci YFiler and 23 loci PPY23 profiles. The YHRD functionality was updated to enable searches in the PPY23 format. The creation of a large global reference database for the PPY23 format assists forensic labs with the implementation of high-resolution Y chromosome diagnostics in their workflow.

A recombination study of 15 X chromosomal short tandem repeat markers using multigenerational family pedigrees

T.M. Diegoli[1,2,3], M. Nothnagel[4], A. Linacre[3], M. Krawczak[5], M.D. Coble[6]

[1]Armed Forces DNA Identification Laboratory, Armed Forces Medical Examiner System, Dover Air Force Base, DE, USA

[2]American Registry of Pathology, Camden, DE, USA

[3]Flinders University, Adelaide, SA, Australia

[4]Cologne Center for Genomics, University of Cologne, Cologne, Germany

[5]Institut für Medizinische Informatik und Statistik, Christian-Albrechts-Universität Kiel, Kiel, Germany

[6]National Institute of Standards and Technology, Gaithersburg, MD, USA

Before practical use of a forensic marker system, questions of independent assortment must be addressed, thereby following the 1991 report of the International Society for Forensic Genetics (ISFG; formerly ISFH (Haemogenetics)) relating to the use of DNA polymorphisms. For autosomal short tandem repeats (STRs), independent assortment ensures that the product rule can be used to obtain the overall likelihood of a profile from the individual marker likelihoods, while haplotype frequencies and the counting method are used for Y chromosomal STRs due to the absence of recombination. Early linkage studies produced a map of the X chromosome that divided 16 STRs into four groups of linked markers [1], and most subsequent studies have employed this model. However, a recent study [2] found evidence for less-than-free recombination between these groups as well as substantial intra-group recombination rates.

We therefore set out to evaluate the organization of 15 X chromosomal STR markers that were typed in 158 families (1156 individuals) at the Armed Forces DNA Identification Laboratory. Both traditional linkage analysis techniques as well as a novel method [2] specific to the analysis of X chromosomal markers and taking mutation rate into account were employed. The authors will present results demonstrating recombination between markers within the same proposed linkage group as well as confirming a mutation rate for X STRs on the order of 10-4. These results support the hypothesis that for these 15 markers, recombination is not a negligible factor in the statistical interpretation of an association between profiles.

Portions of this work have been supported by U.S. National Institute of Justice grant award 2011-DN-BX-K401.

The opinions and assertions contained herein are solely those of the authors and are not to be construed as official or as views of the United States Department of Defense, the United States Department of the Army, or the National Institute of Justice.

[1] R. Szibor, M. Krawczak, S. Hering, J. Edelmann, E. Kuhlisch, D. Krause, Use of X-linked markers for forensic purposes, Int.J.Legal Med. 117 (2003) 67-74.

[2] M. Nothnagel, R. Szibor, O. Vollrath, C. Augustin, J. Edelmann, M. Geppert, et al., Collaborative genetic mapping of 12 forensic short tandem repeat (STR) loci on the human X chromosome, Forensic.Sci. Int.Genet. 6 (2012) 778-84.

Assessing the potential application of X-chromosomal haploblocks in population genetics and forensic studies

V. Pereira[1], C. Tomas Mas[1], N. Pinto[2], A. Amorim[2,3], L. Gusmão[2,4], M. J. Prata[2,3], N. Morling[1]

[1]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[2]IPATIMUP – Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal

[3]Faculty of Sciences of the University of Porto, Porto, Portugal

[4]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

Haploblocks are segments of the genome, usually defined by Single Nucleotide Polymorphisms (SNPs), which have little recombination and have potential interest in forensic and population genetics. In this context and in practical terms the framework described by Ge et al. [1], in which blocks of SNPs are studied together as a transmission unit with multiple haplotypes, is much more appropriate than the consideration of the information provided by each SNP as independent.

Ge and co-workers [1] described criteria to select autosomal

haploblocks with potential use in forensic genetics. The study identified candidate regions that, a priori, met the conditions to be used as forensic markers. Until now, the potential of X-chromosomal haploblocks remains largely unexplored. However, given the lower levels of recombination at the X chromosome, it is expected that the block-like pattern will be more accentuated on that chromosome than on the autosomes.

The present work aimed to provide bases for designing strategies for selection of haploblocks defined by X-SNPs using a Next Generation Sequencing approach. In addition, the potential application in population genetics and forensic studies was assessed. One of the conditions considered in haploblock selection was the simultaneous inclusion of Short Tandem Repeats (STRs) currently used in forensic casework to allow the distinction between identical STR haplotypes and increase the resolution for fine-scale studies. Given that the size of the X chromosome is about 150 Mbps, only four haploblocks could be selected in order to guarantee independence between the blocks.

References

[1] J. Ge, B. Budowle, J.V. Planz, R. Chakraborty, Haplotype block: a new type of forensic DNA markers. Int. J. Legal Med. 124 (2010) 353–361.

FamLinkX – A new software accounting for linkage, linkage disequilibrium and mutations in calculations of relatedness using X-chromosomal marker data

D. Kling[1,2], A.O. Tillmar[3], P. Mostad[2,4]

[1]Norwegian Institute of Public Health, Department of Family Genetics, Oslo, Norway

[2]Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science, Aas, Norway

[3]National Board of Forensic Medicine, Department of Forensic Genetics and Forensic Toxicology, Linköping, Sweden

[4]Chalmers University of Technology and Mathematical Sciences, Department of Mathematical Sciences, Gothenburg, Sweden

Recent progress in the forensic genetic field has expanded the number of markers used in relationship testing. Evidently, the increase in information has brought attention to problems which has not previously been a concern. Several publications address the issues with linkage and linkage disequilibrium when using multiple X-chromosomal markers, for example those included in the Investigator Argus X-12 kit from Qiagen. We have developed a user-friendly software simultaneously accounting for linkage, linkage disequilibrium (LD) and mutation on X-chromosomal data in relationship testing.

The software is based on a new algorithm, where Markov chains are used to handle both linkage and LD. In addition, and important in a forensic setting, we allow for genetic inconsistencies by modelling mutations. The results are convincing and we will present some simulations and examples from case work to substantiate the importance of our work.

The algorithm is implemented in the software FamLinkX, which provides a user-friendly GUI for windows systems. The software is, to our knowledge, first of its kind and is freely available at http://www.famlink.se. Moreover, in addition to our implementation the algorithm opens up for the general use of any number of linked markers in forensic genetics. Similar to the previously published FamLink [Kling et al. 2012], the software adopts predefined pedigrees, but future developments include implementing the possibility to define arbitrary pedigrees.

Cluster analysis of Y-chromosomal STR population data using discrete Laplace distributions

M.M. Andersen[1], P.S. Eriksen[1], N. Morling[2]

[1]Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark

[2]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

We present a new method based on the discrete Laplace probability distribution that approximates the distribution of short tandem repeat (STR) alleles when assuming a haploid Fisher-Wright model of evolution with a single-step mutation model. Both simulated data and the European Y-chromosomal STR haplotype distribution were analysed using the discrete Laplace method. The analyses can be performed on a laptop computer. The simulation study consisted of 9,000 data sets with 500, 1,000 or 5,000 Y-STR-profiles sampled from 60 different populations of size 20,000,000. The average deviation of the estimated probabilities of the Y-STR-profiles from the true population frequencies using the discrete Laplace method was smaller than those calculated with the naïve count estimate method (like 1/n or 1/(n+1) for data set size n) and Brenner's kappa method. When analysing the European Y-STR haplotype distribution, we identified two sub-clusters of the Eastern and Western European Y-STR haplotypes similar to results of previous studies. We also compared pairwise distances (between geographically separated sampling locations) to those obtained using the AMOVA method and found good agreement with previous results. Furthermore, we investigated the homogeneity in two different ways and found that the Y-STR haplotypes from Finland were relatively homogeneous as opposed to the relatively inhomogeneous Y-STR haplotypes from Lublin, Eastern Poland and Berlin, Germany. We demonstrated that the observed distribution of alleles at each locus was similar to the expected one.

Understanding Y haplotype evidence

C.H. Brenner[1,2]

[1] Human Rights Center, UC Berkeley, Berkeley, California, USA

[2] DNA-VIEW, Oakland California, USA

The Y haplotype population-genetic terrain is better explored from a fresh perspective rather than by analogy with the familiar autosomal ideas. The genetic rules for the Y haplotype are different from the autosomal rules and these differences have population genetic consequences. Lack of recombination is one genetic difference. A consequence is that most crime scene target haplotypes are previously unseen, hence the familiar crutch of sample frequency approximating population frequency approximating matching probability is not available and must be replaced by a sounder, more careful derivation of matching probability. Near lack of convergent evolution is another difference. The consequence is to turn upside-down the autosomal tradition that theta – the probability that two random types are identical by descent – is a refinement for matching probability calculation. For rare haplotypes, identity by descent is nearly the whole story.

For haplotype matching probabilities as opposed to autosomal matching probabilities, explicit attention to modeling – how evolution got us where we are – is much more important. The Y haplotype domain is reasonably considered as a high-dimensional sparsely populated space, while autosomal profiles are adequately modeled as a combination of densely populated one-dimensional loci. Consequently intuition learned from autosomal practice is not a reliable guide for Y haplotypes. But by starting from basic principles and proceeding carefully there are several sound and reasonable avenues to haplotype evidential calculation.

Population Genetic Theory for Lineage Markers

B.S. Weir[1], J. Ballantyne[2], J-A. Bright[3], J.S. Buckleton[3], J.M. Curran[4], C.A. Laurie[1], T.R. Moretti[5] and S.P. Myers[6]

[1]University of Washington, Seattle, WA, USA

[2]University of Central Florida, Orlando, FL, USA

[3]ESR, Auckland, New Zealand

[4]University of Auckland, Auckland, New Zealand

[5]Federal Bureau of Investigation, Quantico, VA, USA

[6]California Department of Justice, Richmond, CA, USA

Interpreting Y-chromosome forensic profile matches is made difficult by the lack of recombination on the Y-chromosome and the resulting dependencies among profile components. Brenner, Forensic Science International: Genetics 4 (201): 281-291, presented an interesting approach that modified the proportion of profiles in a database, augmented by a target profile, that matched the target. This approach has advantages over simply presenting the database proportion. In this presentation we adopt a population-genetic perspective in a way that has similarity to the Balding and Nichols approach for autosomal profiles.

We explore the appropriateness of Ewens' sampling theory for Y-chromosome profiles by comparing match proportions predicted by that theory and observed proportions in public domain databases of Y-STR and Y-SNP profiles, as well as in forensic databases. We show generally satisfactory agreement. We explore the effect of population and sample size, and mutation rate, with data generated by the Hoppe Urn simulation procedure. We discuss the difficulty of studies such as these when forensic databases are used as there can be clusters of identical profiles from members of the same extended family which can bias match probabilities that use only database proportions. We extend our population-genetic approach to apply to Y-chromosome mixtures. It appears that a population-genetic approach avoids under-estimation of match probabilities in large sample sizes. We note the report of Xue and Tyler-Smith, Forensic Science International: Genetics 4 (2010): 59-61, of Y-haplotypes for men separated by 13 generations matching at 67 Y-STR loci but differing when SNP typing was considered, and we show good fit to Ewens' results when the number of SNPs is of the order of 100. Although there are some data available to address questions of independence between autosomal-STR and Y-STR profiles, the sample sizes are small. Larger samples are available for SNP profiles, and these generally support assumptions of independence, but they suggest dependencies among autosomal, Y-chromosomal and mitochondrial profiles.

This work was supported in part by NIJ grant 2011-DN-BX-K541

Wildlife Forensic Science – Current and Future Directions

R. Johnson[1]

[1]Australian Museum, Sydney

Rebecca will talk about the relatively young discipline of wildlife forensic science. While much of the congress deals with one species, wildlife forensic science deals with a multitude of species from invertebrates to protected timber. She will demonstrate the fascinating diversity of sample types and the forensically-relevant questions that arise through the illegal trade in endangered species of flora and fauna. Further, in 2013 it is just not enough to be a 'well intentioned conservation geneticist' who occasionally dabbles in forensic science. Quality control, quality assurance and accreditation are essential in the discipline of wildlife forensic science; just because the victim is a dead animal does not make the crime less important. So come and listen to how new DNA-based technologies can be used to track the trade in ivory, shark fins, parrots and reptiles and aid in the protection of endemic and exotic species.

NGS metagenomics and metabarcoding: what are the prospects and problems in forensic applications?

M. Bunce[1], N.E. White[1], D.C. Murray[1], M. Coghlan[1], S. Tridico[1] and J. Haile[1].

[1] Ancient DNA Laboratory and Wildlife Forensic Services, Murdoch University, Perth, Western Australia, Australia.

Introduction and Aims: The advent of next generation DNA sequencing (NGS) platforms, and the ever decreasing costs of running them, is having a profound impact on all forms of genetic analyses. In a forensic context the benefits of NGS whole genome sequencing are self evident, however, the use and applications of amplicon sequencing have not been widely considered. Of particular concern for amplicon sequencing on second generation platforms (e.g. Roche 454 and Ion Torrent PGM) is the requirement for two rounds of PCR and controlling the quality of the resulting data.

Materials and methods: In developing a robust NGS amplicon sequencing workflow our empirical data suggests that a heavy reliance on quantitative PCR and unique 'barcoded' fusion primers represents the best way to ensure meaningful deep sequence data can be generated that is largely resistant to cross-contamination. Careful consideration for the set-up of NGS amplicon workflows is required to ensure data fidelity.

Results and discussion: Using this approach we have explored a variety of biological substrates including traditional Chinese medicines (that contain CITES-listed species), food (to explore labelling accuracy), human hair (to examine bacterial 'fingerprints') and the quantitativeness of data when investigating diet. Deep sequencing approaches to type microsatellite loci have also been investigated in a wildlife forensic context. This presentation will showcase NGS data from a variety of contexts and discuss the laboratory and bioinformatic challenges that need to be considered when developing these workflows for forensic applications.

Species determination of DNA-mixtures using next generation sequencing

A.O. Tillmar[1], B. Dell'Amico[1], J. Welander[2], G. Holmlund[1,2]

[1]Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden

[2]Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden

Species determination can be interesting in a wide range of areas. For example, in forensics it can be used to aid police investigations in cases of poaching, trading with endangered species, and for the identification of forensic stains. The vast majority of existing DNA typing methods, developed for species determination, mainly focuses on a single species source. There are, however, many instances where all species from mixed sources need to be determined, even when the species in minority constitutes less than 1% of the sample. The introduction of next generation sequencing (NGS) opens new possibilities for such challenging forensic samples. In this study we present an ultra deep sequencing assay using 454 GS Junior (Roche) chemistry. The assay was designed by in silico analysis of DNA reference sequences from more than 1600 species from the taxon of Vertebrates. Universal PCR-primers were used for the amplification of a ~100 base pairs long mitochondrial DNA sequence. The amplicons were sequenced and the DNA data were evaluated using dedicated bioinformatic tools. Experiments were performed on artificial species-species mixture samples in order to verify the method's robustness and its ability to detect all species within the mixtures. Finally the assay was applied on two forensic case examples. The results showed to be promising with the ability to detect multiple donors within a mixture and also to detect minor components of less than 1 % in a mixed sample.

Combating transnational organized wildlife crime using dna assignment of poaching hotspots

S.K. Wasser[1], Lisa Brown[1], S. Mondol[1], C. Laurie, B. Weir[1]

[1]University of Washington, Seattle, Washington, USA

Introduction Up to 30,000 African elephants are being killed annually, threating the extinction of the remaining 400,000 elephants in Africa. We know the major transit and end-user countries but relatively little is known about the major sources of the poaching. These poaching hotspots can be localized by DNA assignment of the large ivory seizures occurring worldwide in recent years.

Aims We describe a DNA assignment method to locate the major poaching hotspots across Africa, the accuracy of the method and specific applications to large ivory seizures (≥500kg).

Materials and methods A DNA reference map was made using 16 microsatellite loci extracted from dung samples collected at known locations across Africa. Ivory is subsampled from large seizures in a manner that maximizes inclusion of samples from multiple locations. We then extract the same 16 loci from the subsampled ivory. Sample origin is determined by matching the ivory genotypes to the DNA reference map using a Markov Chain Monte Carlo algorithm.

Results and discussion We are able to assign most large seizures to within 265 km of their actual origins, suggesting the method has sufficient accuracy to achieve our goals. Seizure analyses indicate that large poaching operations usually occur repeatedly in the same area and the number of major sources may be fewer than previously thought. Targeted enforcement of these identified source populations could reduce the elephant slaughter and its associated impacts, create opportunities for controlled deliveries to find major dealers, and potentially choke the networks that drive this transnational organized crime.

Application of next generation sequencing for forensic soil DNA analysis.

A. Khodakova[1], L. Burgoyne[1], D. Abarno[2], A. Linacre[1]

[1] Flinders University, Adelaide, South Australia, Australia

[2] Forensic Science South Australia, Adelaide, South Australia, Australia

Soil is a remarkably complex, diverse, ubiquitous, and easily transferred material which can reveal highly useful information to assist forensic investigations. Despite this potential use, the genetic analysis of soil is not used currently in forensic practice due to a lack of reproducible data.

We report on the use of next generation sequencing for the comparison of soil samples based on their entire genetic composition. Soil samples were collected from three different sites across Adelaide and at three times of the year. Total genomic DNA was extracted (ZR Soil Microbe DNA kit, Zymo Research) and then amplified using arbitrary primed PCR. All the PCR products generated were sequenced using an Ion Torrent personal genome machine (Life Technologies).

Analysis of the resulting sequencing data was performed by annotation against reference databases and taxonomy independent sequence comparison. Domain distributions in the samples, determined using the MG-RAST M5NR database (MetaGenome Rapid Annotation using Subsystems Technology), showed the expected dominance of bacteria (> 94%), a small fraction of eukaryotes (2–4%) and generally less than 1% of Archaea, viruses and other unclassified organisms. A detailed comparison of organism abundance at the phylum level correctly grouped the samples according to their origin. The results of the taxonomy independent comparison (crAss, cross-assembly) showed that the samples from the same location were more similar to each other, rather than those collected from different sites. This approach allows soil samples to be compared based on their entire genetic assemblages and a potential match probability reported.

Funding for Anastasia Khodakova and Adrian Linacre was provided by the Department of Justice South Australia.

The use of NGS for enhanced forensic soil DNA analysis.

J. Young, L. Weyrich, L. Clarke, A. Cooper

Australian Centre for Ancient DNA, School of Earth and Environmental Sciences and

Environment Institute, University of Adelaide, South Australia, 5005, AUSTRALIA.

The recent development of Next Generation Sequencing (NGS) has the potential to provide a more detailed picture of the soil community than traditional DNA fingerprinting techniques, which rely solely on variations in fragment length to distinguish between soil profiles. This new technology will enable a better discrimination between soil samples and the identification of individual taxa present in a soil sample can provide information on the likely source of an unknown soil sample in the absence of reference samples. In this study, NGS data, generated using the Ion Torrent PGM, was used to target bacteria, plants, fungi and eukaryotes using common genetic markers (16S, trnL, ITS and 18S). We examined the ability of each target to enable discrimination between two contrasting locations in South Australia, a wetland and a coastal site. We demonstrate how identification of individual species can be utilised to provide information on the likely source of an unknown soil sample in forensic casework.

Concept for reengineering and expanding the EMPOP mtDNA population database

A.W. Röck[1], W. Parson[1,2]

[1]Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

[2]Penn State Eberly College of Science, University Park, PA, USA

The EDNAP mtDNA population database (EMPOP) represents the largest resource of mtDNA haplotypes in forensic genetics (>30,000 sequences) that undergoes routine quality control employing plausibility and phylogenetic checks. So far, EMPOP is the only freely available mtDNA database that performs haplotype queries in alignment free format (string alignment method, SAM) thus allowing unbiased frequency estimates that are crucial to forensic applications. New developments such as refined quality standards, revised statistical methods and alternative sequencing technologies (NGS) necessitated the development of a reengineered version of the EMPOP database and software.

The new concept is based on a haplotype-centered management system providing more flexibility for sample-specific information (geographic and phylogenetic affiliation). This includes extended queries of mtDNA haplotypes with varying sequence range and a graphical representation of search results using interactive maps.

The concept of haplogrouping mtDNA sequences, i.e. embedding haplotypes in their phylogenetic context, has proven to be useful for quality control of mtDNA data. Phylogenetic information from even small mtDNA segments increases the understanding and aids the interpretation of query results in a forensic case. MtDNA sequences are haplogrouped employing customized software using a maximum likelihood approach (EMMA). Haplogroup information is presented for each database haplotype as well as for query haplotypes. Similar to the

graphical representation of matching haplotypes, interactive maps are used to display haplogroup distribution. The developments contribute to significantly improved mtDNA database query performance and results display. This aids the forensic mtDNA community and augments the relevance of mtDNA typing in the forensic field.

We thank the National Institute of Justice (NIJ, grant 2011-MU-MU-K402), the FWF Austrian Science Fund (TR397) and the European Union Seventh Framework Programme (FP7/2007-2013, grant agreement n° 285487, EuroForGen-NoE) for funding.

A multiplex analysis of RNA expression during injury healing in human dermal injuries for injury age estimation

S. Palagummi[1,2], R. Fleming[2], S. Harbison[2], D. Elliot[1,2]

[1] Department of Forensic Science, School of Chemical Sciences, University of Auckland, Auckland, New Zealand.

[2] Forensic Biology Group, Institute of Environmental Science and Research, Mt. Albert Science Centre, Auckland, New Zealand.

The determination of the timing of a human dermal injury is important in forensic work. It helps answer questions such as, the timing of the injury and incident, the order of infliction of the injuries, the survival time after injury and, the relation of the injury to the incident, that are important to the reconstruction of a crime. In spite of its indisputable importance there currently exists no method to accurately estimate human dermal injury age.

Normal healing begins the moment the cell is injured. Immediately after injury, a time-course cascade of pathways to repair the cell is initiated. The repair process is complex and is broadly divided into three phases: inflammation, proliferation and maturation. Within each phase, a methodical and temporal series of healing events occur, which are brought about by timely changes in the expression of genes involved in injury healing. In this study, we developed a RNA based multiplex assay that identifies some of these genes and used it to analyse their expression through the time course of the injury healing process to determine their usefulness in the estimation of human dermal injury age. We analysed samples from human dermal injuries, collected at specific time points during the healing process, with our multiplex and were able to obtain expression patterns for our selected target genes. The expression patterns observed for some of the selected target genes suggest that they have potential to estimate human dermal injury age, especially during the early stages (days) of injury healing.

mRNA degradation pattern analysis in post-mortem normalized using the DNA

L. Zhang[1], M. Lv[1], Y. Li[1], J. Wu[1], H. Luo[1], L. Wang[1], P. Bai[1], W. Liang[1]

[1]Institute of Forensic Medicine, West China School of Basic Science and Forensic Medicine, Sichuan University, Chengdu, PR China

The research mRNA degradation pattern and post-mortem interval (PMI) analysis is decreasing in last several years because of complex influence factors such as RNase, temperatures, and etc. Furthermore the result of most previous researches did not get good correlation between them. We supposed that the circumstance and RNase gives same influence to all mRNA, thus, the degradation can not be showed after normalization against the mRNA genes. Normalization against DNA has been researched in last decade and rarely used because it is difficult to extract with RNA and does not control for RT and PCR efficiencies. We purposed to use the same size of DNA gene profiling result as the target mRNA to normalize the result of mRNA quantification through the new commercial DNA/RNA co-extraction kit without separately.

48 adult C57 male mice were grouped through 8 PMI with two different temperatures of 37 and 4 The brain tissue and cardiac muscle were taken from each group. DNA/RNA co-extraction using a DNA/RNA co-extraction kit (Bioteke Corperation, Beijing, China). The mRNA and the intron of HIF1-αandß-actin were profiling through RT and qPCR analysis. The correlation between mRNA and PMI after normalization using DNA were not found in all tissues under 4 , but for 37 both brain and cardiac muscle showed very good correlations between 2h to 24h. In first two hours of PMI, the HIF1-αandß-actin increased differently maybe because of the expression of mRNA after death.

Standard 28-cycle amplification of X/Y-FISH labelled epithelial cells for DNA profiling

J. Simons[1], L. Arnold[1], A. Gamblin[1], S. Vintiner[1]

[1] ESR (Institute of Environmental Science and Research), Auckland, New Zealand

The examination of sexual assault evidence frequently involves the analysis of samples that comprise mixtures of male and female cells. Separating male and female cells benefits analysis as the results are more likely to be simplified into profiles from single contributors. Some separation methods have focussed on separation of sperm from epithelial cells, but samples without sperm also require separation (vasectomised males, licked skin etc). X/Y chromosome FISH labelling when combined with laser micro-dissection (LMD) is a reliable method to separate male and female epithelial cells, but has mostly been combined with increased cycle PCR to create DNA profiles, limiting its use in many forensic laboratories.

To determine the limits of cell numbers collected by LMD for standard 28-cycle DNA profiling, and to test the effects, if any, on stochastic variation normally caused by sampling effects.

Male and female epithelial cells were stained using the Vysis CEP X/Y DNA Probe kits, and collected using a Leica LMD6000. DNA was extracted and amplified by the ESR in-house one-tube method, using standard 28-cycle PCR with the AmpFISTR Identifiler™ (Applied Biosystems) multiplex kit.

Full Identifiler™ DNA profiles were produced using standard 28-cycle PCR, and partial profiles suitable for submission were produced from even relatively low numbers of cells collected. Profiling results will be compared with low-copy number PCR on low numbers of cells stained and collected in the same manner, and the observed effects on heterozygote balance discussed.

Friday 6 September, 2013

How do we interpret DNA evidence properly?

J. Buckleton[1]

[1] *Environmental Science & Research, New Zealand*

Starting with the NAS report [1], a series of court rulings [2] and several papers [3, 4] there has been a broadly based criticism of forensic interpretation methods. These vary from questioning whether there is a proper basis for forensic science opinion to noting the large variation in interpretations of DNA profiles. In the interpretation of DNA evidence there has been a focus on reproducibility of interpretation: Getting the same answer for the same evidence in the same lab or in different labs. Using a language from my youth this is a concentration on precision. Should there be an equal concern on accuracy, getting the right answer? This paper considers what we mean by the right answer and how we could test a system for this.

1. *Committee on Identifying the Needs of the Forensic Sciences Community, N.R.C., Strengthening Forensic Science in the United States: A Path Forward. 2009, Washington, D.C.: National Academy Press.*
2. *R v T, in Neutral Citation Number: [2010] EWCA Crim 2439. 2010, Court of Appeal.*
3. *Geddes, L., What are the chances? The New Scientist, 2010. 207(2774): p. 8-10.*
4. *Geddes, L., C. King, and C. Stier, Between Prison and Freedom. The New Scientist, 2010. 207(2773): p. 8-11.*

An investigation of software programs using "drop-out" and "continuous" methods for complex mixture interpretation.

M.D. Coble[1], J.M. Butler[1]

[1] *National Institute of Standards and Technology, Gaithersburg, MD, USA*

In 2006, the ISFG commissioned an expert committee to develop guidelines for mixture interpretation (Gill et al. 2006) and these recommendations have been widely accepted. Among the recommendations was a need for the laboratory to establish a stochastic threshold to determine the risk associated with the loss (dropout) of an allele in the mixture.

Recent improvements in both STR chemistry and CE instrumentation have exacerbated interpretation as laboratories try to analyze highly complex mixtures such as "touch" items with (a) more than two contributors and/or (b) low-level contributors with possible dropout. Current strategies to evaluate low-level mixtures with dropout using the binary Likelihood Ratio (LR) are insufficient and may overstate the weight of the evidence.

As a continuation to the 2006 guidelines, the ISFG recently published recommendations for the interpretation of low-level mixtures when dropout is possible (Gill et al. 2012). Recently, a number of software programs are now available that utilize a "drop-model" incorporating a probability of dropout in the LR (e.g. Haned 2011, Lohmueller and Rudin 2012). Other software programs utilize a "continuous" model to incorporate variation within the data (e.g. TrueAllele (Cybergentics), and STRmix (ANZPAA)).

We have developed a set of complex mixtures including low-level contributors and three- and four-person contributors with differing ratios and allele sharing between the samples to investigate the efficacy and reproducibility of these software programs. We examined the gain in information compared to CPI and binary LR statistics along with other parameters such as reproducibility, time of analysis, and ease of use.

Going totally Bayesian: Lab experiences when moving to a continuous DNA interpretation model

D. Taylor[1], J. Bright[2], J. Buckleton[2]

[1] *Forensic Science SA, Adelaide, South Australia*

[2] *ESR Ltd, Auckland, New Zealand*

In 2012 both Forensic Science SA in South Australia (FSSA) and The Institute for Environmental Science and Research in New Zealand began using a fully continuous MCMC based software system for DNA profile interpretation. Moving to such a system has a number of advantages, not the least of which is the ability to provide a statistical weighting for the comparison of any reference to virtually any evidence profile.

With the transition from human-based binary to computer-based continuous systems of DNA profile interpretation we set out to adopt an approach that extended the ethos of software into all other aspects of our interpretation, reporting and workflow. That is, one of being unbiased and objective.

Adopting a continuous approach has had far reaching effects, from the exhibits we accepted for analysis, the profiles we chose to analyse, the way we constructed our likelihood ratios, the way we reported the result and ultimately the way we provide evidence in court.

Competing pressures to complete objectivity were practicality, resources and comprehensibility of results. Large efforts were made to remove binary based interpretational terms such as 'not excluded', 'inconclusive' and 'match' but still present the result in a way that a jury and court room understand.

This presentation highlights some of the methodology adopted at FSSA in an effort to remove as much subjectivity as possible and the challenges faced in doing so.

Statistical building blocks for the continuous interpretation of DNA evidence

J. M. Curran[1], J. Bright[1,2], J. S. Buckleton[2], D. Taylor[3], H. Kelly[1,2]

[1] *University of Auckland, Auckland, New Zealand*

[2] *ESR, Auckland, New Zealand*

[3] *Forensic Science South Australia, Adelaide, Australia*

Continuous likelihood ratio models are the Holy Grail for researchers interested in the modern interpretation of DNA evidence. Continuous models aim to take full advantage of the additional information contained in electropherograms by incorporating it using a statistically sound framework. If these models are correctly implemented, then they can remove the need for the thresholds, guidelines, and subjective decision making that is involved in current interpretation practice. Such models depend on substantial laboratory work and statistical data analysis to have any hope of functioning correctly. In this talk I will describe some of the work our team has been doing to model stochastic PCR phenomena such as stutter and heterozygous balance. This talk will be aimed at a non-statistical audience.

Evaluating forensic dna profiles using peak heights, allowing for dropin, allelic dropout and stutters

R. Puch-Solis[1], I.W. Evett[2]

[1]LGC Forensics, Tamworth, Staffordshire, UK

[2]Principal Forensic Services Ltd, London, UK

The ever increasing sensitivity of DNA profiling technology means that casework samples increasingly reveal complex profiles that present a range of interpretation issues. We present a model based on electropherogram (epg) peaks heights that addresses the following issues: extraneous peaks (dropin), absence of the alleles of putative donors (allelic dropout) and artefactual peaks appearing one repeat less than the allele of a putative donor (stutter). The profiles are assessed using the likelihood ratio (LR) framework where genotype probabilities are calculated in the usual manner. In a previously developed model, probability densities of peak heights in an epg, conditional on putative genotypes, are obtained by combining Gamma probability distributions of peak heights arising from alleles, stutters, where the parameters are estimated from profiles generated under laboratory conditions from known donors. In this talk we also incorporate a peak height dropin model in the probability density of peak heights where the parameters of the dropin Gamma probability distributions are estimated from profiles obtained from extraction negatives. We illustrate the performance of the model using several examples

On the meaning of the likelihood ratio: is a large number always an indication of strength of evidence?

P. Gill[1]

[1]Norwegian Institute of Public Health, Oslo, Norway

There is general lack of awareness that high LRs based on complex propositions involving e.g. three contributors, does not necessarily translate into probative evidence against a suspect. In some cases there is an increased chance of false inclusions. This is an issue for all LR-based models.

One way to address this issue is to further evaluate or qualify the estimated LR by a performance test. Based on simulations, this was achieved by replacing the reference profile of interest (typically the suspect's profile), by the profile of a simulated random man. We introduced a simple discriminatory metric, dm, which translates the ability of the estimated LR to discriminate between the person of interest and the random man.

Consider a scenario such as Hp=S+U+U vs. Hd=U+U+U: irrespective of the size of the LR, we demonstrated that large dm values indicated strong weight of evidence against the suspect, while small values indicate that the LR has to be reported with caution. We show how the dm can assist reporting officers in comparing several sets of propositions, for example, when there is uncertainty about the number of contributors to the sample.

We also developed a method for computing the exact probability that a random man profile would attain a likelihood ratio that is at least the same magnitude as that estimated, thus giving an additional quantitative measure of the strength of the evidence.

We illustrate these principles with three casework examples using the LRmix module of the Forensim package http://forensim.r-forge.r-project.org/.

Estimating drop-out probabilities of STR alleles accounting for truncation, degradation and stutters

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

[1]Department of Mathematical Sciences, Aalborg University, Aalborg, Denmark

[2]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Introduction DNA from forensic crime scene samples is often present in small amounts and partly degraded. In this work, we analysed DNA from fingernail scrapings from crime cases by STR typing, because finger scrapes are known to have only small amounts of DNA that also may be degraded and, therefore, is a challenge for STR typing and the interpretation of the results.

Aims Previous models for allelic drop-out and degradation were refined to also handle truncation of the signal intensities (e.g. threshold at 50 RFU) and stutters.

Materials and methods Samples from 291 fingernail scrapings crime cases victims with known reference profiles were STR typed with the AmpFlSTR NGMSElect kit (Life Technologies). Only data from samples with at most two STR allele drop-ins (excluding stutters below the stutter limits) were included to ensure non-mixed traces. This restricted dataset included 251 STR profiles.

An existing statistical model capturing the degradation effect of decaying STR signal intensities for increased DNA fragment lengths was modified to handle truncation of the STR signals and the effects of stutters.

Results and discussion The modified model showed significant improvement over the previously developed models. This resulted in analysis of strongly degraded samples, which were rejected as being degraded based on the previous statistical tests. By incorporating truncation and stuttering-effects in the modelling of the underlying signal intensity, it was possible to fit a logistic regression model for drop-out probabilities with similar structure as previously published by the authors.

Implementation of probabilistic models in casework; A case study

B. Kokshoorn[1], J. Klaver[1] & H. Haned[1]

[1] Netherlands Forensic Institute, The Hague, The Netherlands

The methodology for evaluating and reporting complex DNA results, i.e. mixtures of more than two individuals and/or low template samples subject to drop-out and drop-in, has promoted discussion within the forensic genetics community these past few years. Recent developments in probabilistic modeling have opened new avenues to evaluate the results of comparative DNA analyses. Several continuous and qualitative probabilistic models are currently available and implemented in casework. These models rely on case specific parameters. It is this specific model that comprises the subjectivity inherent in the evaluative process.

Here we present a case study of the implementation of a probabilistic model in casework. In 2012 the Netherlands Forensic Institute (NFI) has implemented a likelihood ratio based model (LRmix). This probabilistic model is used by reporting officers as a tool to asses the strength of DNA evidence based on a case specific model. The parameters in this specific model are estimated by the reporting officer and/or by simulation and based on empirical data. The results of the analyses are used to support an evaluative expert opinion which is what is ultimately reported to the court.

Using casework examples, we describe the methodological framework that is currently used at the NFI, and we discuss the strengths and weaknesses of the current approach.

Matching mixtures against DNA databases

J.-A. Bright[1], D. Taylor[2], J.S. Buckleton[2]

[1] Institute of Environmental Science and Research Limited, Auckland, New Zealand

[2] Forensic Science South Australia, Australia

DNA databases have revolutionised forensic science. They are a powerful investigative tool for the identification of a person of interest in a criminal investigation. Previously, profiles generated from crime samples could only be searched against a database of individuals if it was from single contributor (single source) or if a sufficiently clear major contributor could be identified. This meant that a small but significant number of samples were unsuitable for database searching.

The advent of continuous methods for the interpretation of DNA profiles offers a solution. Using these methods, each profile on the database may be compared against the mixture and a likelihood ratio (LR) formed. Those profiles producing a large LR serve as an investigative lead.

Empirical studies have been undertaken to inform the decision of what constitutes a large LR. In a pilot study, 11 mixed DNA profiles (previously unsuitable for database comparison) were compared to the database resulting in eight informative links. Effective liaison with investigating officers both at the selection of samples for searching and at the interpretation of output has been shown to be highly beneficial. This paper reports on the implementation of a method for the matching of mixed DNA profiles against the National DNA Profile Databank in New Zealand and the results from early application.

Advances in understanding DNA transfer: the role of persistence, drying time, temperature and sample type

R.A.H. van Oorschot[1], G. Glavich[1,2], R. McArdle[1,3], R.B. Howie[1,3], K.N. Ballantyne[1], W.H. Goodwin[3], R.J. Mitchell[2]

[1]Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Melbourne, Victoria, Australia

[2]Department of Genetics, La Trobe University, Melbourne, Victoria, Australia

[3]School of Forensic and Investigative Genetics, University of Central Lancashire, Preston, United Kingdom

Understanding DNA transfer can significantly assist with interpretation of DNA profiles and evaluation of proposed scenarios of criminal activity, yet knowledge of factors affecting DNA transfer is, however, limited. We conducted systematic experiments to gain insight into the influence of factors such as persistence, drying time, temperature and type of biological fluid on transfer. The results demonstrate that: DNA of a previous user can, in many situations, remain detectable on an object after use by a second individual; Detectability of the prior user's DNA depends on substrate of the object, location on the object, and manner and duration of use by the second user; The profile contribution of the prior user on the item relative to the last user declines in a linear manner; There is substantially more transfer upon contact when the biological sample is wet than dry; The speed at which biological fluids dry depends on temperature and humidity but is relatively quick (~≤30 min). The transfer rate upon contact declines linearly soon after deposit until the sample is dry. Semen behaves similar to blood and saliva. These findings enhance our ability to judge the likelihood of proposed scenarios of how a person's DNA may or may not have been found in a specific location.

Models for Reporting and Integration of DNA Matching in Missing Persons/DVI Identification Projects

T.J. Parsons,[1] Z. Bajunovi,[1] E. Jasaragi,[1] R. Huel,[1] A. Rizvi[1]

[1]International Commission on Missing Persons (ICMP), Sarajevo, Bosnia Herzegovina

DNA testing is becoming a cornerstone of identification efforts in many incidents of Disaster Victim Identification (DVI) and human rights violations/ armed conflict. DNA matching is commonly conducted through kinship analysis, adding complications to the matching/reporting process. While every event has unique characteristics, certain commonalities relating to DNA exist. Nevertheless, incident responders often react with unsettled questions or divergent approaches on how DNA results should be assessed, reported, and integrated into a multi-disciplinary identification process. DNA is unique in having a scientifically established basis for quantitative interpretation of evidentiary significance, and given a properly designed "blind" testing system, has high objectivity. We will discuss a model for reporting DNA results in large scale identification projects, where accredited DNA match reports are provided to external identification authorities in a range of complicated contexts. Blind DNA testing and matching is conducted, and high certainty matches are reported as posterior probabilities based on generic prior probabilities for the event or defined event-subset. Likelihood ratios are also reported to provide identification authorities a means for "updating uncertainty" of non-DNA evidence that is associated with particular cases. However, more fluid communication mechanisms are used to permit information exchange on possible matches, based on DNA or non-DNA evidence, that are designed to nonetheless retain rigor and objectivity, and feed into a multi-disciplinary reconciliation process. Communication loops with sample submitters for re-association and re-sampling will be discussed, as will considerations of population database choice, reporting of multiple DNA systems (e.g. autosomal and lineage markers), and related missing persons.

Beyond STRs: Ancient DNA and advanced forensic identification of missing persons remains

J.Austin[1]

[1]University of Adelaide, Australia

Identifying the remains of missing persons is a global issue with significant social, legal and cultural impacts. DNA is playing an increasingly important role in the identification of human remains, however, the highly-degraded and fragmentary nature of many cases pose major technical challenges due to DNA decay, potential DNA contamination and issues surrounding authenticity of results. In recent years, major, and often parallel, advances have been made in fields of forensic biology and ancient DNA. These have led to the development of improved DNA extraction from small and highly degraded remains and short-amplicon DNA typing methodologies and next generation sequencing protocols that are suitable for degraded DNA, providing great potential for efficient, highly discriminative genetic information for identity, bio-geographic ancestry and phenotype. These new and emerging technologies have broad relevance to missing persons investigations and "cold-case" analyses and provide powerful new tools for forensic intelligence and identification.

A global ancestry SNP panel for next generation sequencing technologies

C. Phillips[1], W. Parson[2], C. Børsting[3] and the EUROFORGEN-NoE Consortium

[1]Forensic Genetics Unit, University of Santiago de Compostela, Spain.

[2] Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria.

[3] Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

The prospects for typing 200 or more Single Nucleotide Polymorphisms (SNPs) in a single sequencing analysis have become much more realistic in the last two years with the emergence of benchtop next-generation sequencers such as Life Technology Ion Torrent and Illumina MiSeq. We decided to build from the ground up, a globally-informative ancestry SNP (AIM-SNP) panel suitable for next-generation sequencing, despite successfully applying in the last six years, smaller-scale forensic AIM-SNP sets (multiplexes of 23-34 loci) typed with SNaPshot, to analyse a range of population differentiations.

The principal goal was to identify the most differentiating SNPs for comparisons of all five major population groups of Africa, Europe, East Asia, America and Oceania. Once a comprehensive ancestry-informative SNP candidate list was established, an important next step was to balance the level of differentiation achievable for each of the five population groups by measuring cumulative I_n divergence values as the set expanded. This process ensures the marker set can provide unbiased assessments of admixture proportions in individuals with detectable co-ancestry – a significant demographic feature of most urban areas and regions with long histories of population movement.

As the EUROFORGEN global ancestry SNP panel is built into a forensically sensitive multiplex for application, in the first instance, to Ion Torrent, we anticipate certain SNPs with context sequence features will underperform or not be suitable for inclusion. Therefore we will use the SPSmart ENGINES 1000 Genomes SNP browser to select suitable alternative sites in the same LD blocks with equivalent allele frequency distributions.

Acknowledgements: This work was funded by the EUROFORGEN-NoE (Grant Agreement No. 285487).

Development and optimization of four multiplex assays with 99 of the aim SNPs from the euroforgen aims set on the Sequenom® Massarray® system

B. Lundsberg, P. Johansen, C. Børsting, N. Morling and the EUROFORGEN-NoE Consortium

Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Ancestry informative markers (AIMs) can be used in forensic genetics to infer the ethnic origin of a person, e.g. from a DNA trace sample found at a crime scene. The European Forensic Genetics Network of Excellence (EUROFORGEN-NoE) (http://euroforgen.com/) selected more than 200 AIM SNPs with the purpose of developing assays for forensic ancestry testing. The SNPs were selected to allow differentiation between the five major population groups (Africans, Europeans, Asians, Native Americans and Pacific islanders). Here, we describe the development of four multiplexes with 99 of the best AIM SNPs from the EUROFORGEN selection panel. The four multiplexes allowed typing of 35, 25, 25 and 14 AIM SNPs, respectively. Furthermore, one X and one Y chromosome SNP were included for sex determination. We evaluated the sensitivity and overall performance of each multiplex on the Sequenom MassARRAY® System by typing a Danish population. The SNP call rates were 96.1%, 93.1%, 97.3% and 97.2% for the four multiplexes, respectively. The typing results were robust when more than 1.25 ng DNA were used in the PCR. The data analysis was performed using a custom made script in the statistical software R. This script called the phenotypes automatically by using information on peak heights, signal to noise ratios and allelic balance. Three SNPs will be sequenced to elucidate skewed allele balances. These results as well as the allelic frequencies for all 101 SNPs will be presented.

This work was funded by the EUROFORGEN-NoE (Grant agreement no: 285487).

Development of ancestry inference systems for subgrouping of asian

Lee. HY[1], Lyoo SH[2], Cho SH[1], Kang SB[1], LEE SD[1,2]

[1]Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

[2]Institute of Forensic Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

Genetic ancestry inference using Ancestry Informative Markers (AIMs) can offer considerable additional information for criminal investigation. SNP (Single Nucleotide Polymorphism) genotyping system has usually been adopted for this purpose, and many studies have been done in this field. Most trials until now have been limited on the continental level. With globalization, people travel not only all over the world but also extensively within Asia. On this instance, it is meaningful to set the system for the ancestry inference for subgrouping within Asia.

First, we gathered the public genotype data at 54,794 autosomal SNPs in 1,868 Asian individuals from PAN-ASIA consortium. After subgrouping into three categories, East Asia (Korea, Japan, China, Taiwan), South East Asia (Thailand, Malaysia, Indonesia, Philippines, Singapore), and South Asia (India), we have selected fifty candidate SNPs to distinguish these regions by estimating Rosenberg's ancestry informativeness metric:In. Again, 31 candidate SNPs were re-selected considering several conditions and finally, 31-plex amplification system was set up using multiplexed microsphere-based suspension array platform by Luminex®200. We are under working for the SNP typing using the above system for the samples from different area within Asia and statistical analysis would be done after the final data collection.

In conclusion, we have optimized the multiplexed genotyping of 31 SNPs in a simple and informative assay. Final results will be presented and discussed.

Predicting skin colour from DNA using a model-based approach

S. Walsh[1], A. Ralf[1], L. Chaitanya[1], J. Koller[1], L. Kovatsi[2], A. Wojas-Pelc[3], W. Branicki[4, 5], F. Liu[1] and M. Kayser[1]

[1]Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, Rotterdam, The Netherlands

[2]Laboratory of Forensic Medicine & Toxicology, School of Medicine, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece

[3]Department of Dermatology, Collegium Medicum of the Jagiellonian University, Kraków, Poland

[4]Section of Forensic Genetics, Institute of Forensic Research, Kraków, Poland

[5]Department of Genetics and Evolution, Institute of Zoology, Jagiellonian University, Kraków, Poland

Forensic DNA phenotyping (FDP) i.e. the prediction of human externally visible characteristics from DNA for forensic purposes has become a fast growing topic of interest. FDP is expected to aid police investigations by providing appearance information on unknown individuals when conventional DNA profiling or other means of investigation are non-informative. So far, DNA test systems have been developed and forensically validated for eye and hair colour prediction, such as the IrisPlex [1] and HIrisPlex systems [2]. However, a practically useful DNA test for skin colour prediction has yet to be developed. In this study, we have gathered information from previously known skin colour associated SNPs, together with a new list of SNPs from a parallel genome-wide association study on Dutch Europeans, and genotyped over 50 SNPs from over 3000 European and non-European individuals with varying degrees of skin colour. Using this

comprehensive dataset, we performed skin colour prediction analyses and selected a subset of the most predictive skin colour SNPs for subsequent development of an efficient and sensitive DNA test system for skin colour prediction. With this data we envision that skin colour can soon be added to eye and hair colour prediction, providing the complete pigmentation portfolio of an individual from crime scene DNA to further aid police investigation.

References:

1. Walsh et al. Forensic Sci. Int. Genet. 2011, 5: 170-180.

2. Walsh et al. Forensic Sci. Int .Genet. 2013, 7: 98-115.

Epistasis in the determination of human pigmentation is common and improves the DNA based phenotype prediction

E. Pośpiech[1], S. Walsh[2], F. Liu[2], M. Kayser[2], A. Wojas-Pelc[3], W. Branicki[1,4]

[1]Department of Genetics and Evolution, Institute of Zoology, Jagiellonian University, Kraków, Poland

[2]Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, Rotterdam, The Netherlands

[3]Department of Dermatology, Collegium Medicum of the Jagiellonian University, Kraków, Poland

[4]Section of Forensic Genetics, Institute of Forensic Research, Kraków, Poland

The available forensic DNA prediction systems already enable the reliable prediction of some eye and hair colours; however, there is still space for improvement. Here we explore the extent towards which SNP-SNP interactions are involved in human pigmentation traits including skin color and their impact on DNA-based phenotype prediction. The model-building set included 718 Polish samples and the model-verification set included 307 independent Polish samples, all with complete information on eye, hair, and skin color. We selected 38 SNPs within 12 known pigmentation genes and the Vitamin D Receptor gene (VDR). Overall we found 23 significant SNP-SNP interactions with 5 of them showing a noticeable effect on phenotype prediction. Interactions between HERC2 rs12913832 and OCA2 rs1800407 as well as TYRP1 rs1408799, raised the prediction accuracy expressed as AUC from 0.673 to 0.697 and increased the prediction sensitivity by >3%, both for green eye colour. Interaction between MC1R "R" variants and VDR rs731236 increased the sensitivity for light skin by >1% and by almost 3% for dark skin. Interaction between VDR rs1544410 and TYR rs1042602 as well as between MC1R "R" variants and HERC2 rs12913832 provided an increase of red/non-red hair prediction accuracy from AUC of 0.902 to 0.930. Our results thus underline epistasis as a common phenomenon in human pigmentation genetics and demonstrate the value of SNP-SNP interactions on DNA-based prediction of human pigmentation traits. Hence, we advocate for including SNP-SNP interactions in future prediction modeling of human pigmentation traits for forensic applications.

Identification of Single Nucleotide Polymorphisms (SNPs) Involved in the Determination of Physical Appearance

M. Barash[1], A. van Daal[1]

[1]Faculty of Health, Science and Medicine, Bond University, Gold Coast, QLD 4229, Australia.

Forensic Molecular Photofitting is a new area of forensic DNA profiling that seeks to obtain visible information, such as skin, eye and hair pigmentation and more recently, facial morphology, about the depositor of a DNA sample.

This study aimed to identify a set of single nucleotide polymorphisms (SNPs) involved in normal craniofacial variation and develop an assay that can predict the phenotype of the person who is the source of the DNA. To achieve this, more than 1200 SNPs in 173 candidate genes, potentially involved in normal embryonic craniofacial development and various malformations were targeted for genotyping. In addition, approximately 700 markers previously shown to be associated with pigmentation traits such as eye, skin and hair colour, along with ancestry lineage markers and identity informative SNPs, STRs and INDELs were chosen for genotyping.

More than 500 DNA samples along with pigmentation phenotype, ancestry information and 3-Dimentional (3D) facial images have been collected. The 3D images were analysed for more than 90 linear and angular craniofacial measurements, including various craniofacial indexes.

The DNA samples were sequenced using a Next Generation Sequencing platform at a set of candidate SNPs and were evaluated for statistically significant associations with pigmentation, ancestry and anthropometric craniofacial measurements.

The results of this study will be presented.

Pacifiplex: a forensic ancestry SNP panel centered on Australia and the Pacific region

C. Santos[1], C. Phillips[1], M. Fondevila[1], R. Daniel[2], R.A.H. van Oorschot[2], M.V. Lareu[1]

[1] Forensic Genetics Unit, Institute of Legal Medicine, University of Santiago de Compostela, Spain.

[2] Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Victoria, Australia

The analysis of human ancestry and population admixture components is an area of considerable interest in the forensic and anthropological fields. Single nucleotide polymorphism (SNP) ancestry-informative assays, such as the 34-plex assay [1,2], allow the differentiation of major global population groups of Africa, Europe and East Asia. However the 34-plex assay does not include markers informative for Oceanian populations – those sited between the continents of East Asia and The Americas. In particular, Oceanians are poorly differentiated from East Asians, the neighboring region to Oceania and most closely related population group.

In this study, a SNaPshot multiplex assay, Pacifiplex, was developed containing 29 Ancestry Informative Marker SNPs (AIM-SNPs) selected to distinguish four ancestral groups: Africans; Europeans; East Asians and Oceanians.

Australian Aboriginal samples from the Northern Territory were genotyped with Pacifiplex and 34-plex SNP assays and compared with HGDP-CEPH panel samples from the four reference groups, including Oceanians from Bougainvillea in Melanesia and Papua New Guinea. This allowed us to evaluate the efficiency of these AIM-SNPs to cluster populations from different continental origins. Results revealed that the 29 AIM-SNPs of Pacifiplex are highly efficient for inferring the ancestry of individuals from Oceanian populations and provides consistent estimates of ancestry proportions at the population level.

[1] C. Phillips, et al., Forensic Sci. Int. Genet. 1 (2007) 273–280.

[2] M. Fondevila et al., Forensic Sci. Int. Genet. 7 (2013) 63–74.

Acknowledgments: CS was awarded a PhD grant by the Portuguese Foundation for Science and Technology (SFRH/BD/75627/2010 and co-financed by the European Social Fund (Human Potential Thematic Operational Program). The work of MF at NIST has been supported by Fundación Barrie de la Maza Postgraduate Grant Program (2010).

Preliminary assessment of the Ion Torrent PGM system for SNP genotyping in forensic analysis

R. Daniel[1], C. Santos[2], C. Phillips[2], M. Fondevila[2], MV Lareu[2], R.A.H. van Oorschot[1], D. McNevin[3].

[1] Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Victoria, Australia.

[2] Forensic Genetics Unit, Institute of Legal Medicine, University of Santiago de Compostela, Spain

[3] National Centre for Forensic Studies, University of Canberra, ACT, Australia.

The current genotyping platforms for forensic DNA analysis are limited by their multiplexing capabilities and therefore their informativeness with respect to data gained from evidentiary samples. The application of Next Generation Sequencing to Forensic DNA analysis will allow the simultaneous analysis of hundreds to thousands of loci from a single template.

Amongst the platforms currently available, the Ion Torrent PGM system (Life Technologies), based on the detection of H^+ release upon the incorporation of each nucleotide, is a flexible platform compatible with medium throughput SNP genotyping.

A preliminary assessment of the Ion Torrent PGM system for forensic DNA analysis has been conducted. Five established autosomal SNP based forensic PCR multiplexes (including identity, ancestry and phenotype SNPs), consisting of a total of 131 amplicons were combined. Standard reference material and human DNA samples were assessed at template amounts of 0.1, 0.2, 0.3, 1 and 2ng of gDNA on a 314 chip. This study reports sensitivity, sequencing bias, concordance with SNaPshot genotyping and considerations for allele call thresholds (i.e. balancing sequence coverage with the call rates observed for each allele in homozygote and heterozygote genotypes).

Initial results, based on a limited number of samples, indicate that the Ion Torrent PGM system is robust and accurate. However, a number of critical factors must be carefully considered prior to use in forensic DNA analysis.

Haloplex and MiSeq NGS for simultaneous analysis of 10 STRs, 386 SNPs and the complete mtDNA Genome.

M. Allen[1], M. Nilsson[2], M. Havsjö[3], L. Edwinsson[3], J. Grånemo[1], M. Bjerke[1]

[1]Uppsala University, Dept. of Immunology, Genetics and Pathology, Uppsala, Sweden

[2]Stockholm County Police, Forensic Unit, Stockholm, Sweden

[3]AlphaHelix Molecular Diagnostics AB, Uppsala, Sweden

Introduction and aims: The rapid development of next generation sequencing (NGS) technologies in the recent years has improved genetic analyses, allowing high throughput analysis of the genome in medical and biological research. The aim of this study is to develop an optimal NGS system for DNA analysis of historical skeletal remains and forensic biological materials.

Materials and methods: A forensic panel for analysis of 10 STRs, 386 SNPs on the autosomes, X- and Y-chromosomes as well as the complete mtDNA genome was designed. Target selection and enrichment is done using Agilent's HaloPlex system for customized panels bases on a capture technology with high sensitivity. The targets and HaloPlex probe selection were done using the SureDesign software (Agilent) for Illumina 150 bp read-lengths and the option for formalin fixed paraffin embedded tissue samples.
The fast and easy to use benchtop DNA sequencer MiSeq system (Illumina) was used for NGS analysis.

Results and discussion: The combination of Haloplex and MiSeq systems is very convenient for rapid and efficient analysis of multiple markers for forensic identification analysis simultaneously. The use of mtDNA, short targets and the highly sensitive target enrichment system Haloplex, allowed analysis of degraded samples as well as small input amounts of DNA. With all the markers in the assay, individual identification with CODIS markers, SNPs, X- and Y-SNPs as well as mtDNA could be performed. In addition, visible characteristics as eye- and hair-colour and ancestry information could be inferred using this promising forensic NGS assay.

Funding support for research: VINNOVA.