11th Haploid Markers Conference
INFERRING ANCESTRY FROM DNA

17-19th May 2018
Bydgoszcz, Poland

THE HONORARY PATRONAGE OF THE 11TH HAPLOID MARKERS 2018 CONFERENCE „INFERRING ANCESTRY FROM DNA"

UNIWERSYTET MIKOŁAJA KOPERNIKA W TORUNIU
Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy

Professor Grażyna Odrowąż-Sypniewska
Vice-Rector of the Nicolaus Copernicus University in Toruń, Collegium Medicum in Bydgoszcz

UNIWERSYTET MIKOŁAJA KOPERNIKA W TORUNIU

Professor Andrzej Tretyn
Rector Magnificus of the Nicolaus Copernicus University in Toruń

11th Haploid Markers Conference
INFERRING ANCESTRY FROM DNA
17-19th May 2018
Bydgoszcz, Poland

Dear Colleagues,

For the first time, and hopefully not the last, **The 11th Haploid Markers Conference (HM2018)** is being held in Bydgoszcz, Poland. The previous editions of the Conference have taken place in plenty of amazing cities like Berlin, Porto, Innsbruck, Ancona, and Brussels. This year, we have organized the conference in Bydgoszcz, founded over 670 year ago in central Poland, situated along the picturesque banks of the Vistula and Brda rivers. HM2018 host is The Department of Forensic Medicine of Nicolaus Copernicus University located in Bydgoszcz, one of the most experienced in haploid markers' research in Poland.

We have picked "Inferring Ancestry from DNA" to be the main theme of this year's conference. For some time now, we have been gathering knowledge on how to decipher DNA sequences for the purpose of extracting information regarding their ancestral origin. During the Conference we will have the opportunity to explore contemporary biogeographical analysis based on genetic testing, not only encompassing Y chromosome or mtDNA, but – of equal importance – autosomal DNA.

This biennial event attracts the most prominent researchers from all over the world. It is one of the most valuable opportunities to exchange ideas, thoughts and build new scientific networks.

HM2018 is a three-day conference with agenda that covers oral presentations, poster sessions and lunch seminars. The scientific programme has been divided into four distinct parts, dedicated to DNA ancestry testing, new generation sequencing, population genetics and casework. This year's conference gathers over 200 registered Participants that will altogether give 44 speeches and present over 60 posters. Also, two of HM2018 Sponsors, namely Thermo Fischer Scientific and Verogen are preparing lunch seminars presenting their new cutting-edge products and technologies.

We hope that the meeting is fruitful and inspiring, actively contributing to sharing knowledge on advances in haplotype markers use in forensics.

Welcome to Bydgoszcz!

Lutz Roewer & Walther Parson – HM2018 Scientific Committee

Tomasz Grzybowski – HM2018 Organizing Committee


Promega


QIAGEN


QiaLab®
Autoryzowany dystrybutor firmy QIAGEN


ThermoFisher
SCIENTIFIC


VEROGEN

SCIENTIFIC COMMITTEE:

Professor Walther Parson

Professor Lutz Roewer

ORGANIZING COMMITTEE:

Professor Tomasz Grzybowski
chairman

Urszula Rogalla-Ładniak, PhD
scientific secretary

Katarzyna Linkowska, PhD

Anna Ewa Radziszewska, MSc

Marcin Woźniak, PhD

11th Haploid Markers Conference

INFERRING ANCESTRY FROM DNA

17-19th May 2018

Bydgoszcz, Poland

Find out more at **thermofisher.com/challengingsamples**

ThermoFisher SCIENTIFIC

VEROGEN

Introducing the world's first company solely focused on forensic genomics

TRUTH FORWARD™

Verogen serves those that pursue the truth. Investigators need insights to advance serious cases, pursue violent offenders and identify the missing. Powered by Illumina technology and working in partnership with forensic laboratories, we deliver massively parallel sequencing tools to help deliver global justice and public security for all.

Learn how we're advancing massively parallel sequencing to unlock the true potential of forensic genomics

Join our Lunch Seminar
at Haploid Markers 2018
Friday, 18th May, 13:30, Manru Hall

VEROGEN.COM

Venue

The 11th Haploid Markers Conference will take place in the **Opera Nova Congress Centre**. It is located in the Opera Nova building, at the Brda River, next to the charming Mill Island – one of the most recognizable places in the city centre.

Transport

Public transport is well-developed in Bydgoszcz, so we strongly recommend its use. As there is a convenient tram stop located next to the **Opera Nova Congress Centre**, the place is easily accessible. Below you can find some useful directions:

- **From Railway Station „Bydgoszcz Główna"**

  Use Tram no. 5 (direction "ŁOSKOŃ") or no. 8 (direction "WYŻYNY (PĘTLA)") – 4th stop called „Focha-Opera" is the one,

- **From Bus Station „Dworzec Autobusowy Bydgoszcz"**

  Use Tram no. 3 (direction „WILCZAK"), 5 or 8 (direction „RYCERSKA") – 2nd stop called „Focha-Opera" is the one,

- **From the Paderewski Airport**

  Use the bus no. 80 (direction "DWORZEC GŁÓWNY") and change on 6th stop („RONDO JAGIELLONÓW") to any of the trams no. 3 (direction „WILCZAK"), 5 or 8 (direction „RYCERSKA") and get off at „Focha-Opera" stop.

JakDojade App

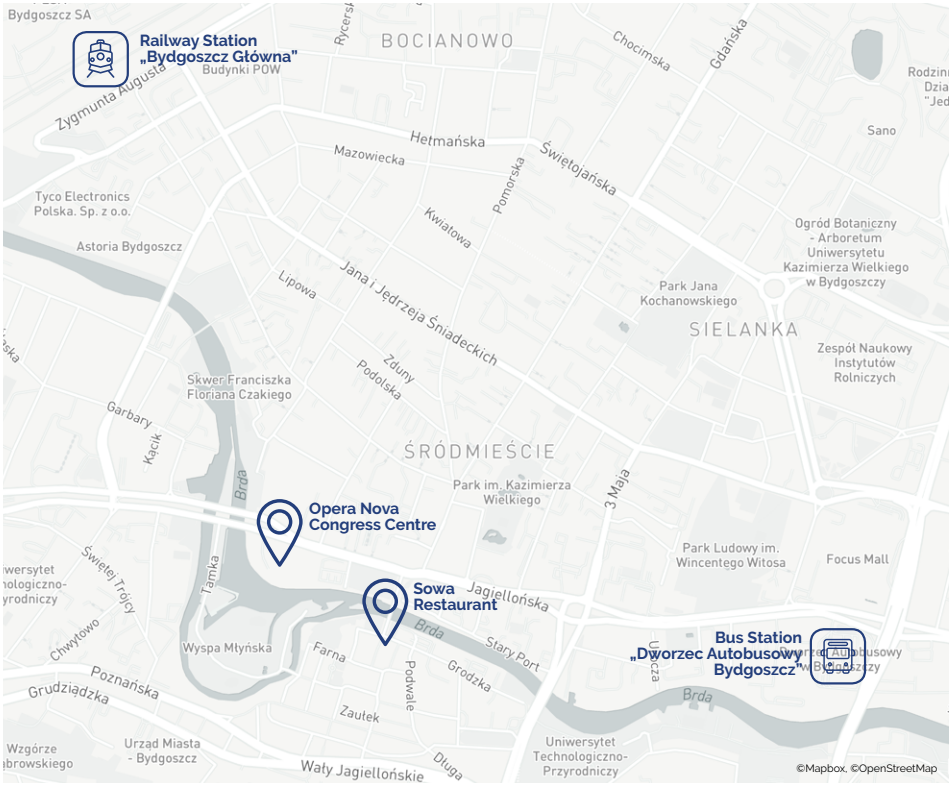
Still not sure if you can easily reach the right place? Try using „Jak dojade" app (for PC, Android and iOS).

ANDROID

IOS

Dinner on 18th May

The dinner after second day of – as we hope – fruity and inspiring conference session, will be quite unusual. We would like to invite you to a concert of a cover band named "4 szmery" that is going to take us into the music world of AC/DC band, a hard rock icon. Don't like this kind of music? No worries. As the whole thing takes place in **Sowa Restaurant**, you can indulge yourself with delicious food, accompanied with a variety of beverages and simply have some rest in beautiful surrounding. The Restaurant itself is located in the city centre within walking distance from **Opera Nova Congress Centre**.



Addresses:

Opera Nova Congress Centre
Marszałka F. Focha 5
85-070 Bydgoszcz

Sowa Restaurant
Mostowa 4
85-110 Bydgoszcz

Phone numbers:

Police, Fire Brigade, Ambulance tel 112 | Municipal Police, tel 986
Urban Crisis Management Centre, tel +48 52 585 98 88 | Tourist Information, tel +48 52 340 45 50

Dr. Christopher Phillips

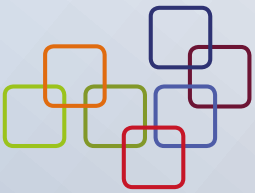
studied Genetics at Birmingham University, UK, between 1974-1977 and in 1978 obtained a MSc in Applied Genetics at the same institute. He started his forensic genetics career in 1979 at the Biochemistry Division of the Metropolitan Police Forensic Science Laboratory, London, then moved to the Forensic Haematology Department of Barts and The London Medical School and worked there till 2001. Since 2001 he has been a full-time researcher in the Forensic Genetics Unit of the University of Santiago de Compostela, Spain. His research interests include SNP analysis applied to medical, population, and forensic genetics, the development of novel forensic polymorphisms, and the creation of open-access online genomics search tools for the genetics and forensic communities.

Prof. Mark A. Jobling

studied Biochemistry at the University of Oxford (UK) and earned a PhD at the Genetics Laboratory of the same University. In 1992 he moved to the Department of Genetics of the University of Leicester, where he has remained until today. His research interests include human genetics, forensics, human evolution, genetic genealogy and ancestry testing. Over the last decade, he has authored many influential papers on genome – and population-level processes shaping human genetic diversity. His work has a translational dimension in genomics and forensics. He is also known for his exceptional educational activities: he co-authored a comprehensive best-selling textbook "Human Evolutionary Genetics" (2nd edition, 2014, Garland Science, New York/London), with Edward Hollox, Toomas Kivisild, Matthew Hurles and Chris Tyler-Smith, another HM 2018 keynote speaker.

Dr. Chris Tyler-Smith

earned a PhD in molecular biology from the University of Edinburgh in 1980. From 1986 to 2003 he worked at the Department of Biochemistry of the University of Oxford. In 2003, he moved to the Wellcome Trust Sanger Institute in Hinxton, UK, where he is a Senior Group Leader of the Human Evolution Team. His work focuses on understanding genetic variation in populations of humans and related species. His team greatly contributed to large international projects, including the 1000 Genomes Projects (concluded in 2015) and African Genome Variation Project. Chris's work mostly involves whole-genome sequencing and is exceptionally well-received by population – and forensic genetics communities.



11th Haploid Markers Conference

INFERRING ANCESTRY FROM DNA

17-19th May 2018

Bydgoszcz, Poland

Massively Parallel Sequencing: From Vision to Reality

In the last 12 months, significant advances in workflows and procedures have enabled a growing number of forensic laboratories to transition massively parallel sequencing (MPS) from a visionary method to a practical tool; validated, implemented and now beginning to influence live casework.

To further this transition and provide unprecedented support and focus to forensic laboratories looking to embrace a new reality, Verogen, the world's first company solelyfocussed on Forensic Genomics, is now the exclusive provider of Illumina-based instruments and chemistry solutions for forensic applications.

Join the Verogen lunch seminar in association with Haploid Markers 2018 to find out more about how MPS is becoming more practical and easy to use including:

- Two laboratory presentations describing alternative approaches to implementation
- The use of MPS data to re-open cold cases
- Generating enhanced results from severely degraded samples
- Handling new marker types including those for ancestry and phenotyping
- Software and chemistry improvements to facilitate casework sample analysis
- Automation of library preparation

Agenda

Konrad Suszynski, Key Accounts Manager, Thermo Fisher Scientific, Poland

Welcome and introduction – 5 min

Philipp Habermeier, Senior Commercial Manager, Western Europe, Thermo Fisher Scientific, Germany

Integrated solutions for human identification, now including IntegenX Rapid DNA Systems – 10 min

Gottfried Weichhold, Field Applications Scientist, Thermo Fisher Scientific, Germany

The Applied Biosystems™ Converge Software – an all-in-one, modular, platform for CE and Next Generation Sequencing Analysis – 30 min

Abstract

When forensic results mean the difference between innocent or guilty, authorities depend on accurate analysis. Thermo Fisher Scientific offer an unmatched combination of innovative technologies, purchasing convenience, and support. As a worldwide leader in forensics, we deliver some of the most comprehensively validated products, expertise, and application support available to the forensic science community. During this seminar at the Haploid Markers conference we'll have a quick overview of our integrated solutions for human identification; these include all the steps to make forensic genetics laboratories successful, from sample collection and extraction to PCR amplification, data analysis and all the services needed to support and validate the workflows to maximise accurate results. As part of the presentation we'll also introduce the IntegenX RapidHIT systems, the newest solutions to integrate into Thermo Fisher Scientific.

The presentation will be completed with a presentation from one of our forensic expert scientists. Data analysis is a very important part of the full forensic genetics workflow. As forensic laboratories begin to adopt sequence analysis of STR markers and mitochondria DNA sequencing by Next Generation Sequencing into casework applications, validation and concordance studies are required. The Applied Biosystems™ Converge Software, part of the Precision ID system, allows for easy comparison of NGS and capillary electrophoresis profiles. As an example, this feature is useful when comparing a crime scene sample analyzed with NGS to a reference sample that has been processed using traditional methods. The presentation will focus on the product details and its applications, including data analysis, and describe how with the adoption of NGS in forensic labs is simpler with the Applied Biosystems™ Precision ID System for Human Identification.

Day 1: Thursday, 17th May, 2018

09:00 - 13:00	**REGISTRATION/POSTER EXHIBITION**	
13:00 - 13:15	**OPENING**	

Session 1: Inferring Ancestry from DNA

13:15 - 13:45	Chris Phillips **Forensic ancestry analysis in 2018 – increasing the genetic detail brings more complexity**	KEYNOTE
13:45 - 14:00	David Ballard **Classification confusion in ancestry prediction**	
14:00 - 14.15	Lutz Roewer **Results of the first European ancestry exercise**	
14:15 - 14:30	Torben Tvedebrink **An exact likelihood ratio test for Ancestry Informative Markers**	
14:30 - 14:45	Carlo Robino **Helping the identification of migrant shipwreck victims: an extended ancestry informative marker (AIM) panel for the Tigray population of Ethiopia**	
14:45 - 15:00	Charla Marshall **An ancestry-informative approach to mitochondrial DNA testing of unknowns from the World War II Battle of Tarawa**	
15:00 - 15:15	**WRAP-UP DISCUSSION**	
15:15 - 16:45	**LUNCH SEMINAR: THERMO FISHER SCIENTIFIC**	LUNCH POSTER EXHIBITION
16:45 - 17:00	Oleg Balanovsky **NGS-Based Haplogroup-Driven Studies Powered With the Population Biobank Helps Infer the Paternal Ancestry**	
17:00 - 17:15	Catarina Xavier **Developmental validation of the VISAGE basic tool Ancestry and Appearance panel in two MPS platforms**	
17:15 - 17:30	Diana Hall **Inferring genetic ancestry with novel haplotype markers**	
17:30 - 17:45	Maria D'Amato **Ethnic and geographic distribution of genetic variation at UniQ-Typer Y-10 TM prototype kit in South Africa.**	
17:45 - 18:00	Marta Diepenbroek **The phylogenetic analyses of the human remains found in the Nazi German death camp as a proof of the Holocaust**	
18:00 - 18:15	**WRAP-UP DISCUSSION**	
18:30	**INFORMAL DINNER**	

Day 2: Friday, 18th May, 2018

Session 2: New Generation Sequencing

09:00 - 09:30	Mark Jobling **Massively parallel sequencing approaches in forensic genetics**	KEYNOTE
09:30 - 09:45	Nicole Huber **Search, Align and Haplogroup – improved forensic mtDNA analysis via EMPOP**	
09:45 - 10:00	Jodi Irwin **Next Generation Sequencing Based Mitochondrial DNA Initiatives at the FBI Laboratory**	
10:00 - 10:15	Kimberly Andreaggi **Impact of sequencing method on mitochondrial DNA length heteroplasmy**	
10:15 - 10:30	Kris van der Gaag **Heteroplasmic variation in hairs and corresponding reference samples by MPS and Sanger analysis**	
10:30 - 10:45	Bomin Kim **Investigation into Point and Length Heteroplasmy in Whole Mitochondrial Genome Using Massively Parallel Sequencing**	
10:45 - 11:00	**WRAP-UP DISCUSSION**	
11:00 - 11:30	**COFFEE BREAK**	COFFEE BREAK
11:30 - 11:45	Yali Xue **Use of the human Y chromosome for understanding the mutational mechanisms leading to structural variation**	
11:45 - 12:00	Arwin Ralf **Yleaf, an efficient software tool for high resolution Y-chromosomal haplogrouping from MPS/NGS data.**	
12:00 - 12:15	Sharon Wootton **Performance of Ancestry Inference using the Precision ID Ancestry Panel**	
12:15 - 12:30	Denise Syndercombe Court **A simplified protocol for high sensitivity mitochondrial DNA analysis.**	
12:30 - 12:45	Maria Szargut **Is MPS always the answer? PCR-based methods for Y-STR haplotyping in challenging bone samples.**	
12:45 - 13:00	**WRAP-UP DISCUSSION**	
13:00 - 14:30	**LUNCH**	LUNCH POSTER EXHIBITION
13:30 - 14:30	**LUNCH SEMINAR: VEROGEN**	

Day 2: Friday, 18th May, 2018

Session 3: Casework

14:30 - 14:45	Sascha Willuweit **Presenting the new interpretation guidelines for Y Chromosome STR typing by Forensic DNA Laboratories in Germany**	
14:45 - 15:00	Charles Brenner **The Y-haplotype Geography Problem**	
15:00 - 15:15	Maria Wróbel **SNP differences in the DNA sequence of Y chromosome's forensically relevant markers.**	
15:15 - 15:30	Chiara Turchi **Evaluation of 34 Y-SNPs on degraded DNA samples using massive parallel sequencing: a GEFI (Italian working group of ISFG) collaborative exercise**	
15:30 - 15:45	Marie Allen **DNA analysis of human remains found on the capsized Vasa warship**	
15:45 - 16:00	**WRAP-UP DISCUSSION**	
16:00 - 16:30	**COFFEE BREAK**	COFFEE BREAK
16:30 - 16:45	Athina Vidaki **Y-chromosome-based epigenetic age estimation: A novel investigative approach for male-female DNA mixtures**	
16:45 - 17:00	Jana Naue **The analysis of age-dependent DNA methylation by massive parallel sequencing in different tissues**	
17:00 - 17:15	María de la Puente **Discovery and selection of microhaplotype markers for forensic identification purposes**	
17:15 - 17:30	Sofie Claerhout **Identification of parallel Y-STR mutations in deep-routing pedigrees: A game of hide and seq**	
17:30 - 17:45	Rebecca Just **Assessing the practical value of STR sequence information for mixture interpretation**	
17:45 - 18:00	Raluca Dumache **Importance of haploid DNA markers in establishing paternity of a child from a brother-sister incest relationship**	
18:00 - 18:15	**WRAP-UP DISCUSSION**	
20:00	**CONCERT "4 SZMERY" AND DINNER**	

Day 3: Saturday, 19th May, 2018

Session 4: Population Genetics

09:00 - 09:30	Chris Tyler-Smith **Whole-genome and Y-chromosomal analyses using high-coverage sequencing**	KEYNOTE
09:30 - 09:45	Mi Hyeon Moon **Massively Parallel Sequencing of 31 Y chromosomal STRs in African Americans, European Americans, Hispanics and Koreans**	
09:45 - 10:00	Amke Caliebe **Comparison of MCMC software for evolutionary analysis of Y-STR data**	
10:00 - 10:15	Thomas Krahn **Exploring the other side of the Native American split by typing haploid markers in depth in samples from Kamchatka**	
10:15 - 10:30	Patricia Villaescusa Urbaneja **New insights in the origin and admixture of Ecuadorian populations**	
10:30 - 10:45	Leonor Gusmão **The maternal inheritance of Alto Paraná revealed by full mitogenome sequences**	
10:45 - 11:00	**WRAP-UP DISCUSSION**	
11:00 - 11:30	**COFFEE BREAK**	COFFEE BREAK
11:30 - 11:45	Maarten Larmuseau **Sex, lies and Y-chromosomes: the secret love lives of our genealogical ancestors**	
11:45 - 12:00	Stephanie Farmer **Examining Viking Ancestry in Irish Surnames**	
12:00 - 12:15	Yahya Khubrani **Exploring population structure in the paternal lineages of Saudi Arabia: from CE to massively-parallel sequencing**	
12:15 - 12:30	Andreas Tillmar **Analysis of regional DNA variation in the northern parts of Sweden reveals genetic substructures of relevance for medical and forensic genetic applications**	
12:30 - 12:45	Marcin Woźniak **101 mitochondrial genomes: diversity and distinciveness of Central European wild boars' mtDNA**	
12:45 - 13:00	**WRAP-UP DISCUSSION**	
13:00 - 13:15	**CLOSING**	

11th Haploid Markers Conference

INFERRING ANCESTRY FROM DNA

17-19th May 2018

Bydgoszcz, Poland

ORAL ABSTRACTS

Forensic ancestry analysis in 2018 – increasing the genetic detail brings more complexity

Author: **Christopher Phillips**

Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain

Forensic ancestry analysis using the DNA recovered from contact traces is making the transition from a very specialised type of investigation to a more mainstream field. This is due in no small part to the emergence of massively parallel sequencing (MPS) as a forensic tool, enabling much more secure genotyping of SNPs from scant amounts of DNA. The typing of mtDNA and Y variation is also firmly established and these markers bring powerful data for analysing maternal and paternal lineages. However, it is only in the last few years that marker sets have been brought together to construct a more detailed picture of an unidentified person's ancestry. The introduction of dedicated autosomal ancestry marker panels for MPS enhances the inferences made from mtDNA and Y loci, and at this moment in time these panels are now ready for expansion and enhancement. This talk will explore recently developed ancestry markers and new approaches to forensic ancestry analysis. New types of ancestry marker can add much more detail to casework analysis and include: X-loci; microhaplotypes; multiple-allele SNPs; and SNPs specifically selected for sub-continental population differentiation. With the increasing genetic detail such markers provide, comes increasing complexity in the genotyping tests they require, the handling of variant data generated and the interpretation necessary to gain a more accurate assessment of a person's ancestry.

Classification confusion in ancestry prediction

Author: **David Ballard[1], Phillips C[2], McNevin D[3], Aliferi A[1], Syndercombe Court D[1]**

[1]King's Forensics, King's College London, London, United Kingdom
[2]Forensic Genetics Unit, University of Santiago de Compostela, Spain
[3]Centre for Forensic Science, University of Technology Sydney

Multiple ancestry marker sets have been described in the literature, along with various commercial implementations of these sets for MPS. However, a comprehensive analysis of how successful the predictions they make can be in a blind test environment is less well described, yet knowledge of error rates and prediction confidence are vital when using these tests in a casework scenario.

Sixty-three samples were typed for ancestry purposes using two different autosomal SNP marker sets. Blind ancestry predictions were undertaken using a range of different algorithms including Bayesian and genetic distance based approaches, along with the HID SNP Genotyper Torrent Suite Server plugin (Thermo Fisher Scientific) commercial ancestry prediction solution. A subset of samples was specifically selected to stress the ancestry prediction algorithms due to the presence of individuals with mixed ancestry or those belonging to populations situated on continental margins.

The results presented here highlight the variation in ancestry prediction that is achieved depending on the analysis method selected and the population reference data used, and underlines the comprehensive validation that is required to assess classification success rates when carrying out forensic ancestry testing. The difference between simplistic versus more complex data interpretation is presented, with specific emphasis on whether increasingly detailed admixture analysis is beneficial or misleading. Lastly, we examine the benefit that is gained from the addition of mitochondrial and Y-STR markers to provide confidence and precision to the ancestry predictions based on autosomal SNP loci.

Results of the first European ancestry exercise

Author: **Lutz Roewer[1], Ansell R[2], Geppert M[1], Gross TE[3], Immel U[4], Kayser M[5], de Knijff P[6], Lutz-Bonengel S[7], Parson W[8], Phillips C[9], Ralf A[5], Schmidt U[7], Schneider PM[3], Tillmar AO[10], Lessig R[4]**

[1]Institute of Legal Medicine and Forensic Sciences, Charité – Universitätsmedizin Berlin, Germany
[2]Swedish National Laboratory of Forensic Science (SKL), Linköping, Sweden
[3]Institute of Legal Medicine, University of Cologne, Germany
[4]Institute of Legal Medicine, Martin Luther Universität Halle, Germany
[5]Erasmus MC University Medical Center Rotterdam, The Netherlands
[6]Department of Human Genetics, Leiden University Medical Center, The Netherlands
[7]Institute of Legal Medicine, Freiburg University Medical Center, Germany
[8]Institute of Legal Medicine, Medical University of Innsbruck, Austria
[9]Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[10]National Board of Forensic Medicine, Linköping, Sweden

We present results of the first exercise on ancestry prediction of an unknown DNA sample conducted by nine European forensic laboratories in five countries. Each lab applied different methods to predict ancestry e.g. Y chromosome SNPs and STRs, mitochondrial DNA and autosomal SNP panels. All labs predict correctly the continental ancestry of the sample donor. The results are discussed in the context of current legal frameworks across Europe.

An exact likelihood ratio test for Ancestry Informative Markers

Author: **Torben Tvedebrink[1], Eriksen PS[1], Mogensen HS[2], Morling N[2]**

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

Ancestry-informative markers (AIMs) are markers that give information about the ancestry of individuals. These markers are of interest to forensic geneticists as they can provide investigative leads to the police regarding the ancestry of an unknown culprit or unidentified human remains.

In forensic genetics, it is recommended to report the weight of the evidence as the likelihood ratio of two competing hypotheses. However, the fact that the likelihood ratios of the data assuming one population compared to the assumption of other populations in a database are large does not necessarily imply that any of the populations are relevant, because the populations might be exclusive, but not exhaustive, i.e. not relevant at all.

To handle this challenge, we derived a likelihood ratio test (LRT) that is a measure of absolute concordance between a profile and a population rather than a relative measure of the profile's likelihood in two populations. The LRT is similar to a Fisher's exact test.

By aggregating over markers, the central limit theorem suggests that the resulting quantity is approximately standard normally distributed. If only a few markers are genotyped or if the majority of the markers are fixed in a given population, the approximation may fail. We overcome this by using importance sampling, and we show how exponential tilting gives an efficient proposal distribution.

The methodology can handle both admixed and non-admixed individuals, where admixture is defined as individuals with different parental ancestry populations. The derived methods were implemented in a freely available interactive, webbased IT-solution (www.genogeographer.org), and as an R package (genogeographer). The interface provides the graphics and tabular summaries needed to analyse SNP based AIMs profiles.

Helping the identification of migrant shipwreck victims: an extended ancestry informative marker (AIM) panel for the Tigray population of Ethiopia

Author: **Carlo Robino[1], Lacerenza D[1], Aneli S[2], Di Gaetano C[2], Kumar H[3], Haddish K[3], Tewelmedhin G[3], Manukonda R[3], Futwi N[4], Fondevila M[5], Álvarez-Iglesias V[5], Lareu MV[5], Phillips C[5]**

[1]Dept. of Public Health Sciences and Pediatrics, University of Turin, Turin, Italy
[2]Dept. of Medical Sciences, University of Turin, Turin, Italy
[3]Dept. of Forensic Medicine, University of Mekelle, Mekelle, Ethiopia
[4]Tigray Health Research Institute, Mekelle, Ethiopia
[5]Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain

In open mass disasters with victims of different nationalities, genetic identification can be facilitated by preliminary analysis of human remains with ancestry informative markers (AIMs).

The most tragic accident in the current European refugee crisis occurred on 18th April 2015, when a migrant vessel capsized in the Straits of Sicily, causing over 800 deaths. According to the testimonies of survivors and evidence later found on board, people from Bangladesh, Syria, West and East Africa (including about 350 Eritreans) were among the victims. Following the recovery of the shipwreck in 2016, post mortem examination of the found bodies was performed, including collection of bone samples for future DNA testing.

This prompted us to create a reference dataset of forensic DNA markers, including AIMs, for the Tigray population (the major ethnic group of Northern Ethiopia and neighboring Eritrea).

Ancestry identification by means of 31 AIM-SNPs ("Global AIMs Nano set"), newly genotyped in the Tigray population in addition to 46 previously tested AIM-Indels, is discussed here. Analysis of population data derived from the literature and public catalogues of human variation showed that the Global AIMs Nano set is highly effective in discriminating between Tigray and sub-Saharan Africans from outside the Horn of Africa. Ancestry inference was also satisfactory with regards to other alleged areas of provenance of passengers of the 18th April 2015 shipwreck, such as South Asia and the Middle East. Northern Africans, although apparently not involved in the 18th April 2015 accident, are another important source of migrants to Europe and potential shipwreck victims. Preliminary comparisons with a small sample set from the Maghreb suggest that Global AIMs Nano could improve the discrimination capacity between Tigray and Northern Africans, which is limited when applying the 46 AIM-Indels panel alone.

An ancestry-informative approach to mitochondrial DNA testing of unknowns from the World War II Battle of Tarawa

Author: **Charla Marshall[1,2], Taylor R[3], Sturk-Andreaggi K[1,2], Hebda L[1,2], Fox A[1,2], Barritt-Ross S[1,2], Berg G[3], McMahon TP[1]**

[1]Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL)
[2]ARP Sciences, LLC; contractor supporting AFMES-AFDIL
[3]Defense Personnel Accounting Agency

The United States (U.S.) government is actively working to identify missing service members from the 1943 Battle of Tarawa, which caused heavy losses of U.S. and Japanese/Korean forces during World War II (WWII). The accounting effort involves the recovery of burials that were recently found on Betio Island, as well as the disinterment of 94 caskets from the National Memorial Cemetery of the Pacific. DNA testing is performed at the Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL) as part of a multidisciplinary approach to identify the missing. Given the aged and degraded nature of DNA from the burials, as well as the generational gap with family reference sample donors, mitochondrial DNA (mtDNA) is the primary line of DNA evidence for WWII unknowns. Sanger and/or next-generation sequencing is performed depending on preservation and

whether the remains were chemically treated. To date, more than 1100 samples have been submitted to AFMES-AFDIL for mtDNA testing, and although DNA testing is ongoing, 206 unique mtDNA sequences have been obtained thus far. Approximately 20% of the reported sequences produced Asian mitochondrial haplogroups, yet historical records indicate that no U.S. service members from the Battle of Tarawa were of Asian descent. Therefore it is most likely that the samples yielding Asian mtDNA sequences represent foreign nationals. In order to expedite the mtDNA testing process, the AFMES-AFDIL has begun leading with sequencing of phylogenetically-informative amplicons within the control region to target diagnostic polymorphisms of Asian haplogroups. Samples are then sorted by maternal ancestry, which determines the downstream processing trajectory. By favoring an ancestry-informative approach for this unique casework context, mtDNA testing is more efficient, which will expedite the resulting identifications and repatriation efforts.

Disclaimer: The opinions and assertions presented hereafter are the private views of the authors and should not be construed as official or as reflecting the views of the United States government.

NGS-Based Haplogroup-Driven Studies Powered With the Population Biobank Helps Inter the Paternal Ancestry

Author: **Oleg Balanovsky**[1,2]

[1]*Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russia*
[2]*Federal State Budgetary Institution «Research Centre for Medical Genetics», Moscow, Russia*

In 2011 we investigated the Y-chromosomal profile of the perpetrator of terrorist act at the Domodedovo airport – that profile is appeared to be unique for populations from Ingushetia Republic of Russia, which allowed fast identification of the perpetrator. However, in many other cases the geographic spread of the haplogroup was too wide. Thus, in the following years we aimed to narrow the zones of paternal origin by in-depth studies of the haplogroup variation – dissecting each haplogroup into tiny subbranches and studying the distribution of the subbranches. Many of them have age within 1,000 years and geographically narrow zone of their spread, thus being extremely informative to indicate the individual ancestry.

This approach includes four steps: (1) sequencing the large portions of the Y-chromosome for the target haplogroup samples; (2) constructing the phylogenetic tree, estimating the ages of branches, selecting the defining SNP for each branch; (3) screening for these SNPs in the multiple populations; (4) creating the frequency distribution maps of the subbranches – the paternal ancestry is expected from a region where branch is found.

For the crucial screening step we used collections of the Biobank of North Eurasia containing DNA samples from 26,000 individuals from 260 populations from Russia and neighboring countries. Thus we focused on the North Eurasian haplogroups C, N3, Q, and R1b. For example, within haplogroup R1b we identified a previously unstudied "eastern" branch, R1b-GG400, found in the ancient Yamnaya populations and presentday East Europeans/ West Asians but absent in West Europe, and studied the R1b-M73 branch likely linked to the Turkic-Kypchak expansion in the Eurasian steppe.

We develop a suite of instruments, including the Y-base containing Y-SNP haplogroups labels for 120,000 individuals worldwide, the Y-Atlas demonstrating world frequency distribution maps of 250 haplogroups, GeneGeo software creating the maps, and NGSConv software calling Y-chromosomal SNPs. The suite helps to analyze the Y-chromosomal variation and locate the individual ancestry.

The study was supported by the Russian Science Foundation grant 17-14-01345

Developmental validation of the VISAGE basic tool Ancestry and Appearance panel in two MPS platforms

Author: **Catarina Xavier[1], Heidegger A[1], Niederstätter H[1], de la Puente M[1,2], Phillips C[2], Kayser M[3], Parson W[1,4] and The VISAGE Consortium**

[1]*Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria*
[2]*Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain*
[3]*Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, Netherlands*
[4]*Forensic Science Program, The Pennsylvania State University, PA, USA*

Molecular intelligence is a growing field of interest within the forensic scientific community and particularly relevant in cases when DNA identification or an STR profile database match is not possible. In such cases, inferring the sample donor's biogeographic origin, appearance traits or even age can raise fundamental investigative leads to the case progress. The ability to characterize the biogeographical origin of a sample is of utmost relevance in modern societies' security and has proven to be very useful in forensic casework [1, 2]. The VISible Attributes through GEnomics (VISAGE) Consortium aims to develop, optimize and forensically validate new tools for molecular estimation of Ancestry, Appearance and Age. The first phase of this project consists in developing a basic tool with state of the art molecular markers for Ancestry/Appearance – SNPs and Age – CpGs estimation. Here we present the preliminary results of an internal validation for the Ancestry/Appearance SNP assay running in two massively parallel sequencing platforms – the Illumina MiSeq and the Thermo Fisher Scientific S5. The results presented here describe a robust method currently undergoing forensic validation tests in an inter-laboratory scenario, that can be used in the future for ancestry/appearance estimation.

Inferring genetic ancestry with novel haplotype markers

Author: **Diana Hall[1], Moriot A[1], Santos C[2], Phillips C[2]**

[1]*Forensic Genetics Unit, University Hospital of Lausanne, Switzerland*
[2]*Forensic Genetics Unit, University of Santiago de Compostela, Spain*

The use of ancestry informative markers to investigate the population of origin of an unidentified DNA sample can provide further information from evidentiary specimens. Although excellent results have been obtained with lineage markers, Indels and SNPs, and several validated marker-sets exist, the current geographic resolution can potentially be improved by combining different types of markers through haplotype analysis. Here, we aim to evaluate the contribution to ancestry inference of autosomal DIP-STR markers. DIP-STR haplotypes offer the advantage of combining low mutation rate Indels (DIPs) able to assign individuals to continents; with high mutation rate STRs informative of populations of origin (1-2). We assessed the ability of an initial set of 23 DIP-STRs to cluster the HGDP-CEPH reference populations, applying the STRUCTURE Bayesian clustering algorithm. The results at K=5 clusters, after excluding geographically close populations (Middle East and Central South Asia), show a clear pattern of five clusters corresponding to the major geographic regions of Africa, Europe, East Asia, Oceania and Native America. The results at K=7 analyzing the complete HGDP-CEPH dataset, show that Middle East and Central South Asia form less distinct clusters and Europe loses clear definition from partial membership with additional inferred clusters. Individual ancestry assignment using the Snipper likelihood-based system shows classification success rates higher than 99% in the five-group analysis, while, in the seven group analysis, samples with Eurasian origin were more difficult to classify, with about 4% of Europeans, 13% of Middle Eastern individuals and 18% of Central South Asians miss-classified to neighboring Eurasian populations. These data confirm the promising contribution of the DIP-STR haplotype approach, and show that further improvement is possible through an accurate marker selection.

Ethnic and geographic distribution of genetic variation at UniQ-Typer Y-10 TM prototype kit in South Africa.

Author: **Maria Eugenia D'Amato[1], Mohaimin K[1], Bendou H[2], Gamieldin J[2]**

[1]University of the Western Cape, Biotechnology Department, Forensic DNA Laboratory. Private Bag X17, Bellville 7535. South Africa
[2]University of the Western Cape, South African National Bioinformatics Institute. Private Bag X17, Bellville 7535. South Africa

The Y-STR profiling prototype UniQ-Typer Y-10TM is composed of DYS710, DYS518, DYS385ab, DYS644, DYS612, DYS626, DYS504, DYS481, DYS447 and DYS449 loci. The distribution of the haplotype and allele variation was investigated in a total of 1478 South African males. Thus far samples represent native groups (Xhosa, Zulu, Pedi, Tsonga, Swati, Venda; N=782), the major immigrant ethnic groups (Dutch descendants, English, Indian; N=312) and admixture in three different geographic regions (N=384). The overall DC was 0.899 with population values ranging from 0.98 (Tsonga) to 0.81 (Xhosa). A reference database under construction allows for the inference of profiles' frequency and geographic location (ystr.sanbi.ac.za). Alleles sequence analysis revealed a total of 20 unreported Repeat Pattern Variants (RPV) that presented homoplasy for loci DYS710, DYS518, DYS447 and DYS449 with correspondence to population group. Furthermore, novel sequence variants at DYS644 observed among Venda and Pedi males (Limpopo province) may require the update to this locus nomenclature. Other novel observations are: triplication at Loci DYS385, duplications at DYS710, DYS481 and DYS518. The DYS710 duplications are often linked to novel DYS644 microvariants, with observations limited to Pedi males.

Population genetic analysis is in progress. Factorial correspondence analysis of individual profiles highlights the potential of this system for ancestry and bioanthropological studies.

The phylogenetic analyses of the human remains found in the Nazi German death camp as a proof of the Holocaust

Author: **Marta Diepenbroek[1], Strobl C[2], Niederstätter H[2], Zimmermann B[2], Szargut M[1], Zielinska G[1], Ossowski A[1], Parson W[2,3]**

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

In 2013 human remains were found in Sobibór, one of 3 death camps built for the secretive Operation Reinhardt, which was the deadliest part of the Holocaust. The camp was destroyed after the revolt that took place in October 1943. All the evidence of its existence, besides the survivors testimonies, was destroyed too. To uncover the history of the place, a group of archaeologist began field work at the site. Besides finding some personal belongings they discovered also the remains of gas chambers and human ashes – the most speaking evidence of what have happened there. During the field work the research group found also skeletons, which no one expected. According to historical data all victims of Sobibór were cremated.

Therefore the remains were subjected to molecular genetic testing in order to shed light into their possible ethnic background. This study included the analysis of entire mitochondrial genomes obtained from the remains by Massively Parallel Sequencing as well as Y-chromosomal markers, both of which are known to carry phylogeographic signatures. Lineage markers analyses revealed that both mitochondrial haplogroups and individual haplotypes represented by all the remains discovered in Sobibór can be found among modern Ashkenazi Jews populations.Also the Y-chromosome haplogroups, confirmed with SNPs analysis can be linked with modern Ashkenazim. Our results, in addition to Nazi confessions and eyewitness testimonies, are the first solid proof of the Holocaust crimes.

Massively parallel sequencing approaches in forensic genetics

Author: **Mark A. Jobling**

Department of Genetics, University of Leicester, United Kingdom

Massively parallel sequencing (MPS) has great potential in forensic genetic analysis, thanks to its ability to simultaneously analyse very large numbers of loci in a single test, and to provide increased (sequence-level) resolution of individual loci. I will review MPS-based approaches to STR and SNP variation, give examples of specific studies (including some from my own lab), and speculate on the limitations and possibilities of these technologies in the forensic arena.

Search, Align and Haplogroup – improved forensic mtDNA analysis via EMPOP

Author: **Nicole Huber[1], Dür A[2], the dna.bases Consortium, Parson W[1,3]**

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

The analysis of human mitochondrial DNA (mtDNA) has proven to be extremely useful in several forensic and population genetic applications and also gains importance in disease studies. Conventionally, the obtained nucleotide sequences are aligned relative to the corrected version of the first sequenced human mitogenome and only the differences at a given position are noted and determine the haplotype.

However, a major limitation of this very common approach is the fact that difference-coded haplotypes may be ambiguous because often more than one single alignment is feasible. Consequently, database searches for forensic frequency estimates that are performed with difference-coded haplotypes are susceptible for biased results, typically the underestimation of the frequency of an mtDNA haplotype. The use of string-based search algorithms that convert query and database profiles into position-free nucleotide strings constitute an important alternative to avoid such database search bias.

The new search software presented here provides this functionality with the additional feature that sequence strings can be translated into phylogenetically aligned difference-coded haplotypes, why the tool can also be used to harmonize nomenclature between different laboratories. In addition, the presented software allows for haplogroup estimation based on Phylotree, representing a comprehensive phylogenetic tree of worldwide human mtDNA variation that is regularly updated.

The new concept extends a previously presented string search algorithm that was implemented in EMPOP in 2010 and represents a comprehensive and user-friendly tool for human mtDNA analysis based on Phylotree Build 17.

Next Generation Sequencing Based Mitochondrial DNA Initiatives at the FBI Laboratory

Author: **Jodi Irwin, Brandhagen M, Loreille O, Just R, Onorato A**

DNA Support Unit, FBI Laboratory, Quantico, VA, USA

This presentation will describe efforts underway at the FBI laboratory to develop and implement Next Generation Sequencing (NGS) assays for mitochondrial DNA testing. As a first step towards introducing NGS technology in routine casework, near-term initiatives are focused on validation of a NGS-based control region assay as a full replacement for the Sanger sequencing protocols presently in use for missing persons and criminal casework. The scope of this work includes, among other things, optimization and validation of chemistry, development of automated (robotic) workflows and design of NGS-specific laboratory information management system (LIMS) modules. Though targeting the control

region alone does not take full advantage of the benefits that NGS brings to mtDNA testing, it will firmly establish the technology in routine operational casework and serve as a vital foundation for future implementation of additional NGS assays. Over the long-term, and given the substantial benefits that complete mtGenome data bring to the discriminatory power of mtDNA evidence, we are also targeting methods that take full advantage of NGS to efficiently produce mtGenome data from both high and low-quality specimens. In addition to these efforts geared towards routine casework operations and based on targeted PCR amplification, we are also actively exploring alternate NGS workflows for the recovery of mitochondrial DNA from the most damaged and limited forensic specimens. Successful application of such assays to specimens that exceed the limits of the laboratory's current DNA testing methods will improve laboratory capabilities and greatly expand the lower range of sample quality from which probative DNA data may be recovered.

Impact of sequencing method on mitochondrial DNA length heteroplasmy

Author: **Kimberly Andreaggi[1,2,3], Parson W[4], Allen M[3], Marshall C[1,2]**

[1]ARP Sciences LLC, Rockville, MD, USA
[2]Armed Forces Medical Examiner System, Armed Forces DNA Identification Laboratory, Dover, DE, USA
[3]Uppsala University, Uppsala, Sweden
[4]Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

Mitochondrial DNA (mtDNA) analysis is an integral tool for forensic laboratories performing decades-old missing persons identifications. In recent years, sequencing of mtDNA has begun to transition from the traditional Sanger chemistry to next generation sequencing (NGS) methodologies. The quantitative nature of NGS offers the ability to characterize heteroplasmic length variants with greater reliability, potentially increasing haplotype resolution. Due to the various chemistries involved in sample preparation and sequencing platforms, differences in length heteroplasmy (LHP) may be detected depending on the approach. To investigate the impact of the sequencing workflow, four different methods were used to generate mtDNA control region (CR) data from 24 samples: PCR enrichment of the entire ~1200-bp CR followed by 1) Sanger and 2) Illumina sequencing, 3) QIAseq Targeted Human mtDNA Panel with Illumina sequencing, and 4) Precision ID mtDNA Whole Genome Panel with sequencing on the Ion S5 System. Sanger sequences were visually inspected by an experienced analyst and LHP was manually characterized in the polycytosine (C) stretches of the mtDNA hypervariable regions 1 and 2 (HV1 and HV2). STRait Razor was used to quantify molecule lengths independent of alignment in the NGS data. Although PCR parameters had minor effects on length variation, the sequencing method had the greatest impact. In the HV1 C-stretch, NGS platforms generated similar data for samples exhibiting LHP (e.g., motifs that include a T16189C). However, the major molecule in the Sanger data consistently had one more C than the NGS data. In the absence of the T16189C, nearly all (>99%) of the sequence data represented a single molecule across the HV1 homopolymeric region. Conversely, even without substantial LHP observed, NGS data generated with the Ion S5 System averaged more than twice as many molecules in HV2 than Illumina approaches and Sanger sequencing. These data highlight how the applied sequencing method, more than the sample preparation, can affect LHP including the major molecule that is typically represented in the haplotype. This evaluation will further inform both the analysis and interpretation of NGS data in homopolymeric regions.

Disclaimer: The opinions and assertions presented hereafter are the private views of the authors and should not be construed as official or as reflecting the views of the United States government.

Heteroplasmic variation in hairs and corresponding reference samples by MPS and Sanger analysis

Author: **Kris van der Gaag[1], Prieto L[2], Desmyter S[3]**

[1]Netherlands Forensic Institute (NFI) The Hague, The Netherlands
[2]Instituto de Ciencias Forenses. Universidade de Santiago de Compostela, España
[3]NICC - Belgian Institute for Forensic Science and Criminology, Brussels, Belgium

Regularly, DNA analysis of hairs uses mtDNA profiling since shed hair tend to contain insufficient genomic DNA. While mtDNA analysis does not provide the same level of discriminating power as autosomal STR profiling, it can provide useful information regarding maternal lineage. Interpretation of mtDNA sequencing data can become complicated by the occurrence of heteroplasmy, which refers to the presence of more than one type of mtDNA in an individual. Heteroplasmy can result in a different sequence for samples from the same individual, such as different tissues or separate hairs.

Previous (Sanger-based) sequencing studies of the mtDNA Control Region in several samples revealed a number of positions in which intra-individual variation between head hairs occurred regularly. Most of the corresponding positions showed detectable levels of heteroplasmy in buccal cell (BC) reference samples by Sanger sequencing but positions with lower frequency intra-person variation often displayed a homoplasmic BC Sanger sequence.

Massively Parallel Sequencing (MPS) offers a new level of resolution for the analysis of low level variation, thereby enabling new possibilities for studying the occurrence and level of heteroplasmy in different tissues or separate hairs.

In a forensic casework context, mtDNA profiles of single shed hairs are often compared to reference profiles derived from buccal swabs. In order to better understand forensic interpretation of mtDNA results of hairs, we have investigated the heteroplasmy level in separate hairs and in corresponding BC reference samples by MPS and Sanger. Since MPS allows the detection of low-level heteroplasmies, an increased understanding about the expected variation in hairs is achieved.

Using MPS by MiSeq, we re-analysed formerly Sanger-sequenced DNA extracts of hairs and BC from randomly chosen individuals [1] and studied additional individuals selected for specific variants in previously identified hotspots of heteroplasmy by using both sequencing approaches.

Investigation into Point and Length Heteroplasmy in Whole Mitochondrial Genome Using Massively Parallel Sequencing

Author: **Bo Min Kim[1,2], Lee EY[2], Shin KJ[1]**

[1]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea
[2]Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea

Mitochondrial DNA (mtDNA), which has a high copy number, is a practical forensic tool that can be analyzed if nuclear DNA is a trace amount or highly degraded. Heteroplasmy, a unique feature observed in mtDNA, is often an issue in the interpretation of forensic casework. Recently, whole mtDNA sequencing has been facilitated by massively parallel sequencing (MPS) technology; therefore, heteroplasmy analysis of whole mitochondrial genome (mtGenome) also becomes more feasible. To investigate point and length heteroplasmy of the whole mtGenome in forensic relevant samples, 20 paired samples including blood, buccal swabs and two hairs from unrelated Korean males were collected, and the whole mtDNA sequences of total 80 samples were analyzed using MPS. For blood and buccal swab samples, the whole mtGenome was amplified with two long-range PCR products; MPS libraries of these amplicons were prepared by using the Nextera XT DNA library Prep Kit and sequenced on a Miseq System. For hair shaft samples stored at room temperature for

more than one year, DNA was extracted from 2 cm of hair shafts separated from the root and amplified whole mtGenome using a Precision ID mtDNA whole genome panel. And then, the MPS library was constructed and sequenced on an Ion S5 system. By aligning MPS results with rCRS, we were able to identify mutation motifs and assign mtDNA haplogroups of each sample; moreover, a point heteroplasmy (PHP) and a length heteroplasmy (LHP) of each tissue from the same individual were scrutinized and compared.The observed frequency of PHP was not significantly different among the tissues by the individual. Fourteen individuals have different PHPs across the tissues; furthermore, 4 cases of the PHPs were reversed in the main base between the tissues.In the case of LHP that was confirmed with STRait razor v3.0, all the LHPs were observed in the hypervariable regions and 5 cases of the LHPs were different across the tissues.The results of this study will provide useful information for the application of MPS of the whole mtGenome to forensic genetic practice.

Use of the human Y chromosome for understanding the mutational mechanisms leading to structural variation

Author: **Yali Xue**

Wellcome Sanger Institute, Cambridge, United Kingdom

The Y chromosome is rich in palindromic and other repeated sequences, which show high levels of structural variation, and novel features of the history of these mutations can be elucidated. We present three such examples investigated in detail using techniques that include molecular combing fibre-FISH and 10x Genomics linked-read sequencing, illustrating different aspects of the mutational mechanisms.

1. A common copy number variant overlapping with the TTTY22 linc RNA within the inverted repeat sequence 3 (IR3) shows 0 to 3 copies in the general population (Shi et al. Hum Genet.137:73-83 (2018)). The 0-2 copy variants arise by gene conversion, and represent the largest gene conversion events (>30kb) reported thus far in the human genome. We therefore see a link between two mutational processes in humans – copy number variation and gene conversion – that are generally considered distinct.

2. The reference sequence carries 2 copies of the VCY gene in the P8 palindrome, but men with 1-4 copies of VCY are present in the population. These originate via processes that include deletion, tandem duplication and the generation of entire de novo palindromes, apparently in a single step.

3. The reference sequence also carries 6 copies of the RBMY gene in four clusters, but the copy number in a worldwide sample can vary from 4 to 11, and most men carry more than 6. Copy number variation arises in diverse ways, some of which appear to involve only tandem duplications of the RBMY genes. The mutation rate is high, and published work suggests that the copy number may influence sperm mobility, but we see no evidence of positive selection.

We see here that the Y chromosome is particularly suitable for understanding general mutational history and processes, and is the locus of choice for some such studies.

Yleaf, an efficient software tool for high resolution Y –chromosomal haplogrouping from MPS/NGS data

Author: **Arwin Ralf, Gonzalez DM, Zhong K, Kayser M**

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

With the rise of massively parallel sequencing (MPS) / Next Generation Sequencing (NGS) comes along an avalanche of human genomic data, including Y-chromosome sequence data. This has led to a dramatic increase of the number of known Y-SNPs and the complexity of the Y-chromosome phylogenetic tree known today. While the Y tree from Karafet et al.

2008 contained 311 haplogroups defined by ~600 binary markers, the current Y trees entail thousands of haplogroups and include tens of thousands of phylogenetically informative SNPs. Depending on the geographic region of interest, these highly refined haplogroups allow improved paternal bio-geographic ancestry inference, which will increase further by more population data for highly derived haplogroups being generated in the future. However, the dramatically improved Y-tree topology together with the high complexity of NGS/MPS data provides challenges for Y haplogroup assignment as prerequisite for paternal ancestry inference. To deal with these challenges and to enable researchers from a broad range of fields, including forensics, to perform high resolution Y chromosomal haplogrouping from NGS/MPS data in a routine way, we developed Yleaf. Yleaf is an automated software tool for human Y-chromosomal haplogroup assignment from MPS/NGS data by using the Y tree made available by the International Society of Genetic Genealogy (https://isogg.org/tree). For any type of MPS/NGS library containing Y-chromosomal sequence reads i.e., whole genomes and whole exomes (both low and high coverage), large genomic regions, high numbers of targeted amplicons, and capture enrichment, Yleaf has been shown to produce reliable haplogroup calling. The software tool was designed to be highly user-friendly, where a single command produces an easy to interpret output file from raw sequencing reads. A special effort was made to make the tool able to efficiently deal with the large data files generated by MPS/NGS. Currently, Yleaf considers over 41,000 Y-SNPs and can define 5358 unique Y-haplogroups. Comparative MPS/NGS data analysis using other haplogrouping software demonstrated the superior performance of Yleaf. As the Y chromosomal haplogroup tree will keep growing, this publically available tool will be updated regularly. We expect Yleaf to become the state of the art tool in Y chromosome haplogrouping useful for paternal ancestry testing in forensics as well as other fields of research where Y haplogroup knowledge is essential.

Performance of Ancestry Inference using the Precision ID Ancestry Panel

Author: **Sharon Wootton, Hasegawa R, Chang CW, Vijaychander S, Kumar A, Mehta S, Lagacé R**
Thermo Fisher Scientific, South San Francisco, CA 94010

Ancestry-informative markers (AIMs) can be useful alongside phenotype informative markers to limit a suspect pool or provide investigative leads when STR profiles are either incomplete or fail to provide a database match. On next-generation sequencing (NGS) platforms, these markers can be multiplexed by the hundreds and can present a comprehensive depiction of the unknown individual's ancestry and lineage using a consolidation of both autosomal and haploid markers.

The Precision ID Ancestry Panel was commercially released with 165 AIMs (1, 2) to run on Ion TorrentTM platforms, and recent software efforts have been made to improve biogeographic prediction accuracy and error estimation. Allele frequencies for 7 groups (Africa, Europe, Southwest Asia, South Asia, Southeast Asia, Americas, Oceania) were calculated using samples from the ALFRED database (3).

We propose a bootstrapped maximum likelihood approach to ancestry admixture prediction and assess its ability to assign biogeographical ancestry to samples from the 1000 Genomes Project and from samples typed for the ALFRED database (3). Admixed profiles were then created by simulating inheritance from randomly selected profiles, and predictions were run on the simulated offspring. Using a higher confidence interval, we demonstrate correlation between reported error and uncertainty in the prediction due to lack of differentiation. We conclude that overall, predictions are generally consistent with self-reported ancestry, and for populations with predictions of higher uncertainty, we propose inclusion of region-specific markers that can further discriminate these populations.

A simplified protocol for high sensitivity mitochondrial DNA analysis

Author: **Denise Syndercombe Court**

King's College London, United Kingdom

The analysis of bone remains that have been subject to burning is a particular challenge for forensic scientists. Mitochondrial DNA (mtDNA) sequencing is the most likely tool to provide genetic information from severely burnt human bones, yet this is problematic and many studies report a failure to analyse DNA from bones exposed to a high heat. Advances in massively parallel sequencing technology have provided new opportunities to enhance mtDNA analysis, both by simplifying the analysis of whole genome sequencing and by improving the sequencing success of compromised samples. Limiting the number of analysis steps in these high sensitivity approaches is also important to minimise the risk of sequencing mtDNA contaminants. We report here the results of a validation study on the prototype PowerSeq Mito Nested System kit from Promega for amplification of control region mitochondrial DNA with good quality sequences obtained at a sensitivity of down to 100 mtGE (the equivalent of less than 1pg of nuclear DNA) and successful amplification of highly degraded samples.

Is MPS always the answer? PCR-based methods for Y-STR haplotyping in challenging bone samples.

Author: **Maria Szargut, Diepenbroek M, Zielinska G, Arciszewska J, Cytacka S, Ossowski A**

Department of Forensic Genetics of the Pomeranian Medical University in Szczecin

The role of Y-chromosome haplotyping for the purpose of forensic genetics cannot be overestimated. Especially in personal identification cases, determining the Y-STR profile of the evidence sample is one of the primary tasks that eventually leads to identification. Although biotechnology companies offer many kits for amplification of Y-chromosomal STR markers, followed by fragment analysis by capillary electrophoresis, obtaining full Y-STR profile can still be challenging in degraded material, such as bone samples. With the dawn of MPS era, new tools for Y-STR calling arose.

In our everyday work with bone samples at the Department of Forensic Genetics of the Pomeranian Medical University in Szczecin, we have been using at first AmpFLSTR Yfiler PCR Amplification Kit, then Yfiler Plus PCR Amplification Kit (Thermo Fisher Scientific) for Y-STR haplotyping. The research which we conducted and hereby present aims at verification of utility of ForenSeq DNA Signature Prep Kit and MiSeq FGx Reagent Kit on MiSeq FGx Instrument (Illumina Inc.) for amplification and sequencing of Y-STR markers for the purpose of personal identification. We examined 90 bone samples, varying in DNA concentration and degradation index (determined by Quantifiler Trio DNA Quantification Kit, Thermo Fisher Scientific), coming from different time periods, throughout the whole XX century and recent times (2015-2017).

Presenting the new interpretation guidelines for Y Chromosome STR typing by Forensic DNA Laboratories in Germany

Author: **Lutz Roewer, Sascha Willuweit**

Institute of Legal Medicine and Forensic Sciences, Charité – Universitätsmedizin Berlin, Germany

Currently the Commission of Criminal Science and Technology of the federal and state forensic science institutes in Germany has approved interpretation guidelines for Y chromosome STR typing by forensic DNA laboratories. A panel of scientists and practitioners from police laboratories and university institutes was entrusted with the task of drafting a regulation. The recommendations are intended to assist the forensic experts in the assessment of Y-STR results, the selection of a suitable biostatistical calculation approach and the reporting.

The guidelines acknowledge the mature state of the Y-STR technique and the databasing efforts and point out that different methods to estimate Y-STR haplotype frequencies have been proposed. The recommended procedure includes specifications on the marker panel, the reference database and the frequency estimation method. In this review will discuss the differences between the guidelines in Germany and the USA as well as the initiative of the ISFG board to update their recommendations on the use of Y-STRs in forensic analysis.

The Y-haplotype Geography Problem

Author: **Charles Brenner**
DNAVIEW & UC Berkeley, USA

Usual approaches to calculating forensic DNA weight-of-evidence for a crime partly rely on suitable population reference data which is used to estimate allele or haplotype probabilities. "Suitable" ideally means the reference data is a random sample from the "universe of alternative suspects."

For this talk assume simple stain matching. Assume it's clear what the population of interest is from which we wish we had reference data. But suppose we don't. For example there is no suitable Y-haplotype data on individual Native American tribes, though there is pooled Native American data. A typical U.S. tribal reservation population is largely geographically and genetically isolated. An individual tribe is younger and smaller than the 15000 year old Native American pool. An older population has accumulated more diversity, by mutation, and a larger population better retains diversity. Therefore both influences imply that the tribe poorly represents the pool, and vise versa.

Hence evaluating DNA evidence on a tribal reservation is a problem. Peter de Knijff has suggested that the problem may even apply to villages in the Netherlands. Therefore I call it the "geography problem" although it might arise due to separation other than by geography, for example if (Asian) Indian sect or caste members recognize one another and avoid genetic invasion.

For autosomal DNA the problem is less severe. There's a reasonable argument that a non-ideal but related reference population will suffice by using a "theta" adjustment. But the theta theory is essentially drift: two populations with a common history have the same alleles whose frequencies are merely different samples from the parent population. By contrast Y-haplotypes of multiple STR loci, all rare, both drift out of existence and mutate rapidly usually to a new unique type. A thousand years after a tribe splits into two, there will be few shared YFiler haplotypes. A theta-type formula cannot work.

What can? Tribe-specific reference data probably, but that may be impractical. Perhaps we need to rethink from the beginning: look for approaches not reliant on reference data. If we could know enough about the beginning and history of an isolated tribe we could model its evolution and calculate the genetic diversity – the typical matching probability. How much is "enough"? Recent research by Anderson and Balding suggests that exacting details may not be necessary.

SNP differences in the DNA sequence of Y chromosome's forensically relevant markers.

Author: **Maria Wróbel, Parys-Proszek A, Marcińska M, Kupiec T**
Prof. Dr. Jan Sehn Institute of Forensic Research, Kraków, Poland

Y-chromosome markers are routinely used in forensic analysis. They are particularly useful in cases of sexual assault, when DNA mixtures with overbalance of female's biological material is present, in paternity cases and in identification of human remains. When there are more

men involved in a case or insufficient amount of data are available, it is recommended to use Massively Parallel Sequencing (MPS) technology which provides high degree of uniqueness of the sample's profile. In cases of frequent Y chromosome STR haplotypes, SNP analysis can give a chance to distinguish two male samples.

The aim of this study was to investigate DNA sequence variability in the Y chromosome STR markers. Two groups of samples were examined: population group which included reference samples from Poland (over 90 samples), and chalenging sample group consisting of male bone samples with different stages of degradation (over 40 samples).

All of the samples were processed with Illumina's ForenSeq kit on the MiSeq instrument using mix A or mix B chemistry. Samples were analyzed according to the standard threshold set by the company. In order to analyze additional variability with Y-STR markers, SNP sequence analysis and therefore its frequency calculation was performed.

Integrating Y-STR and Y-SNP data provide information which can increase discrimination power of the results. Moreover, Y-SNP analysis allow obtaining more relevant information in low-quantity and low-quality, degraded samples. Hence, MPS technology offers an opportunity to provide significant amount of genetic information, and in consequence, strengthening the statistical power of the evidence of a match or a kinship identification.

Evaluation of 34 Y-SNPs on degraded DNA samples using massive parallel sequencing: a GEFI (Italian working group of ISFG) collaborative exercise

Author: **Chiara Turchi[1], Onofri V[1], Bini C[2], Previderè C[3], Carnevali E[4], Robino C[5], Pelotti S[2], Fattorini P[6], Tagliabracci A[1]**

[1]Section of Legal Medicine, Polytechnic University of Marche, Ancona, Italy
[2]Legal Medicine, University of Bologna, Bologna, Italy
[3]Experimental and Forensic Medicine, University of Pavia, Italy
[4]Section of Legal Medicine and Forensic Science - S. Maria Hospital, University of Perugia, Terni, Italy
[5]Department of Public Health Sciences and Pediatrics, University of Turin, Turin, Italy
[6]Department of Medicine, Surgery and Health, University of Trieste, Italy

MPS is becoming available in forensic genetic laboratories, due to its capability in genotyping markers for identification, inference of genetic ancestry and prediction of EVCs. The Precision ID Identity Panel is designed to detect 34 Y-SNPs together with 90 autosomal SNPs. Experiments were carried out to discover improvements and limitations of this panel, including studies of specificity, sensitivity, mixtures, case-type samples and degraded samples. Six Italian forensic genetic laboratories organized a collaborative exercise to improve our understanding over the performance of the panel in casework analysis by testing real forensic stains. The laboratories participating to the study collected 90 samples consisting of 19 bone samples, 20 buccal swabs, 7 cadaveric bloods, 2 cadaveric muscle tissues, 2 fingernails, 1 fresh blood, 1 muscle tissue, 15 paraffin-embedded tissues, 16 touch DNA together with a set (n=7) of artificial depurinated DNAs. An evaluation of degradation index was performed by Quantifiler™ Trio DNA Quantification Kit. Also, a sensitivity test was assessed by a set of 2800M DNA dilutions, up to 12 pg DNA. Libraries were prepared by using the Precision ID Library Kit, testing different PCR cycles, from 21 to 26. 8 pM dilution of each library was prepared. MPS was performed by Ion Torrent PGM platform using the HiQ View Sequencing kit and 318 v2 chips. Y-SNP haplotypes were called by the HID SNP Genotyper plugin. This study investigated the concordance typing results observed in the sensitivity test; the performance of the panel was evaluated by checking the concordance between the DNA quantity, the degradation index and the typing results. Furthermore, an evaluation of the effectiveness of the 34 Y-SNP panel was assessed for forensic laboratories in routine caseworks.

DNA analysis of human remains found on the capsized Vasa warship

Author: **Marie Allen[1], Buś MM[1], Hocker F[2]**

[1]Department of Immunology, Genetics and Pathology, Biomedical Centre (BMC), Uppsala University, Uppsala, Sweden,
[2]The Vasa Museum, National Maritime Museums, Stockholm, Sweden

The Swedish Royal Warship Vasa capsized on her maiden voyage in August 1628, shortly after she left the harbour of Stockholm. In 1961, Vasa was salvaged from the Baltic Sea, and human remains were found scattered in and around the ship. After an osteological analysis, the skeletons were buried in a cemetery in 1963. The grave was reopened in 1989 for further osteological and archaeological analysis. Unfortunately, many of the remains were comingled in the ship, and the total number of individuals has remained unclear. Teeth and bone samples have been taken from several skeletons for DNA analysis. The aim of the analysis is to allow a realistic estimate of the number of individuals that died with the Vasa warship. A second objective is to investigate if some of the individuals were possible siblings, as indicated by morphological features, certain anomalies of the skulls and similar age. Also, information about visible traits as well as biogeographic ancestry is of interest, as some of the individuals have been given a facial reconstruction at the Vasa museum.

As the samples were likely to be highly degraded, at first the HVI and HVII regions of the mitochondrial D-loop region were sequenced using Sanger technology. Furthermore, Massive Parallel Sequencing (MPS) was performed on a subset of samples using the HaloPlex target selection system and the Illumina MiSeq platform. Our custom MPS panel allows simultaneous analysis of the entire mtDNA genome and over 300 autosomal and Y-chromosome markers. The markers will provide individual identification, predictions of phenotypes such as eye, hair and skin colour, as well as ancestry. The success rate for the entire MPS panel was highly correlated with the quality of the DNA samples, but at a minimum mtDNA data was obtained for the tested samples. Despite the age of the samples, the fact that they have been in seawater for 333 years, and long deposition of remains in the soil, DNA amplification, Sanger sequencing and NGS produced data of good quality.

Y-chromosome-based epigenetic age estimation: A novel investigative approach for male – female DNA mixtures

Author: **Athina Vidaki, Montiel D, Kayser M**

Department of Genetic Identification, Erasmus University Medical Center Rotterdam, Netherlands

Inferring an unknown individual's age from crime scene traces can provide important leads for police investigations in search for unknown perpetrators unidentifiable with forensic DNA profiling. Current attempts to estimate a person's age using DNA methylation are all based on autosomal markers, and thus, not suitable for analyzing commonly found mixed samples. Targeting Y-chromosomal DNA methylation markers could allow for the analysis of the male fraction in male-female DNA mixtures, therefore estimating the age of the male contributor. When successful, this would provide investigative leads for finding the unknown male perpetrator, such as in sexual assault cases, and would allow differentiating between male relatives in cases of Y-STR haplotype matches. In this study, we investigate the age predictive value of the 416 Y-chromosomal CpG sites included in the Illumina Methylation Beadchip 450K microarray. For this, we obtained 450K data from blood of several hundreds of male individuals of a wide age range (>18 years old) from public databases and normalized using standard procedures. Forward stepwise selection was implemented to successfully identify a subset of most statistically significant age-associated Y-CpGs using Bayesian information criteria (BIC) and 10-fold cross validation. For prediction modelling, we used both multiclass logistic (young-, middle-, old-aged) and multiple linear (continuous age) regression. Using the preselected, limited number of Y-CpGs (n<30), categorical age classification was achieved with high accuracy (AUC>0.85). Using linear regression, a mean

absolute deviation between true and predicted age of ~5 years was achieved. Investigation of forensically relevant tissue types other than blood, such as saliva, as well as attempts to increase the sample size for age prediction modelling are ongoing. To the best of our knowledge, this is the first study investigating age-associated DNA methylation patterns on the Y-chromosome, proposing an epigenetic age estimation strategy suitable for mixed male-female forensic samples.

The analysis of age-dependent DNA methylation by massive parallel sequencing in different tissues

Author: **Jana Naue[1,2], Sänger T[2], Hoefsloot HCJ[1], Lutz-Bonengel S[2], Verschure PJ[1], Kloosterman AD[3,4]**

[1]University of Amsterdam, Swammerdam Institute for Life Sciences, Amsterdam, The Netherlands
[2]Institute of Forensic Medicine, Medical Center – University of Freiburg, Forensic Molecular Biology, Freiburg, Germany
[3]Netherlands Forensic Institute, Biological Traces, Den Haag, The Netherlands
[4]University of Amsterdam, Institute for Biodiversity and Dynamics, Amsterdam, The Netherlands

The use of DNA methylation (DNAm) for chronological age determination has been investigated widely over the last years for application within the field of forensic genetics. The majority of studies were based on blood samples. The found age related changes cannot be directly assumed to be present in other tissues due to the cell-type-specificity of DNAm. Although bloodstains are often present at crime scenes and reference blood is available for age prediction of living individuals, the analysis of other tissues can play an important role.

In case of unidentifiable bodies we can sometimes only rely on remaining soft tissue e. g. brain and muscle, or hard tissue such as bone. Within this pilot study we investigated, if 13 previously selected age-dependent loci for blood are also informative in other forensic relevant tissues.

The DNAm levels for these 13 markers were analyzed by massive parallel sequencing (MPS) in tissue samples from brain, bone, muscle, buccal swabs, and whole blood of 29 deceased individuals with age range 0 – 87 years. The analysis included amongst others ELOVL2, TRIM59, F5, and KLF14.

Our data show the potential of DNAm analysis for age determination in tissue. A number of DNAm markers that have been shown to be informative in blood are also informative in other tissues, although another CpG sites within the loci showed a higher age-dependency in some cases. Other markers however show a tissue-specific behavior. Insight in these differences is essential when DNAm analysis for age prediction is applied in different tissues.

Discovery and selection of microhaplotype markers for forensic identification purposes

Author: **María de la Puente[1,2], Phillips C[1], Amigo J[3], Parson W[2,4], The VISAGE Consortium, Carracedo Á[1,3], Lareu MV[1]**

[1]Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[2]Institute of Legal Medicine, Medical University of Innsbruck, Austria
[3]Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain
[4]Forensic Science Program, The Pennsylvania State University, PA, USA

Massively Parallel Sequencing (MPS) technologies allow the genotyping of microhaplotype markers (MHs), comprising multiple SNPs in close physical linkage showing contrasting allele frequency patterns. Microhaplotypes offer key advantages in forensic genetics: (i) up to a hundred or more loci can be analysed simultaneously, maximizing the information obtained from typical forensic trace evidence; (ii) the sequence strand phase of the

component SNPs can be reliably detected, so the haplotype combinations amongst MH SNPs reach a higher level of polymorphism than possible from their individual genotypes; (iii) a proportion of MH loci are relatively short, so are likely to match the efficiency of MPS genotyping of degraded DNA currently seen with short-amplicon SNP analysis.

We aimed to identify a set of highly polymorphic MHs that can be implemented in MPS as a forensic tool for identification purposes, focusing on degraded DNA. First, a set of short MH loci with 2 or more SNPs was compiled from 1000 Genomes Phase III data. Second, ~100 subsets were selected from systematic searches of autosomal chromosomes. This strategy intended to obtain a final selection of ~100 MHs ensured a minimum distance between syntenic MH loci. Finally, MH candidates within each subset were ranked and prioritized according to their heterozygosity values. MHs were excluded if they failed to fulfil several sequence quality criteria.

A set of ~100 new, highly informative MHs was compiled. This set has particular relevance for identification of highly degraded DNA such as skeletal remains of mass graves in regions of conflict or identification of missing persons and disaster victims. In these cases, STR profiling can fail or yield incomplete data; while the informativeness of binary SNP genotyping with MPS panels is usually insufficient. It is envisaged that the panel will be useful as a supplementary set for analysis of distant or complex pedigrees and, since a large proportion of the MHs have more than three haplotypes, mixture analysis.

Identification of parallel Y-STR mutations in deep-routing pedigrees: A game of hide and seq

Author: **Sofie Claerhout[1], Van der Haegen M[1], Van Geel L[1], Larmuseau MHD[1], Decorte R[1,2]**

[1]KU Leuven, Forensic Biomedical Sciences, Department of Imaging & Pathology, Leuven, Belgium
[2]UZ Leuven, Laboratory of Forensic genetics and Molecular Archaeology, Leuven, Belgium

Y-STRs are very useful for lineage identification, paternity testing, kinship analysis and familial searching. It is therefore important to identify all Y-STR mutations and to have knowledge of Y-STR mutation rates in order to correctly estimate the time to the most recent common ancestor (tMRCA) between two paternally related individuals. Through the genealogical pair approach, we previously compared Y-haplotypes obtained by capillary electrophoresis (CE) of paternally related males with known judicial kinship in order to determine individual Y-STR mutation frequencies (Claerhout et al., 2018). When CE is performed to analyze genealogical pairs there is a possibility that certain Y-STR mutations will not be detected as it only characterizes the amount of repeats through size separation, referring to hidden mutations. By analyzing multiple males in extended deep-routing family pedigrees, we were able to detect parallel mutations (PM, independent mutation events to an identical number of repeats in different pedigree lineages) in several families. In total, we observed eleven PM in eight different Y-STRs with a mutation rate of at least 5.94x10-3mutations/generation (DYS449, DYS458, DYS518, DYS570, DYS576, DYS627, DYS635 and DYS724a/b). These PM are currently analyzed by Sanger sequencing in order to investigate possible sequence composition differences as four of these Y-STRs have a complex repeat motif. Until now, sequence differences have been observed within one PM detected in the complex Y-STR DYS518. Insertions on different repeat regions made it possible to reveal the presence of two different mutations leading to an identical number of repeats. This observation supports the fact that use of sequencing technology like next generation sequencing (NGS) will reveal much more information concerning the identity of Y-STR alleles with identification of hidden Y-STR mutations, and more accurate differentiation possibilities between two close relatives.

Assessing the practical value of STR sequence information for mixture interpretation

Author: **Rebecca S Just[1], Bazinet AL[2], Irwin JA[1]**

[1]DNA Support Unit, FBI Laboratory, Quantico, VA, USA
[2]National Biodefense Analysis and Countermeasures Center, Fort Detrick, MD, USA

One of the most anticipated benefits of NGS for STR typing is the potential of sequence level information to improve the interpretation of mixtures. In theory, the ability to differentiate some same-length alleles by sequence should help in both the determination of the number of mixture contributors as well as the resolution of distinct contributors. To facilitate near-term probabilistic genotyping of NGS-based autosomal STR typing results, we have proposed the addition of the longest uninterrupted stretch (LUS) length to the repeat unit as a simple method to designate sequence alleles [1]. We will present a proof of concept study, performed in an open-source probabilistic genotyping program, using this method. The results of the study provide insight into the degree to which utilizing sequence information may improve the interpretation of mixtures relative to using repeat unit information alone. Additionally, we will present the results of mixture simulation studies – performed using the same LUS concept for allele designation – that aimed to examine the potential impact of sequence information on contributor number estimation. Lastly, we will describe how these investigations may be extended to Y-STRs, given the availability of NGS assays that permit simultaneous interrogation of both marker systems.

Importance of haploid DNA markers in establishing paternity of a child from a brother-sister incest relationship

Author: **Raluca Dumache[1,3], Enache A[1], Nicolae R[2], Girbea G[2], Ciocan V[1], Muresan CO[1], Barbarii L[2]**

[1]Victor Babes University of Medicine and Pharmacy Timisoara, Romania
[2]Laboratory of Genetics-National Institute of Forensic Medicine "Mina Minovici" Bucharest, Romania
[3]Laboratory of Forensic Genetics-Institute of Forensic Medicine Timisoara, Romania

In this abstract, we present a sibling incest case involving a father, his two minor sons and his daughter with psychiatric disorders, as the victim. The 19 years old victim got pregnant and the pregnancy was interrupted at 7 weeks. The sexual intercourses were without victim's consent due to her psychiatric disorders. To establish the alleged father of the 7 weeks old female fetus, DNA from saliva samples from the victim, her father and 2 brothers was obtained by buccal swabbing, and from the aborted fetus DNA was obtained from the tissue sample. The DNA extraction was done using PureLink Genomic DNA and the DNA samples were quantified with Quantifier Duo DNA Quantification Kit on 7500 Real-Time PCR System with HID Analysis software v1.2 incorporated. The polymerase chain reaction (PCR) amplification of the DNA samples was performed on ProFlex PCR System using the AmpFLSTR Identifier Plus PCR Amplification kit. Because after amplification on STR markers, the younger brother was confirmed as biological father with a probability of paternity of 99,99999% but their father was excluded on D5S818 but confirmed on the rest of the STR markers, we supplied the analysis with the markers included in PowerPlex 16 HS and AmpFLSTR NGM SElect PCR Amplification Kit. Because the result was the same, with their biological father being excluded on D5S818 but the younger brother confirmed on all STR markers, we used AmpFLSTR Y-filer PCR Amplification Kit to confirm or exclude a mutation in case of victim's father. There was no mutation on the Y-STR markers. Further, we used the markers included in Investigator Argus X-12 QS kit to establish the father of the victim's child. The amplified products were separated and detected by using ABI 3500 analyzer.

In this case, the younger brother was confirmed as biological father of the child and the victim's father was excluded. The DNA data analysis of 23 STR markers, Amelogenin plus 12 X-STR markers, confirmed the suspicion of brother-sister incest.

In conclusion, the paternity could be proved by the analysis of 23 STR markers, Amelogenin and 12 X-STR markers even in the case of brother-sister incest.

Whole-genome and Y-chromosomal analyses using high-coverage sequencing

Author: **Chris Tyler-Smith**

Wellcome Sanger Institute, Cambridge, United Kingdom

We are using high-coverage whole-genome sequencing on the Illumina X10 platform to investigate worldwide human genetic diversity in a number of collaborative projects. (1) We have sequenced the Human Genome Diversity Project (HGDP-CEPH) panel, adding genomic sequence data to the extensive datasets already available for these samples. This is allowing us to investigate worldwide demographic history including population splits and mixtures more thoroughly using autosomal sequences, and construct a phylogeny of the Y chromosomes which shows the expected structure but reveals a number of additional rare deep-rooting lineages. (2) We have also sequenced 87 individuals belonging to 16 populations from the poorly-studied Himalayan region, revealing rapid recent male-driven expansions among high-altitude populations. (3) Finally, we have re-investigated rare African DE* Y chromosomes reported by Weale and colleagues in 2003, confirming their deep divergence from other D and E chromosomes and are considering their implications for expansion of humans out of Africa around 60,000 years ago.

Massively Parallel Sequencing of 31 Y chromosomal STRs in African Americans, European Americans, Hispanics and Koreans

Author: **Mi Hyeon Moon, Shin KJ**

Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea

Y chromosomal STR (Y-STR) is useful to aid resolving sexual crimes and to identify paternal lineage in kinship tests among males. For Y-STR analysis, a lot of capillary electrophoresis (CE)-based commercial kits are now available, but there are only a few massively parallel sequencing (MPS) assays capable of sequence based analysis of Y-STRs with high diversity. Also, it is necessary to collect the sequence variations for Y-STRs analyzed in various population groups in order to give a statistic for MPS analysis of Y-STRs in forensic practice. Therefore, we constructed an in-house MPS panel, which enables simultaneous analysis of 31 Y-STRs by adding 8 new Y-STRs (DYF387S1, DYF399S1, DYF404S1, DYS449, DYS460, DYS518, DYS612 and DYS627) that are highly polymorphic multi-copy loci and/or included in Yfiler Plus kit to the previously developed MPS panel that can analyze the PowerPlex Y23 loci and M175 marker (Kwon et al. Forensic Sci Int Genet. 2016). The new MPS panel was designed to generate small sized amplicons (85 to 274 bp) and uniform depth of coverage for each marker. For more than 200 male DNA samples from African Americans, European Americans, Hispanics and Koreans, the barcoded libraries were prepared for sequencing on a MiSeq System using MiSeq Reagent Kits v3. Then, length-based allele and sequence-based allele for each marker were obtained from the resulting FASTQ data of MPS run using the STRait Razor v3.0. The genotypes that were determined by CE and MPS methods were compared to confirm a concordance and to investigate the sequence variations of Y-STRs. In this presentation, we will show how sequence variation obtained by MPS analysis increases

Y-STR diversity calculated by CE analysis, and also the characteristics of the sequence structure observed in the Y-STRs among 4 different population groups. The results of this study are expected to enhance the applicability of Y-STRs that were analyzed by MPS method to forensic caseworks.

Comparison of MCMC software for evolutionary analysis of Y –STR data

Author: **Amke Caliebe[1], Gundlach S[1], Junge O[1], Wienbrandt L[2], Krawczak M[1],**

[1]Institute of Medical Informatics and Statistics, Kiel University, Kiel, Germany
[2]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

We investigated software for Bayesian inference of population demographics and mutation rates for Y-STR data via Markov Chain Monte Carlo (MCMC) methods. For this, coalescent trees are simulated proportional to their posterior probability by a Markov Chain moving through the tree space. Population parameters such as the scaled mutation rate theta or the exponential growth rate can then be estimated. For the analyses, datasets comprising 8 Y-STRs with a mutation rate of 0.003 per generation and 100 to 1000 samples were simulated for constant and exponentially growing population size. For each data set and programme, the length of the simulated Markov Chain was 2.2 x 107 from which 100,000 trees were analysed for the estimation of the posterior distribution of population parameters. We evaluated the four programmes BATWING, BEAST, IMa2 and LAMARC. Running time and precision and accuracy of estimated population parameters were compared for informative as well as non-informative Bayesian priors. An additional focus was laid on the variety of population genetics models incorporated in the programmes and technical features such as parallel execution, documentation, interface and the programming language.

Exploring the other side of the Native American split by typing haploid markers in depth in samples from Kamchatka

Author: **Thomas Krahn**

YSEQ

East Siberian populations have been studied broadly by their linguistics and by genetic testing. However most haplotyping studies were restricted to Y-STR profiles and the hypervariable region of the mtDNA. In our privately funded project we cooperated with local Russian historians and collected 52 samples with documented indigenous origin from Kamchatka. After typing the HVR1 and Y-STR markers we predicted the haplogroups and selected all non-European haplotypes for in depth analysis. The Y chromosome lines were followed down to their last known branch at the YFull tree by sequencing PCR amplicons around the SNP position. Next to classical Sanger sequencing a new nanopore technology based pipeline was established were more than 40 amplicons were amplified in multiplex PCR reactions and sequenced in parallel with 12 barcodes on a single flow cell. Likewise the complete mtDNA was amplified with 3 LRPCR reactions and the products were sequenced in full length on the MinION instrument. The results allowed us to clarify the Q-L53 root in depth and disclosed an A2b1 mitotype that must have been derived directly before the split of the Native American population. Y-SNP testing has disclosed a spectrum of different branches within haplogroup C-M217 and a large cluster of N-B202 samples. We think that the nanopore sequencing technology will open up new opportunities for on-site sampling and sequencing of populations in the field during the excursion.

New insights in the origin and admixture of Ecuadorian populations

Author: **Patricia Villaescusa Urbaneja[1], Geppert M[2], de Pancorbo MM[1], González-Andrade F[2], Roewer L[2]**

[1]BIOMICs Research Group, University of the Basque Country UPV/EHU. Vitoria-Gasteiz, Spain
[2]Department of Forensic Genetics, Institute of Legal Medicine and Forensic Sciences, Charité-Universitätsmedizin Berlin, Germany
[3]Central University of Ecuador, Faculty of Medical Sciences, Quito, Ecuador

Ecuador is a multicultural country, with a complex history defined by migration and admixture processes. The population of Ecuador is composed of three main ethnic groups, namely Amerindians, Mestizos and Afro-Ecuadorians. The present study has a dual aim. First, to refine the genetic characterization of the male lineages from different ethnic groups of Ecuador and analyzing the admixture state. Second, to further define the distribution of the exclusive haplogroup C-M217 in the Ecuadorian Amerindian population groups.

A total of 523 male individuals from 7 populations of Ecuador, including Kichwas (from three different locations), Tsáchilas, Mestizos and Afro-Ecuadorians (from two different locations), were analyzed for at least 23 Y-STRs and for Y-SNPs.

Our results reveal the presence of Y-SNP M217 (which defines a subhaplogroup of Asian hg C) in 17 Kichwa individuals; 13 from Amazonia, two from Salasaca and two from Otavalo. Hitherto, this haplogroup was only found in South America in 14 individuals from the Waorani and Kichwa in the Amazonian Ecuador. Two other C-M217 findings were reported from indigenous people of Peru and Venezuela. Our analyses confirm the presence of C-M217 in other Kichwa groups and outside of the Amazonian region. Its restricted radiation and absence in the other ethnic groups (including the indigenous Tsáchilas) suggest a strong founder effect in the Northwest of South America. Y-STR analysis reveals a hidden diversity among the different Kichwa populations, as the Kichwas from Salasaca display statistically significant differentiation from the other two Kichwa groups (as well as from the Kichwa from Pastaza collected from previous studies).

In conclusion, our study provides a detailed perspective of the genetic structure and variability of the main Ecuadorian ethnic groups and adds new insights into the distribution of the haplogroup C-M217 in South America, as we confirmed its presence in additional Ecuadorian Amerindian groups.

The maternal inheritance of Alto Paraná revealed by full mitogenome sequences

Author: **Simão F[1], Strobl C[2], Vullo C[3], Machado P[4], Catelli L[3], Carvalho EF[1], Leonor Gusmão[1], Parson W[2,5]**

[1]Laboratório de Diagnóstico por DNA (LDD), Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil
[2]Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[3]DNA Forensic Laboratory, Argentinean Forensic Anthropology team (EAAF), Córdoba, Argentina
[4]Legislación Farmacéutica, Universidad Católica, Hernandarias, Paraguay
[5]Forensic Science Program, The Pennsylvania State University, PA, USA

Alto Paraná is the second most populated department of Paraguay, a South American country that became a Spanish colony in the beginning of the 16th century. Due to a weaker economic power compared to other European colonies, Paraguay converted into an isolated region. A reduction in population size after the Paraguayan war in the 19th century prompted an immigration from Europe and, later, from the surrounding countries (mainly Argentina and Brazil). Mitochondrial DNA typing is a frequently used marker to characterize the genetic composition of populations. In recent years, massive parallel technologies (MPS) have been applied in different scientific fields, allowing a more in-depth and high volume genotyping. We aimed to contribute to the characterization of the maternal genetic inheritance of Paraguay using the Precision ID mtDNA panel and Ion S5™ (Thermo Fisher Scientific) technology. Mitogenomes from 78 non-related males living in Alto Paraná were sequenced. The haplotype diversity was 0.9967 (± 0.0026), which is below the values that

are usually obtained when investigating full mitogenome diversity. Indeed, apart from 58 unique haplotypes, 10 were shared between two individuals. It is worth noting that in six cases of shared haplotypes, the two individuals were born in different departments, and only one shared haplotype belong to individuals born in the same city. Considering the CR alone, the value of haplotype diversity decreases (0.9940 ± 0.0033) with 12 haplotypes being shared between two individuals and one haplotype between four. The Native-American haplogroups A, B, C and D represent 82% of the maternal ancestry in Alto Paraná. West Eurasian and African haplogroups are present in smaller percentages, 10% and 8%, respectively. The mismatch distribution profile reveals a multimodal representation with a high mean number of pairwise differences (50.14 ± 21.93). The reduction of approximately 60% in population size during the Paraguayan war seems to have not significantly affected the diversity of Native American maternal lineages. This can be explained by the much higher proportion of males that were lost during the war (almost 90%).

Sex, lies and Y-chromosomes: the secret love lives of our genealogical ancestors

Author: **Maarten H.D. Larmuseau[1,2,3], van den Berg P[2], Calafell F[4], Claerhout S[1], Gruyters L[1], Vandenbosch M[1], Nivelle K[1], Peeters A[1], Decorte R[1], Wenseleers T[1]**

[1]KU Leuven, Forensic Biomedical Science
[2]KU Leuven, Laboratory of Socioecology and Social evolution
[3]Familiekunde Vlaanderen
[4]Pompeu Fabra University, Department of Experimental and Health Sciences

In many pair-bonding species, fathers may be cuckolded into raising children that genetically are not their own. In human populations, however, the incidence and driving factors of such "extra-pair paternity" (EPP) remain contentious. Here we use a large-scale genetic genealogy approach based on Y chromosomal genotyping to reconstruct spatio-temporal patterns of human EPP rates. Using patrilineal genealogies from the Low Countries spanning a period of over 500 years and Y chromosal genotyping of living descendants, our analysis reveals that although EPP rates were low on average, they were 4-fold higher and reached 5% among the lower socioeconomic classes in the larger cities of the late 19th century. In addition, EPP rates were found to be higher if the legal fathers could exert less social control by working away from home. Together, these results suggest that human extra-pair paternity varied in relation to its potential benefits, opportunity and prevailing levels of social control.

Examining Viking Ancestry in Irish Surnames

Author: **Stephanie Farmer, Breslin K, Muralidharan C, Smyth WJ, McCarthy TV, Kayser M, Walsh S**

Indiana University Purdue University Indianapolis, USA

Ireland is a country rich in history. Individuals all over the world try to trace their heritage back to Irish roots. After the start of the Iron Age, it is believed that small tribes of Celtic people began to slowly infiltrate Ireland where their language and customs were integrated. In 795 AD, Nordic Vikings began to invade the coasts of Ireland and over time founded major cities around the country. Following this, Norman armies began to invade Ireland starting in the 12th century. These historical events helped shape both the culture of Ireland and the ancestry seen in the Irish population today. In Ireland, quite like many countries around the world, the male's surname is patrilineal, just as the Y chromosome is passed down from father to son. The relationship between Irish surnames and their corresponding Y haplogroups was examined to determine if common and rare Irish surnames can be genetically linked to the historical events listed above. The surnames chosen for this study were selected based on their prevalence in Ireland, rare or common, and their proposed historical origin, Celtic, Norse or British. To discover any possible patterns in surnames and Y chromosomal DNA, Y haplogroups were generated from the DNA of 631 Irish male subjects

using an assay specifically developed for the region. The assay contains twenty single-nucleotide polymorphisms (SNPs) that were selected to further resolve the R1b-L21 group for Irish ancestry, the most prevalent haplogroup in Western Europe, and Ireland in particular. Additional Y-STR data was also generated to examine recent surname history within the collected individuals. Each surname was examined to determine whether one haplogroup occurred more frequently and with this method, eighteen surnames investigated were found to have unique patterns. In addition to resolving Y-Surname history patterns, it is also believed that this assay may be beneficial in determining if an unknown DNA sample is of Western European origin and even in some cases, if a more specific Irish origin can be predicted.

Exploring population structure in the paternal lineages of Saudi Arabia: from CE to massively – parallel sequencing

Author: **Yahya M Khubrani[1,2], Wetton JH[1], Jobling MA[1]**

[1]Department of Genetics & Genome Biology, University of Leicester, University Road, Leicester, United Kingdom
[2]Forensic Genetics Laboratory, General Administration of Criminal Evidence, Public Security, Ministry of Interior, Kingdom of Saudi Arabia

Saudi Arabia's indigenous population is organized into patrilineal descent groups, and has a high frequency of consanguineous marriages, which may affect the diversity of uniparentally-inherited forensic markers in different sub-populations. Previously we have used the 27-STR Yfiler® Plus kit to generate haplotypes in 597 unrelated Saudi males, classified into five geographical regions (North, South, Central, East and West). Comparison of the five divisions reveals striking differences that likely reflect the geographical isolation of the desert heartland of the peninsula, and the proximity to the sea of the Eastern and Western areas, and consequent historical immigration. Here, we use massively parallel sequencing (MPS) via the ForenSeq™ DNA Signature Prep Kit (Illumina, CA) to analyse sequence diversity in Y-STR repeat and flanking regions in a subset of 92 males from our previous sample. Data are concordant with CE results, and show higher resolution as expected, particularly among haplotypes shared by males belonging to the high-frequency J1 haplogroup. Since sex-biased processes are likely to influence Saudi Arabian population structure, we also analyse maternal lineages within the same sample using Sanger sequencing of the mitochondrial DNA control region, comparing these data with information on paternal lineages. Our study shows the importance of population structure and differential effects of male and female migration, that may affect the utility of forensic databases.

Analysis of regional DNA variation in the northern parts of Sweden reveals genetic substructures of relevance for medical and forensic genetic applications

Author: **Kling D[1], Wallerström T[2], Andreas Tillmar[3]**

[1]Oslo University Hospital, Department of Forensic Sciences, Oslo, Norway
[2]Norwegian University of Science and Technology (NTNU), Department of Historical Studies, Faculty of Humanities, Trondheim, Norway
[3]National Board of Forensic Medicine, Department of Forensic Genetics and Forensic Toxicology, Linköping, Sweden

The aim of this study was to analyze regional DNA variation in the coastal areas of northern Sweden. Population studies of regional DNA variations are quite rare, and the outcome could give answers to questions related to the origin and the demography of the present day populations, be the foundation for studies in epidemiology and medical genetics, as well as increasing the quality of assessments in forensic genetic applications.

Three northern Swedish regions (Tornedalen, Piteå and Luleå) were selected. All which represent areas that, historically, haves been markedly isolated as late as the 1930's, and could be the reason for the presence of locally specific, and genetically caused, disorders.

DNA from 108 individuals were analyzed for an excess of 900,000 autosomal SNPs, 23 Y chromosomal STRs and 12 Y chromosomal SNPs. We employed an array of different statistical approaches to reveal patterns in the data sets, including population admixture, linkage disequilibrium, runs of homozygosity and multidimensional scaling.

The results reveal that all studied regions clustered, as expected, among European populations, with shortest genetic distances to other Swedish, Norwegian and Finnish regions. Interestingly the population from Piteå showed more genetic similarities with Norwegian regions than with other Swedish regions. Furthermore, individuals from Tornedalen showed to have comparably high levels of linkage disequilibrium and long homozygous runs which could indicate an inbred population with a historically small effective population size.

In summary, this study of the regional DNA variation in northern Sweden illustrates that it is important to take the observed degree of substructure into account for various medical and forensic genetic applications.

101 mitochondrial genomes: diversity and distinc tiveness of Central European wild boars' mtDNA

Author: **Gerc J[1], Radziszewska AE[1], Niedziałkowska M[2], Plis K[2], Kusza S[3], Jędrzejewska B[2], Rogalla-Ładniak U[1], Grzybowski T[1], Marcin Woźniak[1]**

[1]Division of Molecular and Forensic Genetics, Department of Forensic Medicine, Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
[2]Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland
[3]University of Debrecen, Hungary

We have sequenced 101 full mtDNA genomes of wild boars from Poland, Belarussia, Ukraine and Hungary. We have found 81 polymorphic positions, including 78 transitions and 3 transversions. Estimated haplotype diversity was 0.9174. We were able to identify 30 distinct haplotypes in our population sample, with 17 singletons. The three most frequent haplotypes comprised 46% of the population (17%, 16% and 13% respectively). Thus, the wild boar population of Central Europe seems to be extremely homogenous, as far as mtDNA is regarded.

Comparison of our full mtDNA genomes to the available GenBank sequences of European wild boars and pig breeds revealed genetic distinctiveness of the Central European wild boar population, with majority of samples forming a separate phylogenetic clad. This clad was most closely related to some of the Western European wild boars' haplotypes (Fig 1). We have also found a fraction of mtDNA haplotypes similar to commercial pig breeds, such as Mangalitsa and Pietrain. Additionally, we have identified significant structural differences in the D-loop region between our full genomes and other wild boar and pig mtDNA sequences avialable in the GenBank. The origin of this structural differences is unclear and has to be explained, as they may influence both phylogenetic and forensic analyses results.

Altogether, our 30 new mtDNA haplotypes enrich S. scrofa mtDNA genomes content in GenBank by ca. 10% and when only European pigs/wild boars are regarded, our results will constitue more than 50% of available sequences.

11th Haploid Markers Conference

INFERRING ANCESTRY FROM DNA

17-19th May 2018

Bydgoszcz, Poland

POSTER ABSTRACTS

01. Intra – and inter-population analysis of haplotype diversity in Y – filer Plus system of Kazakh population

Author: **Aliya Sarkytbayeva[1], Tazhigulova I[2], Yerezhepov D[1], Li S[2], Alibek Y[3], Sabitov Z[4], Akilzhanova A[1], Zhabagin M[1,5]**

[1]National Laboratory Astana, Nazarbayev University, Astana, Republic of Kazakhstan
[2]Centre of Forensic Expertise of the Ministry of Justice of the Republic of Kazakhstan
[3]Ministry of Internal Affairs of the Republic of Kazakhstan
[4]L.N.Gumilyov Eurasian National University, Astana, Republic of Kazakhstan
[5]National Center for Biotechnology, Astana, Republic of Kazakhstan

The short tandem repeats from Y-chromosome is a powerful tool in forensic and population genetics. Nevertheless, today the National Database „Kazakhstan" of YHRD consists of 441 minimal haplotypes, including only 41 Yfiler Plus haplotypes. In order to increase data from Kazakhstan, three laboratories passed the Quality Control Test of YHRD (YC000225, YC000343, YC000346). We updated Database by 300 new haplotypes from Kazakh population and analyzed intra – and inter-population diversity using 27 Y-STR. For comparative purposes, 8545 East, Central and West Asian 17 Y-STR haplotypes available from YHRD were used.

The obtained results showed variation of genetic diversity from 0.442 for DYS391 to 0.85 for DYS481. Highest gene diversities were found in RM Y-STR markers, which is concordant with previous results obtained by other authors. Investigation of the power to differentiate males from same clan by Yfiler Plus shows the highest values of forensic parameters among all three systems tested (Minimal Haplotype, Yfiler, Yfiler Plus). Yfiler Plus system provides unprecedented possibilities of high resolution for forensic applications. Diversity of the samples at inter-population level (RST) showed lowest RST values from Uighur and Mongolian populations. In conclusion, this study adds new information on Y chromosome diversity in Kazakh population. The study was supported by MES RK grant AP05134955.

02. The databases on Y-chromosomal and mtDNA variation across the globe and within North Eurasia

Author: **Anastasia Agdzhoyan[1,2], Zaporozhchenko V[1,2], Balanovsky O[1,2], Balanovska E[2]**

[1]Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russia
[2]Federal State Budgetary Institution «Research Centre for Medical Genetics», Moscow, Russia

The populations of North Eurasia are very genetically diverse, which creates potential for easier addressing an individual to an ancestral population. To implement this potential into the practice, the large databases and haplotype software are needed which we developed for both, mtDNA and Y-chromosomal databases.

The Y-base which we launched since 2006 includes Y-SNP data for 120,000 individuals and 40,000 STR-haplotypes. Importantly, the database includes more than 10,000 unpublished 17-STR profiles generated by our lab for the East European, Caucasus, Central Asian, Siberian, and Far East populations.

To search Y-STR databases we published free downloadable Haplomatch software, which identifies haplotypes within the selected number of mutation steps from the haplotype in question. Haplomatch works with a customer's database and allows import and export in the Excel. We monthly receive Y-STR profiles from forensic labs and using Haplomatch and Y-base report the likely region of paternal origin of the DNA samples from the crime scene.

For example, in 2014 it distinguished between Britain and Finno-Ugric ancestry of the sunk boat crew.

The mitochondrial DNA variation database Murka was launched in 2003 and consists of the two parts. Murka POP contains 225,000 profiles collected from the literature in such way that one can correctly estimate haplotype/haplogroup frequency in any particular region/country/ethnic group. The sequenced regions vary from HVS (often supplied with SNP/RFLP) to complete sequences. Murka HAP includes 30,000 profiles lacking sufficient information about their source populations and thus not directly used for searching, however this portion of the database provides data for detailed phylogenies, indel alignments etc. Both parts in ensemble, accompanied with the software, allow highly sophisticated database queries helping to solve the unusual forensic cases. Also, Murka is routinely used for estimating the probability of Ashkenazy ancestry for persons seeking the Israel citizenship.

The study was supported by the RFBR grant 16-06-00303, State Tasks 0112-2016-0006 and 0120-1363-639.

03. A Targeted Ancestry Panel for Australian and Japanese WWII Military Remains Recovered in the Asia-Pacific

Author: **Andrew Ghaiyed A[1], James KA[1], Chaseling J[1], Bernie A[4], Brownlie JC[1], Haupt LM[3], Lea RA[3], Griffiths LR[3], Wright KM[1,2,4]**

[1]School of Environment and Science, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia
[2]Royal Australian Air Force (RAAF) No 2 Expeditionary Health Squadron, Williamtown, New South Wales 2318, Australia
[3]Genomics Research Centre, Institute of Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Queensland 4059, Australia
[4]Unrecovered War Casualties-Army, Australian Defence Force, Russell Offices, 2600, Canberra, Australia

The Unrecovered War Casualties – Army (UWC-A) is an organization within the Australian Defence Force (ADF) that investigates the thousands of Australian soldiers in the Asia-Pacific that are yet to be accounted for. DNA is a critical means used to discriminate ancestry of remains recovered from complex battlespaces. Global ancestry panels are unlikely to sufficiently resolve cases for Australian and Japanese WWII military remains with membership proportions often assigned to five or seven major population groups. As a result, this approach is unlikely to provide the required resolution to individual cases exhibiting latent genetic admixture, or from populations that are not well represented in available DNA databases. A two-stepped approach is currently being developed that will allow for binary group-level resolution and which will increase discrimination power for likelihood ratio calculations. This targeted approach will ensure that only a limited number of SNPs are required, whilst also resolving latent genetic admixture, and incorporating redundancy for highly degraded remains.

1000 Genomes project data was used in simulations to demonstrate that only several SNPs are required for maximal group-level resolution. Ancestry calculations rarely include the prior probability incorporated from the investigative circumstances for the case, for example, the proportion of estimated Japanese to Australian soldiers unrecovered remains expected in a given battlefield. The inclusion of additional SNPs can then increase discrimination power for individuals in the two populations to accommodate this effect.

This research provides an ancestry/phenotype informative DNA panel that specifically targets the Australian and Japanese WWII populations, and provides a reporting and statistical framework that can be easily understood and implemented by the UWC-A. If proven to be sufficiently reliable, such a method can lead to the accounting of Australian and Japanese historical military remains with further applications for a broader forensic context.

04. Mitochondrial control region analysis by massively parallel sequencing using Powerseq™ Mito Control Nested System

Author: **Hermanson S, Ewing M, Shaw J, McLaren RS, Downey L, Andrew Hopwood, Storts DR, Promega Corporation**

Promega Corporation

Massively parallel sequencing provides the laboratory a tool to access mitochondrial DNA (mtDNA) analysis in a high-throughput workflow and with a higher level of sensitivity compared with traditional sequencing methods. Increased mixture deconvolution and heteroplasmy resolution are achieved by deep sequencing coverage and digital read counts. Additionally, the use of small amplicons to sequence the mitochondrial control region improves sequencing results from degraded samples. Traditionally, library preparation for massively parallel sequencing workflows require multiple enzymatic and purification steps that are time consuming and often a source of variability and sample loss. The prototype PowerSeq™ Mito Control Nested System utilizes a nested amplification protocol that greatly reduces the number of steps and time required to produce libraries ready for sequencing. A single PCR step both amplifies the target amplicons and incorporates indexed sequencing adapters. The system generates 10 small amplicons (adapted from Eichmann and Parson1) covering the control region of the mitochondrial genome in a single multiplex. We will demonstrate this improved workflow for the nested amplification of mitochondrial HVI and HVII control regions.

05. Evaluation of the casework direct kit for efficient screening of sexu al assault and touch DNA samples

Author: **Loten ML, Graham EK, Andrew Hopwood, Thompson JM, Promega Corporation**

Promega Corporation

Standard DNA extraction methods for forensic samples are often time-consuming and include multiple wash steps which can introduce opportunities for DNA loss. The Casework Direct Kit, Custom provides a simple, fast DNA extraction method without purification steps. The lysate generated from the Casework Direct Kit, Custom is compatible with Promega quantification and amplification systems.

As well as providing a protocol for processing a variety of DNA casework items, two specific applications of this extraction method include

1. Y-screening of sexual assault kit samples and

2. Extraction of trace DNA samples.

In the 2016 SWGDAM recommended a quick screen consisting of a non-differential extraction followed by quantification including a Y-chromosome target be applied to samples from backlog rape kits. The Casework Direct Kit, Custom used in conjunction with the PowerQuant® System performs this requirement and information related to male DNA quantity, autosomal DNA quantity, degradation, and inhibition are quickly obtained from sexual assault kit samples, thereby predicting the success of STR profile generation.

Casework Direct Kit, Custom is also compatible with a variety of casework sample types, including low-template or touch DNA samples. While DNA purification is not routinely required for subsequent STR analysis, the lysate may be purified by extraction using DNA IQ™ chemistry if IPC data in the PowerQuant® System indicates inhibition. We present data evaluating the effectiveness of using Casework Direct Kit, Custom for Y-screening and a variety of casework-type samples.

06. Developmental validation of GT-YDetector 29, a novel 5-dye multiplex kit for amplifying 29 Y-STRs for human profiling

Author: **Atefeh Joudaki, Patel RP, Shojaei Z, Rahiminejad F, Grail A, Zeinali S**

GENETEK BIOPHARMA GmbH, Berlin, Germany

Y-STR profiling is used for male profiling in sexual assault, forensic casework sample, paternity cases and genealogy studies.

The GT-YDetector 29 kit is the only commercial kit available with 5 dyes composition with the highest power of discrimination consisting 29 markers (i.e. DYS458, DYS392, DYS448, DYS576, DYS460, DYS627, DYS437, DYS393, DYS389 I, II, DYS439, DYS533, DYF387S, DYS518, DYS391, DYS456, YGATAH4, DYS390, DYS438, DYS481, DYS385 a & b, DYS570, DYS635, DYS19, DYS449, DYS643, DYS549), and AMXY in a multiplex format. Based on the YHRD database, these markers are the maximal loci which are examined in the Y-STR analysis.

We followed the SWGDAM guideline for preparing developmental validation tests, such as mixed DNA samples, PCR inhibitor, and stability studies. The GT-YDetector 29 was tested for concordance study along with other commercial kits. Obtained results were comparable with the results obtained using other commercial kits. Another advantage is that the rapid mutating markers included in this kit, are beneficial for analyzing complicated cases. The AMXY marker is also included to check whether it is a female sample or there is a PCR failure. The peak height ratio of AMXY will indicate the presence of male-female DNA mixture, hence the users can peruse follow-up test by complementary female DNA markers. In addition, the kit has been optimized to be used on filter paper-based DNA samples directly and analysis using 5-dye set platform are other key advantages.

07. The paternal legacy of San Basilio de Palenque, Colombia

Author: **Beatriz Martínez[1], Gomes V[2,3], Carvalho EF[4], Marrugo J[1], Gusmão L[4]**

[1]Molecular Genetics Lab., Institute for Immunological Research, University of Cartagena, Cartagena, Colombia
[2]Institute of Pathology and Molecular Immunology from University of Porto (IPATIMUP), Portugal
[3] Instituto de Investigação e Inovação em Saúde, i3S, Portugal
[4]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil

San Basilio de Palenque is a small town near Cartagena de Indias, Colombia, founded by escaped slaves in the sixteenth century, and is also considered the first free Afrodescendant population in America. The aim of this study was to investigate the paternal ancestry of this population and contrasting the results obtained with the information available from autosomal ancestry informative markers (AIMs). A sample from 93 unrelated males living in San Basilio were selected for this study. All samples were genotyped for the 27 Y-STR included in the Yfiler® Plus kit. A low diversity of Y-STR haplotypes was found (0.9881±0.0038), with just 76 different haplotypes being present in the studied samples. A subset of 76 samples were also genotyped for 42 Y-SNP loci, which were selected to characterize the main lineages in Native American, European and sub-Saharan African populations. A high diversity of Y-SNP haplogroups was observed, with 15 different haplogroups being present: 62% of African origin [Y-MRCA*(xM13,SRY10831.1), B2a-M150, E1a-M33, E1b1a-M2*(xM191), E1b1a-M191 and E1b1b-M35*(xM78,M81,M123)], 37% European [E1b1b-M123, E1b1b-M81, I-M170*(xM26), I2a2-M26, R1a-SRY10831.2, R1b-P25, R1b – S116*(U152,M529,M153,M167), R1b-U152, R1b-M529], and a single sample belonging to a Native American haplogroup [Q1a2-M3*(xM19,M194,M199)]. The results of a previous study accessing ancestry using autosomal AIMs showed thatthis population has 81.2% African, followed by 10.6% European, and 8.2% Native American contributions. This difference is characteristic of almost all admixed population in South America that suffered from a biased mating between European males and African or Native

females. However, considering that San Basilio de Palenque was isolated for more than two centuries, a lower frequency would be expected. It remains to be investigated at which point in the history of the Palenque these European lineages would have been introduced.

08. Genetic diversity at three palindromic sequences of the human Y chromosome

Author: **Bonito M[1], D'Atanasio E[1], Novelletto A[2], Cruciani F[1], Beniamino Trombetta[1]**

[1]Dipartimento di Biologia e Biotecnologie "C. Darwin," Sapienza Università di Roma, Rome, Italy.
[2]Dipartimento di Biologia, Università di Roma „Tor Vergata", Rome, Italy.

One of the most striking structural features of the male specific region of the human Y chromosome (MSY) is the presence, within the ampliconic sequences, of eight massive palindromes (P1-P8). Each palindrome is composed of two large inverted repeats (arms) separated by a small "spacer" sequence at the centre. These elements, ranging from 30 kb to 2.9 Mb, contain many testis-specific genes and typically exhibit > 99.9% intra-palindromic (arm-to-arm) sequence identity. It has been hypothesized that the high observed sequence similarity is due to abundant gene conversion events between the arms of each palindrome. Although the occurrence of arm-to-arm gene conversion has been clearly demonstrated, the effect of this molecular mechanism on the genetic diversity of palindromes, as well as its rate and extension, remain largely unexplored. To gain new insights into the evolutionary history of the human Y chromosome palindromic sequences and to shed light into the dynamics of intra-palindrome gene conversion, we analysed by high-coverage next-generation sequencing (50x) the shortest known palindromes (P6, P7 and P8) and their relative spacers (for a total of about 0.3 Mb) in 158 samples chosen to represent most of the independent evolutionary lineages (haplogroups) of the MSY. By this analysis we identified several gene conversion events and a peculiar mutational pattern of the palindrome arms with respect to the spacer. Moreover, we found few phylogenetically conserved paralogous sequence variants (PSVs), suggestive of a high arm-to-arm gene conversion activity. Because Y chromosomes are clonally inherited from father to son, it has been possible to capture their evolutionary relationships in a robust phylogenetic tree with known age of each node. By mapping gene conversion events across a Y tree based on thousands of stable mutations obtained from 3.3 Mb of single copy MSY sequences, we were able to calculate a precise Y-Y gene conversion rate for each of the palindromes here analysed.

09. Haplotype data for 17 Y-STR and 12 X-STR loci in nine ethnic groups from Eritrea

Author: **Carla Bini[1], Tangorra E[1], Iuvaro A[1], De Fanti S[2], Sarno S[2], Yohannes T[4], Luiselli D[2,3], Pelotti S[1]**

[1]Department of Medical and Surgical Sciences, Unit of Legal Medicine, University of Bologna, Italy
[2]Department of Biological, Geological and Environmental Sciences - Lab. of Molecular Anthropology & Centre for Genome Biology, University of Bologna, Italy;
[3]Department of Cultural Heritage, Ravenna Campus, University of Bologna, Italy;
[4]Jomo Kenyatta University and Agriculture and Technology (JKUAT), Kenya

Eritrea is a multi-ethnic country located in the Horn of Africa with nine recognized ethnic groups in its population of around 5 million of people. The existing political condition of Eritrea, governed by a secretive dictatorship accused of human-rights violations, plays an outsize role in determining a dramatic run to Europe causing one of the biggest global migration crisis. From 2012, the exodus keeps increasing especially towards Italy where in 2015 was registered the highest number of Eritrean refugees disembarked. The consequence of this desperate massive human migration event often results in tragedies like shipwrecks, rapes and death requiring forensic DNA analysis to resolve disaster victim

identification, unwanted pregnancies and missing person as well as reunification cases. In particular, markers located on the X and Y chromosomes can offer a complementary information compared to autosomal DNA typing especially in kinship testing or for inferring the paternal bio-geographic ancestryof unknown trace donors or missing persons by Y-STRs analysis. Due to the lack of data on STR frequencies, the genetic polymorphisms of 17 Y-chromosome STR loci included in the PowerPlex® Y23 kit and of 12 X-chromosome STR loci included in the Argus X12 kit were evaluated in 250 unrelated males from 9 Eritrean ethnic groups of different linguistic origin (Semitic, Cushitic and Nilotic speakers) in order to generate a relevant reference database useful for statistical evaluation. Here, we present STR allele frequencies, haplotype frequencies and further population genetic parameters of forensic interest for the two set of Y and X STRs markers.

10. Fluorescent CE-based multiplex assay for the analysis of 15 Y – STRs including rapidly and moderately mutating loci

Author: **Carla Bini, Donati B, Tangorra E, Iuvaro A, Mazzotti MC, Pelotti S**

Department of Medical and Surgical Sciences, Unit of Legal Medicine, University of Bologna, Italy

Currently available commercial Y-STR kits allow the reconstruction of the Y haplotype exploited in evolutionary and genealogical studies as well as in forensic field for identification purposes. More recently, Rapidly Mutating Y-STRs (RM Y-STRs) have been introduced in forensic analysis as useful genetic markers to improve the resolution of male lineage differentiation and to allow close male relative separation. The RM Y-STR set is useful to reduce the probability of adventitious matches and may be also applied in presumptive family rape casework to verify the hypothesis of related males involved. The present study aimed to develop a PCR multiplex system for simultaneous amplification by the automated single-capillary genetic analyzer of 15 Y-STRs including the 13 RM Y-STRs previously identified and analyzed (DYF387S1, DYF399S1, DYF403S1a/b, DYF404S1, DYS449, DYS518, DYS526 a/b, DYS547, DYS570, DYS576, DYS612, DYS626, DYS627) together with two additional STRs. The DYS533 is included in two commercial kits and the DYS534 is a moderately mutating STR and both are suitable for fluorescence detection CE instrumentation with 5-dye chemistry. The proposed multiplex assay displayed sensitivity and reproducibility, and can be of great practical use accessible also for forensic genetics laboratories not equipped with last generation technologies.

11. Evaluation of the Precision ID Whole MtDNA Genome Panel for forensic analyses

Author: **Christina Strobl[1], Eduardoff M[1], Bus MM[2], Allen M[2], Parson W[1,3]**

[1] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[2] Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden
[3] Forensic Science Program, The Pennsylvania State University, University Park, Pennsylvania, USA

In this study we present the application of the Precision ID Whole MtDNA Genome Panel (Thermo Fisher Scientific) and the Ion Personal Genome Machine (PGM) for mitochondrial DNA (mtDNA) amplification and Massively Parallel Sequencing (MPS). The panel is composed by 162 amplicons distributed in two multiplex pools with an average targeted fragment size of 175 bp. Due to these features the kit is especially suitable for typical forensic and highly compromised samples. The aim of this study is to evaluate the Precision ID panel using forensic samples that were earlier typed with conventional Sanger-type Sequencing and Capillary Electrophoresis. Our study includes 15 DNA samples obtained from hairs and hair shafts, swabs and ancient solid tissue samples (bones and teeth) that were stored in the freezer for several years. The DNA extracts produced an array of partial and full mitogenome

sequences that matched the previously obtained results and brought additional information in the control and coding regions. This data significantly increased the discrimination power and led to more specific haplogroup estimates. The potential of the panel applied to forensically relevant samples was confirmed by our results. These findings raise the appetite for further experiments to validate and apply this new technology in forensic practice.

12. A compilation of X chromosome microhaplotypes

Author: **Christopher Phillips**

Forensic Genetics Unit, University of Santiago de Compostela, Spain

The X chromosome has certain characteristics which make it a rich source of highly informative forensic markers useful for: kinship tests examining deficient pedigrees; more detailed de-convolution of mixed DNA in sexual assault cases; and, given the highly stratified distribution of X variation between populations, as well as a 75% effective population size, forensic ancestry analysis. X variation can be particularly informative when analysing patterns of co-ancestry in admixed individuals.

We report a screen of the X-chromosome for candidate microhaplotype loci, comprising 3, 4, 5 and 6 SNP combinations in sequences short enough to be efficiently amplified in forensic MPS tests. Amongst several thousand candidates obtained by criteria-led searches of 1000 Genomes data, we selected twenty with suitable characteristics for forensic use and these are detailed.

13. Typing of Syrian individuals with the Precision ID Ancestry Panel

Author: **Petersen MN, Ditte M Truelsen, Pereira V, Børsting C, Morling N**

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

With the geographical location in the intersection between Asia, Africa and Europe, the Middle East was the focus of major human expansions and population admixture events. Several population groups, especially from the Levant, contributed substantially to the genetic diversity of European populations throughout the Neolithic period. Despite the importance of the region, surprisingly few studies have investigated the autosomal genetic diversity and ancestry of the populations in the area. It is crucial to study as many populations from the region as possible and establish reference population databases in order to improve ancestry inference of individuals from the Middle East. This study focused on the genetic characterisation of 94 Syrian individuals using the Precision ID Ancestry Panel (Thermo Fisher Scientific). This commercial panel amplifies 165 ancestry informative markers (AIMs) using the AmpliSeq protocol. Library building was performed on the BioMek® 3000, and samples were sequenced on the Ion S5. The population data were analysed using Principal Component Analysis (PCA) and STRUCTURE. Furthermore, the software GenoGeographer was used to calculate population likelihoods and likelihood ratios, and to evaluate whether the Precision ID Ancestry Panel could be used to identify the ancestry of Syrian individuals.

14. Rapidly mutating Y-STRs population data and mutation rate for 13 RM Y – STRs loci in the population of Serbia

Author: **Dragana Zgonjanin[1,2], Alghafri R[3,4], Almheiri R[3], Stojiljković G[1,2], Petković S[1,2], Milić A[2]**

[1]Institute of Forensic Medicine, Clinical Center of Vojvodina, Novi Sad, Serbia
[2]Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia
[3]General Department of Forensic Sciences and Criminology, Dubai Police G.H.Q., Dubai, United Arab Emirates
[4]College of Science, Biology Department, United Arab Emirates University, United Arab Emirates

Recently, the interest of the forensic community has been focused on new Y-chromosomal short tandem repeats (Y-STRs), termed Rapidly Mutating Y-STRs (RM-YSTRs), which is able to differentiate between close male belonging to the same paternal lineage due to their high mutation rates. The aim of this study was to develop an allelic frequency database for the population of Serbia in order to evaluate the resolution power of 13 RM Y-STRs. A total of 279 unrelated males from the population of Serbia were typed with 13 RM Y-STRs markers: DYF387S1, DYF399S1, DYF403S1a/b, DYF404S1, DYS449, DYS518, DYS526a/b, DYS547, DYS570, DYS576, DYS612, DYS626 and DYS627. A high Y-STR haplotype diversity was found (0.999897) in our sample. As expected, the RM Y-STR loci showed high genetics diversity (GD) values (>0.70) in the Serbia population. The highest GD was observed for the locus DYF399S1 (0.991), followed by loci DYF403S1a (0.976), DYF387S1 (0.907) and DYF404S1 (0.899). In addition, in this study we have estimated a mutation rate for 13 RM-YSTR in 85 pairs of male relatives in the population of Serbia. Results showed that, in the 74 father-son pairs 23 mutations were detected of which 22 were one-step mutations and 1 was two-step mutation, while in the 11 twin pairs 1 mutation was observed in one dizygotic twin pair. Five father-son pairs were found to have mutations at two loci, while one pair was found at four loci. Overall, the most mutable markers were DYF399S1, DYF387S1, DYF403S1a and DYS612. Based on the results of this study, the RM Y-STR markers showed remarkable haplotype resolution power in the population of Serbia, high genetics diversity demonstrating, therefore, their usefulness in forensic identification cases.

15. Allele frequencies of the new European standard set (ESS) loci in the population of Serbia

Author: **Dragana Zgonjanin[1,2], Sojiljkovic G[1,2], Antov M[3], Petković S[1,2], Vuković R[1,2], Milić A[2]**

[1]Institute of Forensic Medicine, Clinical Center of Vojvodina, Novi Sad, Serbia
[2]Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia
[3]Faculty of Technology, University of Novi Sad, Novi Sad, Serbia

The population study was conducted to evaluate the usefulness of the five new STR loci (D10S1248, D22S1045, D2S441, D1S1656 and D12S391) included in the new European Standard Set(ESS) and to establish the allele frequencies. 400 unrelated individuals from the population sample of Serbia were genotyped using the AmpFlSTR® NGMTM PCR Amplification Kit (Applied Biosystems).Power of discrimination was the highest for the two new ESS loci, D1S1656 (0.97635) and D12S391 (0.97305). No significant deviation from the Hardy-Weinberg equilibrium was detected, except for the D2S441 loci. Calculated forensically relevant population statistics shows that the investigated loci are highly discriminative in the population of Serbia, with a combined discrimination power of 0.999999449. A comparison with previously published allele frequency data from other populations is presented.

16. Analysis of 55 Ancestry SNPs for Qatari Population

Author: **Eida Almohammed[1,2], Hadi S[2]**

[1]Ministry of Interior, Doha, Qatar
[2]University of Central Lancashire, School of Forensics and Investigative Sciences, Preston, UK

SNPs are good predictors of ethnicity and serval panels have been published (1). The ForenSeq Signature kit (Illumina) offers coverage of 230 different markers including 55 ancestry SNPs (AISNPS). The ForenSeq Universal Analysis Software (UAS) provides the capability to analyse the sequencing data, visualise results and perform statistical estimates of biogeographic ancestry. The ancestry prediction capabilities in UAS are based on Principal Component Analysis (PCA) built on several reference populations included in the 1000 Genomes project. This set does not include Qatari population. Therefore, the ancestry prediction capabilities of the ForenSeq kit through sequencing on the MiSeq FGx were evaluated by profiling 124 Qatari population samples. The samples were collected from native Qatari population from different regions in Qatar. The data was analysed using STRUCTURE software. These data serves as an addition to the existing Middle Eastern population data for the 55 AISNPS. The results of this study are presented herein.

17.A study of degraded skeletal samples using ForenSeq DNA Signature™ Kit

Author: **Eida Almohammed[1,2], Zgonjanin D[3], Hadi S[2]**

[1]Ministry of Interior, Doha, Qatar
[2]University of Central Lancashire, School of Forensics and Investigative Sciences, Preston, UK
[3]Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia

Recent advances in massively parallel sequencing (MPS) has become a very promising technology for massive genetic sequencing [1]. In this study, Illumina ForenSeq™ DNA Signature Prep Kit was tested to determine if MPS offers a more comprehensive evaluation of degraded samples than the traditional fragment analysis/capillary electrophoresis based method. The Illumina® ForenSeq™ DNA Signature MPS Kit, includes 200 genetic loci [2]. The use of NGS would therefore reduce the analysis time and augment the identification of human remains. In this context, we aimed to analyse the hard tissue degraded samples using Illumina® ForenSeq™ DNA Signature MPS Kit. These samples had given partial profiles with dropout at several loci with GlobalFiler™ kit previously. The MPS kit showed that it is highly sensitive, aids in higher allele recovery for STR loci and provides valuable information about biogeographic ancestry, identity and phenotypic features from a single analysis. The work resulted in highly successful amplification and sequencing of 30 degraded bone/teeth samples using MPS method.

18. Applying the HID-Ion AmpliSeq ™ Identity Panel and genome – wide panels for tracing ancestry from North Eurasian populations

Author: **Elena E Lukianova[1], Lavrov AV[2,3], Nekrasov SV[4], Balaganskaya OA[1], Adilgereeva EP[2], Balanovska EV[2], Balanovsky OP[1,2]**

[1]Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russia
[2]Federal State Budgetary Institution «Research Centre for Medical Genetics», Moscow, Russia
[3]Pirogov Russian National Research Medical University, Moscow, Russia
[4]Thermo Fisher Scientific, Moscow, Russia

For tracing the ethno-geographic ancestry of an individual there is commercially available HID-Ion AmpliSeq ™ Identity Panel kit, including 165 autosomal SNPs. Also, genome-wide profiles were shown to be strongly correlated with the geographic place of origin. The precision of ancestry estimation using both, small and large sets of SNPs obviously depends on the availability of the reference databases on the various ancestral populations.

The objective of this study was to create such databases of populations from Russia and neighboring countries and estimate the precision of ancestry estimations in this region.

We sequenced 200 DNA samples from nine populations – Buryats, Kazakhs, Karaims, Crimean Tatars, Russians, Tajiks, Tatars, Circassians, and Chechens – by HID-Ion AmpliSeq ™ Identity Panel kit on the Ion Torrent ™ platform and created the database of frequencies of 165 alleles. When other samples from the same populations were sequenced, their population of origin was correctly estimated in 80% cases for Asian populations but only in 40% cases for the European population.

In parallel, we compiled the database of genome-wide genotypes of populations from across North Eurasia including published datasets and data we generated using Illumina OmniExpress chip and exome sequencing. To reveal the population structure we applied a broad range of methods: Principal Component Analysis, ADMIXTURE, EEMS (migration barriers detection), Identity-By-Descent, D-statistics. We found that gene pool of populations depends not only on their position in the geographic space but also on the location of geographic barriers, migration corridors, and population size. Virtually all studied individuals clustered with their population of origin, highlighting that scans of hundred thousands SNPs allow estimate ancestry much more precisely than sets of hundreds SNPs.

19. Rapidly mutating Y-STRs in rapidly expanding populations: discrimination power of the Yfiler Plus multiplex in northern Africa.

Author: **Eugenia D'Atanasio[1], Iacovacci G[2], Pistillo R[1], Maria Bonito M[1], Dugoujon JM[3], Moral P[4], El-Chennawi F[5], Sellitto D[6], Trombetta B[1], Berti A[7], Cruciani F[1,6]**

[1]Dipartimento di Biologia e Biotecnologie "C. Darwin," Sapienza Università di Roma, Rome, Italy
[2]Banca Dati Nazionale del DNA, Direzione Centrale di Polizia Criminale, Rome, Italy
[3]Centre National de la Recherche Scientifique (CNRS), Université Toulouse-3-Paul-Sabatier, Toulouse, France
[4]Department of Animal Biology-Anthropology, Biodiversity Research Institute, University of Barcelona, Spain
[5]Clinical Pathology Department, Faculty of Medicine, Mansoura University, Mansoura, Egypt
[6]Istituto di Biologia e Patologia Molecolari, CNR, Rome, Italy
[7]Carabinieri, Reparto Investigazioni Scientifiche di Roma, Sezione di Biologia, Rome, Italy

The male-specific northern African genetic pool is characterised by a high frequency of the E-M81 haplogroup, which expanded in very recent times (2-3 kiloyears ago). As a consequence of their recent coalescence, E-M81 chromosomes often cannot be completely distinguished on the basis of their Y-STR profiles, unless rapidly-mutating Y-STRs (RM Y-STRs) are analysed. In this study, we used the Yfiler® Plus kit, which includes 7 RM Y-STRs, to analyse 477 unrelated males coming from 11 northern African populations sampled from Morocco, Algeria, Libya and Egypt. The Y chromosomes were assigned to monophyletic lineages after the analysis of 72 stable biallelic polymorphisms and, as expected, we found a high proportion of E-M81 subjects (about 46%), with frequencies decreasing from west to east. We found low intra-population diversity indexes, in particular in the populations that experienced long-term isolation. The AMOVA analysis showed significant differences between the countries and between most of the 11 populations, with a rough differentiation between northwestern Africa and northeastern Africa, where the Egyptians Berbers from Siwa represented an outlier population. The comparison between the Yfiler® and the Yfiler® Plus network of the E-M81 Y chromosomes confirmed the high power of discrimination of the latter kit, thanks to higher variability of the RM Y-STRs: indeed, the number of chromosomes sharing the same haplotype was drastically reduced from 201 to 81 and limited, in the latter case, to subjects from the same population.

20. Ancestry of ancient humans using GenoGeographer

Author: **Helle S. Mogensen[1], Tvedebrink T[2], Eriksen PS[2], Morling N[1]**

[1]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark
[2]Department of Mathematical Sciences, Aalborg University, Denmark

Ancestry informative markers (AIMs) are genetic markers that give information about the ancestry of individuals. AIMs have been widely used in the study of human genetic history and migrations. Two panels of AIMS has recently been introduced to the forensic community: the 55 AIMs from the Kidd lab and the 123 AIMs from the Seldin lab. We investigated the ancestry of the Ötzi , the Ust'-ishim, and the Denisovan Hominin individuals. We extracted genotype information for as many as possible of the Kidd and Seldin markers from public available repositories and analysed the AIMs profiles with the GenoGeographer IT-tool. The Ötzi–individual was a man, who lived between 3,400 and 3,100 BC. The mummified body was found in the Ötzal Alps. We collected genotypes of 164 markers of Ötzi from published data. The "population in database–score" indicated that the true population of origin was included among the reference populations. The AIMS-profile of Ötzi had the highest likelihood in the Italian population. However, other European, North African, and Middle Eastern populations could not be excluded.

The remains of the Ust'-ishim individual was found in western Siberia. It is estimated to be 45,000 years old. We collected genotypes for 29 of the Kidd markers and 59 of the Seldin markers. Using exclusively the 59 Seldin markers, the "population in database-score" indicated that the true population of Ust'-ishim was included among the reference populations. The highest likelihood was found with the Han Chinese population (HCS), but other East Asian populations could not be excluded.

The remains of a Denisovan Hominin, who lived 41,000 years ago was found in the Denisovan Cave in the Altai Mountains in Siberia. We collected genotypes from 16 Kidd markers and 50 Seldin markers. Using exclusively the 16 Kidd markers, the "population in database–score" indicated that the true population of origin was included among the reference populations. The highest population likelihood was found in a Nigerian population; however, other Sub Saharan populations could not be excluded.

21. Ancestry prediction capacity of the 165 ancestral informative markers in the Precision ID Ancestry Panel

Author: **Helle S. Mogensen[1], Pereira V[1], Tvedebrink T[2], Eriksen PS[2], Morling N[1]**

[1]Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark
[2]Department of Mathematical Sciences, Aalborg University, Denmark

Ancestry informative markers (AIMS) are genetic markers that gives information about the ancestry of an individual. Prediction of the ancestry are valuable information in a police investigation of crime scene samples and in identification cases. However, ancestry prediction from genetic markers isdifficult due to the complex history of humans.

In the present study, we investigated the prediction capacity of the 165 AIMS included in the Precision ID Ancestry Panel from Thermo Fisher Scientific. Firstly, we optimized and configured the reference population database. Based on genetic similarity, the reference populations were grouped into six metapopulations: Europe, Middle East, North Africa, South/Central Asia, East Asia, and Sub Saharan Africa. We also included Greenlandic and Somali reference populations, which are mixed populations, because of their frequencies and importance in Denmark. Secondly, we derived a likelihood ratio test that is a measure of the absolute concordance between an AIMs profile and that of a population rather than

a relative measure of the profile's likelihood in two populations. With this test, we were able to determine if an individual was included among the reference populations.

We investigated AIMs profiles of 578 individuals from 109 countries with genotypes of 74 or more of the 165 AIMs and analysed the results with the Genogeographer – tool (Genogeographer.org). We evaluated the results of the likelihood ratio tests in combination with the assignment of metapopulation of origin. A true population of origin was found for 413 of the 578 individuals. The population of origin of 370 of the 413 individuals (90%) were correctly assigned. Correct assignments of the populations were obtained for 86% of the European, 96% of the Middle Eastern, 77% of the South/Central Asian, 98% of the East Asian, 88% of the Sub Saharan, as well as all North African, all Greenlandic, and all Somali individuals.

22. Peopling of Jammu and Kashmir-India, a story with perspectives to Y chromosomal and mitochondrial DNA

Author: **Indu Sharma[1], Sharma V[1], Singh G[1,2], Kumar P[2], Rai E[1], Vilar M[3], Sharma S[1]**

[1]Human Genetics Research Group, Shri Mata Vaishno Devi University, Katra, Jammu and Kashmir, 182320
[2]Department of Human Genetics, Guru Nank Dev University, Amritsar
[3]Institute of Human Genetics, University of Jammu, Jammu and Kashmir, 180006, India
[4]Human Genographic Project, National Geographic Society, USA

Jammu and Kashmir (J&K) is situated in the Shivalik hills and population of state is grouped into three territories: Jammu, Kashmir and Ladakh region which are geographically isolated. This region has been suggested to have served as a corridor for various migrations and immigrations in the mainland India and Eurasia due to its location, which might have impacted the social structure of the valley. To find out the distribution of Y chromosome haplogroup in J&K, we genotyped 133 markers of non-recombining region of Y chromosome (NRY) in384 males of J&K. Genotyping was done by Agena Massarray Platform. 83 individuals were subjected to complete mitogenome (mtDNA) sequencing, on Illumina sequencing paltform (Hiseq. 2500).Our Y chromosome analysis showed distribution allocated the studied samples into eleven major haplogroups R, Q, P, L, K, J, H, G, E, F and C highlighting the genetic diversity in the region. mtDNA analysis revealed presence of 35 sub-lineages of M and 48 sub-lineages of N haplogroup, also various deep in time maternal lineages never reported before are found. The huge paternal and maternal diversity provides genetic evidence that this region has acted as corridor deep in time as well to many recent events.

23. Mutation rates for 3 Y-STR loci in a population from São Paulo state, Southeast, Brazil

Author: **Isabela B Ambrosio, Polverari FS, Valim SB, Braganholi DF, Cicarelli RMB**

Laboratory of Paternity Investigation, NAC-FCFAr - São Paulo State University, Araraquara, Brazil

Short tandem repeats (STR) are the genetic markers most often used in human identification and paternity tests. In parentage cases, the analysis of 16–20 autosomal STRs can provide a conclusive result in the majority of cases. To confirm a paternity exclusion, it is necessary to identify at least three inconsistencies between parent and child in the marker set, but sometimes one or two inconsistencies can be identified and, in these cases the analyzes of sexual chromosomes, as well as other factors such as mutations should be considered. Mutation is considered an important evolutionary force and it can be identified by comparison of genotypes between offspring and their parents. In particular, Y-chromosomal STRs (Y-STRs) are useful for deficiency paternity testing. They are transmitted without recombination from fathers to sons so they represent paternal lineages. However, spontaneous germline mutations lead to differentiation of Y-STR haplotypes between a father and his sons and

can result in an erroneous exclusion of biological. The Y-chromosome analyzes is part of the routine of our laboratory; mainly in genetic reconstruction cases which the presence of men from same patrilineal lineage of the putative father are available, or even to confirm paternity exclusion. In these particular cases it becomes necessary a table of the mutation frequency analysis on the Y-STR markers for our population (São Paulo – Brazil). Since 2011 until nowadays, our laboratory has an agreement with the public defenders of the state of São Paulo, processing 80 cases per month. The analyzes have been made by PowerPlex Fusion 6C Kit (Promega Corporation TM, Madison, WI) which contains 3 YSTRs (DYS391, DYS570 and DYS576). In this study, 242 father/son were analyzed with a total of 726 allelic transfers and 8 mutations were found. The higher mutational frequency marker was DYS576 (2.07x10-2), followed by DYS570 (8.26 x10-3) and DYS390 (4.13x10-3). Overall mutation rate across the 3 loci were 1,10x10-02 , which is a high value, since DYS576 and DYS570 are considered rapid mutation markers. In all cases, there was only one locus mutated with gain/loss of repeats in the son. When comparing the mutation rates of our population for each marker using YHRD database, they are in agreement, being the DYS576 the highest rate and the DYS391, the lowest.

24. The use of phylogenetic analysis in cases of human trafficking

Author: **Jakub Czarny, Powierska-Czarny J**

Institute of Forensic Genetics, Bydgoszcz, Poland

Forensic research in cases of human trafficking are especially challenging, since the reference material originating from victims is rarely available.We describe a case of suspicion of participation in human trafficking which included a phylogenetic analysis of results of genetic polymorphism in the range of HV1 and HV2 regions of mtDNA and Y-STR loci.Hand towels and suspenders were collected from the house belonging to people suspected of trafficking Vietnamese, Albanian and former USSR citizens from the area of a major city in Poland to Western Europe. Traces of faeces were found on the hand towels, whereas the suspenders contained traces of blood. Based on the analysis of autosomal STR markers, the traces originating from the owner of the house were discarded and biological material belonging to two males and one female was identified. Polymorphism typing in the range of HV1 and HV2 mtDNA as well as Y-STR loci was conducted in case of traces originating from males using the Y-Filer kit. The genetic material belonging to the female was subjected to typing in the range of mtDNA. Based on the results of the analyses, the maternal as well as fraternal kinship between these people was discarded. The results of polymorphism typing in the range of mtDNA were subjected to further analysis using the HaploGrep application. Based on the Y-STR typing, the haplogroup was established using Haplogrup Predictor whereas the geographical origin was determined using the yhrd.org database. It was established that the female belonged to mtDNA haplogroup M7b1a+, which occurs in Easters Asia.In case the material originating from the two males, the first was identified as the member of haplogroup N in the range of Y chromosome and mtDNA haplogroup N, whereas the second belonged to haplogroup I2a, which occurs mainly in the area of south-eastern Europe, and mtDNA haplogoup T2, which occurs in Southern Europe and the Caucasus.The presented results show that the use of phylogenetic analysis in standard forensic studies, which is realized in the range of Y-STR loci and/or mtDNA may provide valuable data regarding the probable geographical origin of people, whose traces were secured during the course of cases associated with human trafficking.

25. Identification of the Second World War victim using autosomal and Y – STR markers

Author: **Zupanič Pajnič I, Bajželj M, Gornjak Pogorelc B, Jezerka Inkret, Balažic J**

Institute of Forensic Medicine, Faculty of Medicine, University of Ljubljana

Over 100,000 victims of interwar and post-war killings are still buried in hidden mass graves all over Slovenia and remain unidentified. Here, a case of molecular genetic identification of the Slovenian castle man Rado Hribar using autosomal and Y-STR markers is presented. The castle man was liquidated along with his wife in January 1944 near his castle (Strmol). The hidden grave with skeletal remains of the spouses Hribar was found in 2015 and only incomplete remains of a male and female skeleton were excavated. The part of the skull, ribs and vertebrae were missing from the male skeleton and for the female skeleton only the skull without the lower jaw and foot were found. We suppose that the bones of the shallow grave were taken away by animals. The living relatives were traced only for Rado (two paternal nephews and niece) and since spouses did not have children the genetic identification of Ksenija was not possible. We analysed the left third molar from the female skeleton and the left second molar, right femur and right tibia from the male skeleton. We cleaned the bones and teeth, removed surface contamination, and ground them into powder. Prior to DNA isolation using Biorobot EZ1 (Qiagen), 0.5 g of tooth and bone powder was decalcified. The nuclear DNA of the samples was quantified with the PowerQuant kit and STR typing performed using different autosomal andY-STR kits. From 0.09 to 9.36 ng DNA/g of powder was obtained. The gender of both skeletons was confirmed by amelogenin and Y-STR typing and identical profiles were obtained from left second molar, right femur and right tibia. Full autosomal and Y-STR profiles allowing the identification of the Slovenian castle man Rado Hribar by comparison with family references and the relationships between males (uncle and nephews) were confirmed by Y-STRs. The product rule was used to estimate a combined LR for autosomal and Y-STRs and the statistical analyses showed a high confidence of correct identification with posterior probability of 99.997%. For traceability in the event of contamination, we created an elimination database including genetic profiles of all persons that had been in contact with the skeletal remains and no match was found. After more than 70 years, the skeletal remains of the spouses Hribar were returned to the living relatives and buried in a family tomb.

26. Victims of the biggest Second World War family killing in Slovenia identified via autosomal and haploid markers

Author: **Zupanič Pajnič I, Obal M, Bajželj M, Gornjak Pogorelc B, Jezerka Inkret, Balažic J**

Institute of Forensic Medicine, Faculty of Medicine, University of Ljubljana

WWII killing with nearly one hundred thousand victims stands as one of the major conflicts in the recent history of Slovenia. Here, we present the genetic identification of victims of the biggest family killing happened in Slovenia where 10 members of the same family were killed in 1942, and seven of them were buried in a hidden mass grave (father, three uncles, two aunts and a wife of the oldest uncle). In March 2015, the excavation of the remains began, but only 3 incomplete female skeletons were excavated. Only 20 meters away, relatives encountered bones later, and in August 2016, a burial site of at least 3 males was excavated. The victims were buried in the forest in shallow graves and the excavated skeletons were incomplete in both graves. A total of 12 bones and teeth were analysed and compared to two living relatives (son and daughter). We analysed the left second molar, femur and tibia from one female skeleton and femur and tibias from another two victims from the female grave. From the male grave 6 femurs were analysed. Prior to DNA isolation 0.5 g of powder was decalcified. The DNA was purified in a Biorobot EZ1 (Qiagen) device. The nuclear

DNA of the samples was quantified with the PowerQuant kit and STR typing performed using different autosomal andY-STR kits. Up to 2.7 ng DNA/g of powder was obtained. We managed to obtain nuclear DNA for successful STR typing from 7 bones and one molar. From the female grave profiles were obtained only for one victim (identical profiles of left second molar, femur and tibia) and from the male grave from five femurs, among them there were two pairs of femurs. Full autosomal profiles allowing the identification of 4 family members; one of the aunt from the female grave and two uncles and the father of two children used as a family references from the male grave. The relationships between males (father/son, uncles/nephew and brother victims) were further confirmed by the analyses of Y-STRs. The product rule was used to estimate a combined LR for autosomal and Y-STRs and the statistical analyses showed a high confidence of correct identification with posterior probability higher than 99.9% for all four victims identified. For traceability in the event of contamination, we created an elimination database and no match was found. After more than 70 years, the skeletal remains were returned to the surviving children (brother and sister) who buried their relatives in a family tomb.

27. Tracing Maternal Lineages on the Maltese Islands

Author: **Joanna Vella[1,2], Bodner M[3], Huber N[3], Zimmermann B[3], Borg J[2,4], Felice A[1,2], Parson W[3,5]**

[1]Department of Physiology and Biochemistry, University of Malta
[2]Malta BioBank (BBMRI.mt), Centre of Molecular Medicine and Biobanking, University of Malta
[3]Institute of Legal Medicine, Medical University of Innsbruck
[4]Department of Applied Biomedical Science, University of Malta
[5]Forensic Science Program, The Pennsylvania State University

The population of the Maltese Islands is one of the densest, with over 450,000 residents in 316 km2. The archipelago has a rich demographic history. Although historical records trace population origins to the Temple people, contemporary Maltese are descendants from those who re-populated the islands at the turn of the first millennium AD.

Maltese mitochondrial DNA (mtDNA) data is not publicly available. A new high-quality mtDNA control region (CR) dataset was set up to evaluate Maltese maternal lineages. A total of 798 samples were collected randomly with associated ancestry data from Malta and Gozo. This population collection is archived in the Malta BioBank (BBMRI.mt). The EMPOP protocol was used to amplify and sequence a subset of 300 samples with a minimum of four EMPOP sequencing primers according to forensic quality guidelines.

So far, 256 full CR sequences were typed of which 168 were unique (66%). mtDNA haplotypes were checked on EMPOP and Phylotree and haplogroup frequencies were calculated. The majority of the observed Maltese mtDNAs (76%) could be attributed to Westeurasian haplogroups: H (34%), T (18%), K (13%), J (5%), U (5%), X (1%), W (1%). African mtDNA lineages were also present: L2 (10%), L3 (1%), M1 (0.4%). Other lineages were observed at frequencies of 0.8% – 5.9%. The frequency of haplogroup H in Malta was similar to that in Sicily and Southern Italy (38%). Sub-Saharan L2 and L3 lineages were present at higher frequencies than found in Sicily and Southern Italy. The observed sub-clade L2a1c6 was previously reported in Spain (Galicia) and Portugal and is estimated to be 1,700 years old. The presence of L lineages in Malta represents the sub-Saharan gene flow into Europe.

This population genetics research provides a first insight into the matrilineal origins of the Maltese. The dataset can be used as the first national reference database for mtDNA applications in forensic and missing persons casework and population genetic studies.

28. Updated phylogenetic tree of Canis lupus familiaris mitogenomes

Author: **Katarzyna Skonieczna, Duleba A, Hołub K, Grzybowski T**

Division of Molecular and Forensic Genetics, Department of Forensic Medicine, Faculty of Medicine, Collegium Medicum, Nicolaus Copernicus University

Mitochondrial DNA (mtDNA) analysis of dogs is of special interest for forensic community, being applied to identification of biological traces. Recently published entire mitochondrial DNA sequences of Canis lupus familiaris has greatly contributed to the reconstruction of global mtDNA diversity of dogs. Here, we used 276 new mitogenomes of acceptable quality available from the Genbank database and 555 mitogenomes previously included in dogs' phylogenetic tree to update haplogroup nomenclature of C. l. familiaris. Altogether, 831 mitogenomes were used to reconstruct a global phylogenetic tree. 276 new sequences were classified into the major A, B or C haplogroups, while none of them represented the D, E or F lineages. 154 mitogenome sequences were assigned into the previously known mtDNA branches. The other 122 sequences allowed us to define 50 new subhaplogroups. Majority of them (78%) fall within the A clade. For the B and C haplogroups, six and five new subclades were defined, respectively. The updated phylogenetic tree can be further used for data quality control in forensic and population genetics.

29. Association of somatic mutations in TP53 gene with somatic mitogenome mutations in colorectal cancer cells

Author: **Katarzyna Skonieczna[1], Jawień A[2], Marszałek A[3], Grzybowski T[1]**

[1]Division of Molecular and Forensic Genetics, Department of Forensic Medicine, Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
[2]Department of Vascular Surgery and Angiology, Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland
[3]Department of Tumor Pathology and Prophylaxis, Poznan University of Medical Sciences and Greater Poland Cancer Center, Poznan, Poland

TP53 is a tumor suppresor gene that is involved in many cancer-related proccesses such as apoptosis, cell cycle arrest or activation of many genes involved in DNA repair. Previous reports suggested that p53 plays also an important role in mtDNA maintenance. As somatic mitogenome mutations are observed frequently in colorectal cancer cells it is important to verify whether they arose due to TP53 mutability. Thus, the main aim of our recent study was to investigate whether somatic mtDNA mutations are associated with TP53 mutational status. Here, we analyzed the most frequently mutated region of TP53 gene, which spans the 3 – 9 exons. 134 TP53 sequences (of cancer and matched normal tissues) were determined for 67 colorectal cancer patients using dideoxy method. Altoghether, nine hereditary polymorphisms in TP53 gene were observed in normal colon tissues. Moreover, 42 somatic mutations in TP53 gene were found in about 36% of colorectal cancer specimens. None of the identified hereditary polymorphisms (neither alleles nor genotypes) were associated with clinicopathological features or somatic mtDNA mutations in colorectal cancer cells. Somatic TP53 mutationswere not associated with clinicopathological features of colorectal cancer patients either. Nevertheless, somatic mutations in TP53 gene were significantly more frequent in colorectal cancer cells harboring somatic mitogenome mutations ($p = 0.0069$). We have also shown that somatic substitutions exclusively ($p = 0.0017$) but not solely somatic indels ($p > 0.05$) in mtDNA were associated with somatic TP53 mutations in colorectal cancer cells. Thus, the results of our study suggest that somatic TP53 mutations may lead to the loss of p53 function and further cause accumulation of somatic substitutions but not indels in mitochondrial genomes of colorectal cancer cells.

30. The mitochondrial DNA polymorphisms in “situs inversus” detected by Next Generation Sequencing

Author: **Umino K, Muramatsu H, Katsuya Honda, Ishizawa F, Iwabuchi Y, Sugano Y**

Department of Legal Medicine , University of Tsukuba

As a great output is obtained, the next-generation sequencing provided as a great tool for a DNA investigation. Particularly, in the mitochondrial DNA study, this method is extremely effective because we can analyze whole genome simultaneously as well as D loop. We already reported mitochondrial DNA polymorphism with the sex chromosomal aberration, Turner syndrome. In addition, we analyzed all mitochondrial genomes about case of situs inversus. Situs inversus is the congenital malformation that all internal organs were formed in a mirror image.

Material and methods

Total mtDNA genome in situs inversus were analyzed. Long range PCR were performed with Nextera preparation spanning the entire human mitochondrial genome (16,569 bp). Library sequencing on the MiSeq (Illumina) is followed by data analysis. We compared it with the rCRS sequence by second analysis software of the mtDNA variant analyzer. Furthermore, polymorphism analysis was done by EMPOP system (R11).

Results Mt DNA sequence in Situs inversus was found to be unique, and specific, as it was not found in the normal population (database 0/27399) We detected SNP site, there were, 16209C 16223T 16269G 16324C, 73G 146C 204C 207A 263G 408A 489C. By EMPOP database. The mtDNA from the case of Situs inversus, defined M7a1b1 (missing mutation G207A, private mutation T146C, haplotype with similar sequence is not found).

Discussion Situs inversus is the congenital anomalies that relatively have high incidence to be found in one in approximately several thousand. According to the Kawabe et al., It is estimated that this abnormal incidence in Japan is approximately 2 times than Europe and America. As three patients of ten siblings is already reported in Japan, we cannot deny the genetic transmission of this abnormality (Oshima et al.). Our data show the possibility that mitochondria DNA haplotype has some kind of relations with this abnormality.

31. Analysis of the heredity of Turner syndrome detected by the MiSeq FGx Forensic Genomics System

Author: **Katsuya Honda, Muramatsu H, Umino K, Ishizawa F, Iwabuchi Y, Sugano Y**

Department of Legal Medicine , University of Tsukuba

The MiSeq FGx Forensic Genomics System is a powerful tool for examination of genetic research, because we can simultaneously amplify up to 231 forensic loci in a single multiplexed PCR, targeted loci include 27 common forensic autosomal STRs, 24 Y-STRs, 7 X-STRs, amelogenin and three classes of single nucleotide polymorphism (SNPs). In this study, we used this system for examine the mode of heredity in X chromosomal aberration, Turner`s (XO, female) syndrome. Many polymorphic loci were analyzed by MiSeq FGx system form blood sample in Turner`s (XO, female: including mother and affected daughter). The analytical procedure was carried out following instructions of the defined protocol by Illumina. Turner syndrome are known to be hemizygote of XO. By our experiment, ten sites of homozygotes were detected among 27 autosomal STRs. These data suggest that there is a genetic deviation in this disease in not only sex chromosomes but also autosome STRs. The hereditary between the mother and the child did not have the major contradiction, however, in TPOX, the mutation from an 8 type to a 16 type was suspected. Autosomal STR of D21S11 in chromosome 21, a mixture of the allele of 30/31/33.2/ was found. We already reported that specific polymorphism of mitochondrial DNA in Turner syndrome. Our data

indicate that possible interaction of the sex chromosome and the autosomal chromosome exists. MiSeq FGx Forensic Genomics System was developed for forensic identification, but it was found that there was alternative utility to scan a genome widely, and to be able to screen a genetic disorder. We want to push forward analysis about the SNP mutation in future.

32. Expansion of mitogenome reference data with next-generation sequencing technologies for forensic applications

Author: **Kimberly Sturk-Andreaggi[1,2], Marshall C[1,2]**

[1]ARP Sciences LLC, Rockville, MD, USA
[2]Armed Forces Medical Examiner System, Armed Forces DNA Identification Laboratory, Dover, DE, USA

The Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL) received a 2017 Research and Development in Forensic Science grant from the National Institute of Justice to expand mitogenome reference data with next-generation sequencing technologies. This project will provide a quality-controlled database of 5,000 mitogenome haplotypes from individuals of diverse maternal ancestry for worldwide forensic, medicolegal and public use. The mitogenome databasing project will utilize high-quality, single-source samples from anonymized donors originating from several regions of the United States with known provenance as well as a smaller sampling of global populations. A well-vetted NGS method of long-range PCR, dual-indexed library preparation, Illumina sequencing, and mitogenome-tailored bioinformatic data analysis will be utilized to produce accurate haplotypes. A background signal assessment will determine the minimum coverage and variant frequency thresholds that can be utilized for reliable variant detection. Automation will be incorporated into the wet and dry lab portions of sample processing to generate the haplotypes in a high-throughput and cost-effective fashion. The speed and accuracy of the data generation will be coupled with stringent quality review involving internal and external quality control (QC) checks to ensure the integrity of the haplotypes obtained. Interpretation guidelines for mitogenome profile reporting and comparison will be developed from the data and opened for comment by the broader forensics community. Moreover, nuclear DNA (nDNA) data including autosomal and Y-chromosomal STR alleles will be generated from a subset (~500) of the samples. While this data will be primarily used as a QC measure, the nDNA sequence data will be made publicly available. Altogether, this quality-controlled mitogenome database will enable forensic practitioners to make statistical inferences from casework mitogenome data, thereby facilitating the use of the entire mtDNA locus in forensic casework.

Disclaimer: The opinions and assertions presented hereafter are the private views of the authors and should not be construed as official or as reflecting the views of the United States government.Kimberly Sturk-Andreaggi and Charla Marshall

33. Y-chromosomal STR typing of cell-free DNA circulating in maternal plasma for prenatal paternity testing

Author: **Dziennik A, Krzysztof Rębała**

Department of Forensic Medicine, Medical University of Gdansk, Gdansk, Poland

Different studies demonstrate that foetal DNA constitutes from 3.4–9.7% of total cell-free DNA (cfDNA) circulating in maternal plasma in early pregnancy up to 6.2–20.4% in late pregnancy. Owing to specificity to males, Y-chromosomal STR typing has been proven to detect and discriminate male component in male/female DNA mixtures even at a ratio of 1:1000. Thus, taking into account a fact that the Y chromosome is always inherited paternally, Y-STR typing of plasma cfDNA seems to be a powerful tool to resolve paternal lineage of a male foetus

for prenatal paternity testing. The issue of paternity of a foetus is particularly important from the perspective of the Polish law, which sanctions abortion when a pregnancy results from a crime. The aim of our study was to assess efficiency, sensitivity and specificity of Y-STR typing of plasma cfDNA as a tool for prenatal paternity testing. Peripheral blood samples were collected from roughly five dozen pregnant women at different stages of pregnancy (24–39 week), carrying both male and female foeti. Peripheral blood was centrifuged within 8 hours from collection of the samples to separate plasma, which later underwent cfDNA extraction with the use of a QIAamp DSP Virus Spin Kit (Qiagen). In addition, umbilical cord blood samples of neonates were collected immediately after birth as control specimens. Y-chromosomal microsatellites were amplified with the use of an AmpFlSTR Yfiler PCR Amplification Kit (Thermo Fisher Scientific). Results of Y-STR typing of plasma cfDNA were confronted with results of umbilical cord genotyping. The findings and conclusions drawn from the study are shown and discussed, especially in relation to the gender of the foetus and the stage of the pregnancy.

34. A DNA-Military Ancestry Predictor (DNA-MAP): Software to assist in ancestry prediction of unidentified historical military remains.

Author: **Kyle James[1], Chaseling J[1], Wright K[1,2,3], Bernie A[3], Gabric A[1]**

[1]School of Environment and Science, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia
[2]Royal Australian Air Force (RAAF) No 2 Expeditionary Health Squadron, Williamtown, New South Wales 2318, Australia
[3]Unrecovered War Casualties-Army, Australian Defence Force, Russell Offices, 2600, Canberra, Australia

DNA-Military Ancestry Predictor (MAP) is a simple Knowledge Based Decision Support System (KBDSS) developed to assist the Unrecovered War Casualties – Army (UWC-A) predict ancestry of historical military remains. Ancestry prediction involves the application of the Bayesian formula using conditional probabilities. A theoretical approach to studying the effects of the parameters involved becomes complex and unwieldy with a large number of inputs, each with a different degree of reliability. This approach is also unsuitable for end users; UWC-A Investigators and members of the Australian Defence Force Identification Board.

DNA-MAP uses combined genetic and historical data to evaluate the probability of a set of remains belonging to either an Australian or a Japanese World War II (WWII) soldier. DNA-MAP utilises What-If scenarios developed from realistic possibilities. The prior expectation of relative numbers of Australian soldiers missing in geographical areas may be available from historical army records, but actual values from different sources may vary considerably. Identification of appropriate reference populations is also difficult. The composition of the population of Australian WWII soldiers is quite different from today's Australian population. It is also unlikely the DNA probability estimates from individual countries are appropriate for people of these nationalities during WWII. Migrants to Australia do not represent random samples from their respective countries. Variation and unreliability also stem from huge variation in sample sizes available to estimate DNA probabilities. Emphasis is on using DNA areas which are rare in one nationality and common in another and it is important that samples are of sufficient size, ensuring that if a rare event does exist it will be detected.

The What-If scenarios enable the user to see the effects varying these different parameters have on the outcome, namely, the probability the recovered remains are of Australian ancestry. This paper presents the first stage of DNA-MAP's development.

35. A Methods of detecting a rare event – the rare topic event in forensics

Author: **Kyle James[1], Chaseling J[1], Wright K[1,2,3], Bernie A[3], Gabric A[1]**

[1]School of Environment and Science, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia
[2]Royal Australian Air Force (RAAF) No 2 Expeditionary Health Squadron, Williamtown, New South Wales 2318, Australia
[3]Unrecovered War Casualties-Army, Australian Defence Force, Russell Offices, 2600, Canberra, Australia

Ancestry analysis depends on assigning an individual to a population based on the presence or absence of specific genetic traits believed to be carried by individuals in that population. Previous data of varying sample sizes are used to determine the 'specific traits' relevant for a particular population. An important aspect which appears to receive little, if any, attention is the possibility of rare events. An individual may be assigned toPopulation A because they have a trait commonly seen in that population but not seen at all in Population B. However, failure to observe this trait in Population B may simply reflect the sampling protocol used. The correct ancestry for this individual could be Population B.

In other research areas such as ecology, epidemiology and veterinary science considerable research on detecting rare events has been performed. A review found most methods to be unsuitable for forensic application, but methodology provided by Green and Young1 is appropriate. If a rare event occurs at a rate of 1 in 200 (as seen in Poulsen2), then to be sure of a 95% power of detection, a minimum sample size of 600 is required. In forensic science there are studies using sample data with as few as fifty samples; these should be treated with caution. Other research in a range of areas including the physical, natural and social sciences stresses that increasing the number of variables (in forensic science the number of markers) will also increase the sample size needed for valid statistical analysis.

Many published studies in forensic science are made with samples which have low power to detect the rare events which could be present. The detection of rare events can lead to dire implications if ignored. A greater awareness of methods to account for rare events is required.

36. Genetic Polymorphisms of 26 Y-STR loci in Han Population of China

Author: **Lei Shang, Xiaoting M, Jian Z, Wenhua M, Xin M, Zhengliang Y, Guangshu D, Xi W, Wanshui L**

Beijing Engineering Research Center of Crime Scene Evidence Examination & Key Laboratory of Forensic Genetics, Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China

This study was to determine the frequencies and parameters of 26 Y chromosomal short tandem repeat loci of Han population from China. Private alleles of each area were also studied to better explore the application value of DNATyperY26 Kit. Blood samples of 802 non-related male were collected and genotyped through PCR amplification and capillary electrophoresis based on DNATyperY26 Kit. Genotype data was then analyzed for allele frequencies and parameters. We found 180 alleles with their frequencies ranging from 0.001~0.854. Genetic diversity (GD) value of Y-STR loci distributed from 0.261~0.965 with most of them higher than 0.5. Discrimination capacity (DC, 0.969) of the Kit was very high since 777 haplotypes were detected in 802 samples. The study also showed high haplotype diversity (HD, 0.9999). Moreover, we observed 5, 5, 10, 7 private alleles respectively for each area. In conclusion, DNATyperY26 Kit can be widely applied to the research of forensic science and population genetics. Analysis of private alleles may benefit in source region inference, which still needs further exploration.

37. Genetic characterization of Y and mtDNA lineages in the admixed population of Marajó Island, Northern of Brazil

Author: **Loiola S[1], Simão F[1], Pinheiro LML[2], Leonor Gusmão[1,3], Carvalho EF[1]**

[1]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil
[2]Department of Natural Sciences, Federal University of Pará (UFPA), Brazil
[3]Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal

The Island of Marajó is located in the north of Pará state, about 90 km from the capital, Belem. It is the largest fluvial-maritime archipelago in the world, bathed by the Atlantic Ocean and the Amazon and Tocantins rivers. At the time of Portuguese colonization, it is believed that 30 different indigenous nations inhabited the Island. Currently it is composed by an admixed population, with Native American, European and African genetic contributions. To investigate the genetic background of the Marajó population, a sample was selected from unrelated individuals with local ancestry for at least three generations. A total of 72 males were genotyped for the 27 Y chromosome specific STR loci included in the YFiler Plus kit. Additionally, 55 samples were sequenced for the mtDNA control region. For the 27 Yfiler Plus markers, a haplotype diversity index of 0.9969 ± 0.0032 was observed, with 58 haplotypes being unique, four being shared by two individuals and two shared by three individuals. This value can be considered low when compared with the one obtained for the Brazilian population from Rio de Janeiro (0.99997). Since the Y-STR data available for other populations from the North region of Brazil do not include all markers from the YFiler Plus kit, a new analysis was performed for a subset of 18 Y-STRs. The sample from Marajó showed the lowest value of haplotype diversity, in comparison to Belém, Boa Vista, Macapá, Manaus, Palmas, Porto Velho, Rio Branco, Santarém, or even with a sample of Native Americans from São Gabriel da Cachoeira (Amazonia). Genetic distances were low between Marajó and all otheradmixed populations from the North, but high with São Gabriel da Cachoeira, indicating a predominating European male lineages contribution. In the analysis of mtDNA total control region, a high haplotype diversity was observed (0.9833+/-0.0097). In the comparison with data available for HVI, significant differences were found between Marajó and samples from the South, Southeast, Northeast and North regions of Brazil. The mtDNA haplogroup composition revealed a high contribution of Native American maternal lineages (approximately 66% of the samples belong to haplogroups A, B, C and D). The African haplogroups represented 29% of the samples and only 5% where from European maternal origin.

38. Analysis of 23 Y-STRs in a population sample from eastern Paraguay

Author: **Ribeiro J[1], Vullo C[2], Simão F[1], Almeida APF[1], Romero M[2], Quiroz A[3], Machado P[4], Velázquez V[5], Carvalho EF[1], Leonor Gusmão[1]**

[1]Laboratório de Diagnóstico por DNA (LDD), Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil
[2]DNA Forensic Laboratory, Argentinean Forensic Anthropology team (EAAF), Córdoba, Argentina
[3]Instituto de Previsión Social, Asunción, Paraguay
[4]Legislación Farmacéutica, Universidad Católica, Hernandarias, Paraguay
[5]Hospital Regional, Encarnación, Paraguay

As for most South American populations, historical data point to a high ethnic diversity in Paraguay, resulting from the encounter between Native groups and European colonizers, and the arrival of African slaves. Recent migrations from Europe and other South American countries are also expected to have shaped the current genetic structure of Paraguay. Y chromosome markers are widely used in population genetics, to infer paternal ancestry and male mediated movements between populations. Due to high mutation rates, Y-STRs are suitable to trace recent founder events and disclose genetic differences between closely related populations. The aim of this study was to characterize 23 Y-STR markers in a population sample from Paraguay, since no data is yet available for Y chromosome specific

markers in this population. Therefore, 537 unrelated males, living in seven departments from the eastern region of Paraguay, were genotyped using the PowerPlex® Y23 system, following manufacturer's instructions (Promega). A high haplotype diversity was found (0.9993), with 471 different haplotypes being present in the whole sample. When comparing samples from different departments, no statistically significant differences were found (FST ≤ 0.0043; non-differentiation p-values for 50,000 permutations ≥ 0.0502), pointing to a genetic homogeneity of the paternal ancestry of the populations in the eastern region of Paraguay. Genetic distances (FST) were also calculated between Paraguay and other Admixed populations from South America, as well as with European, African and Native American populations. This analysis revealed no significant differences with Argentina, Rio de Janeiro (Brazil), São Paulo (Brazil) and Costa Rica, all very close to the Iberian populations. Significant differences were found between Paraguay and Bolivia, Ecuador, Peru and Panama, most probably due to the higher Native American paternal ancestry of these populations, which were closer to the Native populations.

39. Y-STRs and autosomal markers in paternity testing. Two curious casework examples

Author: **M Lourdes Pontes, Lima G, Cerqueira J, Porto MJ**

Serviço de Genética e Biologia Forenses, INMLCF, IP – North Delegation

Haploid markers, namely Y-STRs, X-STRs or mtDNA, have been very useful for antropological and ancestry investigations, but are also useful in Forensic Genetics. In Forensic genetics, besides criminal investigation cases, lineage markers are also used in some complex paternity cases.

Here we report two casework examples were we had some difficulties in interpreting the results after the use of Y-markers. In both cases no incompatibilities were observed in the autosomal markers included in the last generation comercial kits such as PowerPlex Fusion, but, incompatibilities were observed in the rapidly mutating Y-STRs included in the same kits.

Cases like the described in this work, must be treated with caution, namely in the form of communicate the results to the Courts.

40. A literature review on the prevalence of pigmentation traits

Author: **Maria-Alexandra Katsara, Nothnagel M**

Department of Statistical Genetics and Bioinformatics, Cologne Center for Genomics, University of Cologne, Cologne, Germany

The distribution and prediction of externally visible characteristics has been of interest for decades, with a strong focus on eye and hair color pigmentation. Numerous recent studies have revealed genetic and environmental factors that contribute to these two traits and that have been used to predict them. However, use of prior knowledge on the geographic distribution of those traits, which may improve prediction accuracy, has so far been barely used. To this end, we conducted a literature review on the geographic prevalence distribution for eye and hair color categories, thereby assessing reasonable prior probabilities for their prediction. Somewhat surprisingly, we found the amount of reliable scientific data available to be quite limited, despite an ongoing interest in this topic for over a century. We will report on the set of those countries for which we were able to compile trustworthy outcomes concerning eye and hair color distribution. We will also present results from the statistical analysis of these data, including the application of interpolation techniques. Our results will hopefully facilitate the improvement of already existing and of novel prediction methods for pigmentation traits.

41. The dna.bases Consortium STRidER, the STRs for Identity ENFSI Reference database and quality control platform

Author: **Martin Bodner[1], Parson W[1,2], The dna.bases Consortium**

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

The statistical evaluation of autosomal Short Tandem Repeat (STR) genotypes is based on allele frequencies compiled in databases of sufficient size that are established in the course of population studies. Haploid marker databases (YHRD and EMPOP) have clearly demonstrated that centralized quality control and data curation is essential to maintain the high quality required in forensic genetics. We present STRidER, the STRs for Identity ENFSI Reference Database (https://strider.online), a publicly available online allele frequency database, quality control and software platform for autosomal STRs. STRidER expands on the ENFSI DNA WG STRbASE and has been developed in agreement with the DNA Commission of the ISFG [1]. STRidER constitutes a significant improvement and innovation and serves the scientific community beyond forensics in multiple inter-related ways:

- the high-quality allele frequency database enables scientifically reliable autosomal STR genotype probability estimates
- allele frequency tables of autosomal STR loci from diverse populations can be downloaded for use in third-party software
- centralized autosomal STR data quality control is provided prior to publication
- accepted datasets become rapidly available online and receive a unique and traceable accession number
- individual STR genotypes are not accessible to comply with privacy regulations
- STRidER serves as a platform for the development of novel tools for STR data
- the curated ISFG STR Sequence Structure Guide can be downloaded in its latest version

More loci and populations will continuously be included based on the availability of high-quality data. Additional features will be offered to allow storage and handling of autosomal STR sequence data.

42. Analysis of Y-chromosomal Polymorphism and the Clan Structure in Kazakh Populations

Author: **Maxat Zhabagin[1,2], Balanovsky O[3,4]**

[1]National Center for Biotechnology, Astana, Republic of Kazakhstan
[2]National Laboratory Astana, Nazarbayev University, Astana, Republic of Kazakhstan
[3]Vavilov Institute of General Genetics Russian Academy of Sciences, Moscow, Russia
[4]Research Centre for Medical Genetics, Moscow, Russia

Kazakhs are patrilocal and patrilineal population: Kazakh society is organized into patrilocal clans and most individuals trace their paternal ancestry up to seven generations. Descent groups claim to have distinctive common ancestors and detailed oral tradition describes the paternal ancestry from the socio-genealogical perspective. The Y-chromosome is a powerful tool to analyze the paternal genealogical lineage from the biological perspective. Parallel studying the Y-chromosomes and clan affiliation enables tracing the association between social and biological relationships and inferring clan ancestry from DNA.

To investigate this association, we have studied 14 clans from 19 districts. About 2000 samples of the clan members were genotyped by 44 SNP and 17 STR markers of the Y-chromosome. In addition, we performed the entire sequencing of the MSY region for nine lineages of Argyn clan and five lineages of Kypchak. The observation of genetic analysis (AMOVA, Mantel test) proved that clan structure shapes gene pool even more evidently than

geographical distances. Phylogenetic analysis demonstrated that the clan members share generally a recent common ancestor. In conclusion, this work has allowed us to make some historical investigations and biogeographical analysis.

The work was supported by the Russian Foundation for Basic Research Nº 17-304-50005 and the Ministry of Education and Science of the Republic of Kazakhstan (Grant Nº AP05134955).

43. A familial search – need for haploid markers confirmation

Author: **Milica Keckarević Marković, Mihajlović M, Tanasić V, Kecmanović M, Keckarević D**

Faculty of Biology, University of Belgrade, Belgrade, Serbia

Since relatives share alleles identical by descent, a familial search has been widely and routinely used in missing persons and disaster victims identifications for many years. By autosomal STR markers familial search one identifies partial matches, and subsequently, the likelihood ratio of possible relationship versus coincidental match is calculated in order to differentiate possible relative candidates from false positives. Considering that male relatives share Y chromosome identical by descent, identical or similar Y chromosome haplotypes strongly points out male-to-male relationship. Mitochondrial DNA, due to maternal transmission, indicates a more complex relationship between men. The familial search of forensic DNA databases is, also, a potent tool for resolving criminal cases and could compensate for missing investigative leads. On the other hand, inadequate and incompetent use of familial search could lead to wrongful assumptions and misconducts in criminal investigations. Here we present a case of familial search match indicating paternity, resolved by haploid markers. A DNA profile from a rope used in a murder committed 24 years ago, unresolved to date, analyzed by AmpFlSTR® Identifiler® kit, was found to be of possible first – degree or close relative of a person involved in another murder, 22 years later. Although crime scenes were geographically distant, there was an intriguing possibility of father-son relation. The combined paternity index (CPI) was 218, not enough for paternity confirmation, but also not unexpected, having in mind that Identifier set of loci contains few loci with a relatively low power of exclusion. The analysis was extended to ESS set of loci and paternity was practically excluded based on D1S1656 and D12S391. Subsequent Y – haplotype analysis of both profiles showed differences that clearly ruled out any possible male – to – male relationship. Further analysis of mtDNA control regions revealed no maternal line relationship, as well.

Analysis and interpretation of forensic data must be regulated and monitored by forensic community, considering that many jurisdictions have no strict regulations that could deal with challenging situations, and everything depends on the competence of forensic data analyst.

44. Diversity of Y-STR haplotypes in Serbs from old Hercegovina

Author: **Pašaljić Đ, Milica Keckarević Marković, Kecmanović M, Keckarević D**

Faculty of Biology, University of Belgrade, Belgrade, Serbia

Since the Y chromosome is passed from father to son mostly unchanged, Y-chromosomal short tandem repeat (Y-STR) haplotyping is widely used in forensic DNA analysis to trace a paternal lineage of a sample donor or in a complex paternal/kinship cases. On the other hand, for reconstructing the Y chromosome phylogeny and identification of the divergence time of different male lineages, Y-chromosomal single-nucleotide polymorphisms (SNPs) are more suitable. In the present study, we analyzed diversity of Y-chromosomal haplotypes and haplogroups from unrelated males from Old Hercegovina (the area includes part of present-day Montenegro, the area around Prijepolje which is currently in Serbia, and the

areas which are today included in the Bosnia and Herzegovina), since the vast majority of today's Serbs originate from that Dinaric settlement population.

DNA was extracted from buccal swabs of 266 unrelated males. Amplification was performed using the PowerPlex Y23 kit and after separation of amplified fragments on the ABI Prism 3130 Genetic Analyzer, samples were analyzed with GeneMapperID-X software.

Haplotype frequencies were calculated by counting method. Gene and haplotype diversities were calculated using the formula HD=n/(n – 1)x(1-Σpi2). Discrimination capacity was determined by dividing the number of observed haplotypes with the number of samples.In silico haplogroup assignments were made using NevGen predictor.

In total, of the 266 analyzed samples, 241 different haplotypes were observed, and 226 of them were singletons. Discrimination capacity was 0.906 and matching probability 0.0048 (1 in 208). The gene diversity values ranged from 0.3138 at DYS393 to 0.884 at DYS481. The haplotype diversity was0.998922. The most prevalent haplogroup was I2a (48%), followed by R1a (22%), E (10%) and I1 (5%).

The present study has shown that haplogroup I2a was the most prevalent, which is also in line with the previously published data on I2a haplogroup found to be overrepresented in Balkan region. Our data should provide a much more precise insight into the processes and flows of historical migrations of the Dinaric settlement population, which is the Herzegovinian region, and in the continuity/discontinuity of the medieval and contemporary Herzegovinian population.

45. Mitochondrial Haplogroup and Haplotype Status in Tumor and Normal tissues of Korean Lung Cancer patients

Author: **Moon-Young Kim[1], Lee JH[1], Deok Lee S[1,2]**

[1]Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Korea;
[2]Institute of Forensic Science, Seoul National University College of Medicine, Seoul, Korea

In some special situations, previously resected tumor tissue could be the only sample for human identification or paternity test, because of death of the subject or rejection to sampling. With de novo somatic mutation and genetic heterogeneity, representative characteristics of malignant tumor, genomic instability of mitochondrial DNA itself might have influence to the determination of haplogroup and individual haplotype.

55 formalin-fixed samples from 13 Korean patients who were treated for lung cancer in Seoul National University Hospital from 2013 to 2017 were collected. They included 14 primary tumors with 14 normal lung tissues and 24 metastatic tumors with 5 surrounding normal tissues other than lung. Precision ID mtDNA Whole Genome Panel (Thermo Fisher scientific, Waltham, MA, USA) was applied for massively parallel sequencing. Initial list of variants was acquired by Torrent Variant Caller ver 5.0 with minimum minor allele frequency (MAF) of 10 %. Haplogroups were determined by MitoTool and PhyloTree build 17.

All samples from each patient were classified to identical haplogroup, regardless of tissue type or organ. 6 of 13 patients were assigned to macrohaplogroup D and 3 patients to macrohaplogroup B, which were frequent in Korean population. Total of 311 private mutations beside haplogroup-defining were observed in 94 positions. Private mutations were more frequent in tumor, with 6.74 mutations per samples in average (4.56 in normal). 22 % of mutation-bearing positions were in D-loop. 13.8 % of positions showed mutation in 2 or more patients and were all located in D-loop. Meanwhile, 46 % were found only in

tumor samples for once each, and mostly located in coding region of various genes. Private mutations were appeared as homoplasmy or heteroplasmy, with MAF from 10 % to 50 %.

The haplogroups were classified consistently in tumor and normal samples. However, numerous private mutations were observed with variable number, location and heteroplasmy level according to patients and tissue types. These variations should be considered in identifying a personal genetic profile and comparing with other profiles.

46. DNA methylation variation among European and Chinese population – a new approach to human population differentiation.

Author: **Patrycja Daca-Roszak[1], Pfifer A[2], Jarza B[2], Witt M[1], Ziętkiewicz E[1]**

[1]Institute of Human Genetics, Polish Academy of Sciences, Poznań, Poland
[2]Department of Nuclear Medicine and Endocrine Oncology, Maria Skłodowska-Curie Institute-Oncology Center, Gliwice Branch, Gliwice, Poland

The aim of our study was to identify CpG positions representing DNA methylation differences between European and Chinese populations.

B-cell lines (Coriell Repository) from European (n=18) and Chinese (n=18) males were bisulfate converted and analyzed with Illumina Infinium Human Methylation 450 Bead-Chip Array. After quality control (Genome Studio Methylation Module), data was preprocessed using the quantile normalization method SWAN (minfi Bioconductor package). To identify differently methylated CpGs, the Student's t-test corrected for multiple testing was performed and absolute values of Mav_diff (the differences between average methylation levels in both populations) were calculated for each of the interrogated CpG positions.

96 CpG loci representing the highest inter-population differences were identified ($q<0.05$; $Mdiff>1$). After in silico and sequencing validation, 82 CpG positions were rejected from further studies due to the limitations imposed by the used Illumina probe (e.g. hybridization to multiple map addresses, targeting genomic sequences containing SNPs) (see Daca-Roszak et al. 2015). This part of the study also indicated the importance of correcting Illumina microarray results with regard to a confounding impact of CpG-located SNPs.

A subset of 10 CpG loci ($q<0.05$; $Mdiff>1$) representing inter-population methylation differences was selected for validation using pyrosequencing technology. In the first step, all the preselected CpG loci are tested in a number of independent B-cell lines. Subsequently, a set of CpG that pass the validation step on B-cell lines will be examined using DNA isolated from the whole blood samples.

Preliminary results of the validation (in progress) in B-cell lines confirm differences in the level of DNA methylation between the studied populations. So far, a significantly higher level of methylation in Chinese compared to European cell lines ($p<0.00001$) was observed in one of the tested loci (cg23669876).

This research was financed by the NCN National Center of Science (Nr 370227).

47. GHEP-ISFG Collaborative Exercise on Mixture Profiles of Y – Chromosome STRs (GHEPMIX05_HAP2): Results and evaluation (mock case #2).

Author: **Pedro A Barrio[1,2], Crespillo M[2], Luque JA[2] on behalf of the GHEPMIX working commission**

[1]INTCF - National Institute of Toxicology and Forensic Science, Department of Madrid, Spain
[2]INTCF - National Institute of Toxicology and Forensic Science, Department of Barcelona, Spain

One of the main goals of the Spanish and Portuguese-Speaking Group of the International Society for Forensic Genetics (GHEP-ISFG) is to promote and contribute to the development and dissemination of scientific knowledge in the field of forensic genetics. In this way, GHEP-

ISFG holds different working commissions that are set up to develop activities in scientific aspects. One of them, the Mixture Commission of GHEP-ISFG, has organized annually, since 2009, a collaborative exercise on analysis and interpretation of autosomal STRs mixture profiles. Until now, six exercises have been organized.

In the 2015 edition (GHEP-MIX05), with 24 participant laboratories from 7 different countries, one of the exercise aims was to give a general vision about mixed profiles of Y-chromosome STRs analysis. Through the proposal of mock cases, edition and statistical treatment were assessed. With this poster, we will show the statistical evaluation results for one of the mock cases (mock case #2).

Regarding laboratories characterization, it can be highlighted that all participants used haploid markers results (Y-chromosome and/or mtDNA) in casework. When they asked if they emitted results for haploid markers in daily casework, only one lab answered negatively. However, in the case of mixed profiles of haploid markers of Y-chromosome, five labs did not emit results for this kind of profiles. In relation to statistical evaluation, 60% of the participants did not perform the statistical evaluation in the case of mixtures of Y-chromosome profiles. Notwithstanding the above, against the mock cases proposed in this exercise, up to 68% of the participants made a statistical assessment through LR calculation, although setting different hypotheses, and obtaining LR values with differences up to 2 orders of magnitude (LR = 6.88E+02 – 2.51E+05).

Taking these data into account, the main conclusions obtained from this exercise module may be summarized in the need for continuing education in the analysis and interpretation of mixture DNA profiles, in particular, in the case of Y-chromosome mixture profiles.

48. Evaluation of homoplasy in Y23 loci: a pilot study on samples with known 400 –yearslong genealogical history in eastern Czechia

Author: **Radim J. Vašut[1,2], Stenzl V[3], Patera J[3]**

[1]*Palacký University in Olomouc, Faculty of Science, Department of Botany, Olomouc, Czech Republic*
[2]*Palacký University in Olomouc, Faculty of Education, Department of Biology, Olomouc, Czech Republic*
[3]*Institute of Criminalistics Prague (ICP), Prague, Czech Republic*

Surname genealogies co-evolve with Y-STR profile in European societies and generally can infer from each other. Close similarities in matrices of 23 Y-STR loci are considered to be related to genetic relationship between persons. With increasing numbers of tested persons in Moravia (1000+ individuals), we are detecting increasing number of similar Y23 profiles in lineages with known genealogies (usually pedigrees of 300-400 yr old). Such similarities that originate in the same geographic region might reflect past genetic relationship (before year 1600). In order to evaluate events of homoplasy in such similarities in Y23 profiles, we sequenced (Illumina MiSeq) profiles, which are known to be genetically related (in past 400 years) as well as similar profiles that are known to be not related in past 400 years. The occurrence, rate and chance of misinterpretation due to homoplasy is discussed on example of genealogies of the east-Moravian clans.

49. DNA research based on haploid markers in different tissues after allo –HSCT from unrelated donor

Author: **Renata Jacewicz[1], Lewandowski K[2], Rupa-Matysek J[2], Jędrzejczyk M[1], Lorent D[1], Berent J[1]**

[1]Medical and Forensic Genetics Laboratory, Department of Forensic Medicine, Medical University of Lodz, Poland
[2]Department of Hematology and Bone Marrow Transplantation, Poznan University of Medical Sciences, Poznan, Poland

The analysis of haploid markers in biological materials derived from person after allogeneic hematopoietic stem cel transplantation (allo-HSCT) from unrelated donor was presented.

The possible implications in the case of investigation such a materials in judicial aspects were considered as well as preventive procedures were proposed.

50. Y-chromosomal SNP analysis in three ethnic groups in Republic of Macedonia

Author: **Renata Jankova-Ajanovska, Risteski S, Jakovski Z, Jovanovic H, Pavlovski G, Cakar L, Rakipi A, Stankov A**

Institute of Forensic medicine, criminalistic and medical deontology, Medical Faculty, University of St."Cyril and Metodius", Skopje, R.Macedonia

The analysis of Y chromosomal SNPs (Y-SNPs) except of its tremendous value in evolutionary research, can be useful in forensic application, especially in cases of body identification or criminalistic identification when there are no reference samples for comparison. In such cases any further information including Y-SNP haplogroup may be of great help. Knowing the structure of most frequent haplogroups in each territory could be a powerful tool in forensic analysis. The aim of this study was to determine the frequency of Y-SNP markers and haplogroups in three biggest ethnic groups in Republic of Macedonia. In this study were analysed 314 unrelated males, 103 Macedonians, 109 Turks and 102 Albanians. SNPs analysis were perforemed in 6 multiplex PCR reactions with 38 different markers. In Macedonian ethnic group the most frequent haplogroups were: I2a-P37.2 (28.15%), E1b1b1-M35 (19,41%) and R1a-M198 (19,41%), in Albanian ethnic group were E1b1b1-M35 (35.29%), R1b1b2-M269 (19.60%) and J2b-M102 (17.64%). In Turks the most frequent haplogroups were J2a-M410 (34.86%), R1b1b2-M269 (20.18%), I2a-P37.2 (19.26%) and E1b1b1-M35 (11.92%).

51. Amelogenin-negative cases and a case with Y-chromosome interstitial deletion involvingsome Y-STRs and Y indel marker identified in Belarusian males during forensic DNA analysis

Author: **Shyla A, Sergey R Borovko, Korban VV**

State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus

The use of the gender marker amelogenin (AMEL), incorporated in multiple commercial multiplex DNA typing kits, is a common practice in forensic DNA analysis. Although amelogenin is believed to be an effective marker for gender determination of biological samples there are reports of AMELY or AMELX (homologues of amelogenin) dropouts detected worldwide.

A total of 17 cases with amelogenin abnormalities have been detected in Belarusian males during 16-year expert practice in forensic DNA analysis: 13 AMELY null and 4 AMELX null cases. Some of them we have been described earlier [1]. Among these 17 cases 3 cases were obtained from paternity testing group, 14 cases from crime casework group. Genetic mechanisms underlying AMELX or AMELY dropouts identified in these cases fall into four categories: 1) deletion involving AMELY and DYS458 loci (n=4), 2) loss of the long arm of the Y chromosome with partial X-Y translocation in XX males (n=3), 3) loss of most of the Y

chromosome in XX males (n=3), 4) mutation in the primer binding region of AMELX (n=4) or AMELY (n=3) loci.

Y-STR dropout has been described in various populations. In current study we report one case discovered during routine paternity testing in which some Y-STRs and one insertion/deletion polymorphic marker on the Y chromosome (Y indel) dropout was observed. Both AMELX/AMELY alleles were detected in the male of Belarusian origin. However, 3 Y-STRs, DYS391, DYS570, DYS576, and Y indel were undetectable in this sample. We assume that null alleles of DYS391, DYS570, DYS576, and Y indel are caused by interstitial deletion of Y chromosome.

52. Characterization of a 41-plex PCR amplification Assay for Male-Specific Databasing Applications

Author: **Siddhita Gopinath, Ge J, Heid C, Carbonaro A, Short M, Mulero J**

Thermo Fisher Scientific Inc.

DNA databases are indispensable tools in forensics to help solve crimes by matching autosomal STR profiles obtained from crime scene samples with known crime offenders. In more recent years the forensic community has been debating the inclusion of Y-STR markers to existing databases to help determine or exclude relationships, identify missing persons, infer ancestry and interpret mixture.

We developed a 41-plex that simultaneously amplifies the 27 Y-STR markers included in the Applied Biosystems Yfiler™ Plus PCR Amplification Kit plus 11 new Y-STRs (DYS549, DYS645, DYS557, DYS593, DYS522, DYS444, DYS596, DYS643, DYS447 and DYS527a/b) and 3 Y-indels, which together can provide extremely high discriminating power. This multiplex is designed to process single-source reference samples using direct PCR amplification from blood samples on paper substrates without the need for sample purification. This 41-plex was built in a 6-dye multiplex format with PCR products ranging from 68-570 base pairs and it is compatible with detections on the 3130xl, 3500xL and 3730xl instrument platforms. Particularly, the Y-indels can be used to quickly exclude male lineages, as the mutation rates of the Y-indels are significantly lower as compared to STR's (i.e., ~10-9 vs 10-3 per locus per generation).

This study shows the feasibility of amplifying 41-markers simultaneously and describes the optimization of the multiplex assay to deliver high first pass success rate and assay performance when amplifying blood samples on paper substrates. In addition, the haplotype diversity and discriminatory capacity calculations with the expanded multiplex will be presented.

53. Study of Y-chromosome STR markers in United Arab Emirates population

Author: **Tareq Ahmed, Alghafri R**

General Department of Forensic Sciences and Criminology, Dubai

The recently introduced 6-dye Yfiler Plus multiplex which includes 27 Y-STR loci (DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS456, DYS390, DYS438, DYS392, DYS518, DYS570, DYS437, DYS385 a/b, DYS449, DYS393, DYS439, DYS481, DYF387S1a/b and DYS533) has been used to study 343 UAE Arab male individuals using Yfiler Plus® amplification kit. This set includes seven rapidly

mutating loci (RM Y-STRs). These RM Y-STRs are useful for discriminating between closely related and unrelated males..

According to measures of genetic diversity the highest diversity were observed at loci DYS385=(0.94984), DYF387S1=(0.930523) and DYS449=(0.895402). Therefore, these loci should be considered the most diverse and polymorphic for forensic testing which can be used to distinguish between male relatives. 313 haplotypes were observed in UAE Arab male population and 15 haplotypes were shared between two individuals. Discrimination capacity for 27 loci among the UAE Arab male population was determined to be 95.43% whereas haplotype diversity was found to be 0.99973. AMOVA results showed that UAE Arab male population was placed at far genetic distance from European populations such as Denmark, Italy, Spain and United States. While it shows closer genetic distance to the regional populations from Iran, Iraq, Egypt, Yemen and Kuwait.

54. Massively-parallel sequencing of uniparentally – inherited forensic DNA markers

Author: **Tunde I. Huszar, Jobling MA, Wetton JH**

Department of Genetics & Genome Biology, University of Leicester, University Road, Leicester LE1 7RH UK

We used massively parallel sequencing (MPS) to analyse uniparental markers with the prototype kit PowerSeq™ Auto/Mito/Y System by Promega in a set of 100 samples carrying diverse Y chromosomes whose phylogenetic relationships are known from previous megabase-scale resequencing. We amplified the Y-STR markers of the PowerPlex® Y23 System, plus the mitochondrial control region using ten overlapping amplicons, within a single multiplex reaction.

The sequence diversity of Y-STRs and their flanking regions was considered within the phylogenetic framework. This allowed the observation of several specific haplogroup associations for SNPs and indels, reflecting the low mutation rates of such variant types, while repeat pattern variants were observed to be less phylogenetically coherent and showing more recurrence, reflecting their relatively high mutation rates. The diversity-based approach allowed the scoring of rare variants, which might not be observed over such a wide range in population-focussed studies; these variants suggested the re-evaluation of the reported region of the repeat sequences for DYS385a,b, DYS481 and DYS390.

This set-up also allowed us to amplify control regions of a diverse set of mitochondrial genomes in a single reaction together with the Y-STRs, and to compare the mitochondrial variants detected with previously collected data.

This study uses MPS analysis to reveal considerable additional diversity at the Y-STRs, demonstrates high concordance with CE data, facilitates nomenclature standardisation and places Y-STR sequence variants in their phylogenetic context. The mitochondrial variants observed highlight the considerations and applicability of such a design in a multiplex approach and the feasibility of combining different haploid markers when analysing samples by MPS.

55. Evaluation of the NovaQUANT™ Human Mitochondrial to Nuclear DNA Ratio Kit in a forensic context

Author: **Vania Pereira, Andersen JD, Børsting C**

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

Analysis of mitochondrial DNA (mtDNA) is an useful tool in forensic genetics since mtDNA is less prone to degradation and present in high numbers in many cell types. Although mtDNA has some limitations when discriminating between individuals, it can provide an investigative lead in cases where the nuclear DNA (nDNA) is degraded or present in low amount. In these situations, it would be beneficial to determine the number of mtDNA copies and evaluate whether the samples may be used for mtDNA genotyping.

The NovaQUANT™ Human Mitochondrial to Nuclear DNA Ratio Kit (Novagen®, Merck) is a qPCR quantitative assay that measures the copy number ratio of mtDNA to nDNA. The kit amplifies four PCR targets (two mtDNA and two nDNA) in singleplex reactions. The ΔCt values from the two mtDNA/nDNA pairs are calculated and used to estimate the ratio between mtDNA and nDNA. The kit was developed for clinical genetic applications, where the amount of DNA is not a limiting factor (the recommended input DNA of the kit is 0.2-2ng per qPCR reaction). Here, we evaluated the NovaQUANT™ Human Mitochondrial to Nuclear DNA Ratio Kit in a forensic context. We tested the performance of the kit by using different nDNA inputs (6.25pg-1ng) and trace samples from several sources.

56. Evaluation of the Precision ID mtDNA Whole Genome Panel for massively parallel sequencing of mitochondrial genomes

Author: **Vania Pereira, Longobardi A, Børsting C**

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

Massively parallel sequencing (MPS) offers a fast and cost effective method for sequencing of the whole mtDNA genome. Several strategies have been described for the amplification and sequencing of the complete mtDNA molecule. Recently, ThermoFisher Scientific released the Precision ID mtDNA Whole Genome Panel that amplifies the entire mtDNA genome in two multiplex PCRs with 81 primer sets.

In this work, the performance of the Precision ID mtDNA Whole Genome Panel was evaluated by testing different amplification methods (2-in-1 or conservative), number of PCR cycles (21, 23, 25), and reagent volume used in the PCR and library building (recommended volume or half-volume). To address the performance of the panel in a forensic casework setting, a dilution series, controlled mixtures and trace samples were also sequenced.

The normalised read depths of the individual fragments were consistent regardless of input, amplification method, or reaction volumes used. The conservative method showed slightly higher library concentrations than the 2-in-1 method. PCR and library preparation with half-volume of the reagents can be applied for standard reference samples without compromising the quality of the results. The panel was highly sensitive. Complete mtDNA genome sequences were obtained from as little as 6.25pg genomic DNA. Analyses of controlled mixtures showed that the sequencing output matched the mixture ratios indicating that the sequencing results were a loyal representation of the input DNA. Results were more challenging in trace samples, and guidelines are needed for the interpretation and reporting of the results to the investigative authorities.

57. Massively parallel sequencing of whole mitochondrial genomes: Croatian population study

Author: **Korolija M, Viktorija Sukser, Rožić S, Račić I, Lipovac K, Barbarić L**

Forensic Science Centre "Ivan Vučetić", Zagreb, Croatia

Mitochondrial genomes are particularly suitable for forensic analysis, due to their stable circular structure, small size (16569 bp), and large copy number per cell. During evolution, mtDNA accumulated variations throughout the sequence, thereby differentiating thousands of mitochondrial haplogroups known today. From these variations it is possible to decipher population's matrilineal origins and historical migrations. In forensics, in order to assign a certain weight of evidence to mtDNA profile, a population study is the necessary prerequisite for establishing haplogroup frequencies. Therefore, we aimed at sequencing whole mt genomes in a representative sample of Croatian population, comprising 299 volunteers. DNA was extracted from buccal swabs, purified and quantified. Long-range PCR approach was adopted for amplification of whole mtDNA in two fragments (9,1 kb and 11,2 kb). Libraries were prepared using Nextera XT Library Prep Kit (Illumina). Normalized and pooled libraries were sequenced on MiSeq instrument (Illumina). Sequencing and quality metrics strongly correlated with manufecturer's specifications, sometimes even exceeding the specified values. Our data were of sufficient quality in order to produce confident haplogroup determination. The composition of Croatian population sample displayed the prevalence of mt haplogroups from branch H (36%). Second most common haplogroups came from U and K branches (29%, combined), and J and T branches (16%, combined). Furthermore, almost 28% of samples contained point heteroplasmies with minor allele frequencies ≥10%, which is comparable with the level of heteroplasmy detection by Sanger sequencing. The most common heteroplasmy was 16093Y (in 10 out of 83 heteroplasmic samples). Detection of heteroplasmies is also a subject of interest in forensics, which could strengthen the power of discrimination in mtDNA analysis. Overall, frequencies of mt haplogroups detected in our samples are concordant with the distribution of haplogroups in Europe. It is our intention that this population study forms the basis of a database to be used for purposes of forensic identification, when such need arises in casework. We also propose that mt haplotypes detected in our population sample contribute to further development and branching of the global phylogenetic mtDNA tree.

58. Identification of human remains from rebels' graves found on the Gediminas Hill in Vilnius city of Lithuania using Y chromosome STR loci

Author: **Vilma Ivanova[1], Bunokiene D[1], Jankauskas R[2]**

[1]The State Forensic Medicine Service, Vilnius, Lithuania
[2]Department of Anatomy, Histology and Anthropology, Faculty of Medicine, Vilnius University, Vilnius, Lithuania

The January Uprising was an uprising which took place in the former Polish-Lithuanian Commonwealth (present-day Poland, Lithuania, Belarus, Latvia, parts of Ukraine, and western Russia) against the Russian Empire. It started on 22 January 1863 and lasted until the last rebels were captured in 1864. For almost a century and a half, Lithuania has been trying to discover the uprising victims' graves, however, unsuccessfully until the beginning of 2017, when the emergency maintenance works started on the top of the Gediminas Hill, following a dangerous landslide, gave a new push to the investigation. Some of the human bodies were covered in lime, buried without coffins, with their hands tied. In the graves, archaeologists found pieces of clothing and buttons from the shirt and trousers. In one grave a silver medallion was found, containing an image of the Blessed Virgin Mary, of the Chapelof the Gates of Dawn, a famous shrine in the old town of Vilnius, on one side and

on the other side – a picture of the Snipiskes Chapel. Also one male's right hand contained a golden ring with an inscription inside: „Zygmont Apolonija 11 Sierpnia/30 Lipca 1862". This evidence enabled to identify that the remains belong to Zigmantas Sierakauskas, the leader of the uprising.

During the excavation, bones of 21 individuals were found on the hill. The skeletal remains of 17 individuals were excavated. Analysis of 17 femurs was conducted. The bones were cleaned, removing the contamination of the surface, and the bones were grounded into powder. 0,7g of the powder of each sample were subjected to DNA extraction. For DNA purification silica-membrane-based spin columns were used. The nuclear DNA of the samples was quantified using the real-time PCR (0.02 to 1.46 ng DNA/μ). Y-chromosome haplotypes were obtained from 17 bones using Yfiler® Plus PCR Amplification Kit (Thermo Fisher Scientific, USA). Two profiles coincided, proving that these individuals belonged to the same male lineage; this find coincided with anthropological data and enabled to confirm identity of two brothers executed at the same day.

59. The haplotype differences of Mitochondrial DNA hypervariable region 1 in Chinese population

Author: **Wenhua Ma, Li W, Chen S, Liu B, Hu L, Xu X, Shang L**

Beijing Engineering Research Center of Crime Scene Evidence Examination & Key Laboratory of Forensic Genetics, Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China

In this study, we built a sequence database of the hypervariable region 1 (HVI) of mitochondrial DNA (mtDNA) from 303 unrelated individuals of China, performed probability computing and preliminarily investigated its significance in forensic science. HVI was amplified through Nested-PCR, and the amplicons were purified withWizard® PCR prep purification system (Promega Company) before they were sequenced using BigDyeTM Terminator cycle sequencing v2.0 Ready Reaction Kit. Then, the product was precipitated using isopropyl with a final concentration of 60%, and detected on ABI Prism 377 sequencer. The electrophoretic results were analyzed using Sequence Analysis and Sequence Navigator software, and were sorted and classified with Excel.

In total, 235 haplotypes were detected in 303 individuals, in which 204 haplotypes were observed only once, 19 haplotypes were found twice, and 6 haplotypes were observed three times. Moreover, the frequency of other 6 haplotypes was 13, 8, 7, 6, 5 and 4, respectively. Statistical calculation using Nei's method indicated a haplotype diversity (h) of 0.957. The most common haplotype in Chinese population is 16223T 16362C, the frequency of which is 0.0429 in this study. Due to the lack of sufficient statistical data, we only give a conclusion that either the haplotypes are identical or not on cases involving mtDNA testing at present, rather than making further evaluation of the identical results. The strength of mtDNA evidence is partly limited by the size of the database. For example, many of the rare haplotypes was found only once in the mtDNA database, but the observed haplotype frequency 1/n (n represents the size of the database) is much lower than the real frequency. As the size of the database increases, this defect will be corrected. The database of 303 mtDNAs in this study is obviously too small, and the ideal size would be 1000-2000. Besides, mtDNA displays significantly geographical and demographic characteristics in our handling caseworks, so an ideal database should be divided by region and population.

60. Forensic characteristics and phylogenetic analysis of both Y – STR and Y-SNP in Li and Han ethnical groups from Hainan Island of China

Author: **Song M[1], Wang Z[1], Zhang Y[2], Wang M, Qian X[1], Yiping Hou[1]**

[1]Institute of Forensic Medicine, West China School of Basic Science and Forensic Medicine, Sichuan University, Chengdu 610041, China
[2]Hainan Provincial Blood Center, Haikou 570311, China

Hainan is the second largest island and the smallest and southernmost province of China. Li ethnical group is the earliest inhabitant of Hainan Island and the total population of Li ethnical group is 1.46 million. To study the relationship between Li and Han ethnical groups, blood samples were collected from 102 unrelated male Li individuals and 200 unrelated male Han individuals in Hainan Island. We combined next generation sequencing (NGS), capillary electrophoresis and pyrosequencing under the term 'NGS+' for typing Y-STRs and Y-chromosomal single nucleotide polymorphisms (Y-SNPs). As results, the high-resolution Y-SNP haplogroup and Y-STR haplotypes with 27 loci were obtained by NGS+. Significant genetic differences were observed between Li and Han ethnical groups in Hainan Island.

61. Development and Validation of 18 X-STR Loci

Author: **Zhengliang Yu, Ma W, Mo X, Yao Y, Bai X, Sun H, Li W**

Beijing Eenineering Research Center of Crime Scene Evidence Examination&Key Laboratory of Forensic Genetics, Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China

A multiplex amplification system of 18 X-STR loci was developed, and the stability, sensitivity and specificity of the system were verified. A total of 18 X-STR lociDXSGATA31E08, DXS10079, DXS10103, DXS7132, DXS9895, DXS7133, DXS7424, DXS7423, DXS6789, DXS9902, DXS6810, DXS8377, DXS101, HPRTB, DXS8378, DXS6797, DXS6804, GATA165B12 were selected from the X chromosome, and the Amelogenin loci was added. Multiple PCR primers were designed and marked with four fluorophores (FAM, HEX, TAMRA, ROX), a total of 140 alleles were detected for the 18 X-STR loci among 413 Beifang Han individuals. The special genetic mode of the X chromosome determines the important role of X-STR is an important complement to other genetic markers. X-STR can be used in identification of father-daughter, mother-son, sisters, half-sister, segregation, incest. The polymorphism information content (PIC) was 0.434-0.908, the discrimination power (DP) was 0.4326-0.9831, the combined discrimination power was 0.99999993 in males, and 0.99999999 in females, the combined mean exclusion chance was 0.99999326 in duo cases (CMECduo), and 0.99999999 in trio case (CMECtrio). The TYPER-X19 kit has achieved the application of forensic evidence requirements, can provide an effective means for the identification of complicated and complex parental cases.

62. Developmental validation of Hungary's first forensic mitochondrial CR sequencing analysis protocol

Author: **Zsófia Prónay**

Hungarian Institute for Forensic Sciences

In this study a simple and reliable analysis method was developed and validated based on SWGDAM and ENFSI validation criteria for forensic mitochondrial DNA control region analysis. The method is a Sanger based Big-Dye sequencing, based on seven overlapping

fragments which results high resolution of the C-stretches as well. The technique proved to be suitable for analysing samples containing low copy number DNA.

In our Institute sensitivity was tested by using two dilution row, from 5-5 blood samples of one person. The autosomal DNA concentrations in the first row have been about 0.4, 2, 4 pg/µl and in the second row about 0.1, 10, 20 pg/µl. The method was proven suitable for testing mtDNA from samples even with very low autosomal DNA concentration.

The method's stability was tested by using five tissue types (bones, hairs, muscles, blood samples and buccal swabs) from five different persons. The process was not tissue-dependent and got the same result every time when the quantity of the DNA in the sample was enough for mitochondrial analysis.

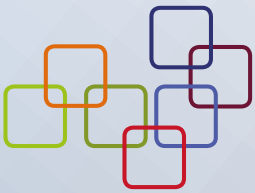
Repeatedly sequencing the same samples five times by the same person and three samples by two different persons shown the method is repeatable and reproducible.

The method is proven human specific by using cat, dog and bird samples which had been sequenced with species specific primers to prove the samples contains mtDNA, then have been tried to analyse with the primers we use for the human method. In the next round of the validation human-animal mixes were sequenced with the method's primers and only the human mitotype have been gained.

For contamination testing negative and positive controls were analysed parallel with the samples. From all the PCRs contaminations appeared only four times, two times from preparing and two times from sample collecting.

Testing with mixed samples proved that we could separate two peoples mtDNA from a mix, if we know their mitotypes.

After completing the described validation procedure and approved by the National Accreditation Authority our protocol become the first validated forensic mitochondrial DNA control region sequencing analysis protocol in Hungary.



11th Haploid Markers Conference

INFERRING ANCESTRY FROM DNA

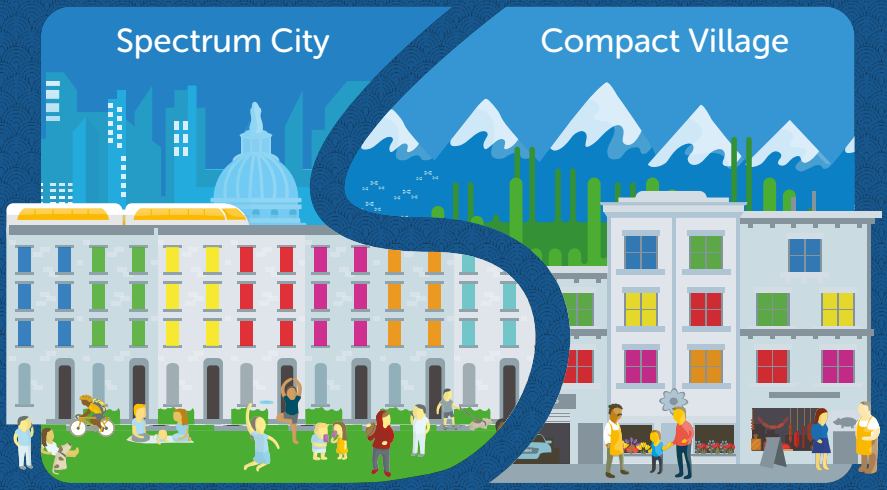
17-19th May 2018

Bydgoszcz, Poland

DAY 1: Thursday, 17th May, 2018	9:00 - 13:00	Registration Poster Exhibition	
	13:00 - 13:15	Opening	MANRU LECTURE HALL, 3RD FLOOR
	13:15 - 13:45	Keynote: Chris Phillips	MANRU LECTURE HALL, 3RD FLOOR
	13:45 - 15:15	Oral Session 1: Inferring Ancestry from DNA, part 1	MANRU LECTURE HALL, 3RD FLOOR
	15:15 - 16:45	Lunch Seminar with Thermo Fisher Scientific	MANRU LECTURE HALL, 3RD FLOOR
		Lunch	IN FRONT OF FIDELIO LECTURE HALL, 1ST FLOOR
		Poster Exhibition	CONFERENCE ROOM NO. IV, 2ND FLOOR
	16:45 - 18:15	Oral Session 1: Inferring Ancestry from DNA, part 2	MANRU LECTURE HALL, 3RD FLOOR
	18:30	Dinner	IN FRONT OF FIDELIO LECTURE HALL, 1ST FLOOR
DAY 2: Friday, 18th May, 2018	9:00 - 9:30	Keynote: Mark Jobling	MANRU LECTURE HALL, 3RD FLOOR
	9:30 - 11:00	Oral Session 2: New Generation Sequencing, part 1	MANRU LECTURE HALL, 3RD FLOOR
	11:00 - 11:30	Coffee Break	IN THE FOYER IN FRONT OF MANRU LECTURE HALL, 3RD FLOOR
	11:30 - 13:00	Oral Session 2: New Generation Sequencing, part 2	MANRU LECTURE HALL, 3RD FLOOR
	13:00 - 14:30	Lunch	IN FRONT OF FIDELIO LECTURE HALL, 1ST FLOOR
		Poster Exhibition	CONFERENCE ROOM NO. IV, 2ND FLOOR
	13:30 - 14:30	Lunch Seminar with Verogen	MANRU LECTURE HALL, 3RD FLOOR
	14:30 - 16:00	Oral Session 3: Casework, part 1	MANRU LECTURE HALL, 3RD FLOOR
	16:00 - 16:30	Coffee Break	IN THE FOYER IN FRONT OF MANRU LECTURE HALL, 3RD FLOOR
	16:30 - 18:15	Oral Session 3: Casework, part 2	MANRU LECTURE HALL, 3RD FLOOR
	20:00	Concert "4 Szmery" and dinner	SOWA RESTAURANT
DAY 3: Saturday, 19th May, 2018	9.00 - 9.30	Keynote: Chris Tyler-Smith	MANRU LECTURE HALL, 3RD FLOOR
	9.30 - 11.00	Oral Session 4: Population Genetics, part 1	MANRU LECTURE HALL, 3RD FLOOR
	11.00 - 11.30	Coffee Break	IN THE FOYER IN FRONT OF MANRU LECTURE HALL, 3RD FLOOR
	11.30 - 13.00	Oral Session 4: Population Genetics, part 2	MANRU LECTURE HALL, 3RD FLOOR
	13.00 - 13.15	Closing the Conference	

Promega

Discover a new world of convenience, flexibility and independence in CE systems



Do you need a high rise for your casework and access to the latest technology?
Or do you prefer to work among the chemistries you grew up using?
Either way, there's a Spectrum System for you.

www.promega.com/SpectrumJourney

Spectrum CE SYSTEMS