AGENDA FOR THE EDNAP MEETING

WARSAW – 26 APRIL 2016

Expected duration: 09.00 - 17.00

Coffee: 10.15 – Lunch: 12.30-14.00 – Coffee: 15.15

Host: Magdalena Jabłonska
Chairman: Niels Morling

Welcome	Magdalena Jablonska
Update on activities concerning	
mtDNA SNP screening – two PCRs, 18 SNPs	Arnoud Kal
Methylated DNA and age exercise	Athina Vidaki
Updates from other groups	
EMPOP	Walther Parson
High quality DNA sequence database	Walther Parson
Nomenclature of STR sequences	Walther Parson
EUROFORGEN-NoE	Peter Schneider
Singapore and DNA Working Group of The Asian Forensic Sciences Network	Christopher Syn
Interpretation of complex DNA mixtures in crime cases	Peter Gill
Outline of a new DNA commission on the evaluation of evidence	Peter Gill
Future activities	
Exercise on mRNA typing with NGS	Cordula Haas
EDNAP meeting in autumn 2016 – 8 November 2016 in Rome?	Niels Morling
(ENFSI DNA WG Steering Group Meeting: 7 Nov 2016 in Rome)	
(Promega Meeting: 9 – 10 November 2016 in Florence)	
Any other business	Niels Morling

First and last name	E-mail	Phone number	Institution / Organisation	Country	Institution`s address
Ankiewicz Hanna	haniaank@gmail.com	507277123	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Ansell Ricky	ricky.ansell@polisen.se	46105628119	Swedish National Forensic Centre	Sweden	NFC 58194 Linkoping 58194 Linkoping Sweden
Banemann Regine	regine.banemann@bka.bund.de	+49 611 5516053	Bundeskriminalamt	Germany	Thaerstr. 11 65193 Wiesbaden Germany
Bastisch Ingo	ingo.bastisch@bka.bund.de	4961116030	Bundeskriminalamt	Germany	Thaerstr. 11 65193 Wiesbaden Germany
Bengs Auli	auli.bengs@poliisi.fi	358504564610	NBI Forensic Laboratory Finland	Finland	PO BOX 285 FI-01301 Vantaa Finland
Boroń Michał	michal.boron@policja.gov.pl	509907991	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Bragoszewska Anna	anna.bragoszewska@policja.gov.pl	512151980	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Camps Lydia	lydiacamps@gmail.com	34696298555	Mossos Esquadra	Spain	Avinguda de la Pau, 120 08206 Sabadell Spain
Daly Dyan	ddaly@fsl.gov.ie	353866031631	Forensic Science Ireland	Ireland	Garda Headquarters Dublin 8 Dublin Ireland
Ferreira Paulo	pamife@hotmail.com	351966843577	Laboratorio Policia Cientifica da Policia Judiciaria	Portugal	Rua Gomes Freire, 174 1169-007 Lisbon Portugal
Gill Peter	peterd.gill@gmail.com	+447786126571	Norwegian Institute of Public Health	Norway	PO Box 4404 Nydalen 0403 Oslo Oslo Norway
Guiness June	june.guiness@homeoffice.gsi.gov.uk	447769887147	Home Office- Forensic Science Regulation	United Kingdom	5 St. Philips Place B3 2PW Birmingham United Kingdom
Hanssen Eirik	eiha@fhi.no	++4741652397	Norwegian Institute of Public Health	Norway	PO Box 4404 Nydalen 0403 Oslo Oslo Norway
Haas Cordula	cordula.haas@irm.uzh.ch	+41 44 635 56 56	University of Zurich, Institute of Forensic Medicine	Switzerland	Winterthurerstrasse 190/52 8057 Zurich Switzerland
Jabłońska Magdalena	magdalena.jablonska@policja.gov.pl	48226012629	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Kadyjewska Ewa	ewakadi@tlen.pl	48 22 60 15855	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Kal Arnoud	a.kal@nfi.minvenj.nl	+31-6-48131812	Netherlands Forensic Institute	Netherlands	PO box 24044 2490AA The Hague Netherlands
Kartasińska Ewa	ewa.kartasinska@policja.gov.pl	48226012625	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland

Kneppers Alexander	s.kneppers@nfi.minvenj.nl	31629623036	Netherlands Forensic Institute	Netherlands	PO Box24044 2490 AA The Hague Netherlands
Kratky Martin	mar.kratky@gmail.com	420974824526	Institute of Criminalistics Prague	Czech Republic	Strojnicka 27 17089 Prague Czech Republic
Laurent Francois-Xavier	francoisxavier.laurent@interieur.gouv.fr	33641561273	Institut National de Police Scientifique	France	31 Avenue Franklin Roosevelt 69134 ECULLY CEDEX France
Makowska Żanetta	zmakowska.72@gmail.com	48 22 60141 44	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Mogensen Helle	helle.smidt@sund.ku.dk	+45 61140212	University of Copenhagen	Denmark	Frederik Vs vej 11 DK-2100 Copenhagen Denmark
Morling Niels	niels.morling@sund.ku.dk	+45 21206110	University of Copenhagen	Denmark	Frederik Vs vej 11 DK-2100 Copenhagen Denmark
MUNGUIA SERGIO	smunguia@guardiacivil.es	+415146000 EXT 48396	GUARDIA CIVIL	Spain	GUZMAN EL BUENO 110 28003 MADRID Spain
Noel Fabrice	fabrice.noel@just.fgov.be	+32 2 243 46 04	NICC	Belgium	Vilvoordsesteenweg 100 1120 Brussels Belgium
Ochocki Mateusz	mateusz.ochocki@policja.gov.pl	602625200	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Parson Walther	walther.parson@i-med.ac.at	+43512900370640	Medical University of Innsbruck	Austria	Muellerstrasse 44 6020 Innsbruck Austria
Pięta Agnieszka	agnieszkapieta2@gmail.com	48516098948	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Porto Maria	m.joao.porto@inmlcf.mj.pt	+351 239854230	National Institute of Legal Medicine and Forensic Sciences	Portugal	Largo da Se Nova 3000-213 Coimbra Portugal
Scheithauer Richard	richard.scheithauer@i-med.ac.at	+43512900370600	Medical University of Innsbruck	Austria	Muellerstrasse 44 6020 Innsbruck Austria
Schneider Peter	peter.schneider@uk-koeln.de	491735916684	Institute of Legal Medicine, University of Cologne	Germany	Melatenguertel 60/62 50823 Cologne Germany
Simonsen Bo	bo.simonsen@sund.ku.dk	+45 28756136	University of Copenhagen	Denmark	Frederik Vs vej 11 DK-2100 Copenhagen Denmark
Syn Christopher	christopher_syn@hsa.gov.sg	6598456812	Health Sciences Authority, Singapore	Singapore	11 Outram Road 169078 Singapore Singapore
Syndercombe Court Denise	denise.syndercombe-court@kcl.ac.uk	+44 20 7844 4155	Kings College London	United Kingdom	Franklin Wilkins Building, Kings College, 150 Stamford Street, London SE1 9NH SE1 9NH London United Kingdom
Tan Jiayu	tan_jiayu@hsa.gov.sg	6596321509	Health Sciences Authority, Singapore	Singapore	11 Outram Road 169078 Singapore Singapore

Vidaki Athina	athina.a.vidaki@kcl.ac.uk	0020 7848 4155	Kings College London	United Kingdom	Franklin Wilkins Building, Kings College, 150 Stamford Street, London SE1 9NH SE1 9NH London United Kingdom
Wysocka Katarzyna	wysockak@o2.pl	2260112601	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland
Zatkalikova Livia	livia.zatkalikova@minv.sk	00421907834948	Ministry of Interior	Slovakia	Pribinova 2 812 72 Bratislava Slovakia
Życka-Krzesińska Joanna	krzejo@tlen.pl	606939391	Central Forensic Laboratory of the Police	Poland	Aleje Ujazdowskie 7 00-583 Warsaw Poland

EUROPEAN DNA PROFILING GROUP (EDNAP) MEETING

Warsaw, Poland

26 April 2016

Host: Magdalena Jablonska
Chairman: Niels Morling

A list of participants is attached.

Welcome
Magdalena Jablonska welcomed members to Warsaw.

Update on exercises
A SNaPshot based method targeting18 common mtDNA mutations Arnoud Kal
Arnoud Kal presented the results of the collaborative EDNAP exercise concerning typing of 18 mtDNA SNPs with the SNaPshot method and whole mtDNA genome sequencing. A draft for a manuscript will be circulated soon (presentation attached).

Methylated DNA and age exercise Athina Vidaki
Athina Vidaki presented the results of the collaborative EDNAP exercise on age estimation by means of measurements of methylation of selected DNA positions (presentation attached).

Updates from other groups
STRidER Walther Parson
Walther Parson gave an update on the High quality DNA sequence database, STRidER (presentation attached).

EMPOP Walther Parson
Walther Parson gave an update on the new version of the database EMPOP. New developments include an event-based search engine (SAM-E), map-based description of database matches, a new tabular layout of search results and the provision of the haplogroup status of mtDNA sequences. EMPOP follows the concept of providing high-quality data and basic results of database searches. The interpretation of search results is the responsibility of the users. There are courses planned for training on mtDNA matters in association with congresses. Individual courses can be organized (presentation attached).

Nomenclature of STR sequences Walther Parson
Walther Parson gave an update on the thoughts about the nomenclature of DNA sequences of STRs (presentation attached).

EUROFORGEN-NoE – General update Peter Schneider
Peter Schneider gave an update concerning the project (presentation attached).

Singapore and DNA Working Group of the Asian Forensic Sciences Network Christopher Syn
(presentation attached).

Other presentations

Interpretation of complex DNA mixtures in crime cases (presentation attached). — Peter Gill

Outline of a new DNA commission on the evaluation of evidence (presentation attached). — Peter Gill

EDNAP website update (www.isfg.org/EDNAP) — Peter Schneider

Members are encouraged to visit the website. Suggestions are welcome.

Future activities

Second exercise on methylated DNA and age — Athina Vidaki

The EDNAP members accepted an invitation from Athina Vidaki and Denise Syndercombe Court for a second collaborative EDNAP exercise concerning forensic age estimation based on investigation of the degree of DNA methylation of select nucleotides. The proposed methods are similar to those of the first exercise making it possible to use equipment from both the MiSeq and the PGM. Kings College will send out invitations, protocols, and 5-10 blood samples that should be investigated. It is the idea to circulate the samples in June 2016 so that the results can be discussed at the next EDNAP meeting (presentation attached).

Exercise on mRNA typing with NGS — Cordula Haas

The EDNAP members accepted an invitation from Cordula Haas for a NGS based study of discrimination between various tissues and body fluids. The investigations can be performed with both the MiSeq and the PGM. Cordula Haas will send out invitations, protocols and some tissue and body fluid samples. It is the idea to circulate the samples in June 2016 so that the results can be discussed at the next EDNAP meeting (presentation attached).

Next meeting — Niels Morling

The next EDNAP meeting will be held 8 November 2016 in Rome in connection with the meeting of the steering group the ENFSI DNA WG Meeting 7 November 2016. The colleagues from Laboratoria Genetica Forense, Universita Cattolica, Rome, will organise the meetings.

Any other business — Niels Morling

There was no other business.

Closing of the meeting

The meeting closed with sincere thanks to Magdalena Jablonska and all other colleagues, who helped to organise the meeting.

Attachments are found at the EDNAP website http://www.isfg.org/EDNAP/Meetings:

- Agenda
- List of participants
- Presentations
  - Arnoud Kal: Report on mtDNA SNP typing
  - Athina Vidaki: Report on methylated DNA and age
  - Walther Parson: EMPOP report
  - Walther Parson: STRidER report
  - Peter Schneider: EUROFORGEN-NoE report.

- Christopher Syn: Singapore and DNA Working Group of the Asian Forensic Sciences Network
- Peter Gill: Interpretation of complex DNA mixtures in crime cases
- Peter Gill: Outline of a new DNA commission on the evaluation of evidence

Netherlands Forensic Institute
Ministry of Justice

Update Exercise mtDNA SNaPshot

Arnoud Kal

Natalie Weiler
Titia Sijen

26 April 2016, Warsaw

A control region-based mtDNA SNaPshot selection tool, integrated into a mini amplicon sequencing method

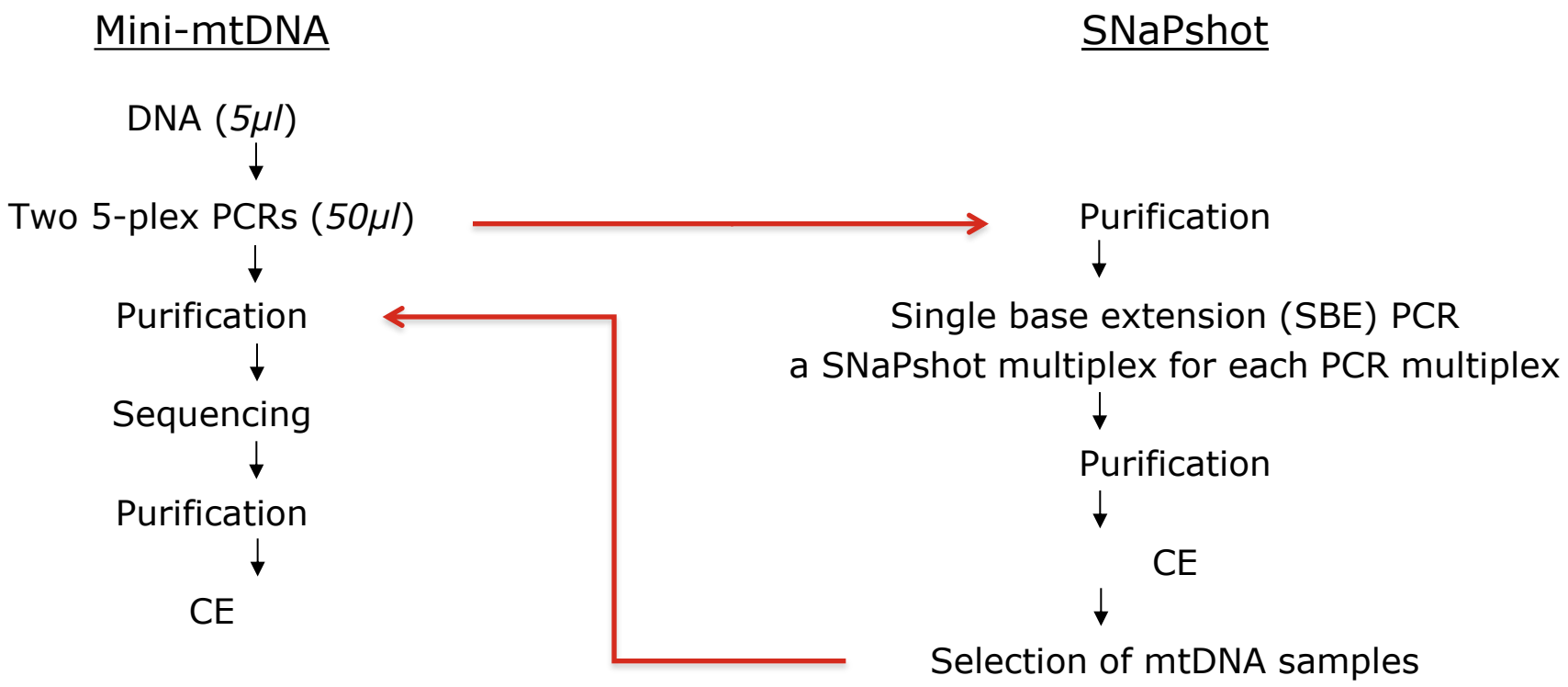
- Targets 18 SNPs in HVS I - II – III

- Degenerate bases in 3’ part primer to cover SNPs at primer binding site positions

- Two SNaPshot multiplexes for PCR products of mini amplicon mtDNA multiplexes (Eichmann et al 2008)

SNP	Base change	Frequency	Haplogroup
73	A>G	0.5551	HV / H / V
146	T>C - T>a	0.0933 - 0.0001	
150	C>T - C>g	0.1028 - 0.0001	
152	T>C	0.2007	
182	C>T	0.0088	
185	G>A - G>t - G>c	0.0541 - 0.0031 - 0.0004	
195	T>C - T>a	0.1986 - 0.0002	
489	T>C	0.1351	M / J
497	C>T	0.0419	K
16126	T>C	0.1799	
16129	G>A - G>c	0.0689 - 0.0111	
16223	C>T	0.1405	
16270	C>T	0.0876	
16278	C>T	0.0646	
16294	C>T - C>a - C>g	0.1071 - 0.0003 - 0.0002	
16311	T>C	0.1676	
16362	T>C	0.0743	
16519	T>C	0.6642	

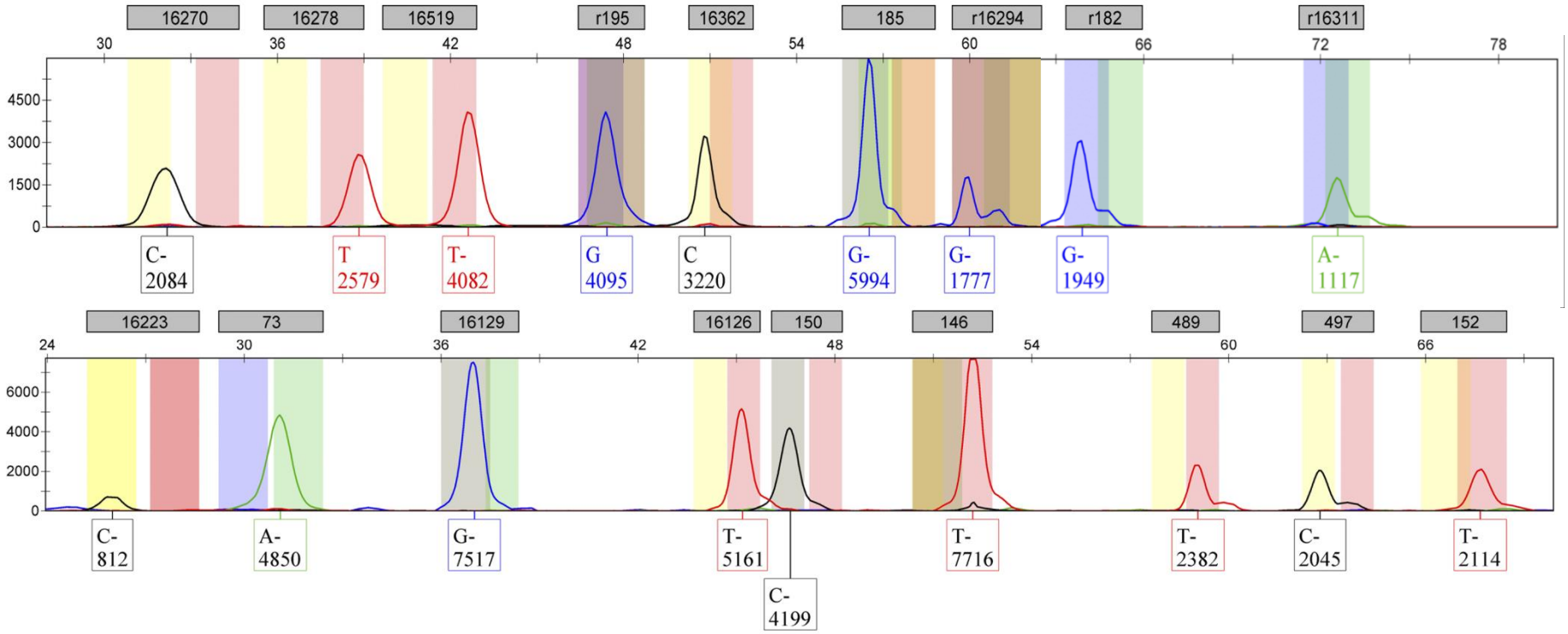
Same PCR product for sequencing and SNaPshot



Example: Case with 30 hairs → 600 sequencing reactions

SNaPshot: Selection of 3 hair samples → 60 sequencing reactions

Optimised SNaPshot assay



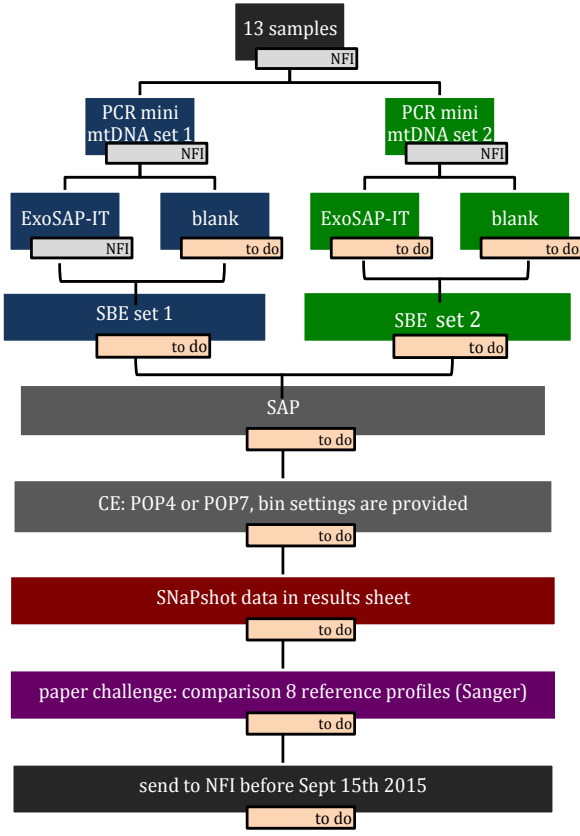
- SNP number preceded by ‘r’: reverse primer
- Allele call followed by ‘-’: rCRS allele

EDNAP Exercise: 3 parts – 14 labs (excl NFI)

① SNaPshot assays on 13 samples for which PCR products are provided

② Paper challenge: compare results 1 to list of 8 references given in standard nomenclature

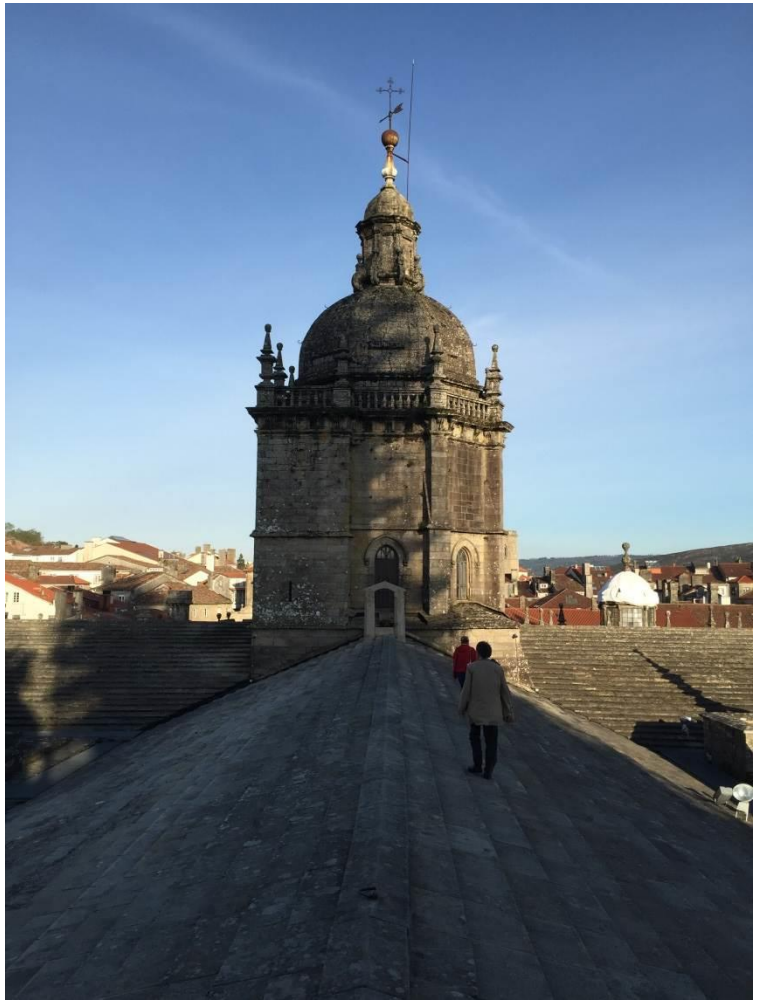
③ Optional: NGS full mtDNA analysis of 2 samples
- » Commercial control DNA sample (cell line)
- » Sample with heteroplasmy



Santiago meeting 2015

Update on results for part 1 and part 2:

- SNP typing, haplogroup inference
- Paper challenge



NGS exercise

Sample A (=sample 11 in part one of the exercise):

- several polymorphisms
- one point heteroplasmy also targeted by the mini-mtSNaPshot (146Y)
- two insertions of an AC-repeat at position 524
- a C-stretch at position 574 for which the exact number of C's remained undetermined by Sanger sequencing (position 574 is a known homopolymeric position)

Sample B

- presumably cell line-derived control DNA (hDNA; Life Technologies)
- several polymorphisms
- a deletion of an AC-repeat at position 523-524.

NGS analysis

3 x MiSeq
1 x Ion Torrent PGM

Full control region (16024 to 16569 and 1 to 576)
Full mtDNA genome

NGS results

The MPS results were highly concordant despite the marked differences in average read coverage between the four laboratories (~1000 - ~85.000)

Exception: an 309.1C insertion in sample B that was not detected by two laboratories and Sanger sequencing, while two other laboratories did observe this insertion.

The ratio between the two bases at a heteroplasmic position was similar for all three laboratories (46-49%)

Manuscript

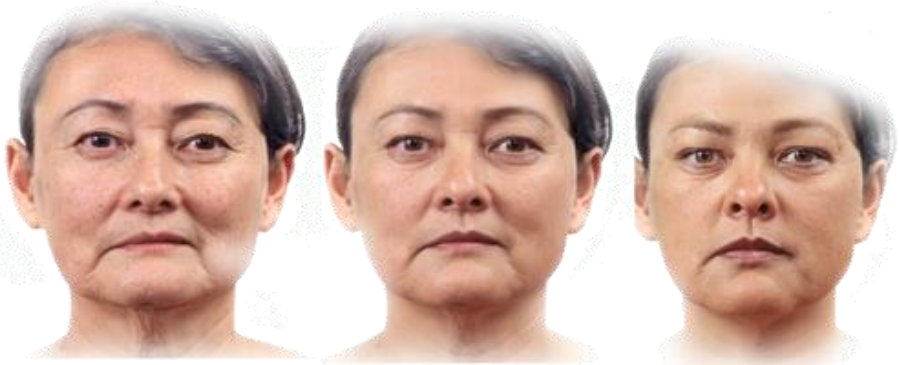
Draft manuscript is allmost ready for sending out to all 14 labs

Big THANK YOU ALL!!

Erasmus MC
University Medical Center Rotterdam

EDNAP Meth-Age exercise Results - part 1

A. Vidaki, D. Ballard, D. Syndercombe Court



EDNAP meeting
Warsaw, 26 April 2016

KING'S College LONDON

Overview

- DNA methylation-based age prediction using NGS - update
- EDNAP Meth-Age exercise part 1 - proposed in Oct 2015
- Participating laboratories
- Samples (part 1) - Dec 2015
- Protocols for both MiSeq and PGM
- Data collection - April 2016
- Questionnaires & NGS run QC
- DNA Methylation results
- Summary
- EDNAP Meth-Age exercise part 2 - Proposal

Sample collection

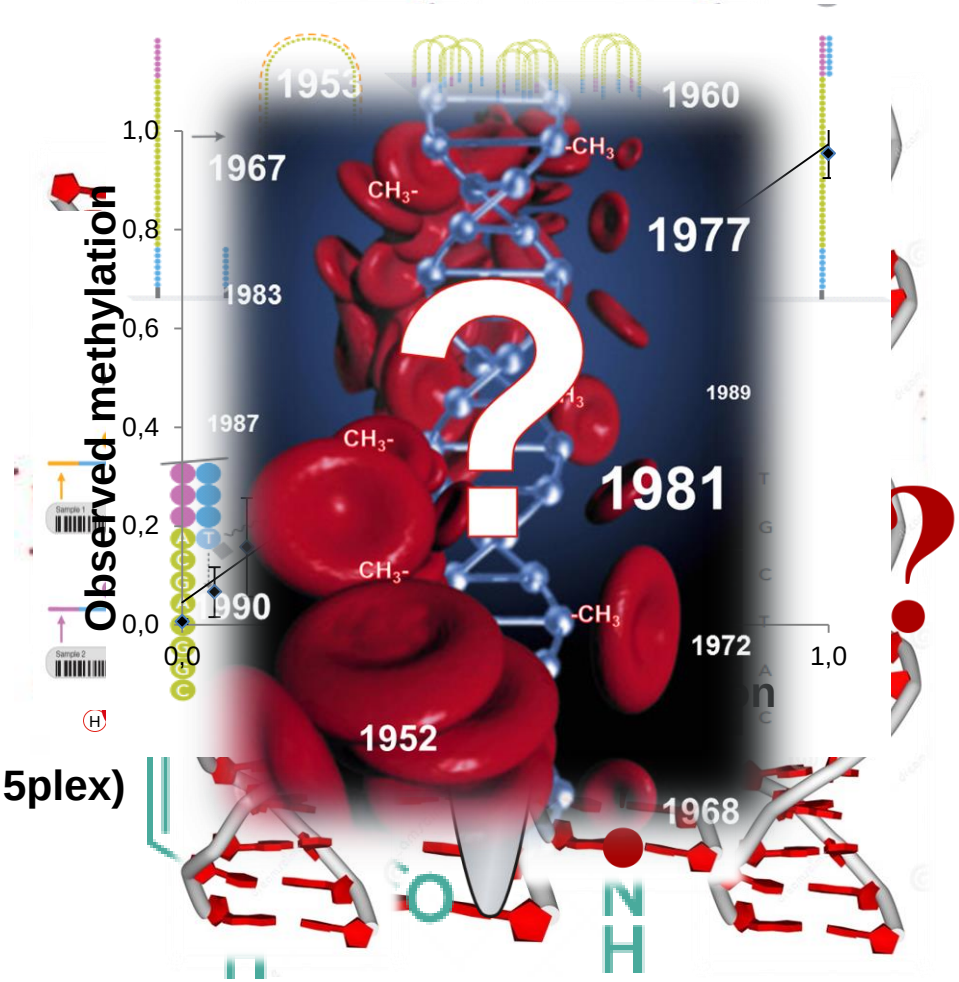
DNA extraction

DNA quantification

Bisulfite conversion

Bisulfite PCRs (7plex & 5plex)

PCR product pooling



PCR product quantification

Library preparation

ary quantification

DNA sequencing

ata normalisation

Age prediction

Task/Part 1 - Method evaluation

Samples

7 commercially available DNA methylation standards (EpigenDx, USA)

0%, 5%, 10%, 25%, 50%, 75%, 100%

To be analysed in duplicate

DNA preparation

100ng of DNA to be converted (5ng/µl)

2 multiplexes (Age 7plex & 5plex) - provided standardised protocol)

PCR products to be combined, purified and quantified

DNA sequencing

Library preparation, amplification and quantification

NGS platform: MiSeq and/or PGM

PGM protocol with the help of Mayra (Innsbruck) & Theresa (Cologne)

Participating laboratories

- Copenhagen, Denmark
- Innsbruck, Austria
- NFI, Netherlands
- Orlando, Florida, USA
- Lyon, France
- NIST, USA
- Victoria, Australia
- Singapore
- London, UK
- Cologne, Germany
- Oslo, Norway
- Zurich, Switzerland
- Santiago de Compostela, Spain

No data/technical problems:

- *Adelaide, Australia*
- *Auckland, New Zealand*

Questionnaires

DNA preparation

- 12X MethylEdge Bisulfite Conversion System, Promega
- 12X Multiplex PCR kit, Qiagen
- 12X MinElute PCR purification kit, Qiagen
- 9X Qubit dsDNA HS assay kit, ThermoFisher Scientific

  2X QuantiFluor, Promega

  1X QIAxcel, Qiagen

Questionnaires

DNA sequencing

MiSeq labs - 9 data sets

- 9X KAPA Hyper Prep kit, Illumina
- 8X KAPA library quant kit for Illumina, KAPA Biosystems

  1X Qubit dsDNA HS kit, ThermoFisher Scientific
- 7X MiSeq reagent kit v2, 300 cycles

  1X MiSeq reagent kit v3, 600 cycles

  1X MiSeq reagent kit v3, 151 cycles (single reads)
- 1X own data analysis tool, TSSV/FDSTools (NFI)

Questionnaires

DNA sequencing

PGM labs - 7 data sets

- 7X Ion Xpress Plus gDNA Fragment library kit, ThermoFisher Scientific
- 6X Ion library quantitation kit, ThermoFisher

  1X Qubit dsDNA HS kit, ThermoFisher Scientific
- 5X Ion PGM Hi-Q OT2 kit/Ion PGM Hi-Q Sequencing kit, ThermoFisher

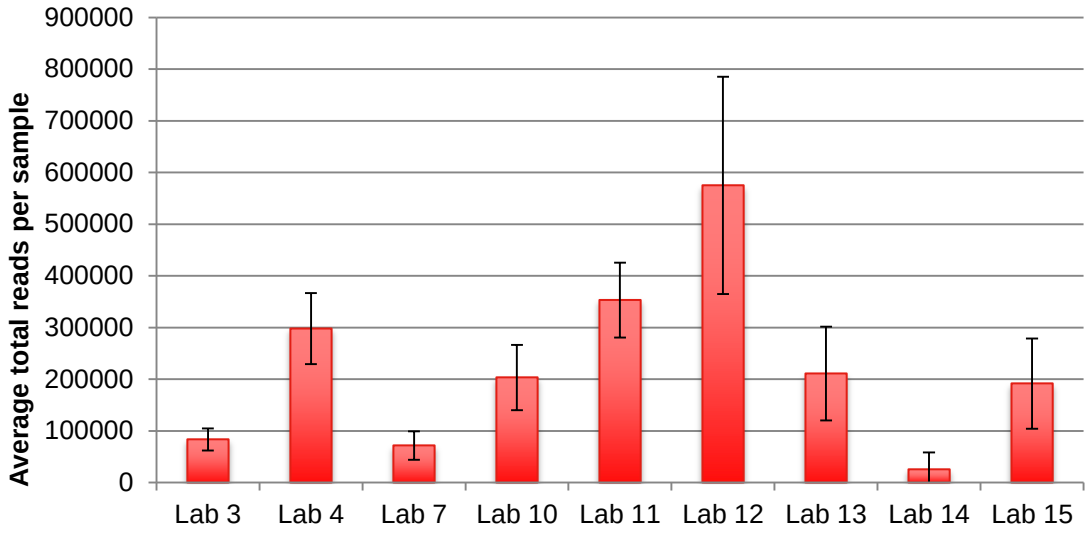
  1X as above plus 8 extra amplification cycles

  1X Ion PGM IC 200 kit/Ion PGM Hi-Q Sequencing kit, ThermoFisher

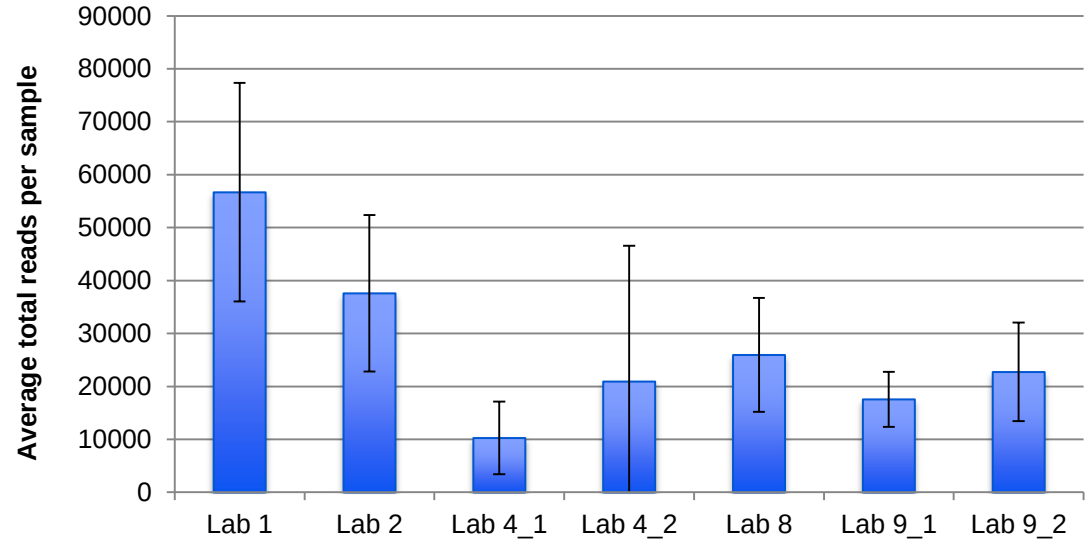
Technical/experimental issues

- There was a mix-up of 2 methylation standards (10% & 25%) (most likely during sample preparation) confirmed by a ‘mixed’ methylation profile (~17%) (Lab 14)
- MiSeq - cluster density was quite high in many labs resulting in lower QC values
- Instructions on library dilution using average fragment length was not clear in the protocol
- Strange higher methylation values for certain CpGs (Lab 13)
- PGM: problems like bad loading, poor alignment, unexpected read length histograms
- A few marker ‘drop-outs’ in PGM data
- Lab 3 did not run the samples in duplicate

Reads - per sample per lab



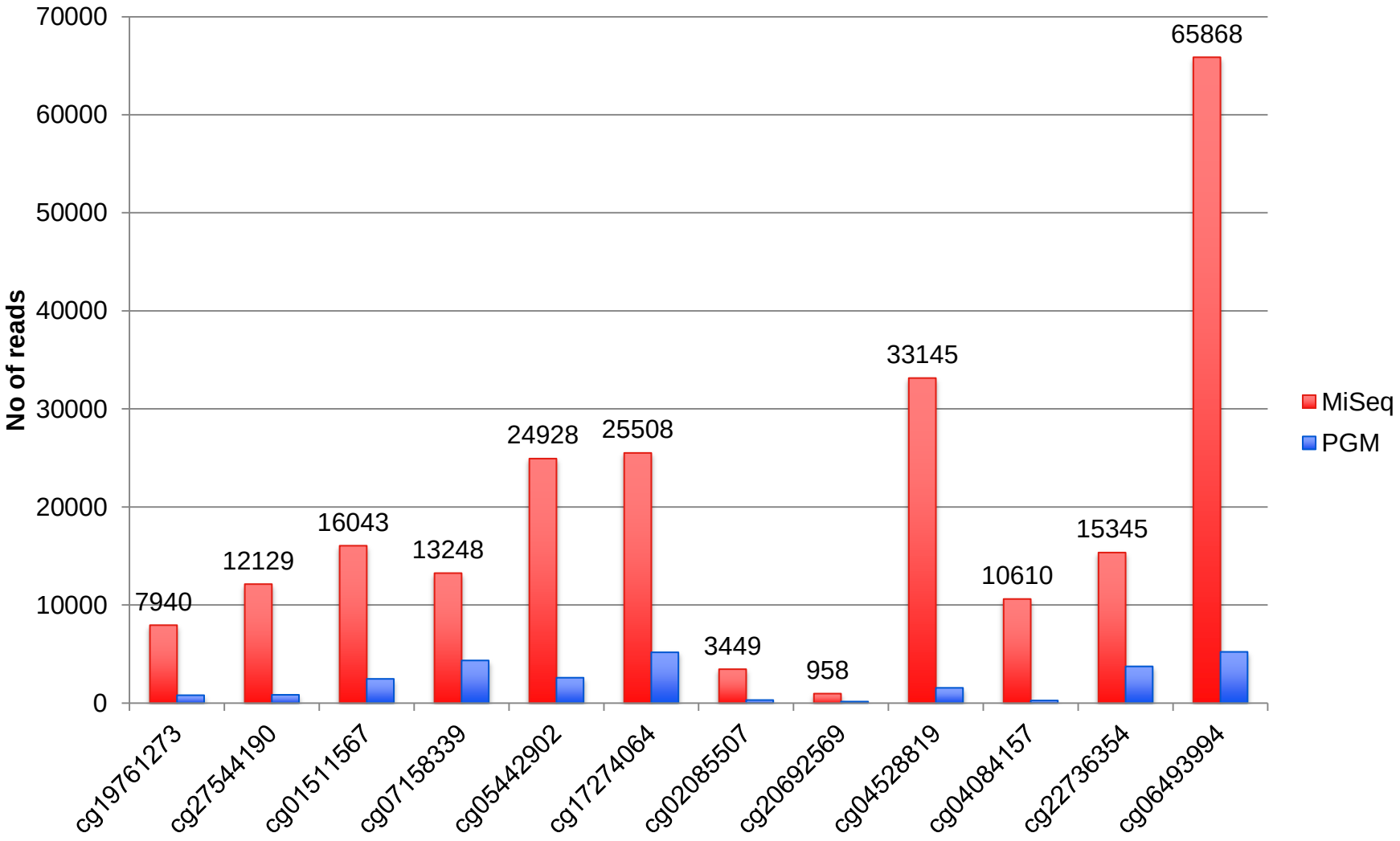
MiSeq ..8X



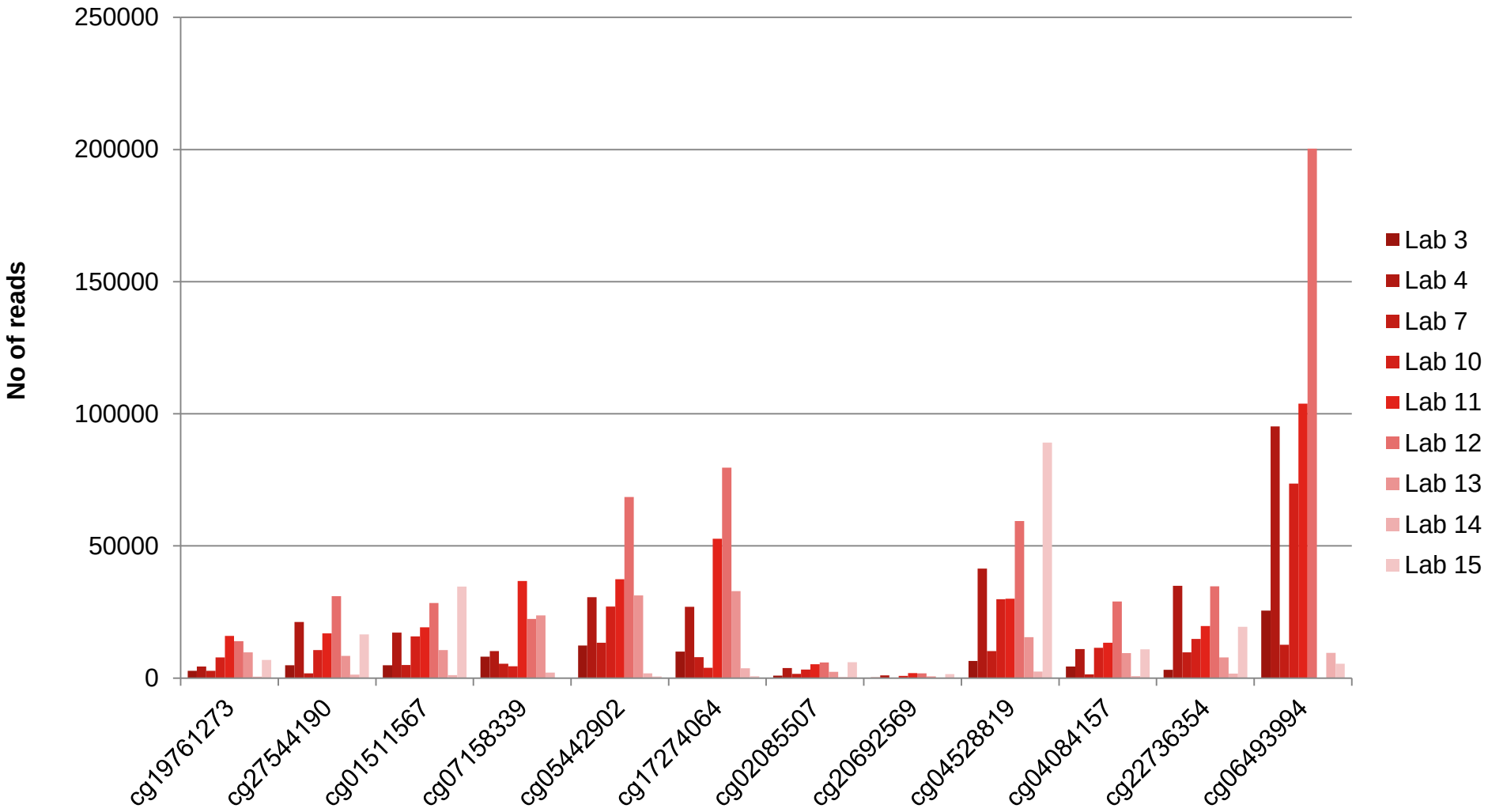
PGM

Reads - per marker

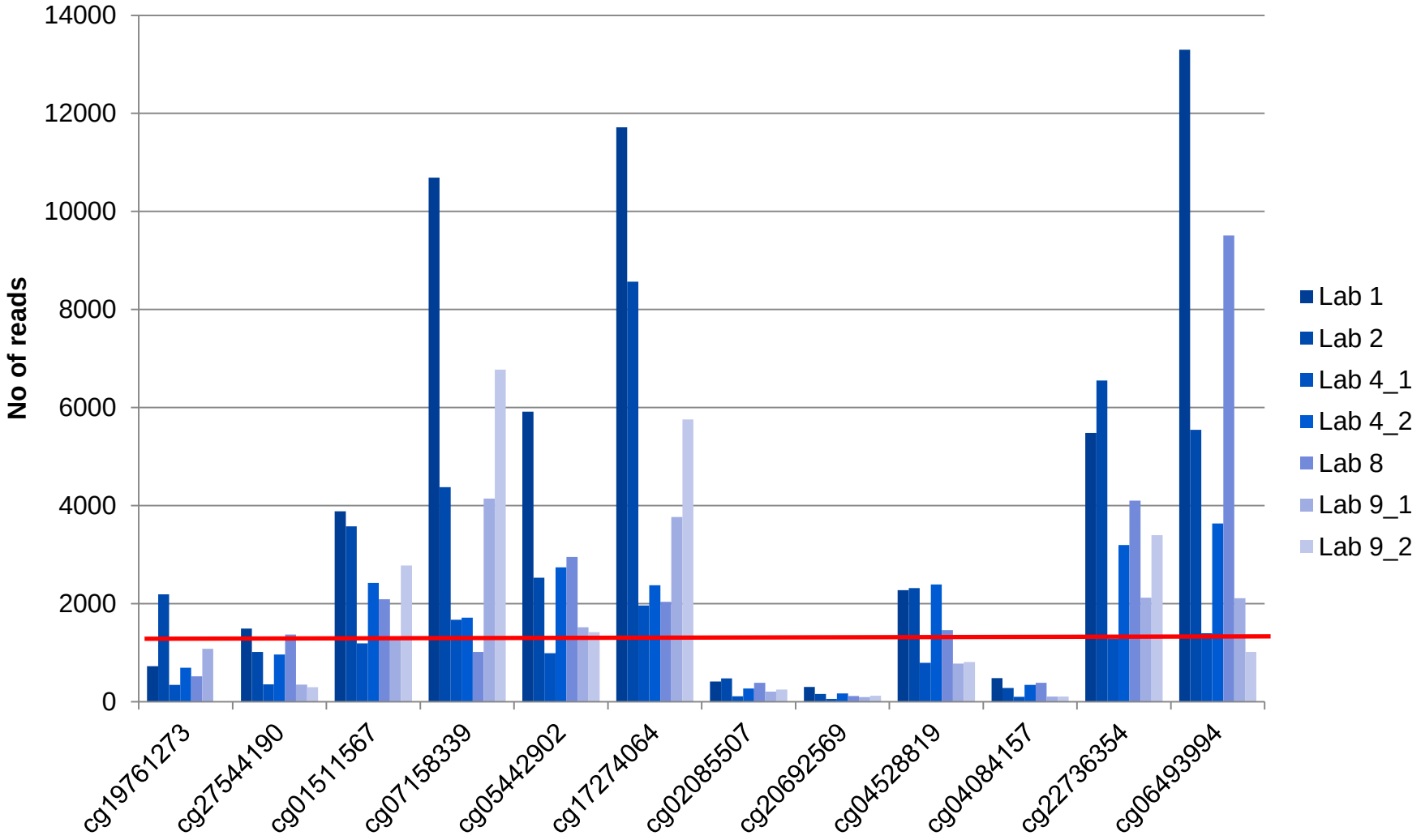
MiSeq ..3X to 41X



Reads - per marker, MiSeq



Reads - per marker, PGM



Methylation data analysis

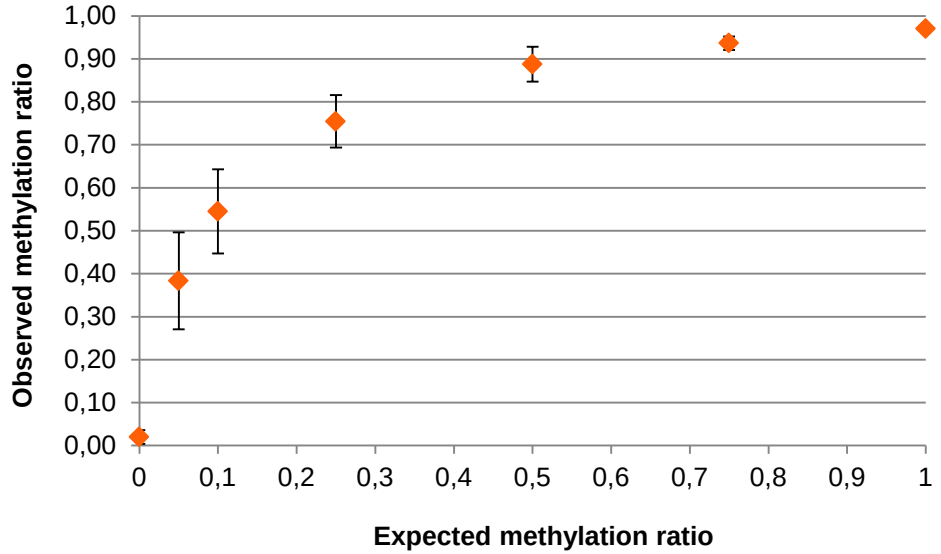
- Laboratories provided raw data in FastQ file format
- Files were analysed using KCL script for generation of vcf and bam files
- The methylation ratio of each site was calculated using the formula:

  % methylation = C reads/(C+T reads)
- The average methylation was calculated from the duplicates and used for analysis

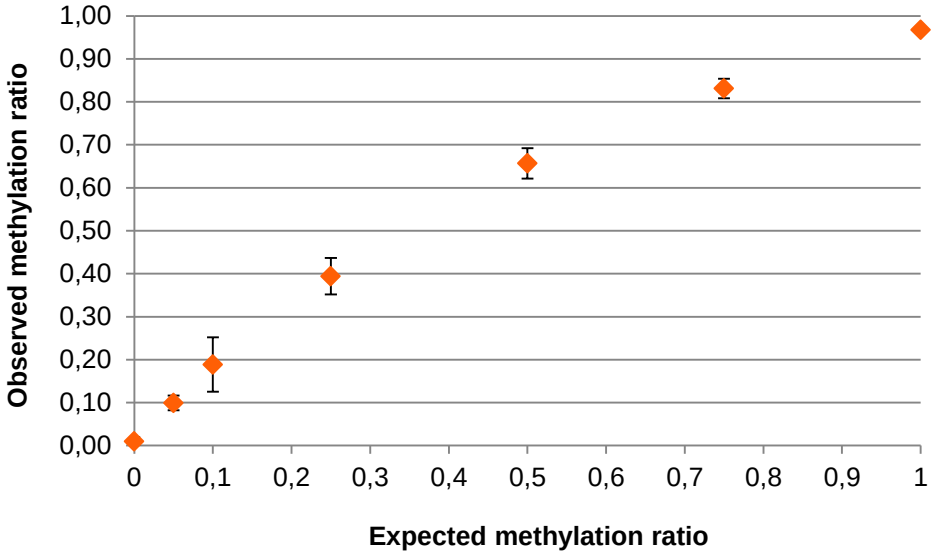
Additional QC:

- DNA conversion rates were also calculated - **>99%** in almost all cases
- Negative samples (PCR-negative, No-conversion control) gave no reads

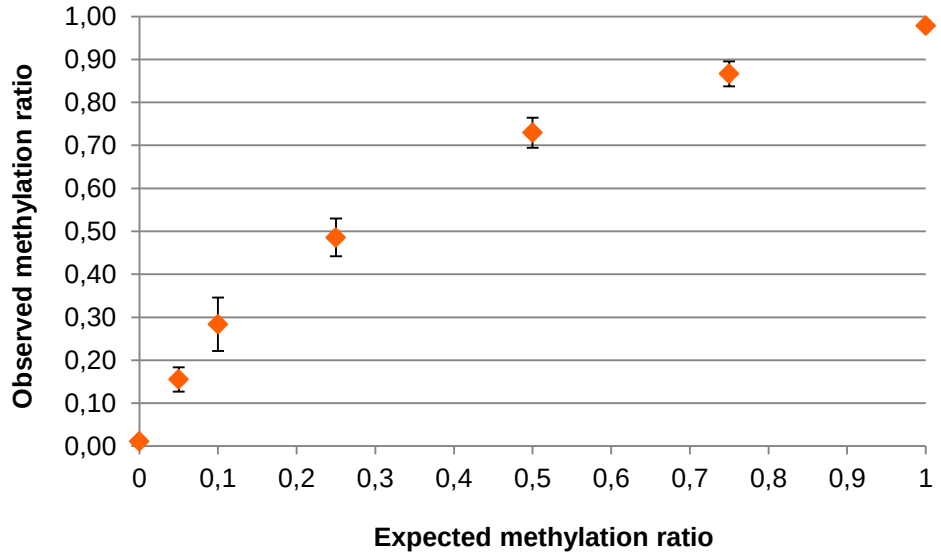
cg19761273



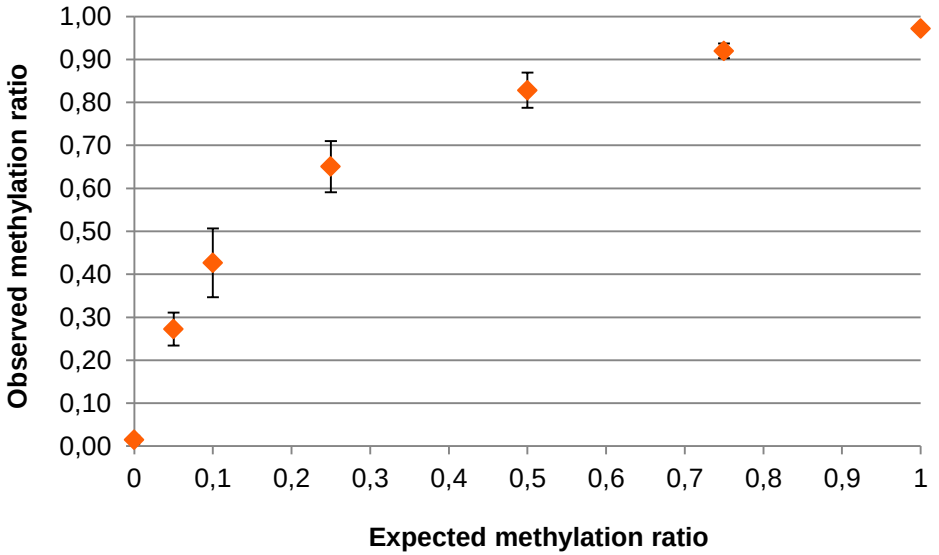
cg07158339



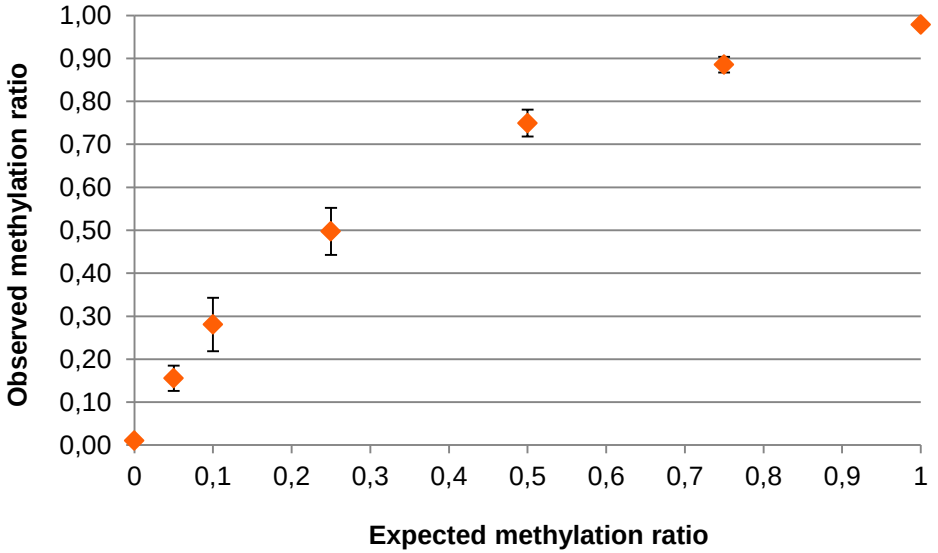
cg27544190



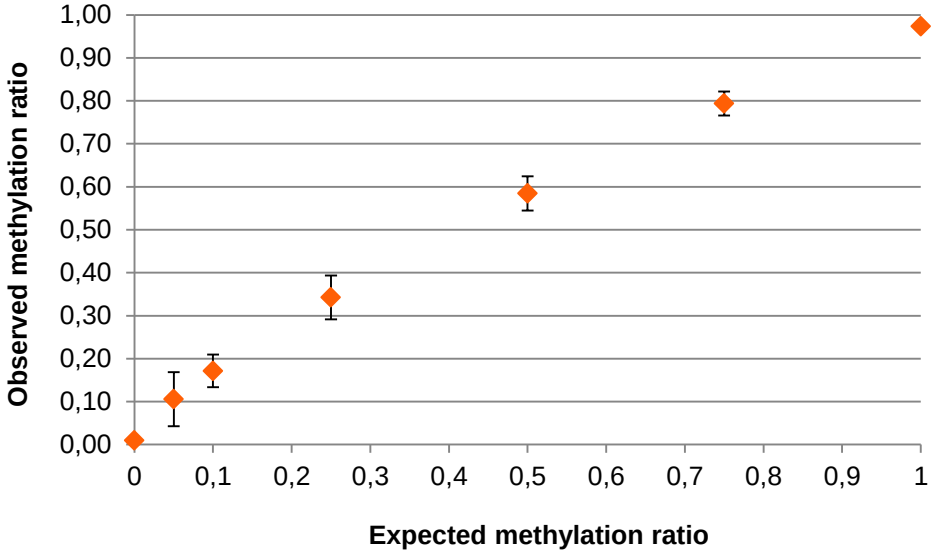
cg05442902



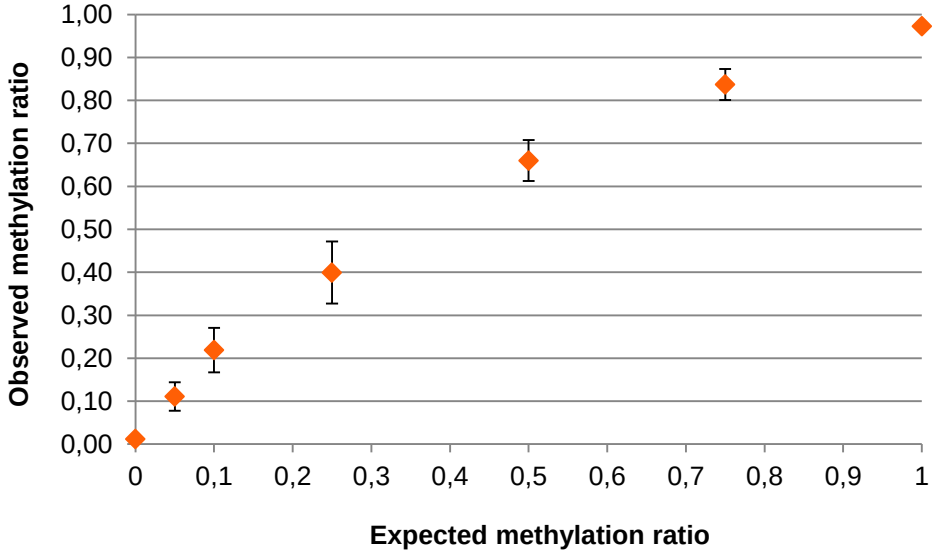
cg01511567



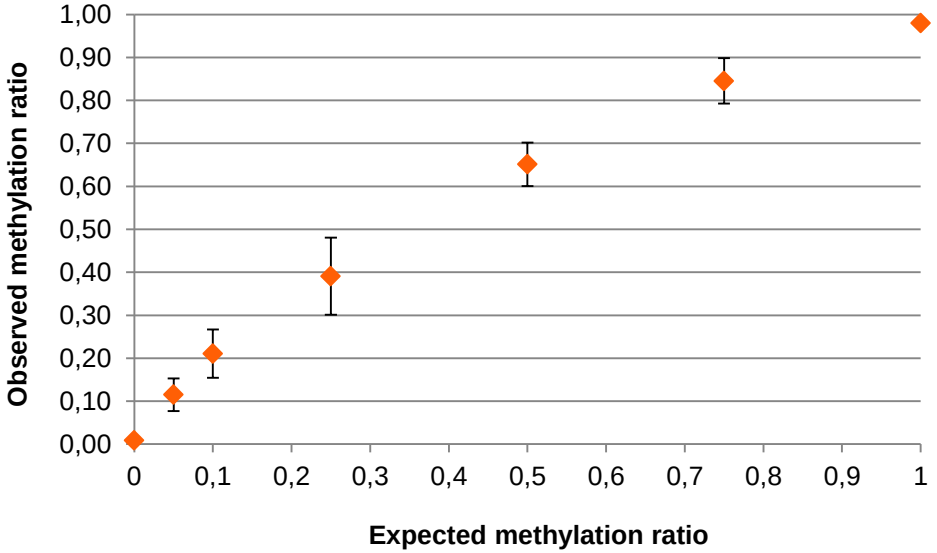
cg17274064



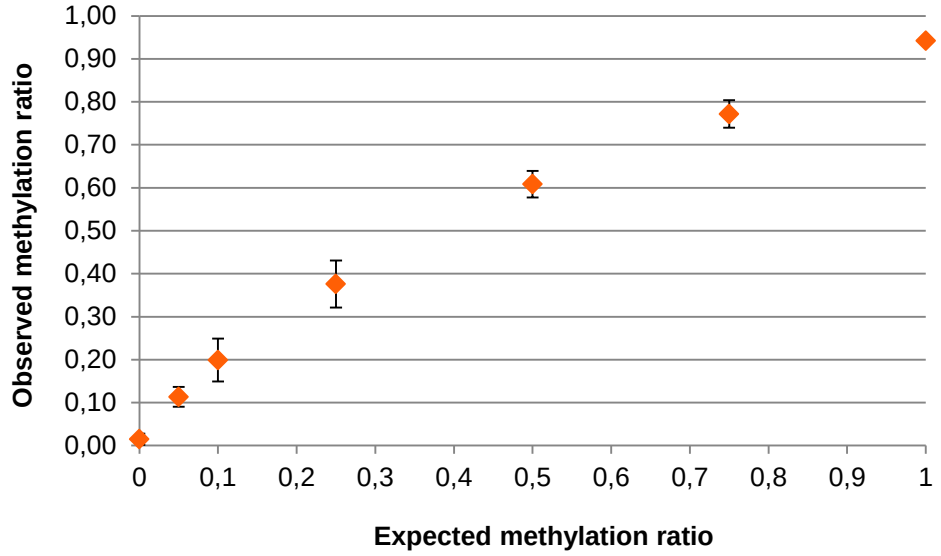
cg02085507



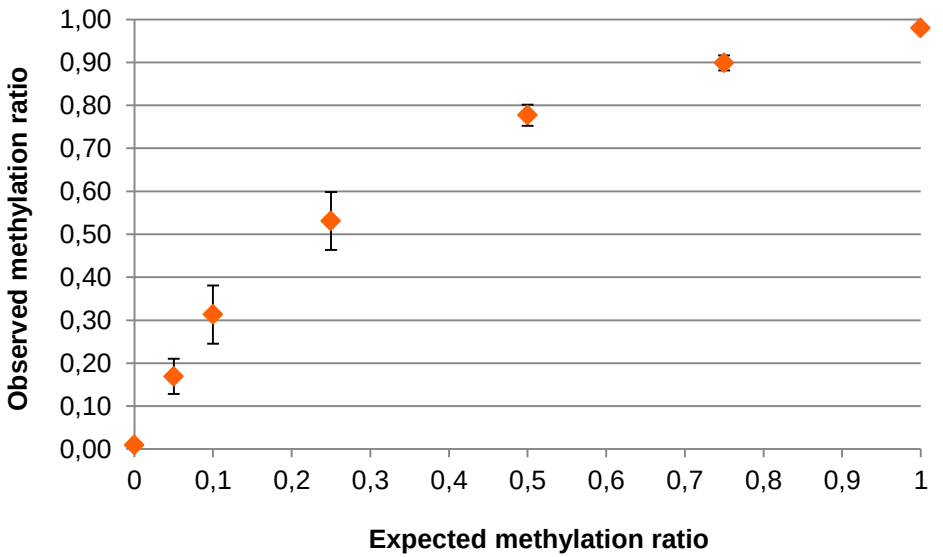
cg20692569



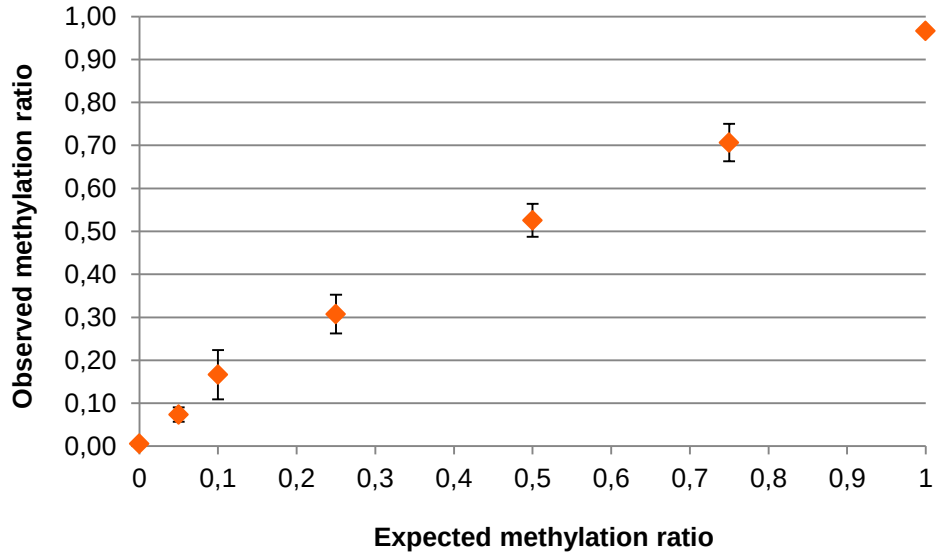
cg04528819



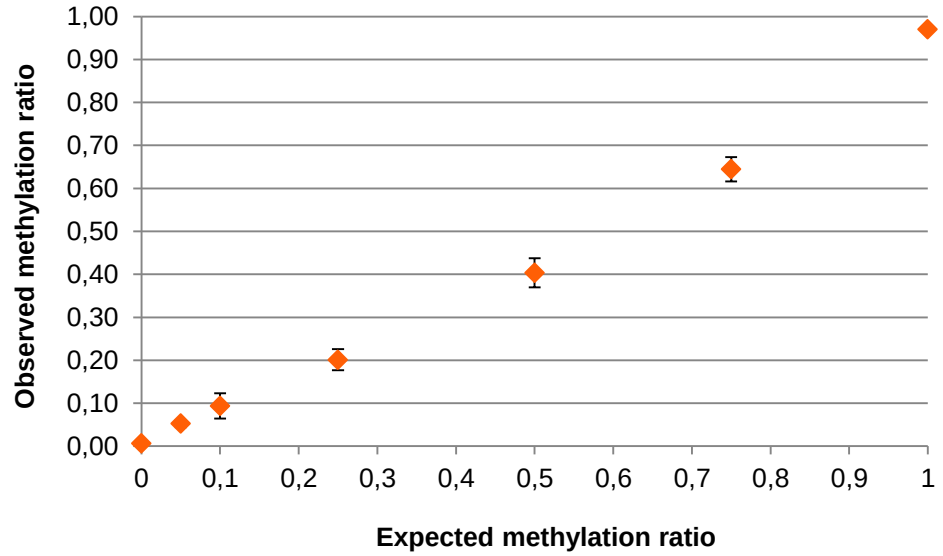
cg04084157



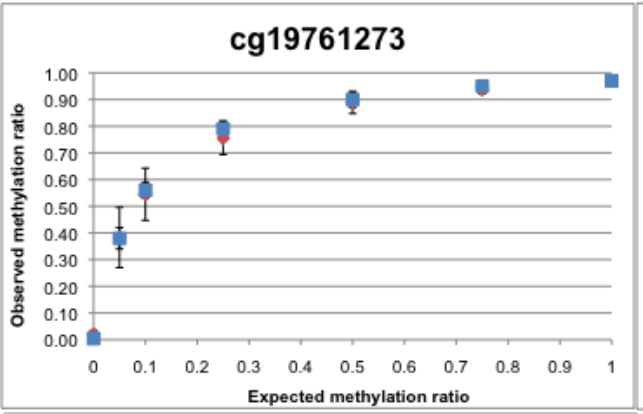
cg22736354

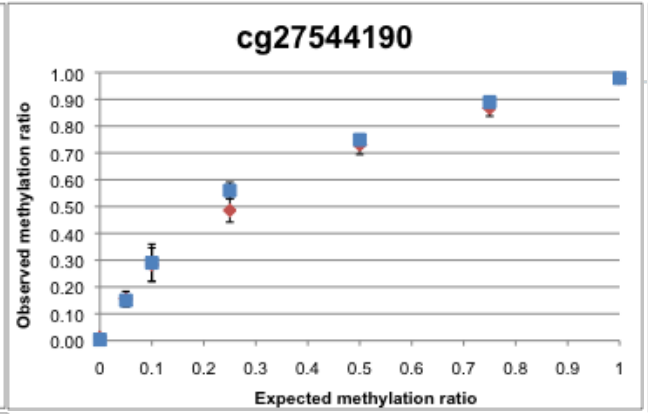


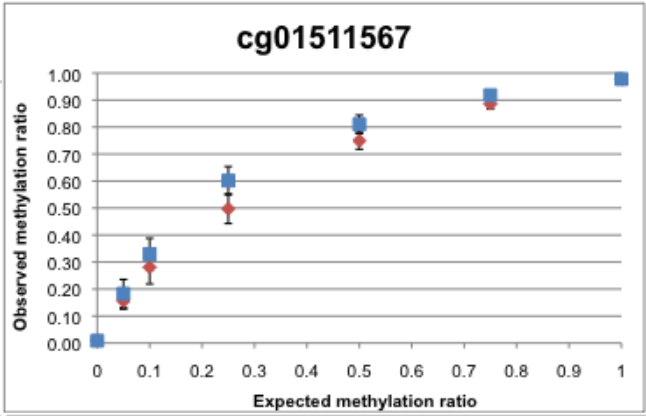
cg06493994

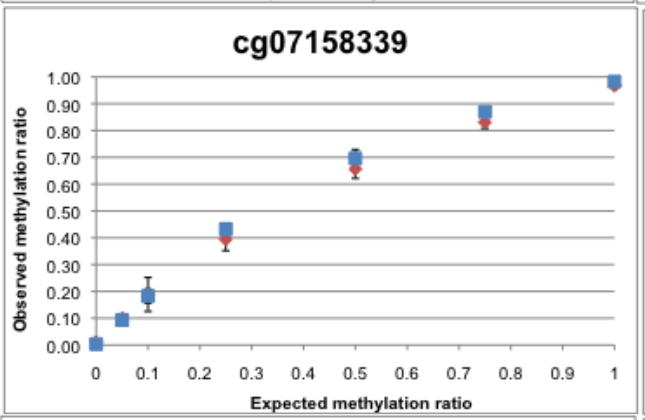


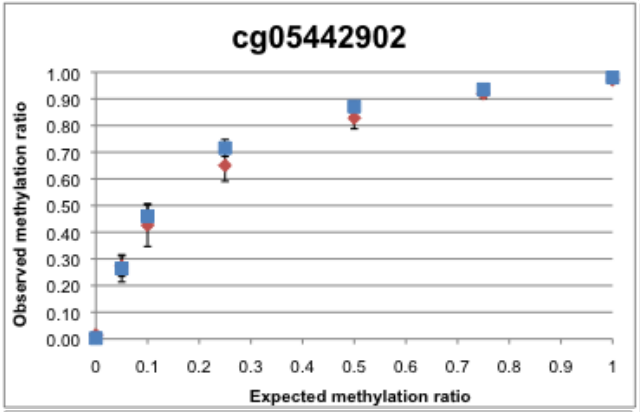
.. and PGM linearity

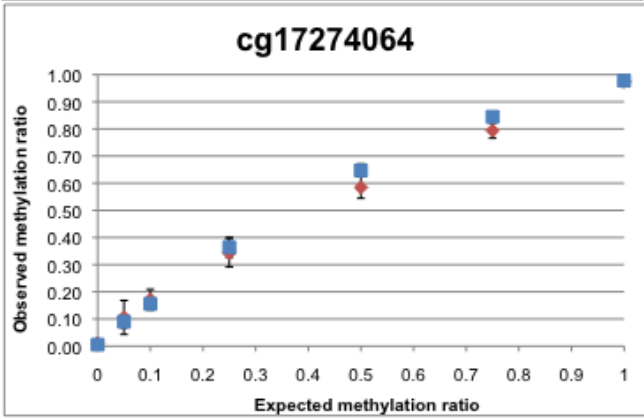


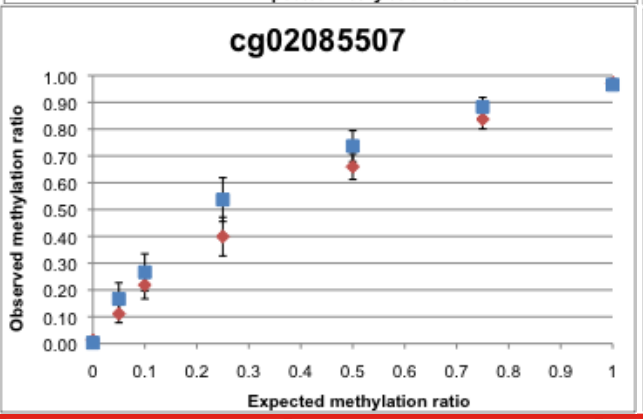


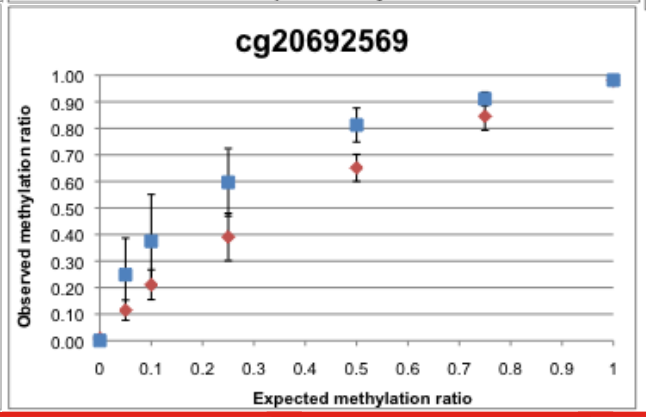


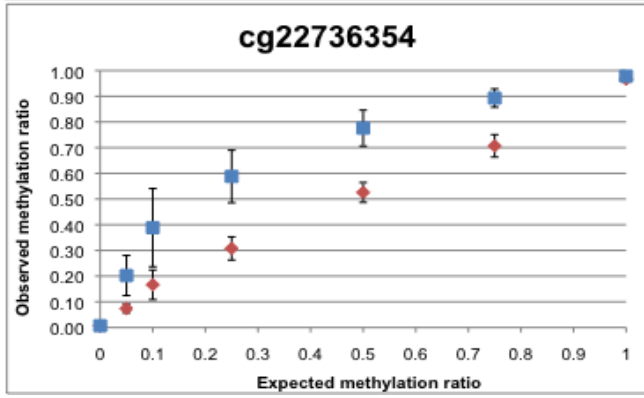












Summary & conclusions

- Generally, considering laboratories' experience with DNA methylation-based NGS, the method worked well
- Optimised MiSeq protocol performed better than PGM
- Various patterns were observed among labs in terms of marker distribution and total reads
- Methylation quantification showed great accuracy
- Mean of each marker for all samples: **SD=3.4%** (range 0-11%)
- As expected, low-methylation samples (5-25%) were the most challenging

Exercise part 2 - Proposal

Samples to be analysed:
5 -10 blood samples (to be analysed in duplicate)
Additional samples (optional)

KCL to provide:
Blood samples/stains
(PCR primers and protocols from part 1)

Participating labs to provide:
Reagents

Potential dates:
Samples to be sent out - Beginning of June 2016
Data to be collected - End of of September 2016
Presentation of results - Next meeting, November 2016

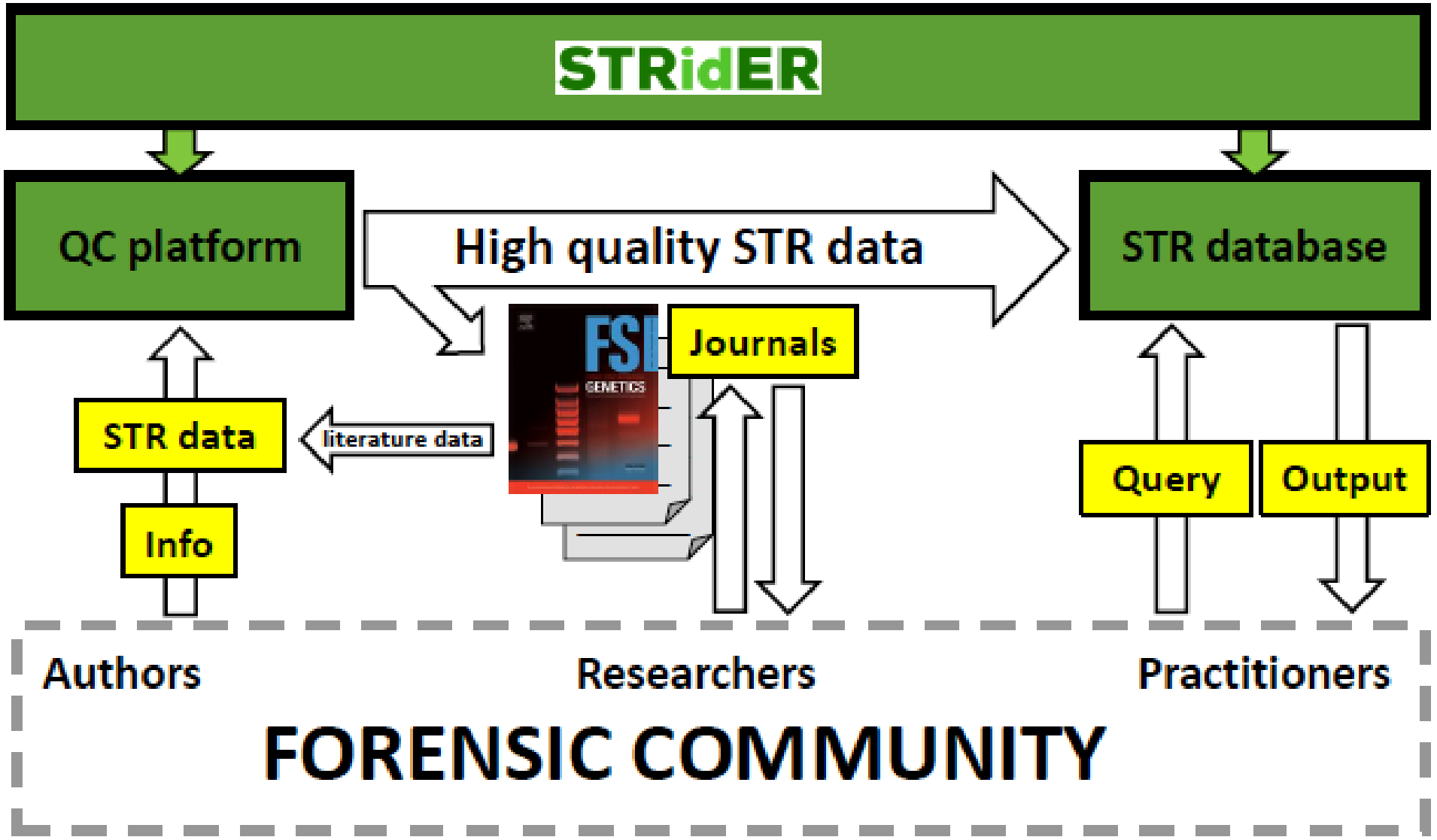
Q&A

Emails:
denise.syndercombe-court@kcl.ac.uk
athina.a.vidaki@kcl.ac.uk

EDNAP Meeting, Warsaw, Poland, April 26 2016

STRidER STRs for identity ENFSI Reference database

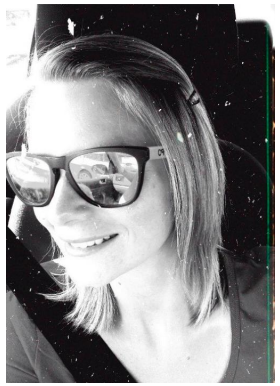
Dr. Walther Parson
assoc. Prof. Institute of Legal Medicine, Medical University of Innsbruck, Austria
adj. Prof. Forensic Science Program, Penn State University, PA, USA
walther.parson@i-med.ac.at



Bodner et al in prep

EDNAP Meeting, Warsaw, Poland, April 26 2016

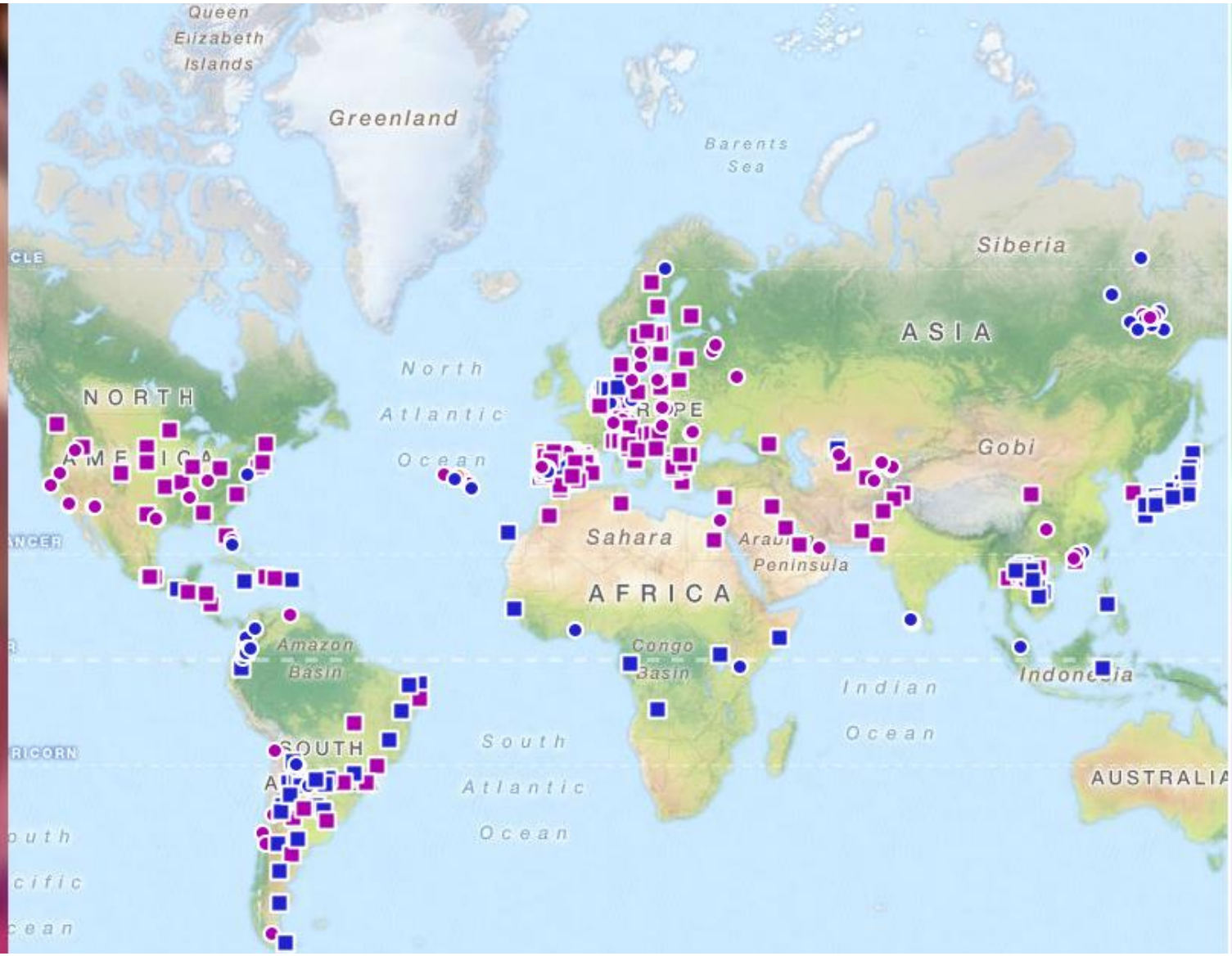
EMPOP mtDNA database



Dr. Walther Parson
assoc. Prof. Institute of Legal Medicine, Medical University of Innsbruck, Austria
adj. Prof. Forensic Science Program, Penn State University, PA, USA
walther.parson@i-med.ac.at

1. **Bodner** Helena, the hidden beauty: Resolving the most common West Eurasian mtDNA control region haplotype by massively parallel sequencing an Italian population sample. FSIG **15**: 21-26.
2. **Gomes** Human settlement history between Sunda and Sahul: a focus on East Timor (Timor-Leste) and the Pleistocenic mtDNA diversity. BMC Genomics **16**(1): 70.
3. **Just** Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. FSIG **18**: 131-139.
4. **Just** Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three U.S. populations. FSIG **14**: 141-155.
5. **Naue** Evidence for frequent and tissue-specific sequence heteroplasmy in human mitochondrial DNA. Mitochondrion **20**: 82-94.
6. **Parson** Molecular genetic analysis on the remains of the Dark Countess: Revisiting the French Royal family. FSIG **19**: 252-254.
7. **Parson** MPS of complete mitochondrial genomes from hair shaft samples. FSIG **15**: 8-15.
8. **Valente** Exploring the relationship between lifestyles, diets and genetic adaptations in humans. BMC Gen **16**: 55.
9. **Xavier** Admixture and genetic diversity distribution patterns of non-recombining lineages of Native American ancestry in Colombian populations. PLoS One **10**(3): e0120155.
10. **Chaitanya** (2016). High-quality mtDNA control region sequences from 680 individuals sampled across the Netherlands to establish a national forensic mtDNA reference database. FSIG **21**: 158-167.
11. **Scheible** (2016). The mitochondrial landscape of African Americans: An examination of more than 2500 control region haplotypes from 22 U.S. locations. FSIG **22**: 139-148.

Haplotype information



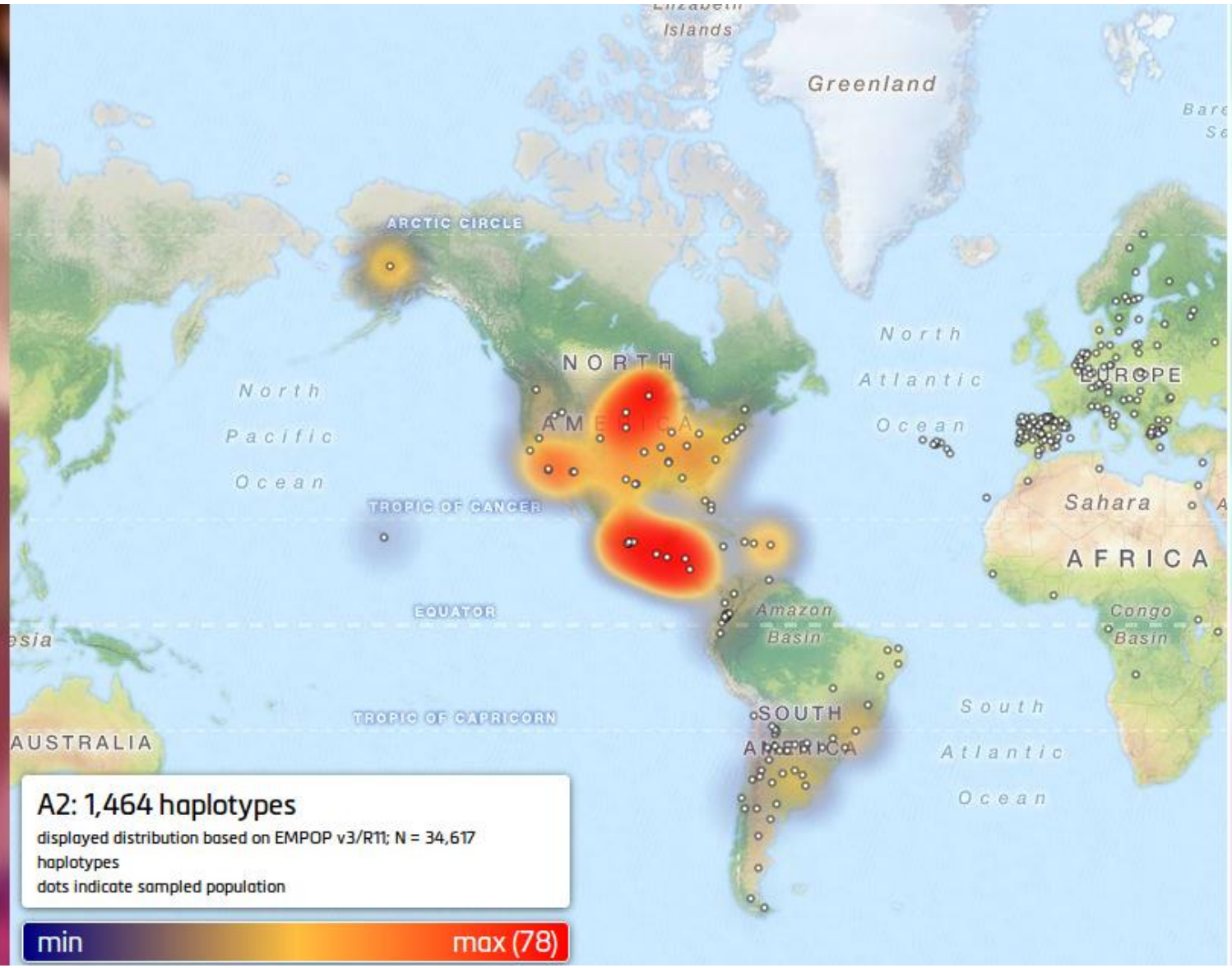
EMPOP holds high quality population data

The EMPOP database aims at the collection, quality control and searchable presentation of mtDNA haplotypes from all over the world.

The scientific concept and the quality control measures using logical and phylogenetic tools were found suitable for forensic purposes, e.g.

- by declaration of the German Supreme Court of Justice (2010)
- the SWGDAM mtDNA interpretation guidelines (2013)
- and the updated ISFG guidelines for mtDNA analysis (2014)

Haplogroup information



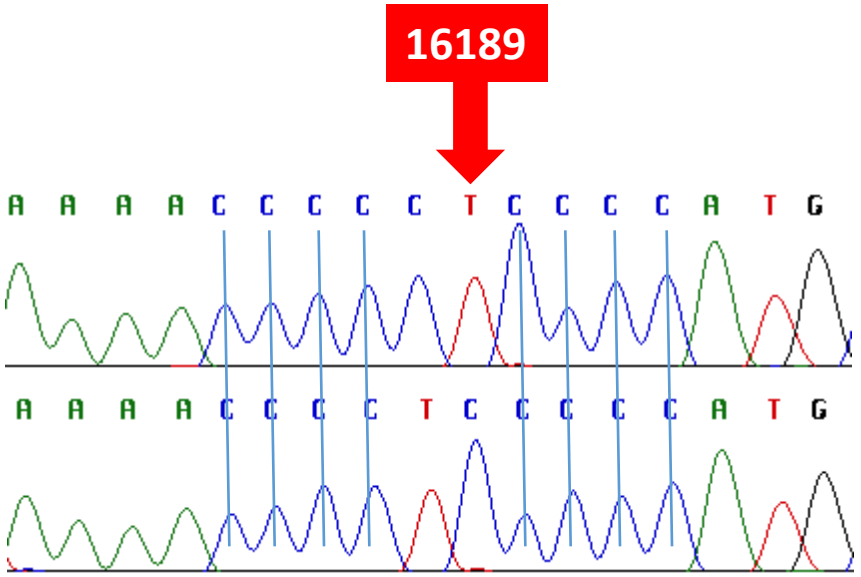
EMPOP generates conservative haplogroup estimates

EMPOP provides automated haplogroup estimates. These are based on maximum likelihood and minimal costs functions. For multiple possible haplogroups most recent common ancestor (MRCA) haplogroups are provided.

The geographical haplogroup patterns are provided via maps to visualize and better understand their distribution.

New developments - "SAM21"

Sequence alignment can be ambiguous



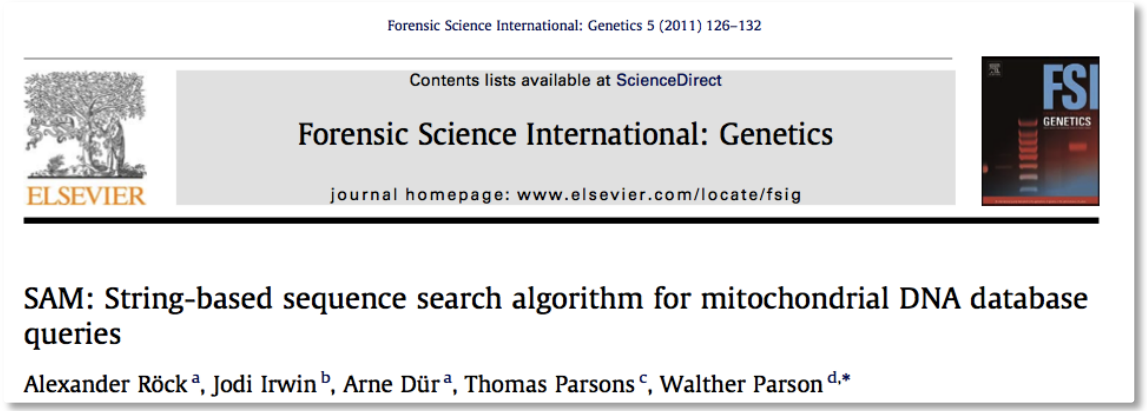
Alignment 1	=	Alignment 2
16188T 16189C		16188- 16193+C

… and many more alignments …

Effect of alignment on database searches

Search method	Alignment 1	Alignment 2
rCRS-coded	28 matches	0 matches

EMPOP V3 R11; N = 34617


Forensic Science International: Genetics 5 (2011) 126–132

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

SAM: String-based sequence search algorithm for mitochondrial DNA database queries

Alexander Röck [a], Jodi Irwin [b], Arne Dür [a], Thomas Parsons [c], Walther Parson [d,*]

Search method	Alignment 1	Alignment 2
String alignment	28 matches	28 matches

EMPOP V3 R11; N = 34617

New developments - "SAM21"

The **S**tring **A**lignment **M**ethod (SAM) guarantees that sequences are found in EMPOP regardless of the alignment

Bandelt and Parson (2008) Consistent treatment of length variants in the human mtDNA control region: a reappraisal, Int J Legal Med 122:1-21

Rule 1. Phylogenetic rule

Rule 2. Anchor 16189 and 310

Rule 3. 3' alignment

2013 Adopted by SWGDAM

2014 Recommended by ISFG

Practical consequences of SAM

Practitioners are free to choose mtDNA alignment and notation of mtDNA haplotypes as EMPOP turns haplotypes into FASTA-like strings and performs the search in **unaligned format**

Reporting is disentangled from **database** searches

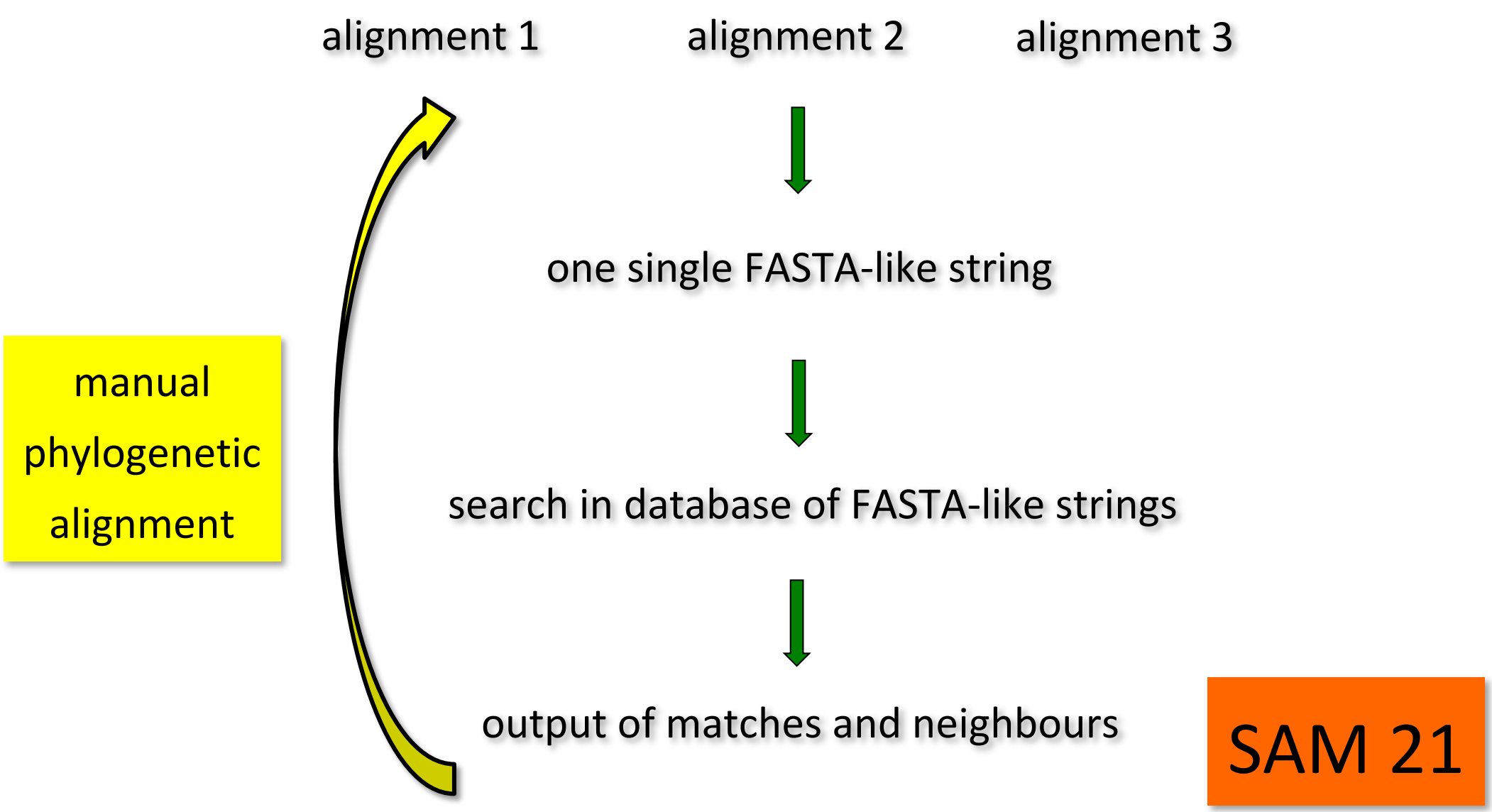
Elaborating the phylogenetic alignment remains an academic task - cognizance of mtDNA phylogeny

BUT

Some labs require consistent alignment for consistent reporting of mtDNA sequences

We are developing a new version of SAM that turns FASTA strings back in phylogenetic alignment

New developments - “SAM21”

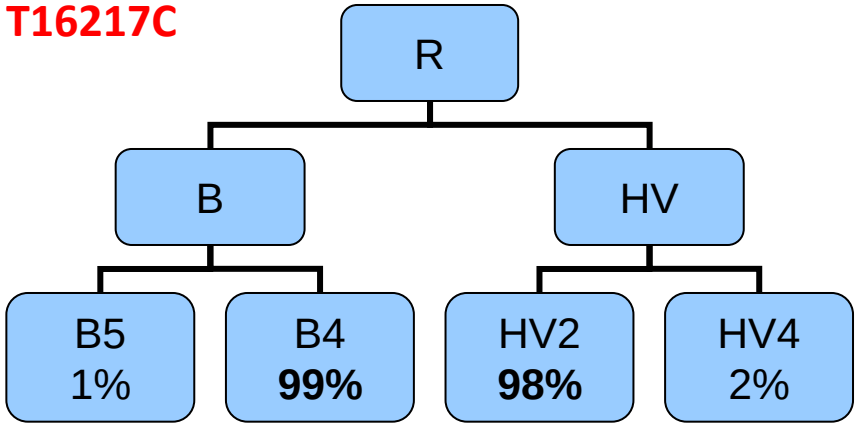


New developments - “SAM21”

SAM21 eats **single** or **batches** of FASTA-like sequence strings and turns these into **rCRS-coded haplotypes** following the **phylogenetic** alignment

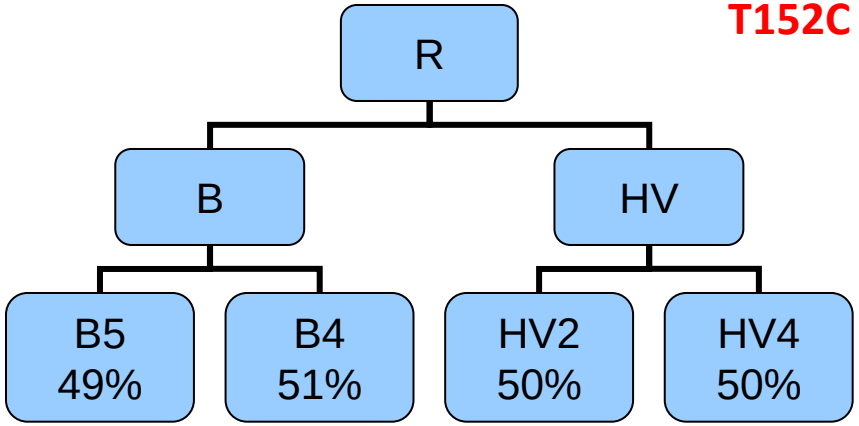
Requirement:
Database of high-quality mtDNA sequences - EMPOP
Weighting scheme for “mutations” - fluctuation rates (Röck et al FSIG 2013)

T16217C



T16217C is a stable marker in hgs B4 and HV2 and therefore a strong signature for hg-estimation

T152C



T152C is strongly fluctuating in all 4 hgs and therefore of little relevance for hg-estimation

New developments - “SAM21”

The new phylogenetic alignment software will be made available via EMPOP

Work flow:
Haplotypes or FASTA-like sequences queried in EMPOP
EMPOP provides database search result
EMPOP provides phylogenetic alignment of haplotype

Advantages:
EMPOP Database releases no more curated by hand but by software
Even large datasets of different background (medical/population genetics) can be automatically aligned into standardized format and directly compared with each other

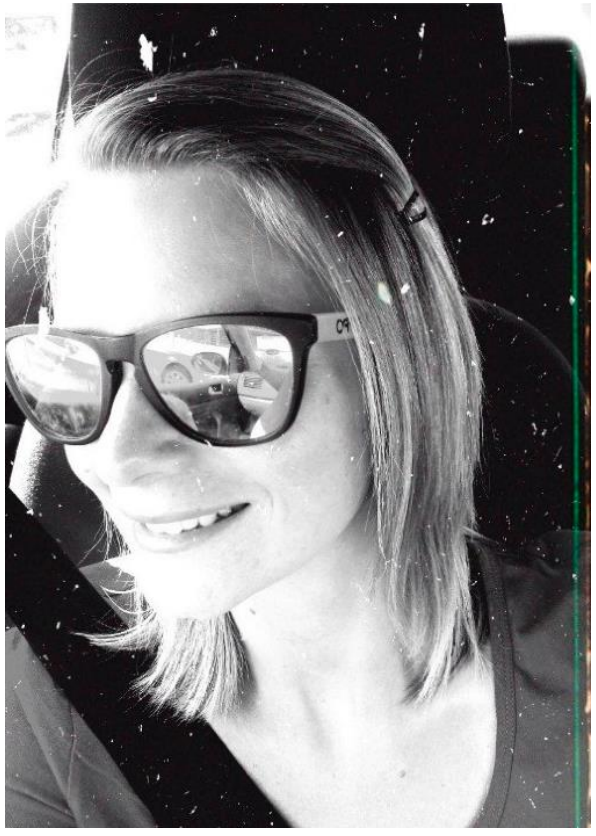
Caveats:
Not all conventions currently used in forensic genetics can be maintained
The new software will come with updated recommendations for mtDNA notation

New developments - “SAM21”

The new software will “soon” be available for testing.

If you are interested in testing and providing feedback, please come see me or write to walther.parson@gmail.com

acknowledgements: Nicole Huber, Arne Dür

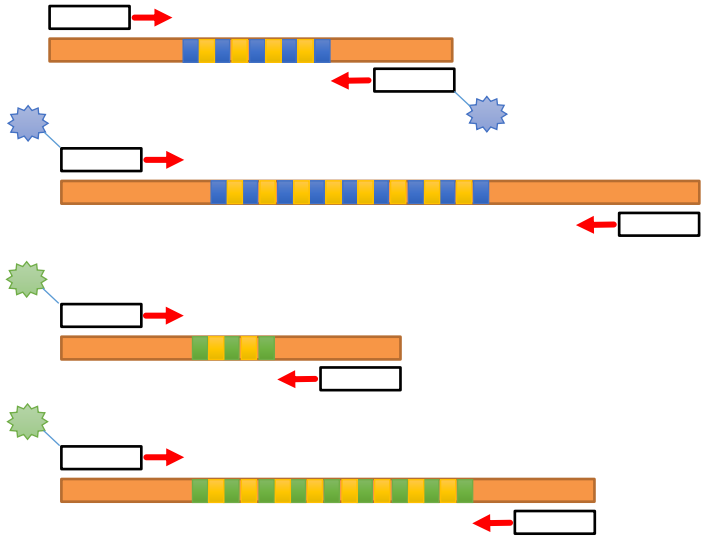


Next Generation Sequencing of STRs

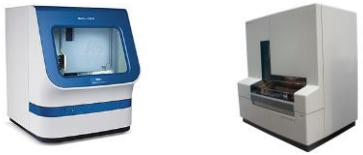
Dr. Walther Parson
assoc. Prof. Institute of Legal Medicine, Medical University of Innsbruck, Austria
adj. Prof. Forensic Science Program, Penn State University, PA, USA
walther.parson@i-med.ac.at

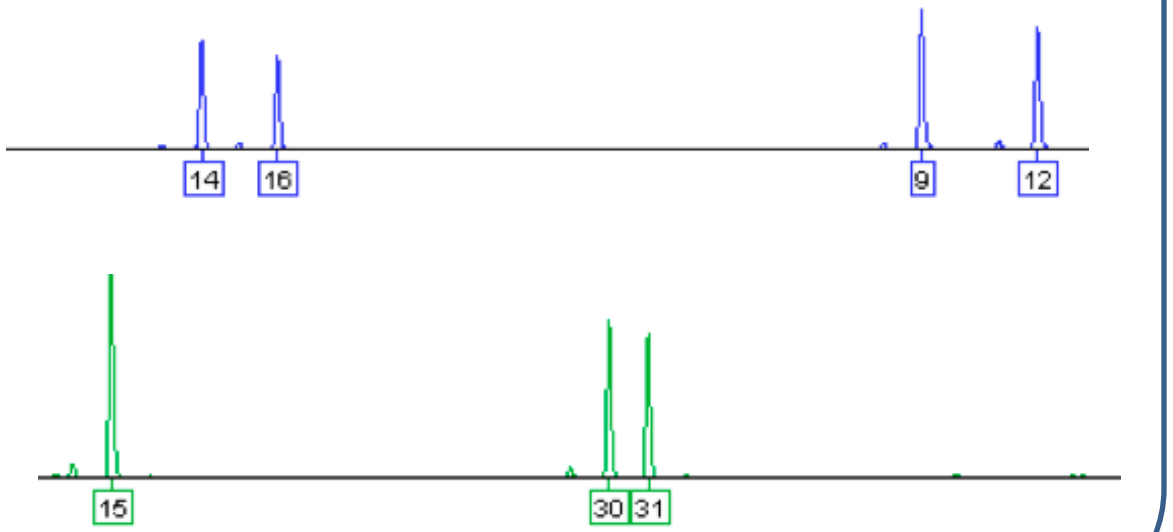
STR Analysis by Electrophoresis

PCR



CE

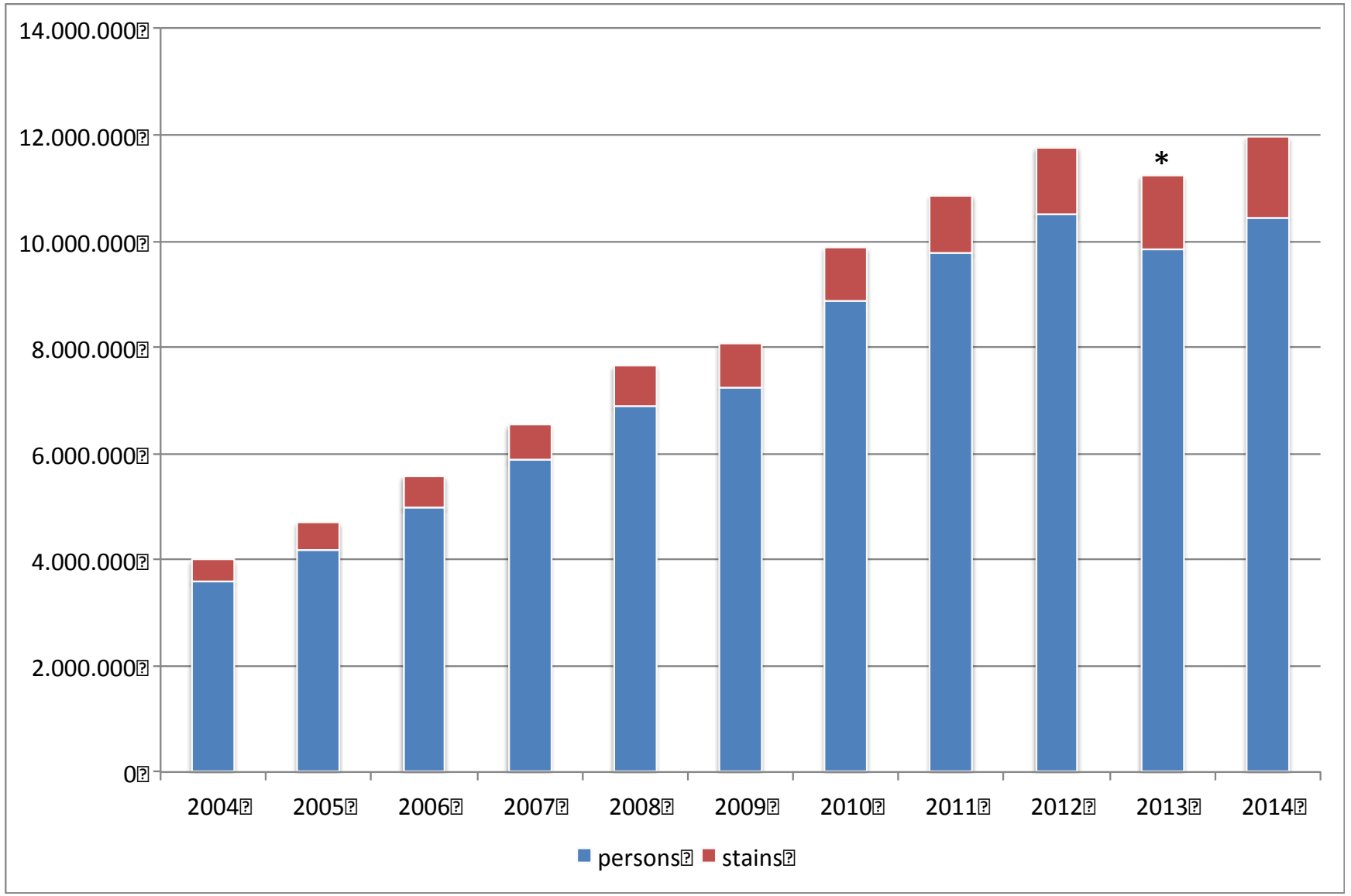




Electropherogram

14/16 9/12 15/15 30/31

National STR databases (ENFSI DNA WG)



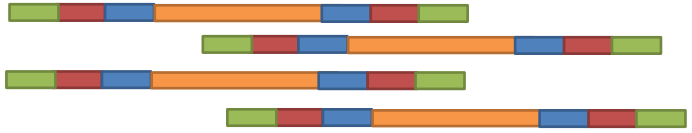
* incomplete data

STR Analysis by Next Generation (Massively Parallel) Sequencing

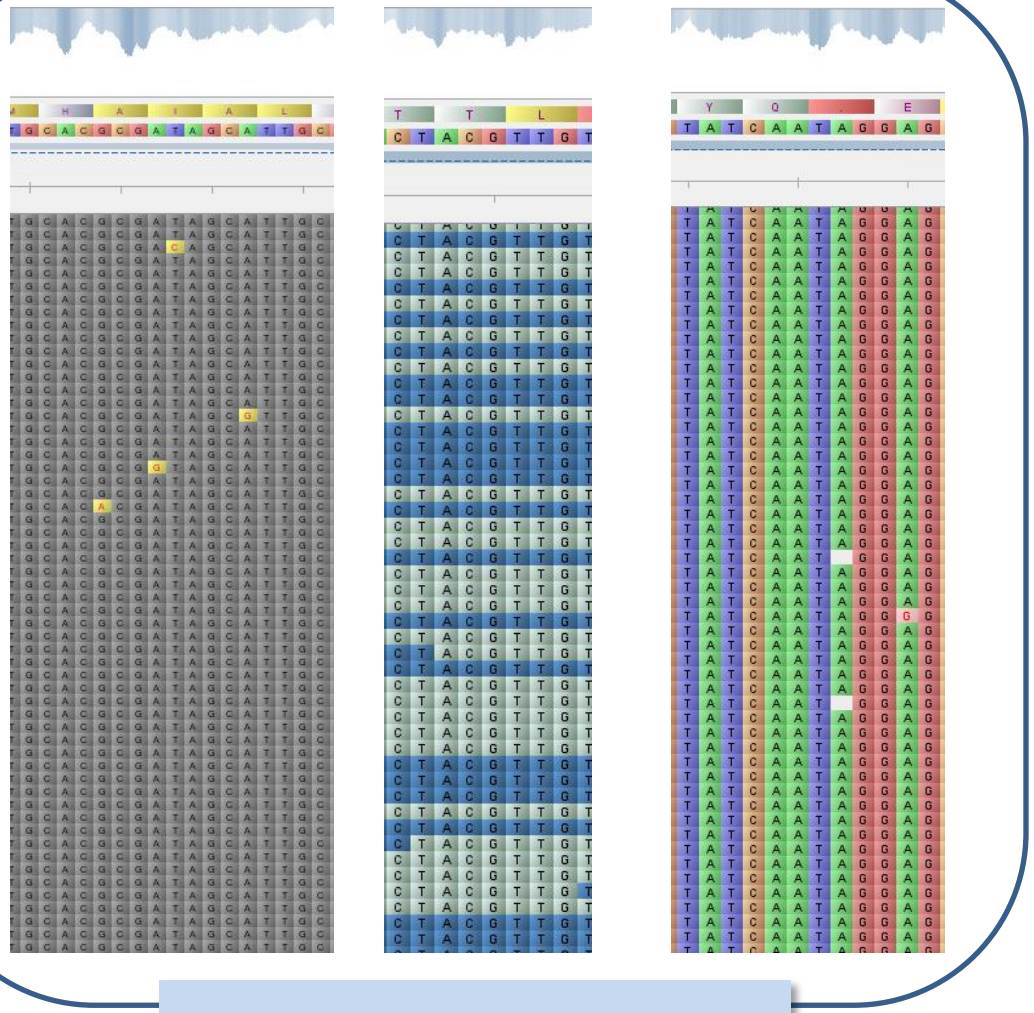
PCR

Enrich

Library

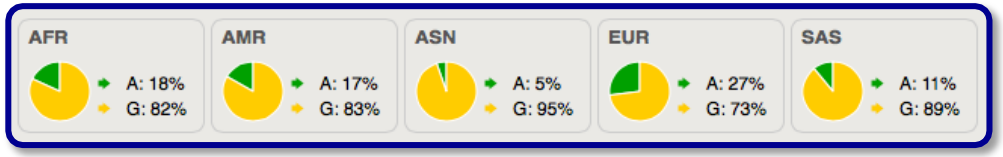


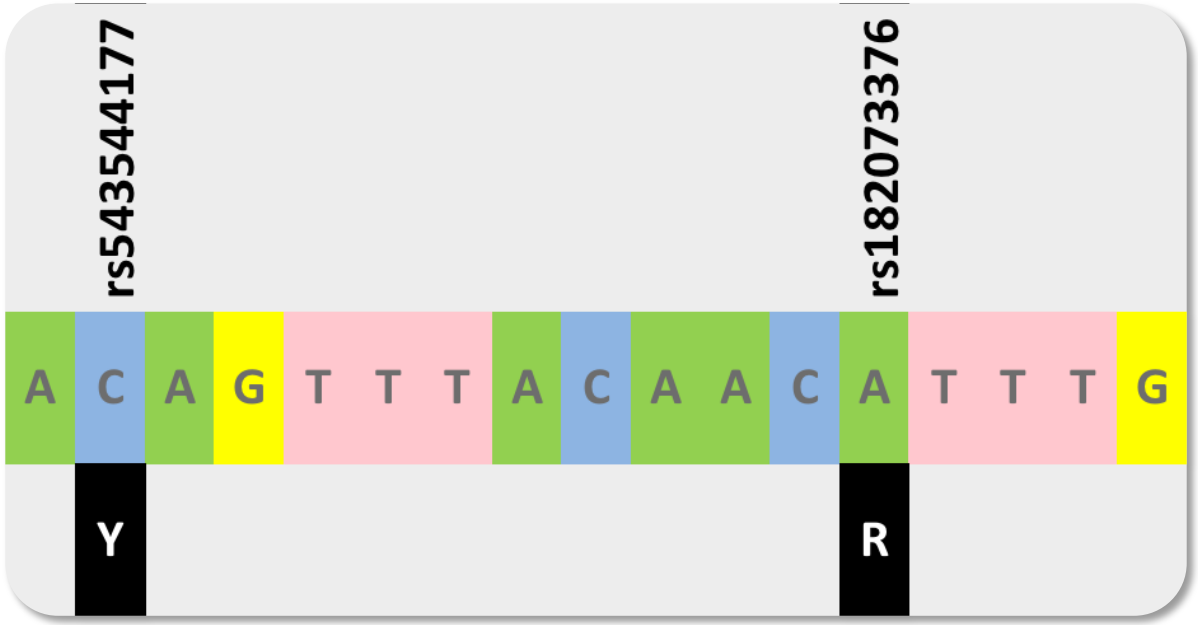
MPS

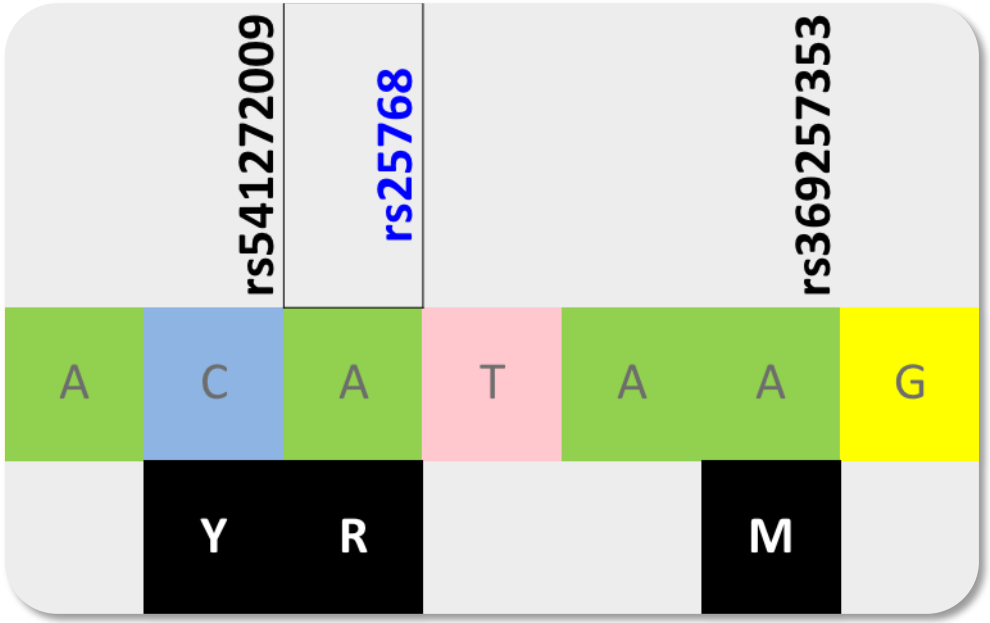


Sequence

STR Analysis by Next Generation (Massively Parallel) Sequencing







Parson et al (2016) ***FSIG***

NGS of STRs: Considerations of the ISFG

Forensic Science International: Genetics 22 (2016) 54–63

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements

Walther Parson[a,b,*], David Ballard[c], Bruce Budowle[d,e], John M. Butler[f], Katherine B. Gettings[f], Peter Gill[g,h], Leonor Gusmão[i,j,k], Douglas R. Hares[l], Jodi A. Irwin[l], Jonathan L. King[d], Peter de Knijff[m], Niels Morling[n], Mechthild Prinz[o], Peter M. Schneider[p], Christophe Van Neste[q], Sascha Willuweit[r], Christopher Phillips[s]

NGS of STRs: Considerations of the ISFG

The executive board of the ISFG introduced a DNA commission to evaluate **initial considerations** regarding STR nomenclature.

The primary goal is to define **minimum criteria** for data analyses and database storage.

Ultimately, this should facilitate compatibility between MPS STR data generated currently and the data that will inevitably follow with wider adoption, while ensuring **backward** and **parallel compatibility** to CE-based STR typing in national DNA databases as well as published population data.

At present, it can be expected that both CE- and MPS-based STR typing methods will continue to **coexist**. Their application to casework will depend on resources, ease of use, speed of analysis, the value of the increased resolution power, and each technique's relevance to complex and challenging cases.

NGS of STRs: Considerations of the ISFG

The adoption of sequenced STR alleles in practical forensic work requires considerations at **three hierarchical levels**:

the full sequence (sequence string),
the alignment of sequences relative to a reference sequence and
the annotation of alleles.

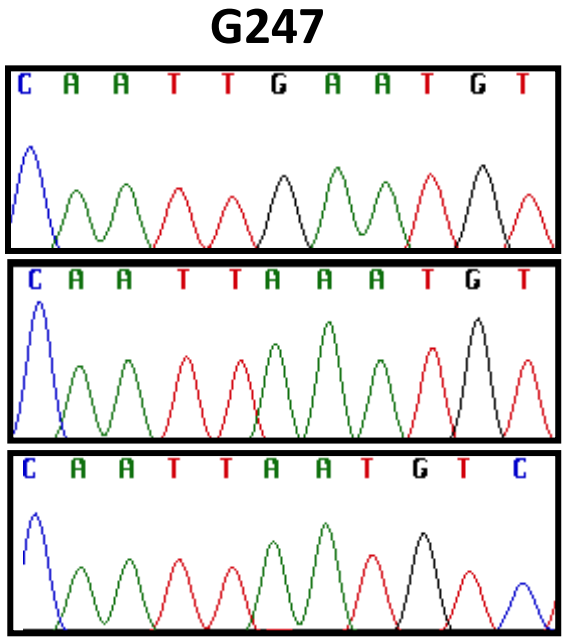
A set of **8 practical considerations** on NGS of STRs (see also appendix of this presentation)

Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: The Full Sequence - Nucleotide String

Consideration 1:
MPS analysis should be performed with software that allows STR sequences to be exported and stored in databases as **sequence (text) strings** to capture the **maximum consensus sequence information**.

<u>Lesson learnt from mtDNA:</u>
Database searches of reference-coded haplotypes/genotypes should be performed in alignment-free format to guarantee that haplotypes/genotypes are not missed due to different alignment/notation.



rCRS

247A — **hg L1'-6'**

247A 249DEL
= 247DEL — **hgs L1c1c, L5b1, L3x2**

'247A 249DEL' and '247DEL' denote the same sequence, but when queried in rCRS-coded format, they would lead to different search results.

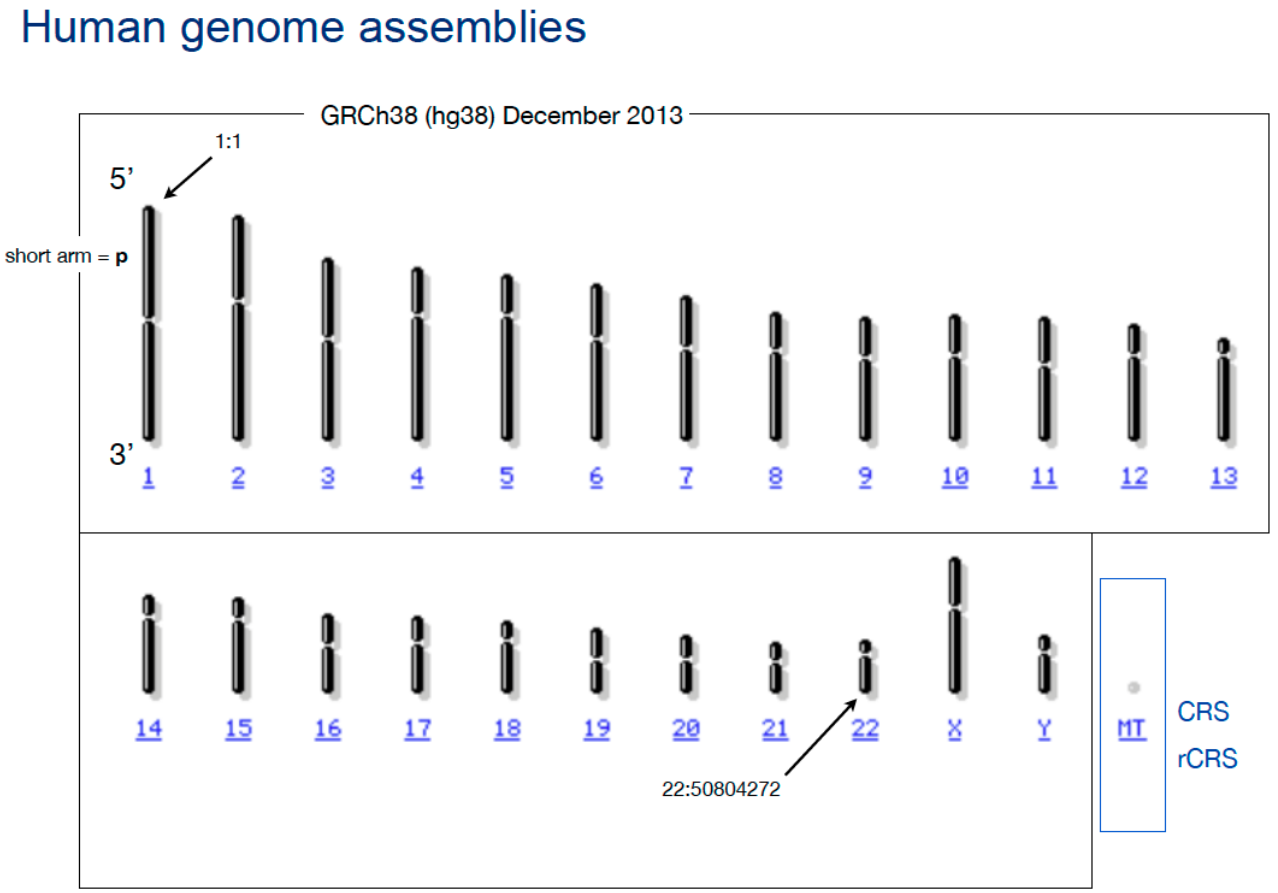
Translation of the haplotypes into nucleotide strings solves this problem (Röck et al (2011) ***FSIG***)

Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Sequence Direction and Genome Assembly

Consideration 2:
The **forward strand direction** assigned in the human genome has been constant for all assemblies published since the first draft in 2001 and can be used to align STR sequences.

Consideration 3:
At the time of writing, **GRCh38** is the most up to date sequence assembly and is recommended as the framework with which to define repeat region structure for sequence alignment and for the mapping of sequence features such as SNPs.



Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Consequences of Harmonizing Alignment

STR	Chr.	Human Reference Genome Assembly GRCh38					Potential Frameshift Exists
		Location of Repeat Region Start	Location of Repeat Region Stop	Repeat No.	Past Repeat Region Sequence Summary	Future Repeat Region Sequence Summary	
D1S1656	1	230769616	230769683	17	[TAGA]16 [TAGG] **[TG]5**	**[CA]5** [CCTA] [TCTA]16	*
D2S1338	2	218014859	218014950	23	[TGCC]7 [TTCC]13 [GTCC] [TTCC]2	[GGAA]2 [GGAC] [GGAA]13 [GGCA]7	
FGA	4	154587736	154587823	22	[TTTC]3 [TTTT] [TTCT] [CTTT]14 [CTCC] [TTCC]2	[GGAA]2 [GGAG] [AAAG]14[AGAA] [AAAA] [GAAA]3	*
D5S818	5	123775556	123775599	11	[AGAT]11	[ATCT]11	*
CSF1PO	5	150076324	150076375	13	[AGAT]13	[ATCT]13	*
D6S1043	6	91740225	91740272	12	[AGAT]12	[ATCT]12	*
D7S820	7	84160226	84160277	13	[GATA]13	[TATC]13	
VWA	12	5983977	5984044	17	[TCTA] [TCTG]5 [TCTA]11 **TCCA TCTA**	**TAGA TGGA** [TAGA]11 [CAGA]5 [TAGA]	*
Penta E	15	96831015	96831039	5	[AAAGA]5	[TCTTT]5	*
D19S433	19	29926235	29926298	16	[AAGG] **AAAG** [AAGG] **TAGG** [AAGG]12	[CCTT]12 **CCTA** [CCTT] **CTTT** [CCTT]	*
DYS19	Y	9684380	9684443	15	[TAGA]3 **TAGG** [TAGA]12	[TCTA]12 **CCTA** [TCTA]3	*
DYS635	Y	12258860	12258951	23	[TCTA]4 [TGTA]2 [TCTA]2 [TGTA]2 [TCTA]2 [TGTA]2 [TCTA]9	[TAGA]9 [TACA]2 [TAGA]2 [TACA]2 [TAGA]2 [TACA]2 [TAGA]4	*
DYS389I	Y	12500448	12500495	12	[TCTG]3 [TCTA]9	[TAGA]9 [CAGA]3	*
DYS389II	Y	12500448	12500611	29	[TCTG]5 [TCTA]12 **48 nt.** [TCTG]3 [TCTA]9	[TAGA]9 [CAGA]3 **48 nt.** [TAGA]12 [CAGA]5	*
DYS390	Y	15163067	15163162	24	**[TCTA]2** [TCTG]8 [TCTA]11 TCTG [TCTA]4	[TAGA]4 CAGA [TAGA]11 [CAGA]8 **[TAGA]2**	*
Y-GATA-H4	Y	16631673	16631720	12	[TAGA]12	[TCTA]12	
DYS385ab	Y	18639713 18680632	18639756 18680687	11 14	[GAAA]11 [GAAA]14	[TTTC]11 [GAAA]14	*
DYS460	Y	18888810	18888849	10	[GATA]10	[TATC]10	*
DYS392	Y	20471987	20472025	13	[TAT]13	[ATA]13	*
DYF387S1ab	Y	23785361 25884581	23785500 25884724	35 36	[AAAG]3 GTAG [GAAG]4 [AAAG]2 GAAG [AAAG]2 [GAAG]9 [AAAG]13 [AAAG]3 GTAG [GAAG]4 [AAAG]2 GAAG [AAAG]2 [GAAG]10 [AAAG]13	[AAAG]3 GTAG [GAAG]4 [AAAG]2 GAAG [AAAG]2 [GAAG]9 [AAAG]13 [CTTT]13 [CTTC]10 [CTTT]2 CTTC [CTTT]2 [CTTC]4 CTAC [CTTT]3	 *

Parson et al (2016) ***FSIG***

D19S433 has been reported on the reverse strand [AAGG] with two uncounted repeats (bold)
Reverse complement results in [CCTT]
Adjusted reverse complement (first possible repeat motive) results in [TCCT] and a 1 bp shift

D19 rev. TGTTG AAGG **AAAG** AAGG **TAGG** AAGG AAGG AAGG AAGG AAGG AAGG AGAGA
D19 for. TCTCT CCTT CCTT CCTT CCTT CCTT CCTT **CCTA** CCTT **CTTT** CCTT CAACA
D19 f.adj. TCTC TCCT TCCT TCCT TCCT TCCT TCCT TCCT **ACCT TCTT** TCCT TCAACA

Potential complications when comparing to earlier sequence data

Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Consequences of Harmonizing Alignment

DYS389 I/II has been reported on the reverse strand

Adjusted reverse complement (first possible repeat motive) results in a shift of 1 repeat unit

=> allele appears one repeat larger

Reverse strand: $[TCTG]_5\ [TCTA]_{12}$ 48 nt. $[TCTG]_3\ [TCTA]_9$

Forward strand: $[TAGA]_9\ [CAGA]_3$ 48 nt. $[TAGA]_{12}\ [CAGA]_5$

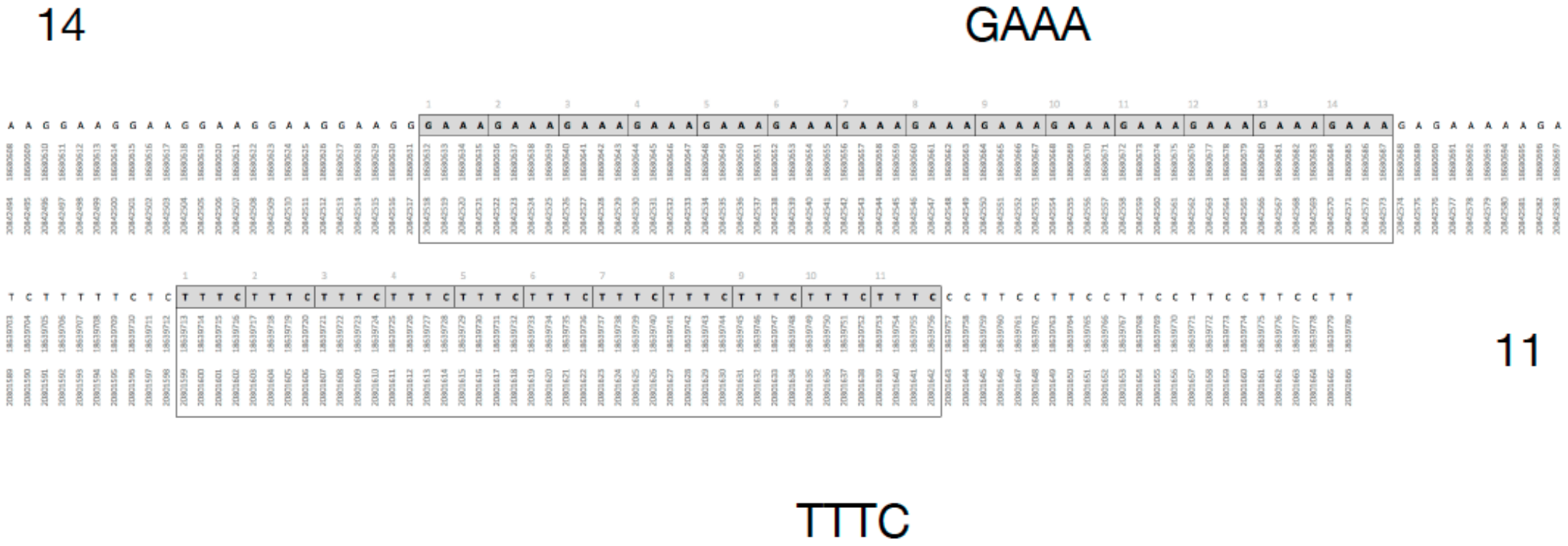
Forward strand, adj.: $[GATA]_9\ [GACA]_3$ 48 nt. $[GATA]_{12}\ [GACA]_6$

MPS data need to be interpreted by software, not manually, to avoid misunderstanding

Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Consequences of Harmonizing Alignment

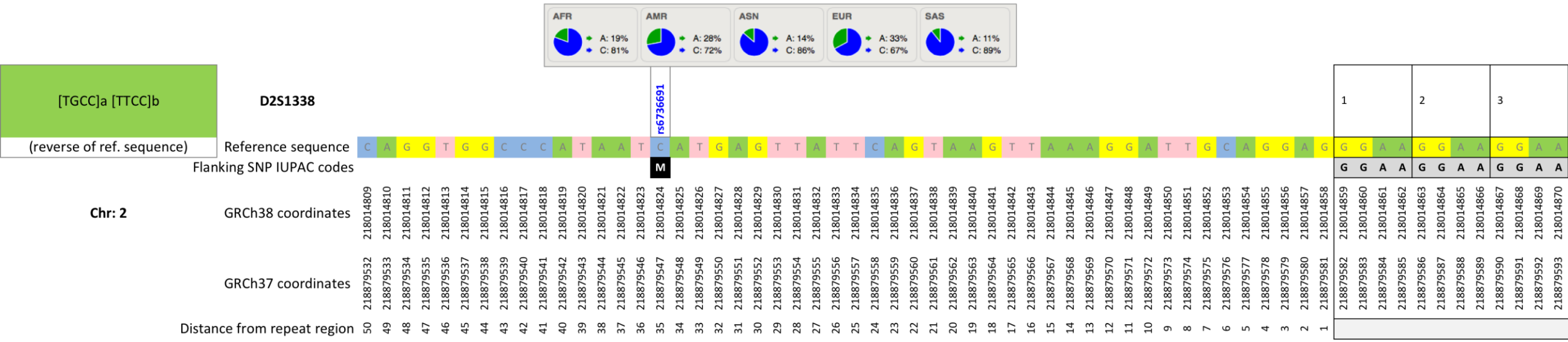
DYS385a/b includes two inversed regions, [GAAA] (rev) and [TTTC] (for)
Using the forward strand, it is not possible to summarize DYS385 a/b repeats by a uniform motif description as was reported in the past.



Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Alignment Examples

Introduction into **alignment of sequences** to a reference including supplementary files for orientation
Supplementary files include the alignment of 36 autosomal, 29 Y- and 7 X-chromosomal STRs including the special case of D5S2500 that is two different STRs located next to each other (1688 bp)



Parson et al (2016) ***FSIG***

The case(s) of D5S2500

Locus details		Human genome reference sequence (5' to 3'): repeat region (centre bold) +/- 200 nucleotides of flanking sequence
NCBI GenBank accession number	G08468 (UniSTS ID: 76230)	AGACCAGCCTGGGCAACATAGAGATATCCTGTCTCTACA AAAATTTTAAAAATTAGCTGGACATGGTGGTGCACACCTG TAGACCTGCACA*CCTGTAGATCGCTGGAGCC*CAAACGTT CAAGGTTACAGTGACCTATGGTCATGCCACTGCACTCCA GCCTGGGCAACACAGACTCTGTTTCTAATACATATATTAG AGA**CTATCTATCTATCTATCTATCTATCTATCTATCT** **ATCTA**TTTGATCCCATGGGGGAGATCACTCCTTTAATAA TGCAAGAGAGTGAATAAGATGGGGGAGGGTGGCGTGCA ATTGAGAGATGAGCTTTATACTGGAATACCAAGTTTATGG TGTGGAACTCTGGGGTATTACACTGAAATGTTAGGGGTT CTCAGTGGGTTATTTGT
Temporary Name	D5S2500.G08468	
Synonyms	Marshfield: GATA67D03; Whitehead-YAC: CHLC.GATA67D03; rs111362704	
Forensic Multiplexes	Qiagen HDplex / Mentype® Chimera® kits	
9947A control DNA genotype	15,16	
Repeat motif in reference sequence	$[CTAT]_{11}$	
GRCh38 coordinates of sequence shown	5:59401244-59401655	
GRCh37 coordinates of sequence shown	5:58697070-58697481	
NCBI GenBank accession number	AC008791*	TTGT*ATCATCCCTGCAAAGTAACG*TTTACTGATAAACCAAA TGATGTGCCATAATTATGTTTTATTTATGGAACAACTTTTTG TTTTTTCTGGAGTTATATATTACCTTCTTTATTTGATTATGT GACATTATCACCAATTTTTCTAGACGTCTCCAAAACATAAT TTTTTAAATTTTAAATTTCTGTTGGTACATAATAG**GTAGGTA** **GGTAGACAGACAGACAGACAGACAGACAGACAGACAGA** **TAGATAGATAGATTGATTGATT**TATGGGGCCCACGAGATA TTTTCATACAGGCAAGCAATGCATAATAATATCAGGGTAAA CAGGGAATCTATTACTGCAAACAT*TTACCCTTTAAGTTACA* *AACAATTCAA*TTATATTCTCCAAACAATTCAATTATATTCTC TTAGTTATTTTGAGATGTATGATAAACTACTGTTGACTGTA GTCACCCTGTTGAGCTATCAAATACCAGATCTTATTTGT
Temporary Name	D5S2500.AC008791	
Synonyms	None identified	
Forensic Multiplexes	NIST miniSTR 26plex / AGCU ScienTech 21-plex	
9947A control DNA genotype	14,23	
Repeat motifs in reference sequence	$[GGTA]_3$ $[GACA]_8$ $[GATA]_3$ $[GATT]_3$	
GRCh38 coordinates of sequence shown	5:59402932-59403399	
GRCh37 coordinates of sequence shown	5:58698758-58699225	

Philips et al (2016) ***in review***

ISFG DNA Commission on STR NGS: Complex Nomenclature Systems

1. Bold segment = the reference genome assembly sequence description
D13S317 Ref (11) -Chr13-GRCh38 82148025-82148068 [TATC]$_{11}$
D13S317[CE12]-Chr13-GRCh38 82148025-82148068 [TATC]$_{12}$ 82148001-A; 82148069-T

2. Locus name and capillary electrophoresis allele name
D13S317[CE12]-Chr13-GRCh38 82148025-82148068 [TATC]$_{12}$ 82148001-A; 82148069-T

3. Chromosome and human genome assembly version
D13S317[CE12]-**Chr13-GRCh38** 82148025-82148068 [TATC]$_{12}$ 82148001-A; 82148069-T

4. STR repeat region co-ordinates (start-end) for reference allele
D13S317[CE12]-Chr13-GRCh38 **82148025-82148068** [TATC]$_{12}$ 82148001-A; 82148069-T

5. Description of STR motifs
D13S317[CE12]-Chr13-GRCh38 82148025-82148068 **[TATC]$_{12}$** 82148001-A; 82148069-T

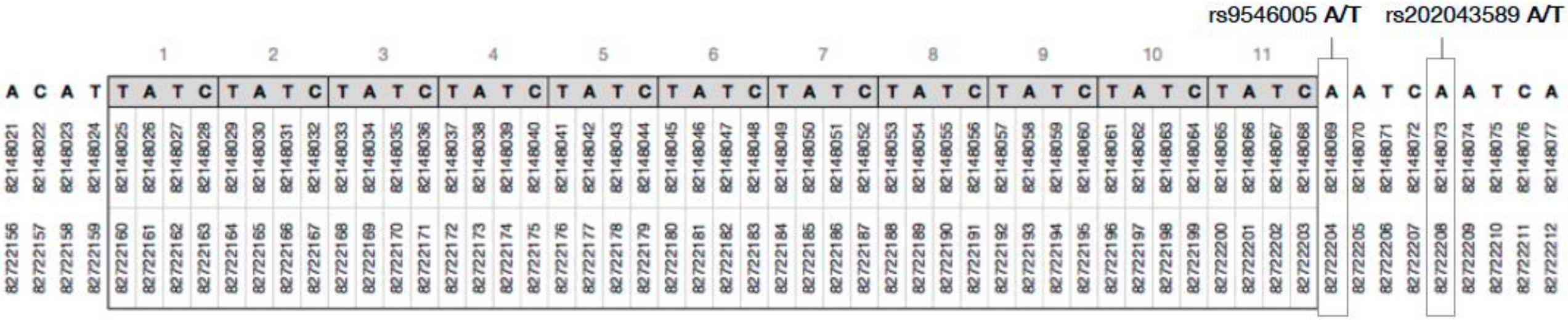
6. Location of flanking region variants
D13S317[CE12]-Chr13-GRCh38 82148025-82148068 [TATC]$_{12}$ **82148001-A; 82148069-T**

Parson et al (2016) ***FSIG***

ISFG DNA Commission on STR NGS: Simple Nomenclature Systems

Simple STR nomenclature systems typically represent easy-to-read unique identifiers consisting of the STR locus name and the operationally-defined repeat-based allele designation derived from CE (e.g. D13S317 13a1a)

D13S317



Parson et al (2016) ***FSIG***

EU collaborative project to investigate STR sequences and software

Acknowledgements

EUROFORGEN Network of Excellence FP7-SEC-2011-285487

FWF Der Wissenschaftsfonds.

Translational Research project L397
"EMPOP–an innovative human mtDNA database"

Research project P22880-B12
"Genetic discovery of an early medieval Alpine population "

NIJ National Institute of Justice 2011-MU-MU-K402

"Maximizing mtDNA Testing Potential with the Generation of High-Quality mtGenome Reference Data"

EMPOP

Mayra Eduardoff
Catarina Xavier
Christina Strobl
Gabriela Huber

Martin Bodner
Cordula Berger
Harald Niederstätter

POPH QUALIFICAR É CRESCER. QREN QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL GOVERNO DA REPÚBLICA PORTUGUESA UNIÃO EUROPEIA Fundo Social Europeu

Collaborators
Arne Dür (Innsbruck)
Jodi Irwin, Rebecca Just (AFDIL/FBI)
Bruce Budowle, Jonathan King (UNTSC)
Lilian Andrea Casas (Colombia)
EMPOP collaborators

ThermoFisher SCIENTIFIC

Robert Lagacé
Sharon Wootton
Joseph Chang

illumina

Nicola Oldroyd
Angela Kalbande

References

Bodner M et al (2015) Helena, the hidden beauty: Resolving the most common West Eurasian mtDNA control region haplotype by massively parallel sequencing an Italian population sample FSI Genet **15**: 21-26

Borsting C et al (2014) Evaluation of the Ion Torrent HID SNP 169-plex: A SNP typing assay developed for human identification by second generation sequencing FSI Genet **12**: 144-154

Brandstätter A et al (2004) Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database Int J Legal Med **118**(5): 294-306

Daniel R et al (2015) A SNaPshot of next generation sequencing for forensic SNP analysis FSI Genet **14**: 50-60

Eduardoff M et al (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM FSI Genet **17**: 110-121

Gettings KB et al (2015) STR allele sequence variation: Current knowledge and future issues FSI Genet **18**: 118-130

Just R S et al (2015) Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three US populations FSI Genet **14**: 141-155

King JL et al (2014) High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq FSI Genet **12**: 128-135

LaRue BL et al (2014) Characterization of 114 insertion/deletion (INDEL) polymorphisms, and selection for a global INDEL panel for human identification Leg Med (Tokyo) **16**(1): 26-32

Marcinska M et al (2015) Evaluation of DNA variants associated with androgenetic alopecia and their potential to predict male pattern baldness PLoS One **10**(5): e0127852

Oberacher H et al (2001) On-Line Liquid Chromatography Mass Spectrometry: A Useful Tool for the Detection of DNA Sequence Variation Angew Chem Int Ed Engl **40**(20): 3828-3830

Oberacher H et al (2001) Analysis of polymerase chain reaction products by on-line liquid chromatography-mass spectrometry for genotyping of polymorphic short tandem repeat loci Anal Chem **73**(21): 5109-5115

Parson W et al (2016) Massively Parallel Sequencing of forensic STRs: Considerations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements FSI Genet

Parson W et al (2015) Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples FSI Genet **15**: 8-15

Phillips C et al (2014) Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set FSI Genet **11**: 13-25

Pitterl F et al (2008) The next generation of DNA profiling--STR typing by multiplexed PCR--ion-pair RP LC-ESI time-of-flight MS Electrophoresis **29**(23): 4739-4750

Pitterl et al (2010) Increasing the discrimination power of forensic STR testing by employing high-performance mass spectrometry, as illustrated in indigenous South African and Central Asian populations. Int J Legal Med **124**(6): 551-558

Pospiech E et al (2015) Evaluation of the predictive capacity of DNA variants associated with straight hair in Europeans FSI Genet **19**: 280-288

Santos C et al (2015) Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise FSI Genet **19**: 56-67

Yang Y et al (2014) Application of next-generation sequencing technology in forensic science Genomics Proteomics Bioinformatics **12**(5): 190-197

Zeng X et al (2015) Selection of highly informative SNP markers for population affiliation of major US populations Int J Legal Med 10.1007/s00414-015-1297-9

Appendix - Considerations of the ISFG on NGS of STRs

Walther Parson, David Ballard, Bruce Budowle, John M. Butler, Katherine B. Gettings, Peter Gill, Leonor Gusmão, Douglas R. Hares, Jodi A. Irwin, Jonathan King, Peter de Knijff, Niels Morling, Mechthild Prinz, Peter M. Schneider, Christophe Van Neste, Sascha Willuweit, Christopher Phillips

Massively Parallel Sequencing of forensic STRs: Considerations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements

Forensic Science International Genetics (2016) in press

Consideration 1:

MPS analysis should be performed with software that allows STR sequences to be exported and stored in databases as sequence (text) strings to capture the maximum consensus sequence information.

Consideration 2:

The forward strand direction assigned in the human genome has been constant for all assemblies published since the first draft in 2001 and can be used to align STR sequences.

Appendix - Considerations of the ISFG on NGS of STRs

Consideration 3:
The choice of reference sequence is crucial for standardizing STR nomenclature systems. At the time of writing, GRCh38 is the most up to date sequence assembly and is recommended as the framework with which to define repeat region structure for sequence alignment and for the mapping of sequence features such as SNPs. Software will be required to handle comparisons between multiple reference sequences, particularly in the short term, where sequence variants listed by 1000 Genomes currently retain GRCh37 coordinates. Continued discussions are necessary to decide whether or not to adapt to novel genome assemblies.

Consideration 4:
Further work is needed to translate the nomenclature of STR loci thus far coded relative to the reverse strand and repeat region start and end points. There is a need to strictly define these and other anchor points to specify the repeat regions.

Consideration 5:
Although simple STR nomenclature systems may be required at some point in the future to facilitate communication and data exchange, comprehensive STR nomenclature systems are preferred for early adopters of STR MPS analysis in order to ensure compatibility with MPS data generated in the future. Backward compatibility to the repeat-based nomenclature derived from CE needs to be maintained to preserve the universal applicability of established national STR databases.

Appendix - Considerations of the ISFG on NGS of STRs

Consideration 6:
To account for relevant genetic variation outside common repeat regions, STR sequences stored as sequence strings should include flanking sequences as well as the genome coordinates of the sequence read start and end points.

Consideration 7:
Updated allele frequency databases will be necessary to take full advantage of the increased power of discrimination offered by MPS generated STR data. A unified nomenclature system is needed to ensure compatibility of worldwide population databases.

Consideration 8:
Future forensic MPS multiplexes would benefit from retention of past markers for backward compatibility and a marker selection process based on population data, molecular biology, sequencing chemistry, and a continued dialogue between the forensic community and commercial suppliers.

EUROFORGEN-NoE Update: EDNAP Meeting Warsaw 2016

Peter M. Schneider

Institute of Legal Medicine

University of Cologne (Germany)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Overview on recent activities

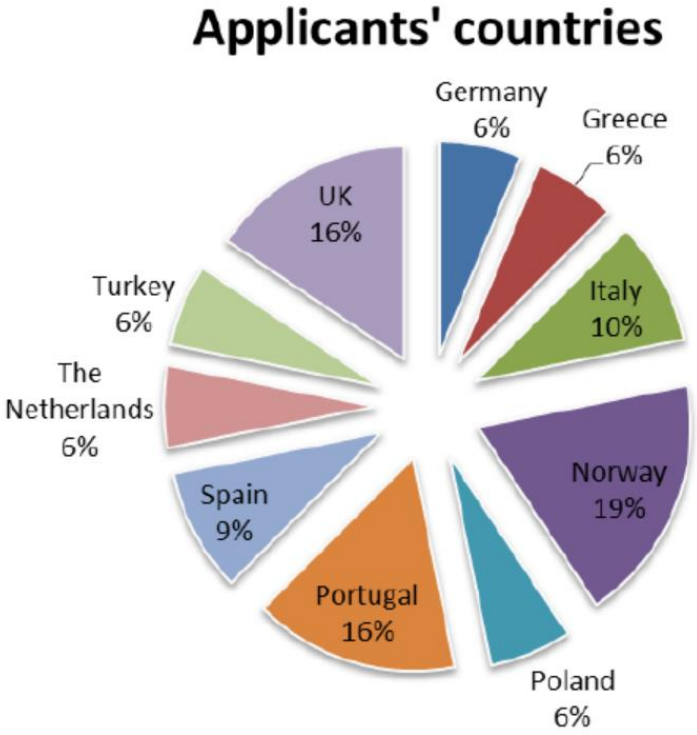
- **The EUROFORGEN short term fellowships**
  - Some statistics and final call 2016
- **Some new publications …**

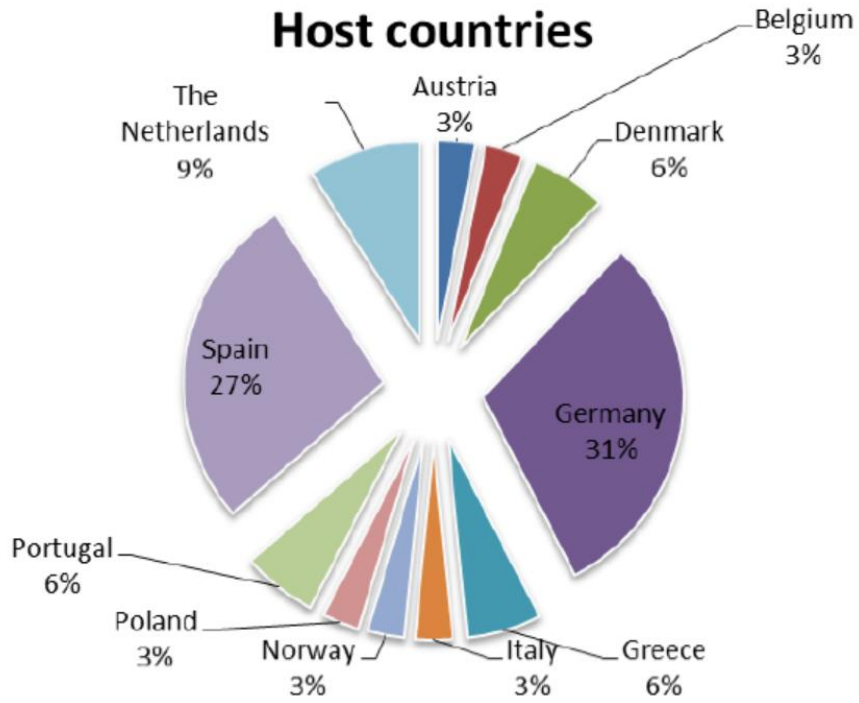
  … from 2016
- **The Virtual Institute**
  - and how to get there
- **The EUROFORGEN Conference**
  - in collaboration with IALM, Venice, 23rd June 2016
- **Training and education news**
  - and a new initiative
- **Dissemination activities**

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

The Short Term Fellowship Program

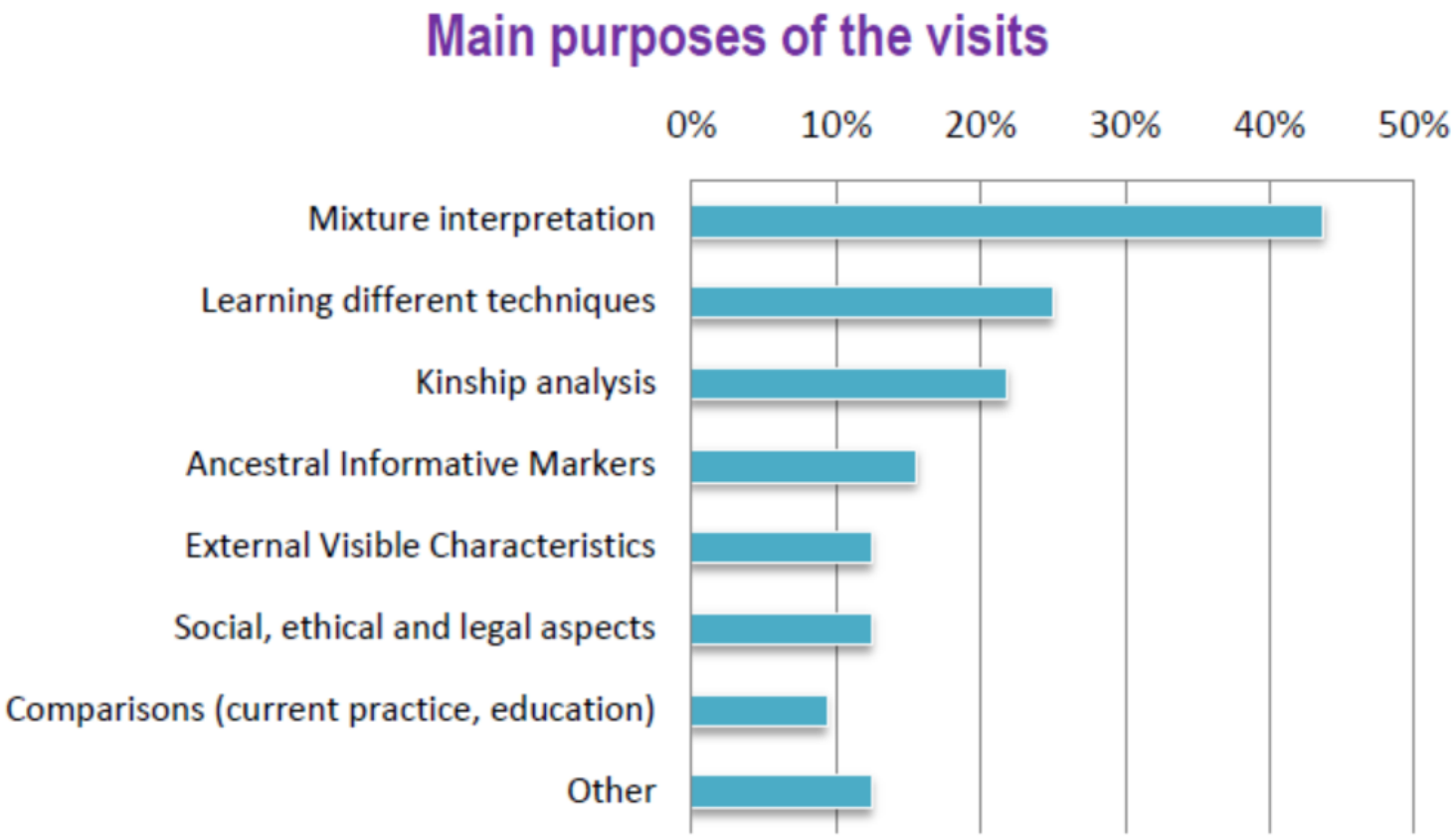
- **First and second calls 2013-2015**
  - 14 fellowships awarded to 28 colleagues from 10 countries visiting host labs in 11 countries





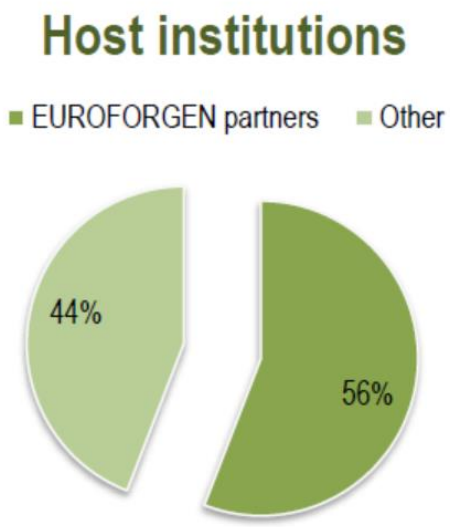
The Short Term Fellowship Program

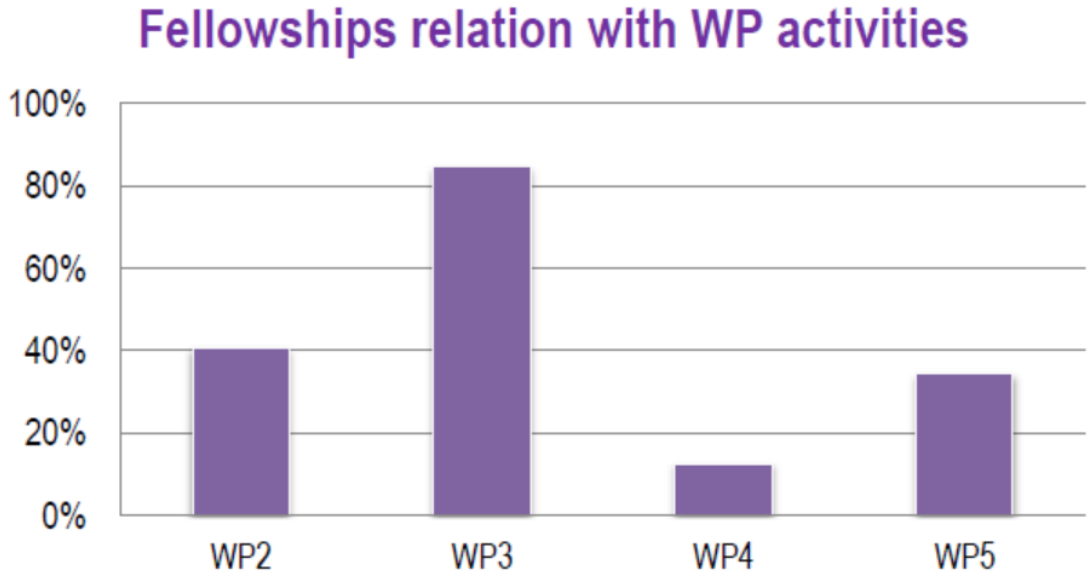
- **First and second calls 2013-2015**



The Short Term Fellowship Program

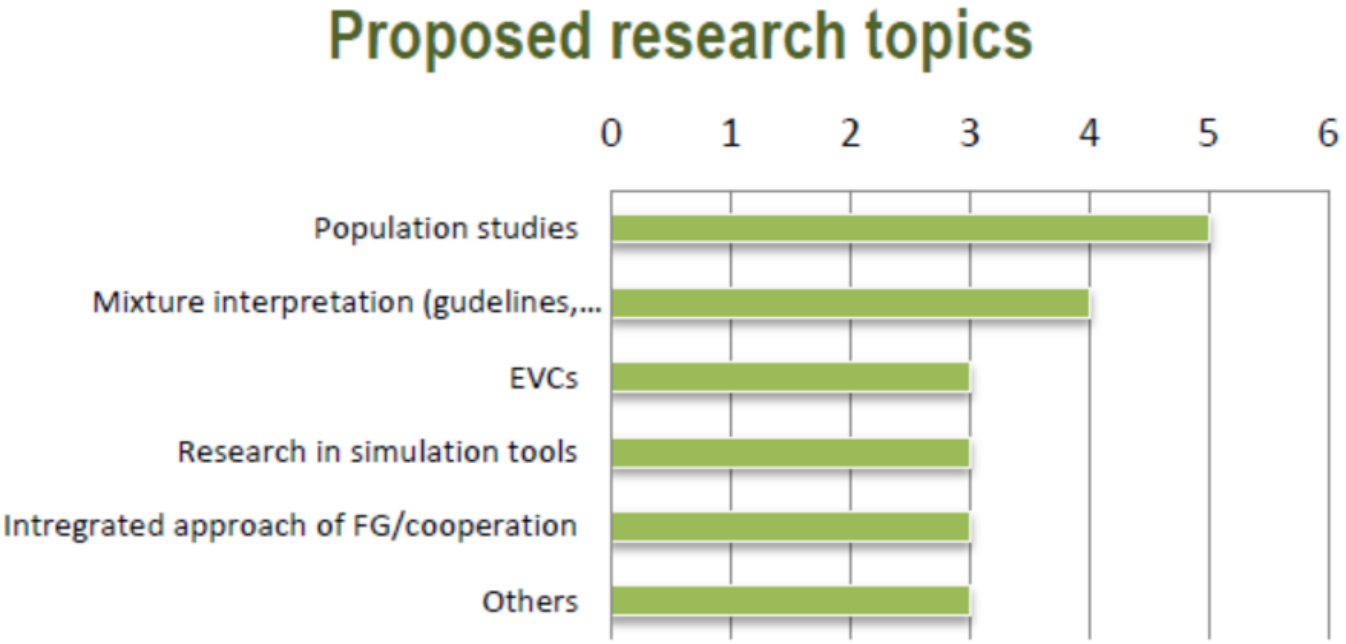
- **First and second calls 2013-2015**





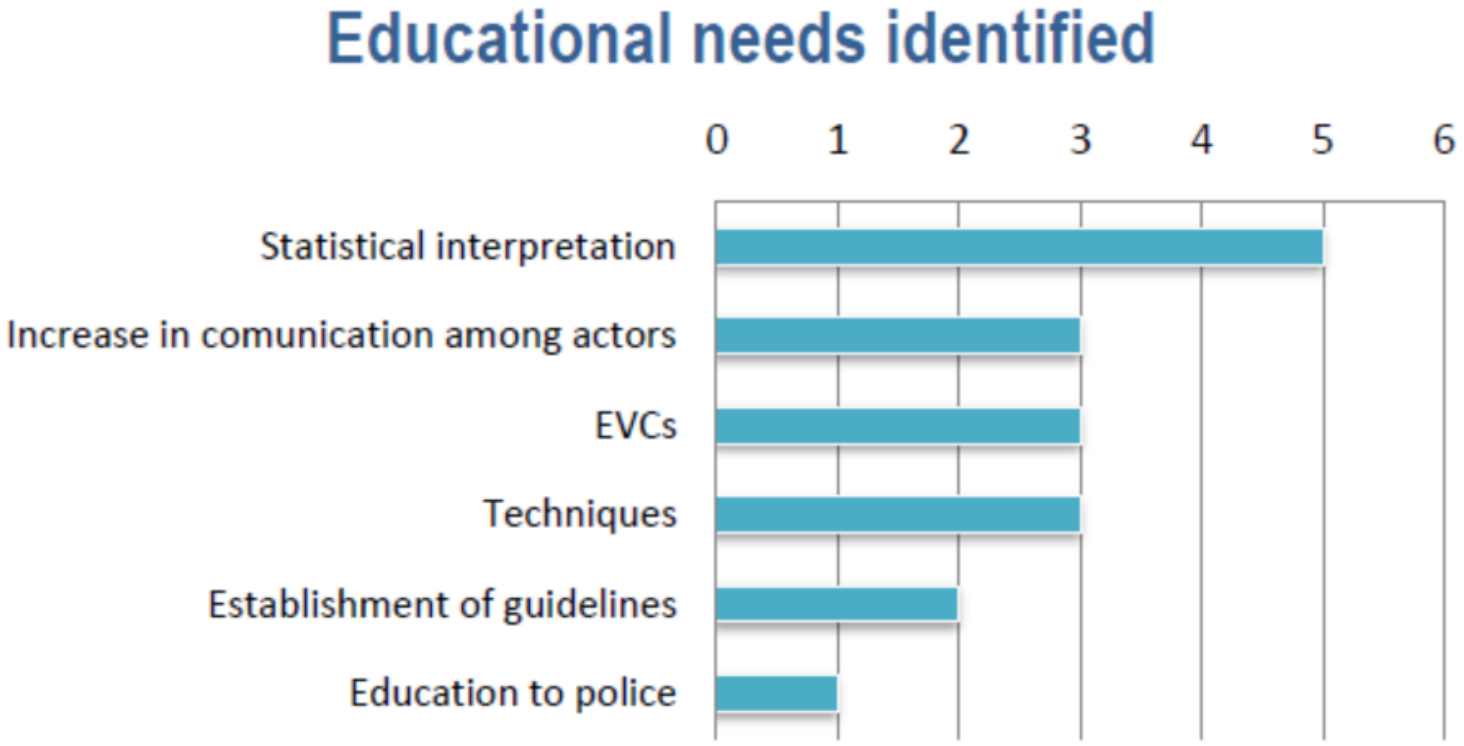
The Short Term Fellowship Program

- **First and second calls 2013-2015**



The Short Term Fellowship Program

- **First and second calls 2013-2015**



The Short Term Fellowship Program

- **Third and final call 2016**
  - 20 new fellowships open,
    also inviting stakeholders from police and judiciary
    - Exchange visits for 3-5 days
    - Active participation in workshops related to EFG aims
    - Other research/training activities related to scope of WPs 2-5
    - Activities to promote closer collaboration between justice, police and scientific community
  - Fellowships must be completed until September 30, to allow for reimbursement of costs within 2016
  - Travel support up to EUR 500
  - Application details on the website

Recent publications

EUROFORGEN Network of Excellence

EUROFORGEN Network of Excellence

Login Search Contact Sitemap Imprint

Home ▸ Dissemination Activities ▸ Research articles

search GO

- About EUROFORGEN-NoE
- The Group
- The Project
- Networking Activities
- Training
- News
- Dissemination Activities
  - Consortium publications
  - Research articles
- Contact
- Login

Tweets by @EUROFORGEN

Project publications

The original articles listed below have been published in scientific journals, and mainly describe results from Work Package 3. In case of co-authorship, the work of one or several of the contributing authors has been funded by EUROFORGEN-NoE.

2016

Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios.

Forensic science international. Genetics. 2016 Mar;21:81-9

Authors: van den Berge M, Ozcanhan G, Zijlstra S, Lindenbergh A, Sijen T

PubMed.gov

Newsletter (1/2016)

Dowload here

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Recent publications

- M. Van der Berge et al.: **Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios.**
  FSI Genetics. 2016 Mar; 21:81-9

- P. Gill: **Analysis and implications of the miscarriages of justice of Amanda Knox and Raffaele Sollecito.**
  FSI Genetics. 2016 Mar 3; 23:9-18

- Ø Bleka et al.: EuroForMix: **An open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts.**
  FSI Genetics. 2016 Mar; 21:35-44

- M. Eduardoff, T.E. Gross et al.: **Inter-laboratory evaluation of the EUROFORGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM™**
  FSI Genetics. 2016; in press

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

The Virtual Institute of Research for Forensic Genetics

EUROFORGEN
Network of Excellence

Logout Search Contact Sitemap Imprint

search GO

- About EUROFORGEN-NoE
- The Group
- The Project
- Networking Activities
- Training
- News
- Dissemination Activities
- Contact
- EUROFORGEN partner area
- EUROFORGEN members area
  - EUROFORGEN course material
  - EUROFORGEN publications
  - Recommended open software
  - Train-the-trainers section

Virtual Institute for Forensic Genetic Research in Europe

Our website will provide a framework for exchange of expertise and data, not only between consortium members but with any other individuals or institutions working in forensic genetics in Europe. It will bring together the knowledge and resources centered on forensic genetics tools and education at a European level, and allow researchers, forensic practitioners, stakeholders and legal experts to interact with the network. Currently, the following resources are available:

- **EUROFORGEN Course Material>: Up-to-date lectures and presentations on major topics of forensic genetics derived from the "Train the Trainers" workshop series.**
- **EUROFORGEN publications>:** Original publications from EUROFORGEN Consortium members available for downloading.
- **Recommended Open Software>:** a list with open software tools is displayed together with a brief description on their applications.
- **Train-the-Trainers Section>:** it contains the "TTT Blog", a discussion forum to post comments and questions related to training issues, to get directly into contact with the EUROFORGEN trainer team.

Please use the blog for your feedback, and your suggestions for improvement. The contents will be regularly updated and expanded.

Your EUROFORGEN-NoE team.

Quicklinks

You can apply for → **short term fellowships**.

The European landscape in forensic genetics: → **Geographical display and contact data**

→ The winners are found! The EUROFORGEN competitive call review led to 3 proposals with highest score.

Newsletter (4/2014)

Dowload here

Consortium

- **Dedicated "for members only" area of website**
  - Can only accessed after individual registration, and obtaining a user name and password
  - All colleagues working in institutions that have submitted their contact data by submitting a questionnaire in the initial inquiry will be admitted
  - Please do not hesitate to inquire if you are not sure about the participation of your lab!

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

The Virtual Institute of Research for Forensic Genetics

EUROFORGEN Network of Excellence

L

Home ▸ Networking Activities ▸ European Virtual Institute of Research in Forensic Genetics

European Virtual Institute of Research in Forensic Genetics - access query

You are interested in becoming a member of the European Virtual Institute of Research in Forensic Genetics?

If you are a scientist working at a forensic genetics laboratory, or a professional working in an institution of the justice system, you are invited to join the Virtual Institute. Please see our Newsletter 3/2014 for further details.

Please enter your personal contact data, and the data of your institution below. We will verify your request and come back to you in the following days.

One requirement to get access to the EUROFORGEN-NoE Virtual Institute of Research in Forensic Genetics is the participation of your institution by submitting the EUROFORGEN-NoE questionnaire.

Your EUROFORGEN-NoE team.

SEVENTH FRAMEWORK PROGRAMME

EU
wit

- **Privileged access to new content:**
  - **Course Material:** Up-to-date lectures and presentations on major topics of forensic genetics derived from the "Train the Trainers" workshop series.
  - **Publications:** Original publications (PDF) from Consortium members available for downloading.
  - **Open Software:** a list with open source / accessible software tools is displayed together with a brief description on their applications.
  - **Train-the-Trainers Section:** a discussion forum to post comments and questions related to training issues, to get directly into contact with the EUROFORGEN trainer team.
- **In preparation:**
  - Training videos and online learning tools

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

EUROFORGEN - International Dissemination Conference

"Forensic DNA analysis in the light of the new security needs"

Venice, June 23rd 2016

Forensic DNA analysis in the light of the new security needs

- **Session 1: FROM CRIME SCENE TO COURT ROOM**
  - **Evidence challenges and advanced interpretation methods**
  - **The interpretation debate: miscarriages of justice**
- Challenges in Forensic Genetics (John Butler, NIST)
- The impact of the Innocence Project on forensic science policy and practice – (Sarah Chu, New York)
- The interpretation debate – how misleading is DNA evidence? (Peter Gill)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Forensic DNA analysis in the light of the new security needs

- **Session 2: FROM GENOTYPE TO PHENOTYPE**
  - **The next step in forensic genetic intelligence**
  - **State-of-the-art and future directions**
- Genetics of human visible traits (Ana Valdes & Timothy Spector; King's College)
- Recent progress in predictive DNA analysis (Wojciech Branicki)
- New technologies and approaches in forensic genetics - application to real cases (Chris Phillips)

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Forensic DNA analysis in the light of the new security needs

- **Session 3: SCIENCE IN SOCIETY**
  - **Ethical and legal aspects, the societal dimension of forensic genetics**
  - **Security issues in Europe from a DNA perspective**
- Forensic DNA Evidence in Sexual Assault Cases – Workflow, Reporting and Information Exchange (Lutz Roewer, Berlin)
- Ethical aspects of the evaluation of the evidence and communication (Erin Murphy, NYU Law School)
- Ethical and social challenges of transnational exchange of DNA data in the EU (Helena Machado, University of Coimbra)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Forensic DNA analysis in the light of the new security needs

- **Session 3: SCIENCE IN SOCIETY**
  - **Ethical and legal aspects, the societal dimension of forensic genetics**
  - **Security issues in Europe from a DNA perspective**
- Ensuring the social life of innovations in forensic genetics (M. Wienroth, Northumbria University)
- Contribution of DNA typing to the resolution of crimes against humanity (Lourdes Prieto, USC)
- European law regulating the forensic application of DNA testing and its use for national databases (Kristiina Reid, London)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Forensic DNA analysis in the light of the new security needs

- **Round table:**
  - **Current challenges in Ethics & Legal & Social aspects in different European countries** -- Robin Williams (chair and UK),
  - Angel Carracedo (Spain and Portugal),
  - Susi Pelotti (Italy),
  - Peter Schneider (Germany),
  - Christian Doutremépuich (France),
  - Gunilla Holmlund (Sweden)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

EUROFORGEN Conference - June 23, 2016

- **Venue and registration details**
  - **www.ialm2016venice.org**
  - Venice Convention Center, Lido Island
  - Reduced registration for IALM partner societies available until May 31st, 2016 (€ 465, after this date only onsite registration)
  - Current members of the EUROFORGEN Virtual Institute will get an additional discount after the conference
  - Includes access to all sessions of IALM Symposium
- **Session on Forensic Genetics and Genomics, June 24th**
  - Speakers: A. Carracedo, M. Kayser, W. Parson, P. Schneider
  - Short presentations

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

WP5: Education, Training and Career Development

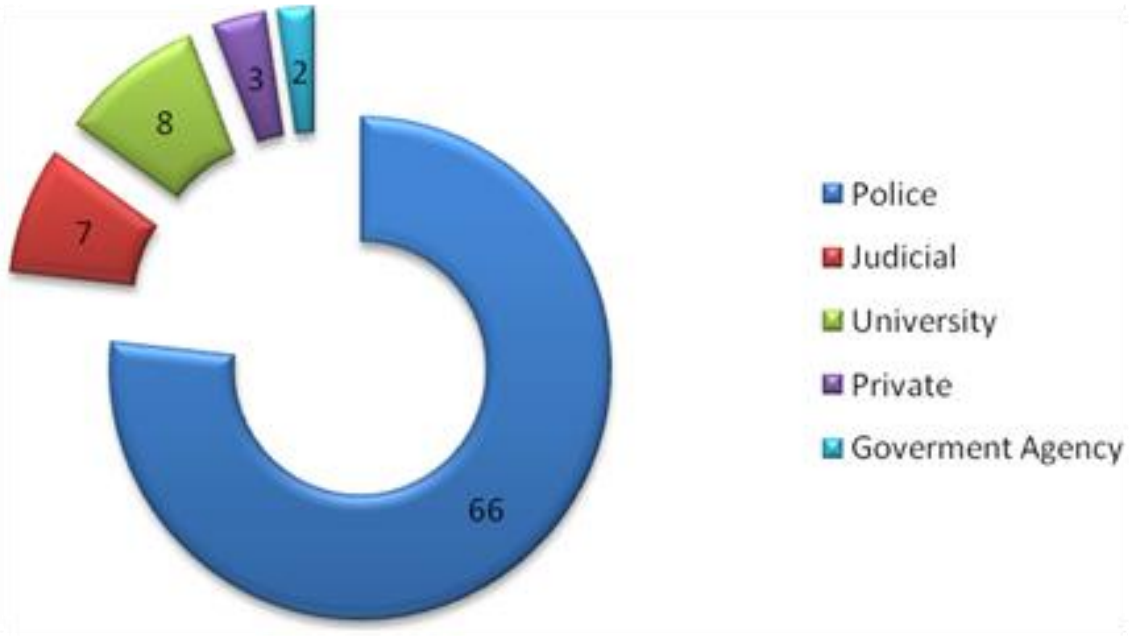
- **New WP5 manager since July 2015:**

  Angel Carracedo (USC)

- **New and ongoing activities planned for 2016**
  - Extend collaboration with CEPOL: support needed from CEPOL national contact points to express interest for training in forensic DNA analysis and interpretation
  - Initiate collaboration with European Judicial Training Network (EJTN) to offer advanced training on forensic DNA to judiciary
  - Support satellite training workshops at national level with EUROFORGEN teachers
  - Develop new contents on education for the Virtual Institute website
  - Introduce a curriculum for post-graduate education in forensic biology and genetics

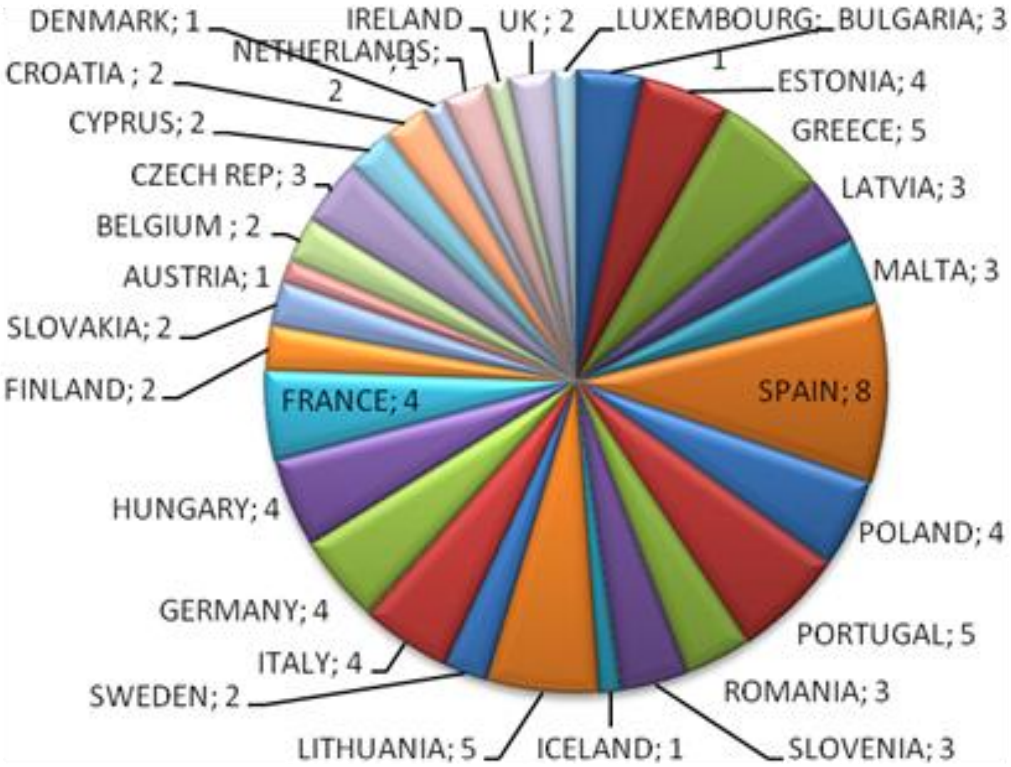
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

- **Three CEPOL courses 2012-2015 in Avila / Spain**
  - Origin of participants by institutions



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

- **Three CEPOL courses 2012-2015 in Avila / Spain**
  - Origin of participants by country



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

- **Introducing the first EUROFORGEN Summer School**
  - scheduled for **July 17-21, 2017**, to take place in **Santiago de Compostela, Spain**
  - Audience: Students of Law and Biomedical Sciences, Judiciary, Police personnel at different educational levels
  - Covering relevant basic and advanced topics in forensic genetics
  - Not funded by EC, moderate tuition fees will be charged
- **The EUROFORGEN Summer School will continue**
  - Taking place annually at changing locations in Europe
  - Addressing the needs of the community
  - Supporting the continuing platform of the Virtual Institute of Research in Forensic Genetics

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

EUROFORGEN
Network of Excellence

EUROFORGEN - European Fore...
Public Group

Joined | Share | Notifications | ...

Discussion | Members | Events | Photos | Files

Search this group

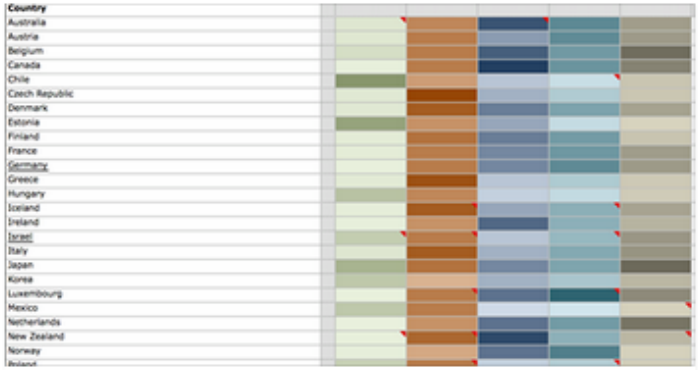
Write Post | Add Photo / Video | Create Poll | More

Write something...

RECENT ACTIVITY

Peter Schneider
Yesterday at 4:34pm

Very useful to create figures displaying quantitiative research data!



Create a Heatmap in Excel | PolicyViz

Tutorial on how to create a Heatmap in Excel using the Conditional Formatting options.

POLICYVIZ.COM

Like | Comment | Share

Thore Egeland, Luis Souto Miranda and 4 others

ADD MEMBERS

Enter name or email address...

MEMBERS 274 Members (3 new)

Message · Invite by Email

DESCRIPTION Edit

The EUROFORGEN Network of Excellence is creating the Europe... See More

TAGS Edit

Forensic DNA · Science · Genetics

GROUP CHATS

No group chats, start one now.

+ Start New Chat

CREATE NEW GROUPS

Groups make it easier than ever to share with friends, family and teammates.

Create Group

RECENT GROUP PHOTOS See All

find us on Facebook

Please do not forget to join our Facebook group!
… already 274 members!

EUROFORGEN-NoE
@EUROFORGEN

#EUROFORGEN is now on @twitter. Follow us to find out more about #forensic genetics & join our group @facebook.

EUROFORGEN
Network of Excellence

RETWEETS 2 LIKES 5

EUROFORGEN
Network of Excellence

Home

- About EUROFORGEN-NoE
- The Group
- The Project
- Networking Activities
- Training
- News
- Dissemination Activities
- Contact
- Login

Tweets by @EUROFORGEN

EUROFORGEN-NoE
@EUROFORGEN

#EUROFORGEN will participate in "Security Research & Innovation Event 2016", The Hague; srie2016.com #SRIE2016 #forensicGenetics

19 Apr

About EUROFORGEN-NoE

The main objective of the European Forensic Genetics EUROFORGEN-NoE -- is the creation of a **Euro Forensic Genetic Research**.

Forensic genetics is a highly innovative field of appl impact on the security of citizens. The project, fu programme of the European Commission under the 7 started in January 2012 and will run for 5 years. It is Schneider of the Institute of Legal Medicine at th Cologne, Germany.

The network includes 16 partners from 9 countries in European forensic genetic research. It aims to crea existing collaborations as well as to establish new specialized field of security. Therefore, all key pla stakeholders and end-users (e.g. police institutions educational centres and scientific societies will be integ Furthermore, the societal dimension of applying fore investigations, potentially affecting privacy rights and addressed. Finally, training and educational activitie establish common European-wide standards for ge genetic data related to biological evidence.

SRIE 2016, June 1& 2, 2016

EUROFORGEN-NoE
@EUROFORGEN

#EUROFORGEN will participate in "Security Research & Innovation Event 2016", The Hague; srie2016.com #SRIE2016 #forensicGenetics

9:41 AM - 19 Apr 2016

17h00– 18h00 **Panel discussion: Forensics (Connection to the NL PRES event)**

Peter M. Schneider (Coordinator of EUROFORGEN project, DE)

Arie Ijzerman (Chair of the COSI)

Jan de Kinder (Chair of ENFSI)

Representative of a Forensic Institute (TBC)

Session Chair: Michele Socco (DG HOME, EC)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Thank you very much for your attention!

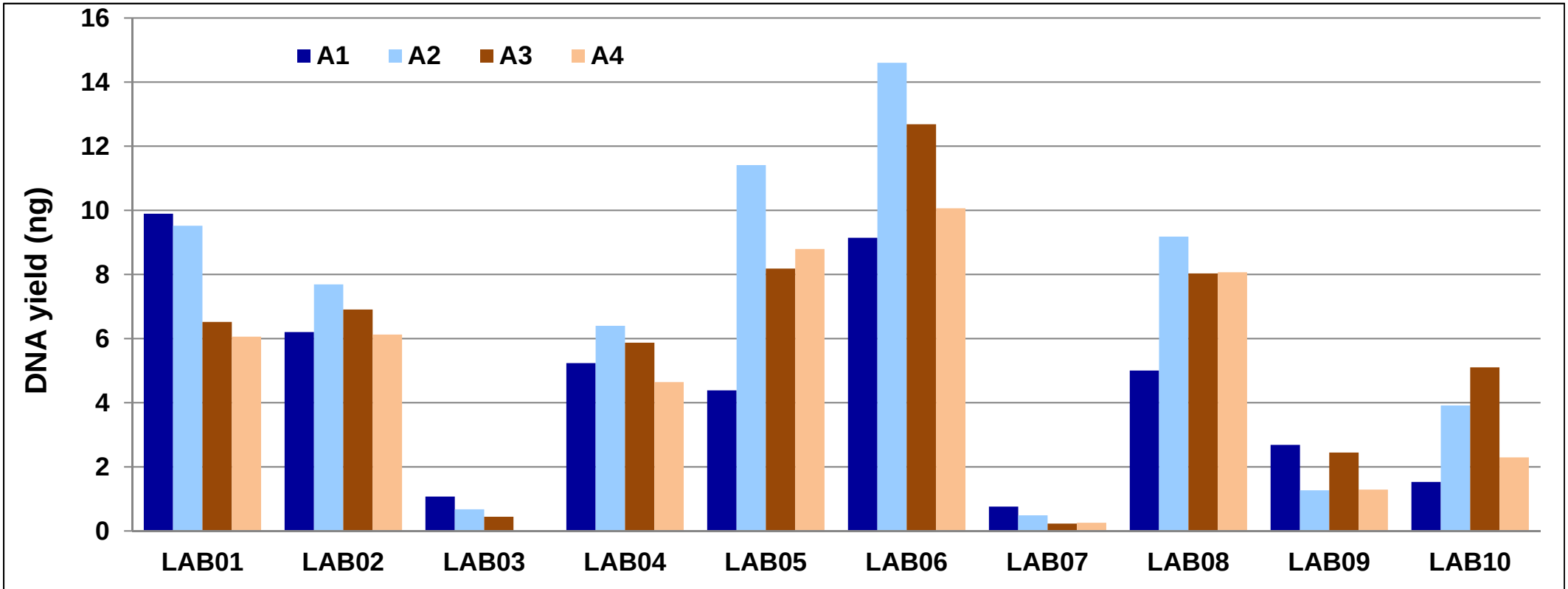
EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Asian Forensic Sciences Network
Inter-Laboratory DNA Collaborative Exercises

- Proposed by HSA at AFSN Meeting 2014 in Seoul
- Objectives
  - Benchmark processes and capabilities of Laboratories
  - Strengthen cooperation among member institutes
  - Enhance quality of forensic services in the region
- 2 exercises conducted in 2015
  - DNA extraction efficiencies
  - Mixture interpretation approaches
- 3rd exercise being planned for June 2016 – Amplification and interpretation of DNA profiles

Asian Forensic Sciences Network Inter-Laboratory DNA Collaborative Exercise 1

- Labs given 2 pairs of replicate blood samples to process.
- Extracted DNA returned to HSA → quant and amp.
- 10 Labs from 6 countries participated.
    - Indonesia, Malaysia, Philippines, South Korea, Mongolia, Thailand.



Asian Forensic Sciences Network
Inter-Laboratory DNA Collaborative Exercise 2

- Labs were sent electropherograms of six 2- and 3-person mixtures with specified analytical and stochastic thresholds.
- Labs interpreted the profiles, and identified the alleles that would be used for matching → much variation observed.
- 10 Labs from 8 countries participated.
  - China, Indonesia, Malaysia, Mongolia, Phillippines, Singapore, South Korea, Thailand.
- Variation in statistical approaches.
  - Modified Random Match Probability, Combined Probability of Inclusion, Likelihood Ratio.
- Report to Labs included suggested interpretations by Dr Bruce Budowle.

Asian Forensic Sciences Network

- 3rd Inter-Lab Exercise planned for June 2016.
- Labs will be sent 5 lyophilised DNA samples – 2 single-source and 3 mixtures (1:1, 5:1, 5:5:1).
- Labs to re-constitute, amplify, interpret according to internal guidelines.
- Identify the alleles that would be used for matching.
  - Lab's AT and ST? Detection and reporting of the minor contributor? Statistical approach? …

- **AFSN 2016 meeting – Bangkok, Aug 2016**
- **AFSN 2017 meeting - China**
- **AFSN 2018 meeting - Singapore**

HSA Biology Division

ASCLD/LAB-*International*

Forensic Samples

Examination for Biological Material

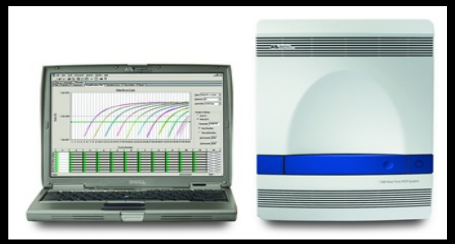
KM
Hexagon OBTI

AP
RSID-Semen & PSA
Sperm HyLiter LMD

RSID-Saliva

Extraction
(8 x Maxwell-16)

Quantitation
(2 x 7500) Quant Duo



PCR (x8)
IDD, ID+, ESX, Y23

Detection
(3 x 3500xl, 1 x 3500)

Detection
(2 x 3500xl)

Volume crime TAT ~ 35 days
Major crime TAT ~ 130 days
DNA databasing ~ 14 days

Technology Translation

Direct Human ID

- GlobalFiler & PP Fusion 6C
- MPS-based
  - MtDNA
  - STRs
  - Identity SNPs

Indirect Human ID

- Y-STR database → familial searching
- Prediction of ancestry, age, phenotype

Beyond Source

- Time of deposition
- Persistence
- Transfer
- Tissue origin

Evolution

- Recovery of material
- Rapid DNA & Direct Amp
- DNA from non-human sources?

Validation of software and courtroom experiences

Peter Gill

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Contents lists available at ScienceDirect

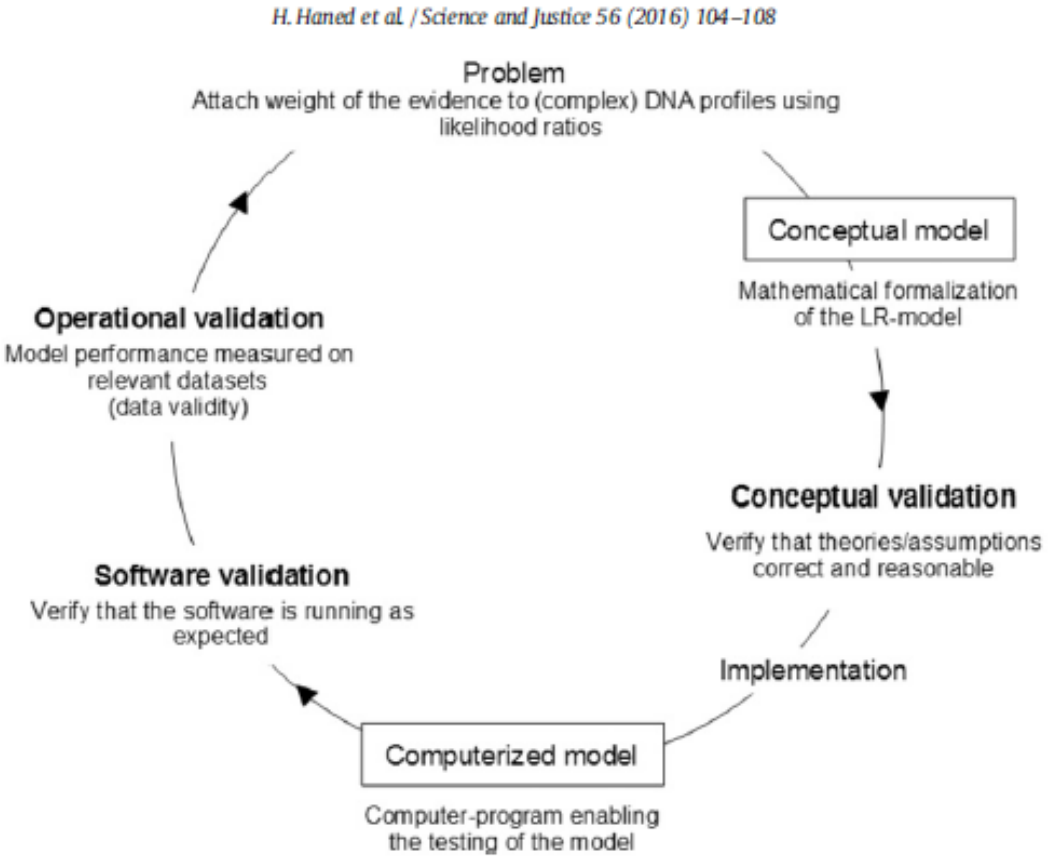
Science and Justice

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

Validation of probabilistic genotyping software for use in forensic DNA casework: Definitions and illustrations

Hinda Haned [a,*], Peter Gill [b,c], Kirk Lohmueller [d], Keith Inman [e], Norah Rudin [f]

[a] Netherlands Forensic Institute, Department of Human Biological traces, The Hague, The Netherlands
[b] Norwegian institute of Public Health, Oslo, Norway
[c] Department of Forensic Medicine, University of Oslo, Norway
[d] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, 621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, United States
[e] Department of Criminal Justice Administration, California State University, East Bay, 4069 Meiklejohn Hall, 25800 Carlos Bee Boulevard, Hayward, CA 94542, United State
[f] 650 Castro Street, Suite 120-404, Mountain View, CA 94041, United States



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Step 1. Conceptual validation

- Model theory and assumptions were explained and justified in a “statistical specifications” report,
- Model theory was formalized and published in peer reviewed journals.
- operational validation document written

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Step 2. Operational validation

- Model output was compared to expert opinion on 20 cases,
- Model output was compared to the following programs: Lab
- Retriever, LikeLTD, FST, GRAPE
- Performance tests using 211 controlled mixtures (of one up
- to five contributors) and 621 (overall-loci) likelihood ratios
- were compared to expected trends based on gold standard
- conditions where parameters were known.

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Validation

Step 3. Software validation

- Software output was evaluated analytically, using the Xcas algebra software
- Model output was evaluated on 77 controlled NGM mixtures, and N1000 LRs were computed and compared to expected trends using known parameters
- LRmix output was evaluated against analytical formulae derived for simple examples
- LRmix output was evaluated against an independent reimplementation

of the model (in the Java language), using
77 controlled NGM mixtures, and 1000 LRs were computed and compared

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Validation statement:

"Over the 1095 LR calculations were submitted to comparisons of LRmix and other software, for all tested samples, the same conclusions were obtained. We therefore concluded that LRmix is validated for use in casework, within the limitations described in the operational validation document."

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Version control of open-source models

- Version control is the same as for commercial models.
- Versions can only be changed by the developers. Version changes are recorded of course.
- Users are of course completely free to carry out research on the code – but of course this would create a new software that cannot be uploaded to the official website
- This is no different to practice in medical genetics where the vast majority of software that is produced is open-source.
- Advantages of open-source:
  - Transparency
  - Availability
  - Standardisation
  - http://lrmixstudio.org/faq/and https://en.wikipedia.org/wiki/Open-source_software

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

EuroForMix

A user-friendly software for evaluating STR/SNP profiles using peak height information

Developed by Oyvind Bleka (FHI)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Forensic Science International: Genetics 21 (2016) 35–44

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

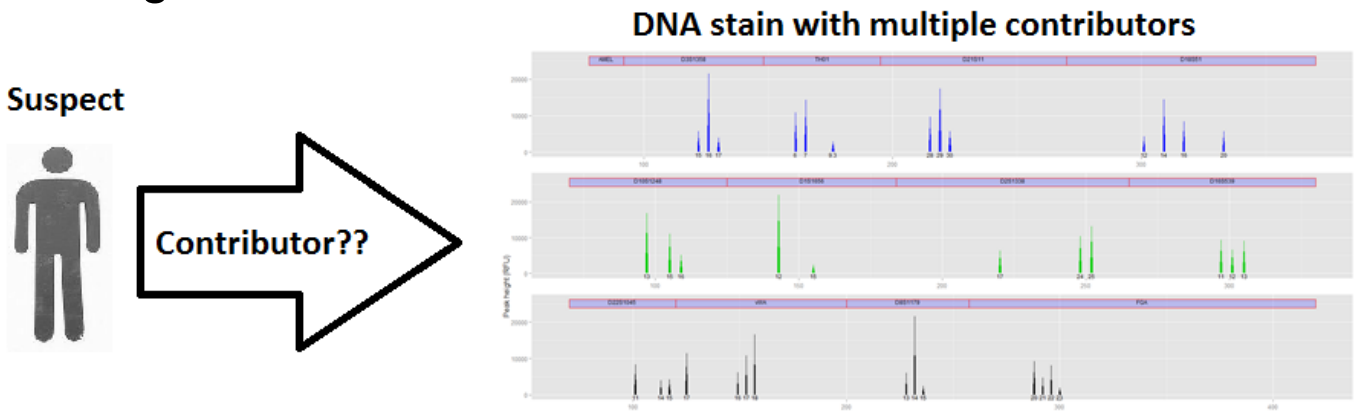
EuroForMix: An open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts

Øyvind Bleka [a,b,*], Geir Storvik [b,1], Peter Gill [a,c,1]

[a] Department of Forensic Biology, Norwegian Institute of Public Health, Oslo, Norway
[b] Department of Mathematics, University of Oslo, Oslo, Norway
[c] Department of Forensic Medicine, University of Oslo, Oslo, Norway

About EuroForMix

- A Graphical User Interface which implements and extends the continuous model from Cowell et.al (2015). DNAmixtures.
- Parameters for mixture proportion, peak height distribution, stutter proportion and degradation are automatically taken into account. Drop in parameter informed from lab observations.
- Unlike some other continuous models there is no need for calibration, but prior information can optionally be specified.
- Weight of evidence (WoE) calculation of crime samples now uses peak height information!



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Benefits of the gamma model

- No requirement to carry out an assessment of stutter, peak height characteristics as this is automatic – it is accommodated by the data actually evaluated by the model.
- MCMC is not used to evaluate the LR – but is used in sensitivity analysis to model the unknown parameters
- This program has been completely rewritten from the original idea from Cowell et al (2015) 'DNAmixtures' which is distributed as open source but requires purchase of HUGIN program.
- Our validation has compared the two programs and we show they produce the same results.
- Open-source is transparent. All of the algorithms are published in sufficient detail for others to implement from scratch

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

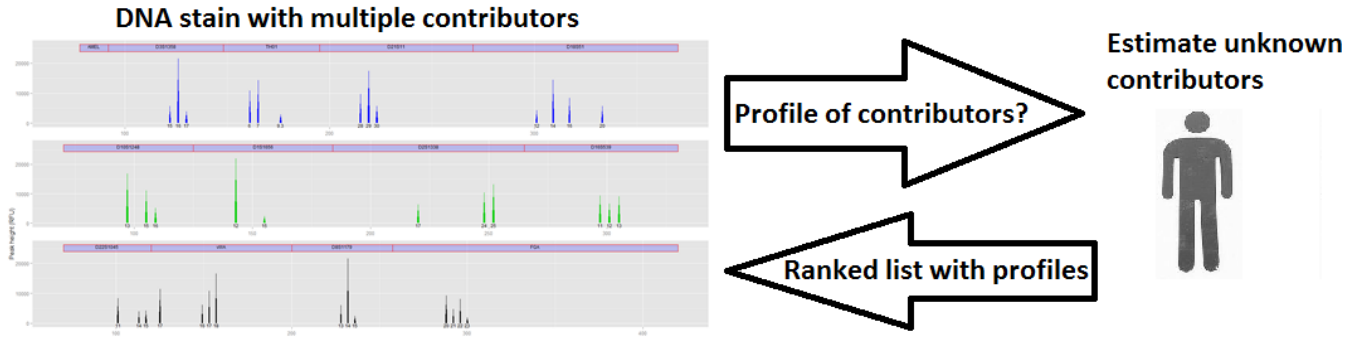
Features

- The continuous model in EuroForMix supports:
  - Multiple contributors in hypothesis
    - Can condition on any number of reference profiles
    - Can specify any number of unknowns (practical limit is 4)
  - Replicated samples
    - Not strictly necessary but will accomodate consensus samples
  - Stutters
  - Allele drop-out
  - Allele drop-in with a peak height model
  - Coancestry effect (Fst-correction)
  - Degradation of peak heights over fragment length

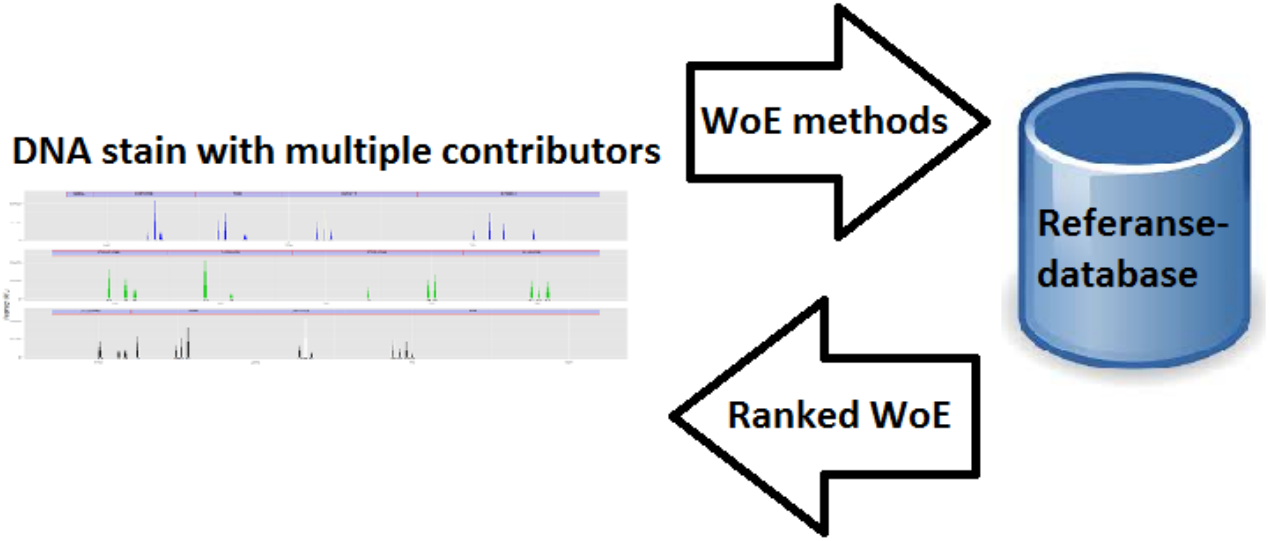
EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Other applications

Deconvolution:

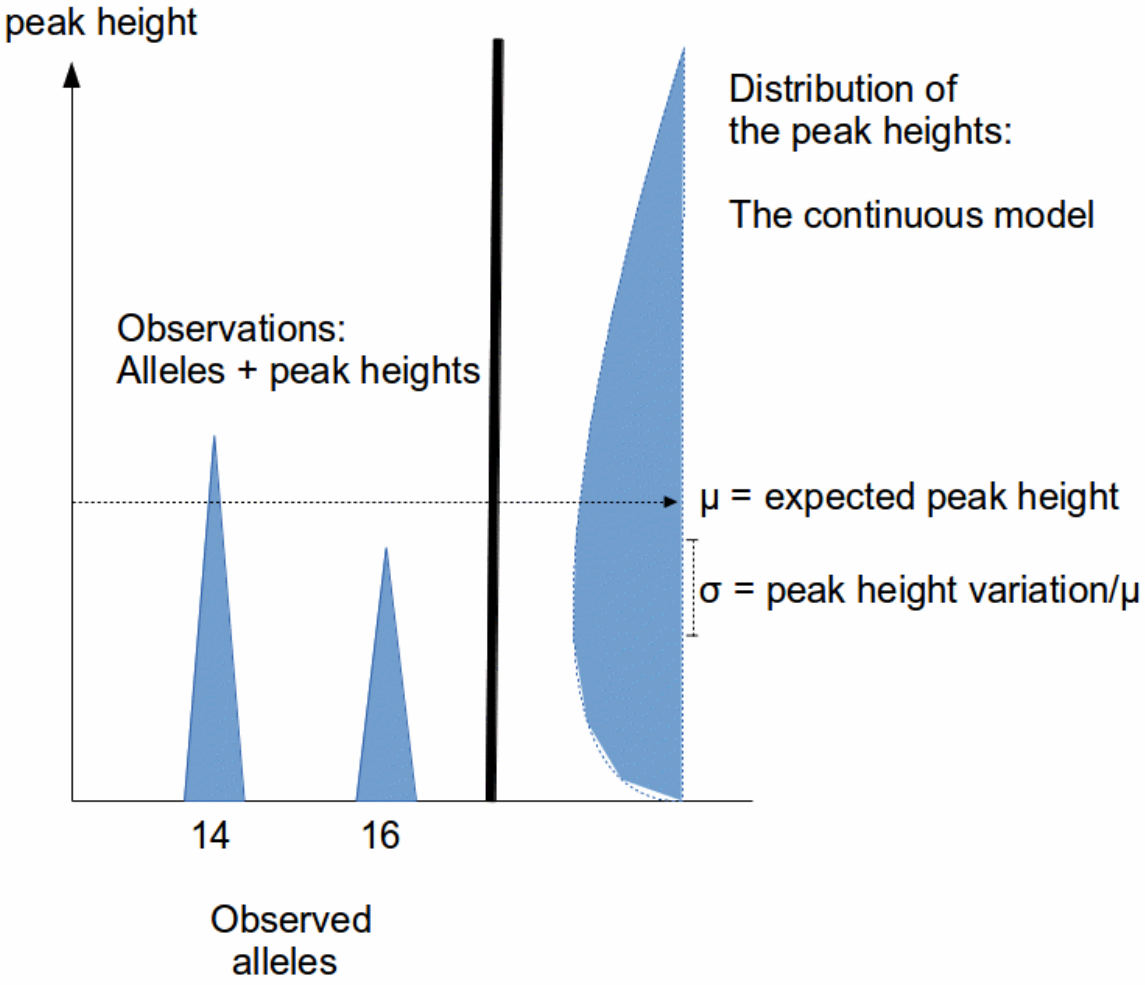


Database search:



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

The continuous model



Three parameters needed:

Expected peak height (μ)
Peak height variation (σ)
Mixture proportion (Mx)

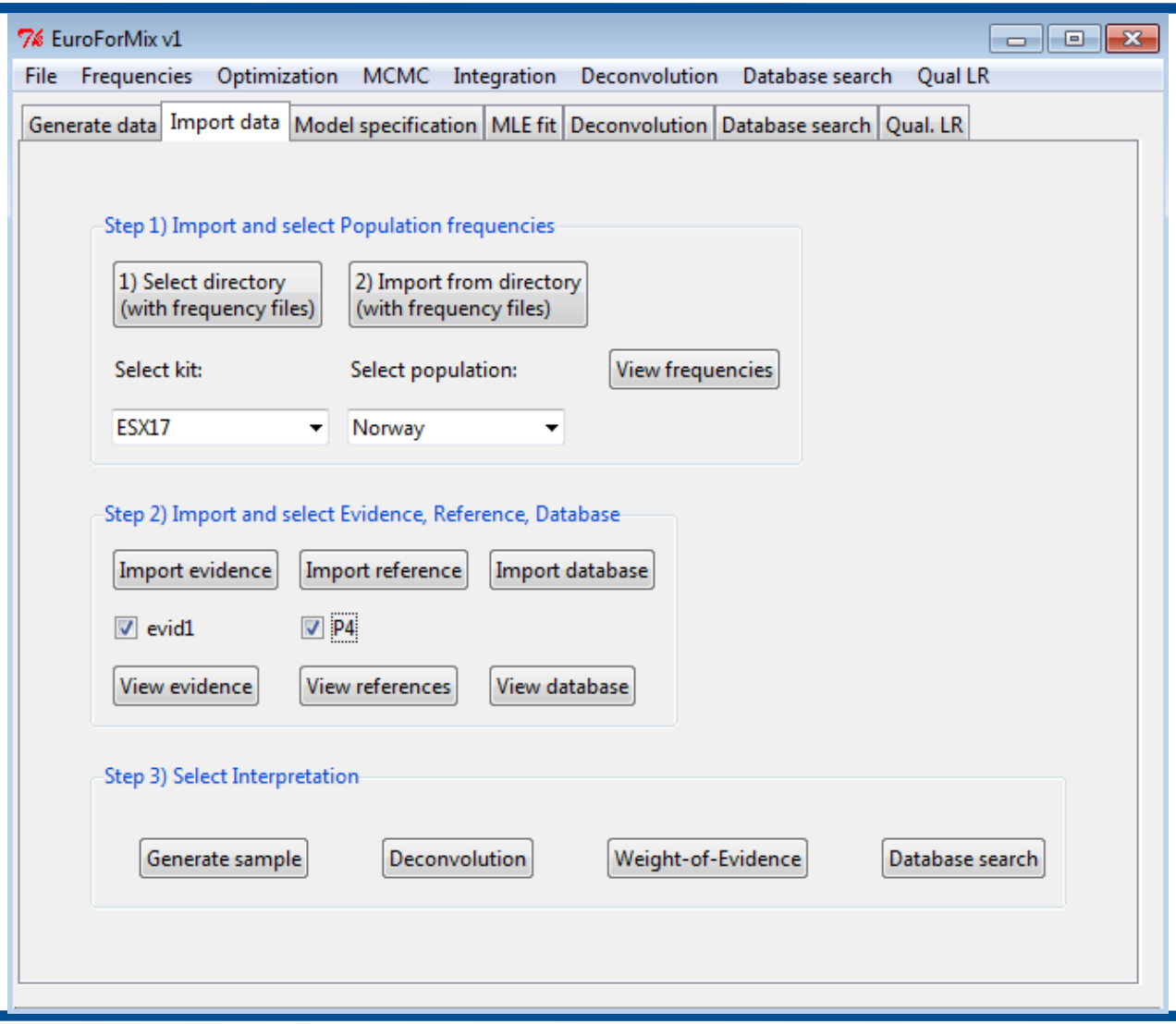
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

A case example

- A robbery was committed. An individual R was apprehended with a balaclava and provided a reference sample which matched sample E.
- R is conditioned under both hypotheses.
- Person S is implicated but he denies wearing the balaclava
- Person S donates a reference sample
- Aim is to determine strength of evidence if S is a possible contributor to the sample
- The likelihood ratio (LR) was calculated using the following alternative hypotheses
- Hp : “Individual R and suspect S both contributed to E"
- Hd : “Individual R contributed to E, but suspect S did not"

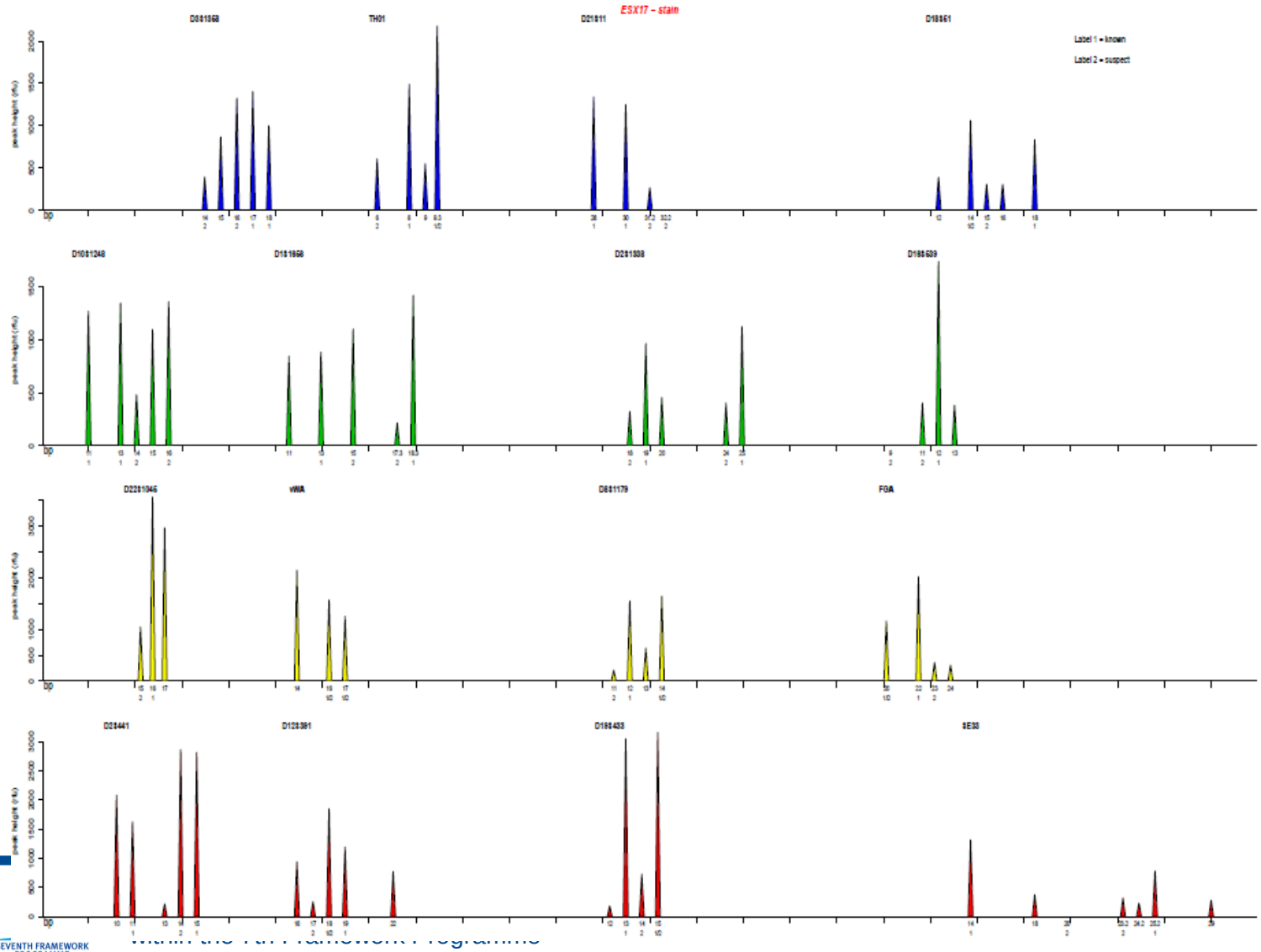
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

The GUI: Import

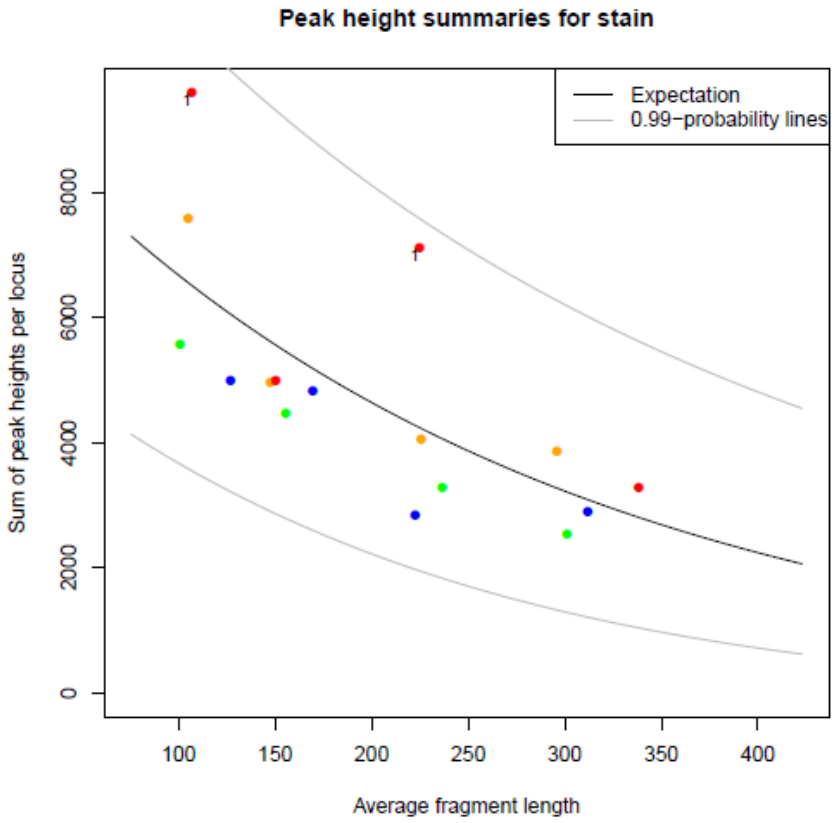


EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Epg generated by Euroformix



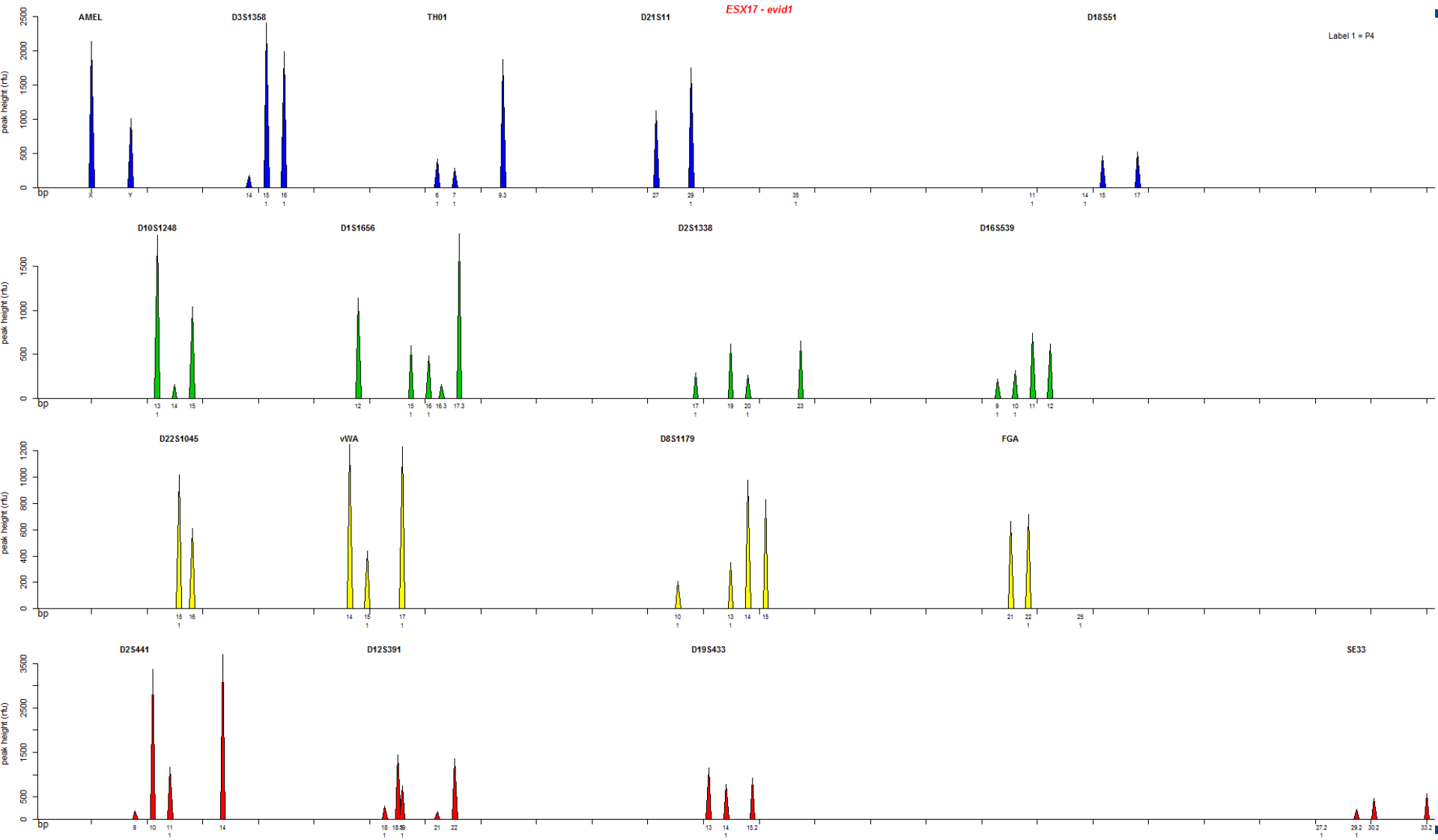
Peak height summary of the crime stain



- Degradation clearly presented.
- D2S441 and D19S433 have much greater sum peak height than the other markers.

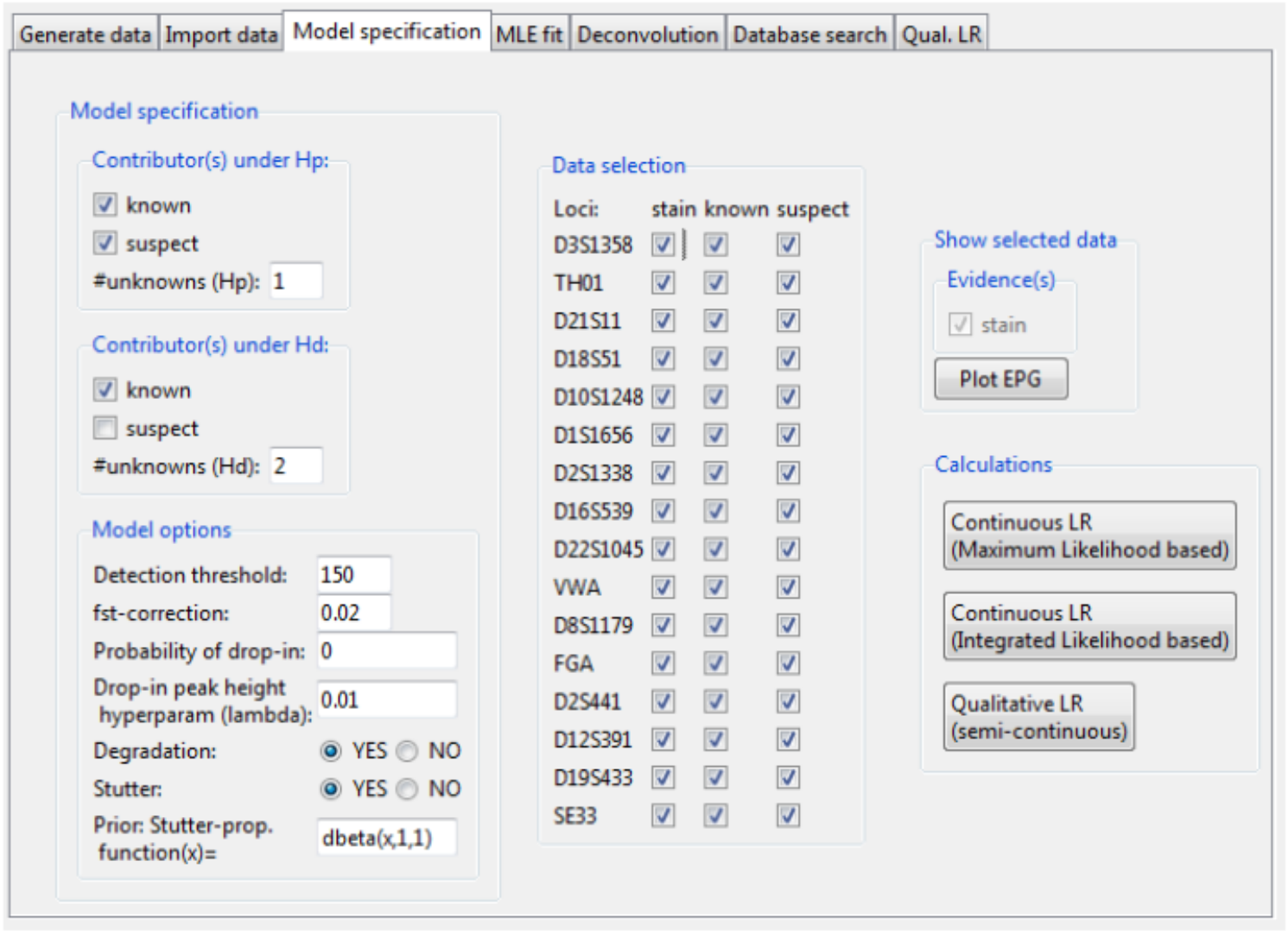
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

The GUI: View data



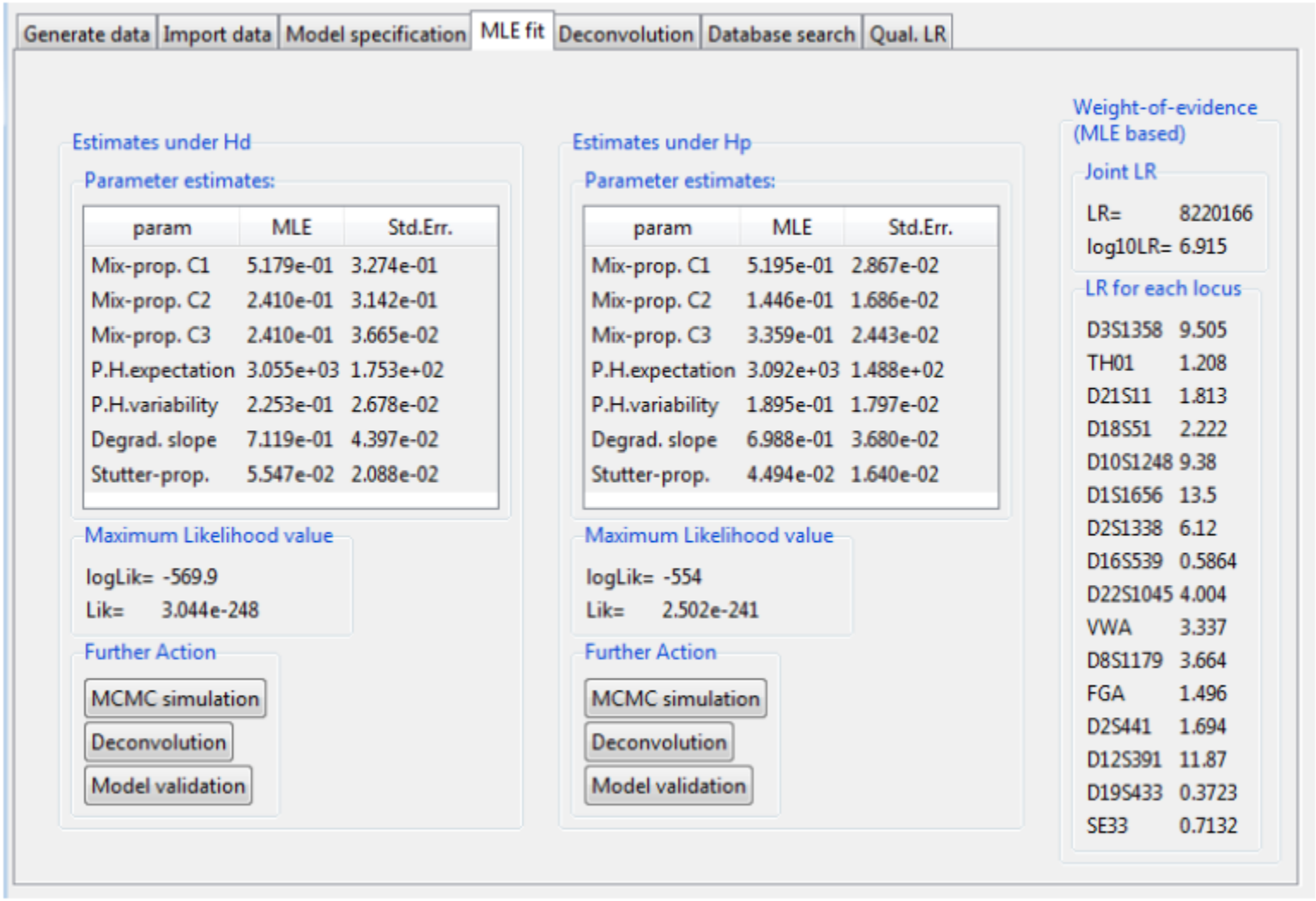
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Model specification



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Results with the ML approach



Time-usage: Less than a minute.

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Model selection

- We investigated the maximum likelihood value (logLik) for several models under the hypothesis H_d.
- We assumed K numbers of contributors.
- Values of $\log_{10} LR_{ML}$ are presented for each model.

K	Stutter	Degrad.	#param	logLik	$\log_{10} LR_{ML}$
3	no	no	4	-587	
3	no	yes	5	-578	
3	yes	no	5	-582	5.1
3	yes	yes	6	-571	7.6
4	no	no	5	-585	4.3
4	no	yes	6	-575	6.7
4	yes	no	6	-582	5.1
4	yes	yes	7	-571	7.6

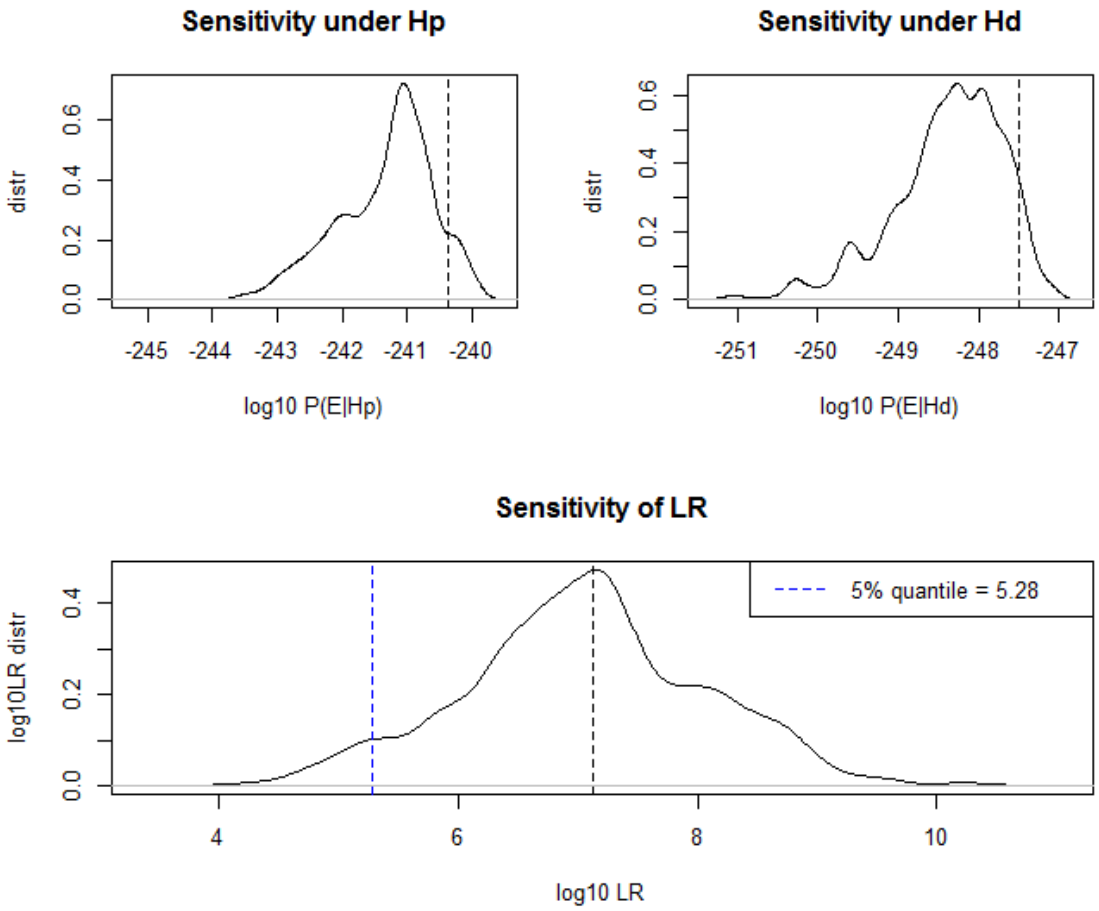
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Model selection

- **The idea is to use the simplest model possible that produces the lowest LogLik parameter**

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

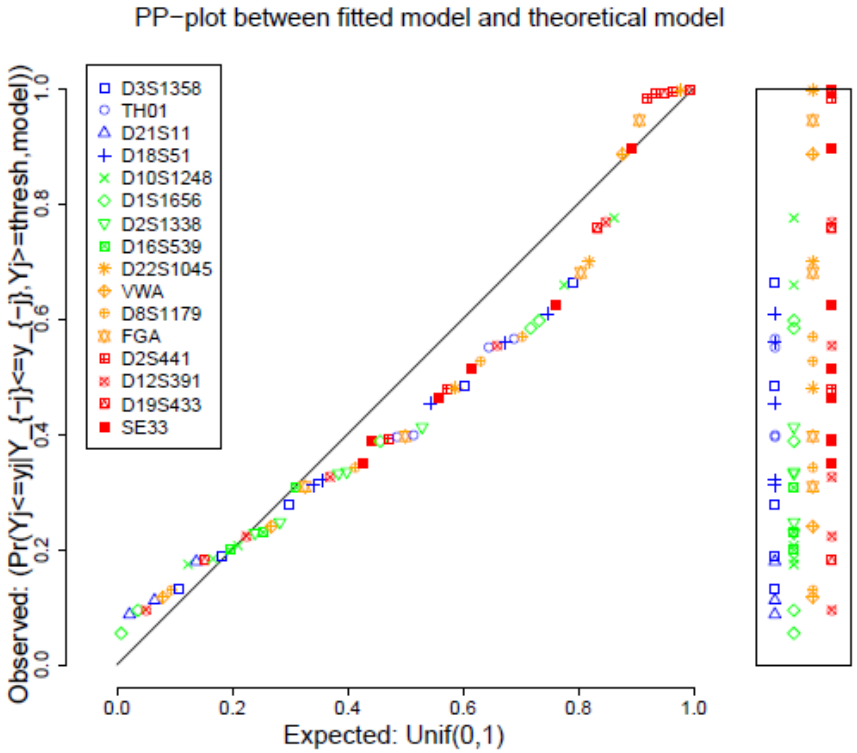
Sensitivity analysis



EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Model validation

- A Probability-Probability plot under H_d was produced to check whether the observed peak heights fits the ML fitted model.



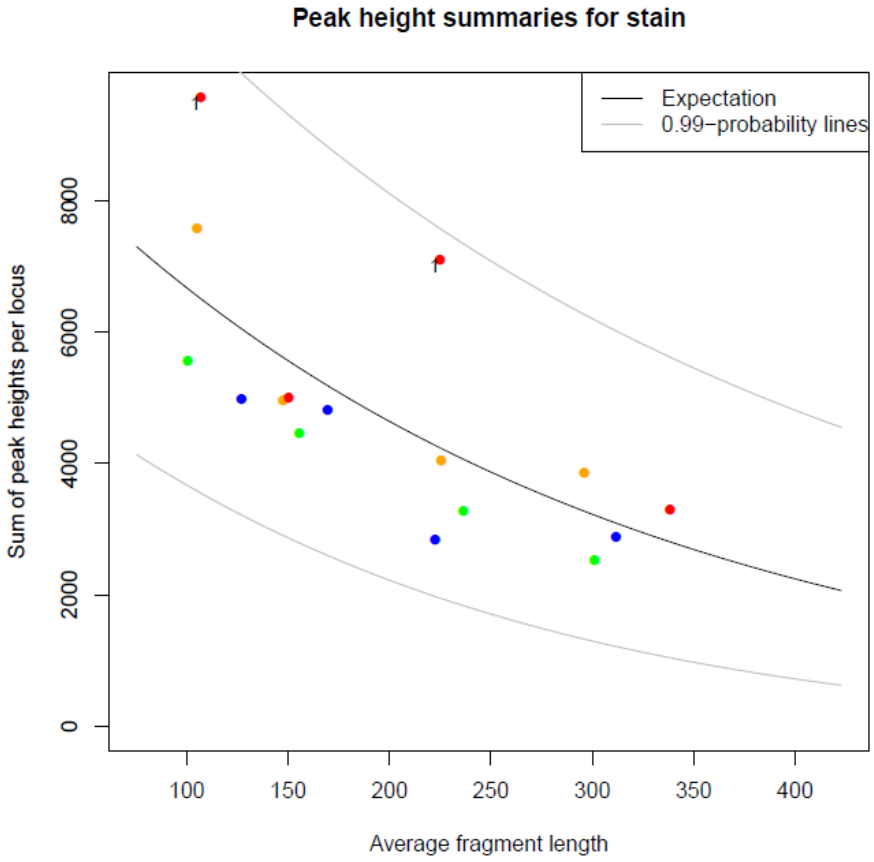
Estimates the cumulative probability of the observed peak heights conditional on the other peak heights

Estimation of the unknown profile

Generate data | Import data | Model specification | MLE fit | **Deconvolution** | Database search | Qual. LR

rank	D3S135	TH01_	D21S11	D18S51	D10S12	D1S165	D2S133	D16S53	D22S10	VWA_g	D8S117	FGA_g	D2S441	D12S39	D19S4	SE33_g	posterior
1	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.1522
2	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.1348
3	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	11/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.07684
4	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	11/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.06801
5	15/16	9/9.3	28/30	12/16	15/16	11/15	20/24	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.02666
6	15/16	9/9.3	28/30	12/16	15/16	11/15	20/24	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.0236
7	15/16	9/9.3	28/30	12/16	15/16	11/15	19/20	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.02306
8	15/16	9/9.3	28/30	12/16	15/16	11/15	19/20	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.02041
9	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/13	18/29	0.01905
10	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/13	18/29	0.01686
11	15/16	6/9	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.01366
12	15/16	9/9.3	28/30	12/16	15/16	11/15	20/24	11/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.01346
13	15/16	6/9	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.01209
14	15/16	9/9.3	28/30	12/16	15/16	11/15	20/24	11/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.01191
15	15/16	9/9.3	28/30	12/16	15/16	11/15	19/20	11/13	17/17	14/14	12/13	22/24	10/14	16/22	13/15	18/29	0.01164
16	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	12/13	22/24	10/14	16/22	13/14	18/29	0.0106
17	15/16	9/9.3	28/30	12/16	15/16	11/15	19/20	11/13	17/17	14/14	13/14	22/24	10/14	16/22	13/15	18/29	0.0103
18	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	11/13	17/17	14/14	12/13	22/24	10/14	16/22	13/13	18/29	0.009612
19	15/16	9/9.3	28/30	12/16	15/16	11/15	20/25	12/13	17/17	14/14	13/14	22/24	10/14	16/22	13/14	18/29	0.009381

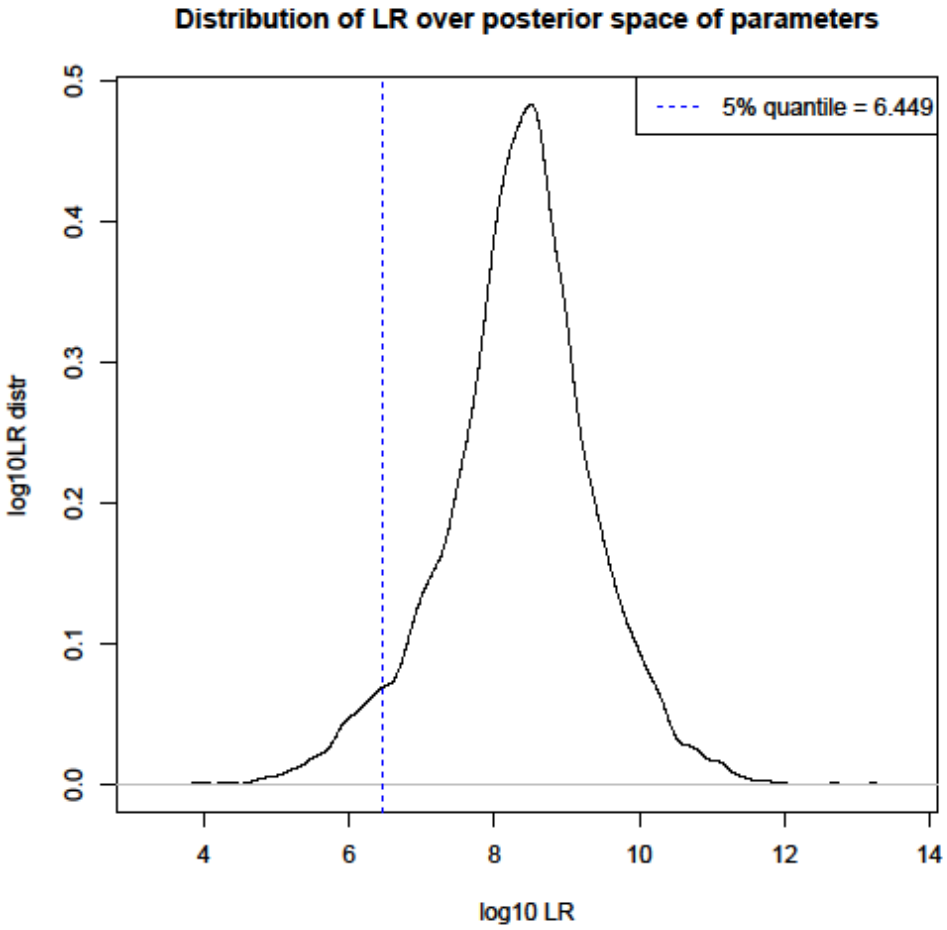
- By conditioning that H_p is true we can estimate the most probable profiles of the unknown individual.
- The method is based on the ML estimated model.
- An individual with reference profile corresponding to rank 4 was later identified.



- Degradation clearly presented.

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Plot from GUI: Sensitivity analysis

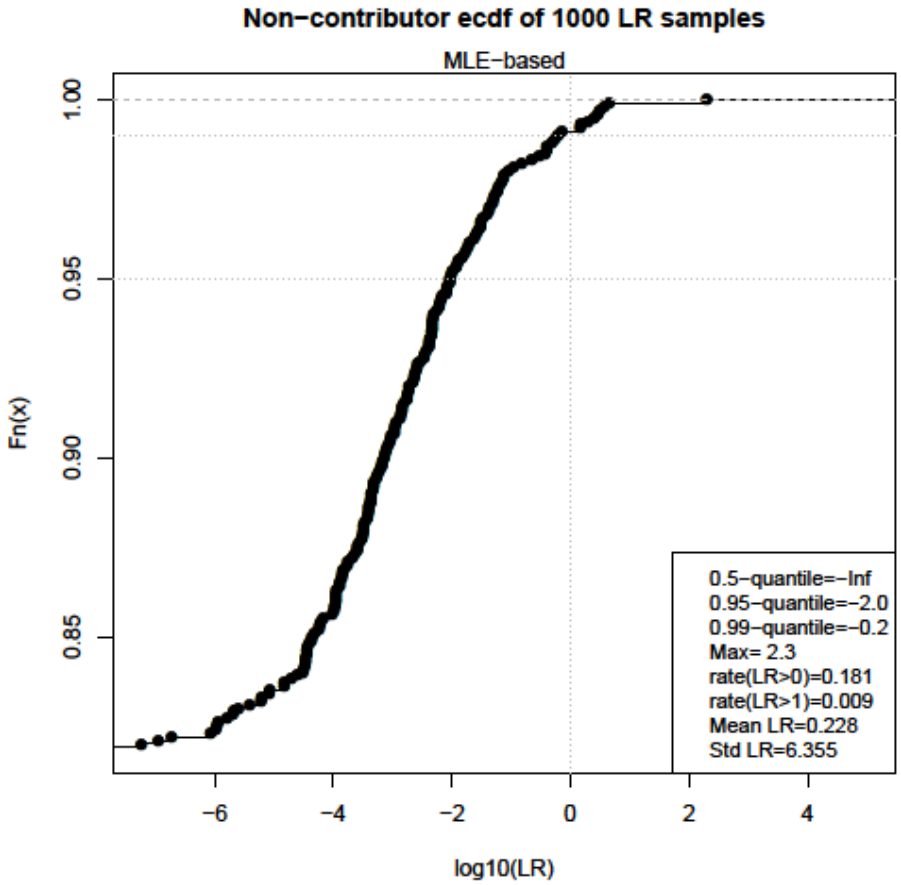


We have shown
that the 5 percentile
minimises the number
of false positive results
(max LR=10 in our validation)

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Plot from GUI: Non-contributor analysis

- Used to investigate that the specified model is not too “Hp-favoring”



Accessibility

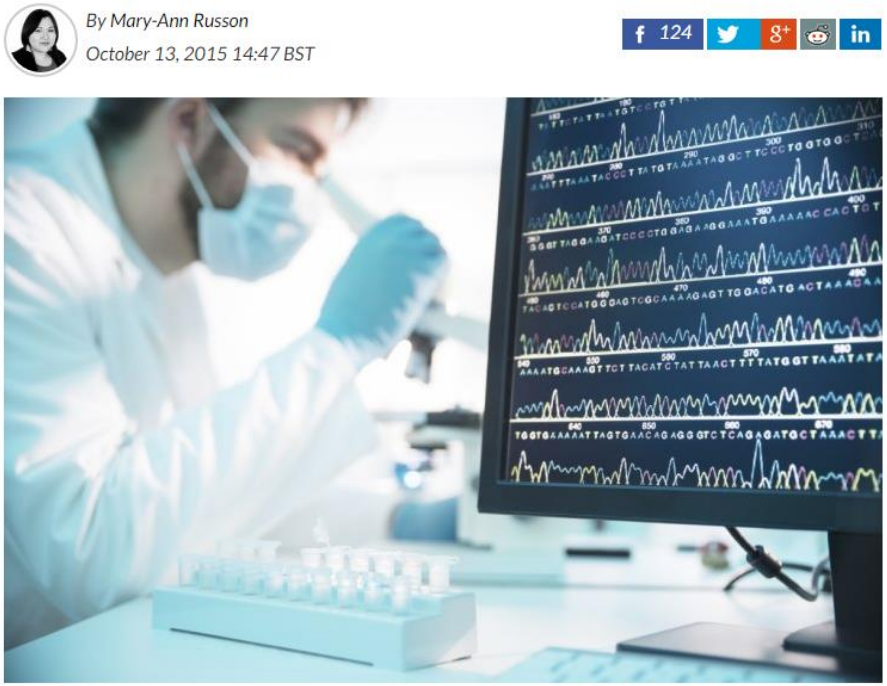
- EuroForMix is open-source and freely available through the R-package **euroformix** which is downloadable from:

- Homepage:
  - www.euroformix.com
  - There is a tutorial, manual and a vignette which explains all technical details.
  - Being worked on to provide more examples

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Court experiences

US: Murder suspect challenging DNA software used to incriminate him by seeking source code



By Mary-Ann Russon
October 13, 2015 14:47 BST

f 124

Lawyers for a US man who faces the death penalty are trying to disprove the DNA identification software that was used to incriminate him (iStock)

The defence lawyers for a man charged with two counts of homicide in the US are trying to gain access to the source code of the DNA software program that incriminated him.

In the US, a court challenge to disclose commercial software was rejected by the judge

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Court experiences
R. v. Mathew Hamlen

- Challenge to the use of a commercial program
- Defence argued that they were unable to evaluate the software/ data in the case (also it had been introduced at relatively short notice by the prosecution)
- The commercial provider agreed to release the software for this purpose
- However there was insufficient time within the trial to allow for a case evaluation to happen.
- The judge ultimately decided to exclude the DNA evidence, based on the inability to challenge it sufficiently in the time frame given.
- One other case in Scotland where there was a successful challenge on the basis that the software could not be challenged

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Court experiences
R v. Waseem Fazal

- Sheffield Crown Court May 2015
- A complex DNA profile was obtained from a gun handle where the following propositions were agreed:
  - Hp = Mr Fazal + 2 unknown and unrelated individuals
  - Hd = 3 unknown individuals unrelated to Mr Fazal
- The original report from South Yorkshire Police gave strength of evidence of 1 in 1 billion using the ‘matching allele count method’
- The data were analysed by the prosecution scientist using LiRa and the LR=2.9 million
- The defence carried out an analysis with LRmix Studio and the LR=300,000

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Procedure under rule 33.6

- A joint report of experts is written that states the areas of agreement and disagreement arising from the conclusions drawn in relation to the DNA profile result obtained from the swab from the gun

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Agreement

- We agree that:
- a) Different solutions have been published in the peer reviewed scientific literature to evaluate the weight of evidence associated with complex DNA profiles. The different solutions are expected to generate different numerical values.
-
- b) There is no recommended method or gold standard. However, the software methods used in this case, namely LiRa (LGC Forensics) and LRmixStudio (as used by the Netherlands Forensic Institute) are accepted as validated methods for use in forensic casework.

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

- **We agree that:**
- a) The statistics calculated by the different methods in this case are:

LiRa	LRmixStudio
2.9 million	300, 000 (three hundred thousand)

- b) To place the significance of the above divergence into perspective, Steele and Balding [1] suggest that a difference of about one order of magnitude, e.g. 300,000 vs 3 million, is negligible.
- c) To aid understanding, and to assist the court, we suggest that it may be preferable to convert the numeric strength of evidence into a verbal scale, which operates on an order of magnitude basis, that has been proposed by the European Network of Forensic Science Institutes (ENFSI) [2] and reproduced in its entirety below:

Verbal scale

10,000 - 1,000,000	...provide very strong support for the first proposition rather than the alternative …are far more probable given… proposition...than proposition...
1,000,000 and above	...provide extremely strong support for the first proposition rather than the alternative …are exceedingly more probable given… proposition...than proposition...

	LiRa	LRmixStudio
Statistical evaluation	2.9 million	300, 000 (three hundred thousand)
Verbal equivalent	Extremely strong	Very strong

Conclusion

- Therefore, in our opinion, the forensic findings provide [very strong to extremely strong] support for the proposition that the DNA profile has originated from Mr Fazal and two unknown unrelated contributors rather than from three unknown contributors (who are unrelated to Mr Fazal).

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

R. v. Bailey

- DNAmixtures (gamma model) reported a LR 10^{21}
- Compared with LRmix where LR> 1bn
- Therefore there was complete agreement on the strength of the evidence

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Conclusions relating to the reporting of complex DNA profiles in general

- **What does recent experience show?**
  - There is no agreed method
  - Experience is showing that ‘equality of arms’ between defence and prosecution is pre-requisite, otherwise cases may be rejected on the basis that they ‘cannot be properly challenged or evaluated’
  - There is no requirement for defence experts to use the same method as the prosecution as there is no best method. Hence we see a developing tendency for open source to be used to challenge expensive commercial software. Time constraints imposed at trial preclude the option of temporary access that may be given by the provider.
  - The difficulty (for the court to understand) is that we expect differences in the results
  - If the methods are validated, then the results are both correctly reported
  - Use of the verbal statement does help to put this into perspective

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Verbal scales

- The verbal scale operates on orders of magnitude, which is useful.
- The verbal scale limited at 1bn - unlikely that greater power is needed.

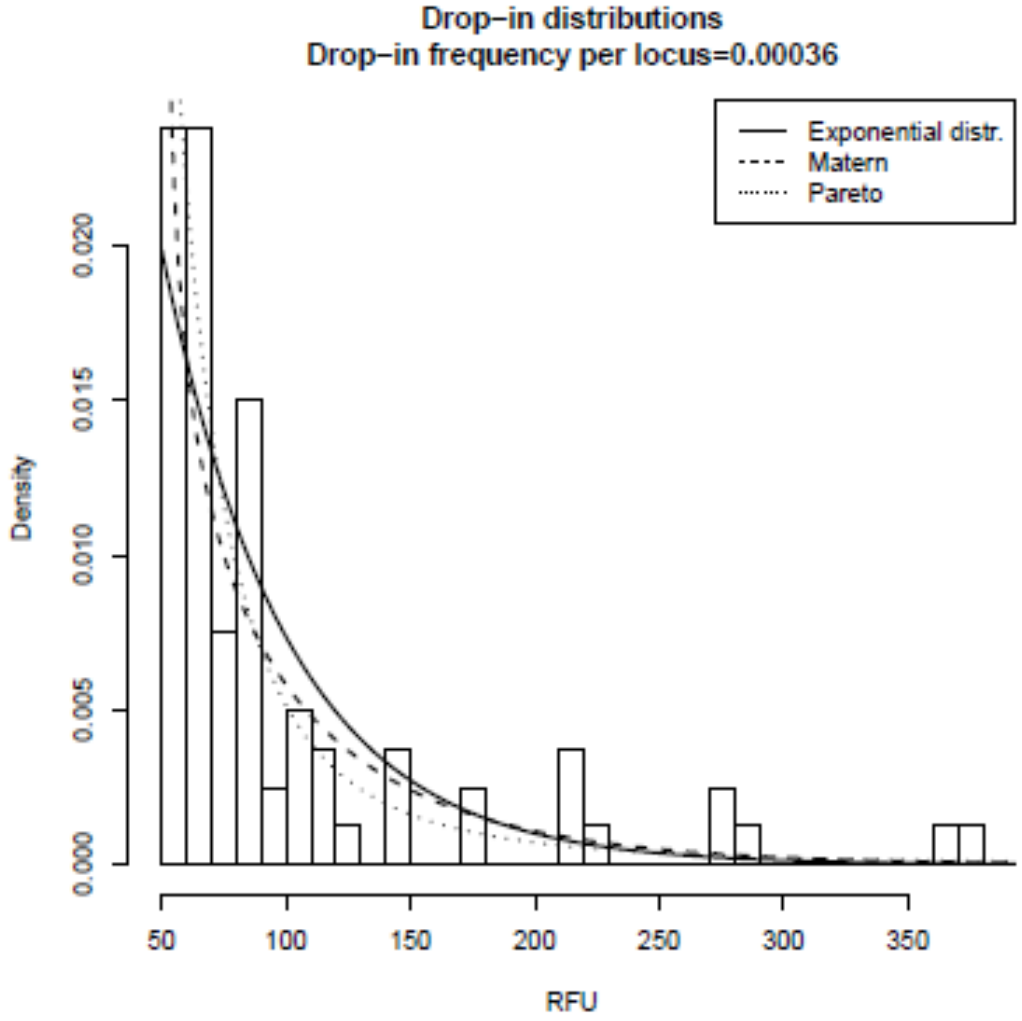
EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Drop in

- NFI NGM dataset were analysed
- A set of N = 14757 negative control samples were generated From this data, a total of x = 80 false positive alleles were found (excluding the amelogenin marker), so that the relative frequency of drop-in per STR marker (out of total
- L = 15 markers) was estimated as $x * L/N$ = 0:00036
- The maximum likelihood estimate for the rate parameter is

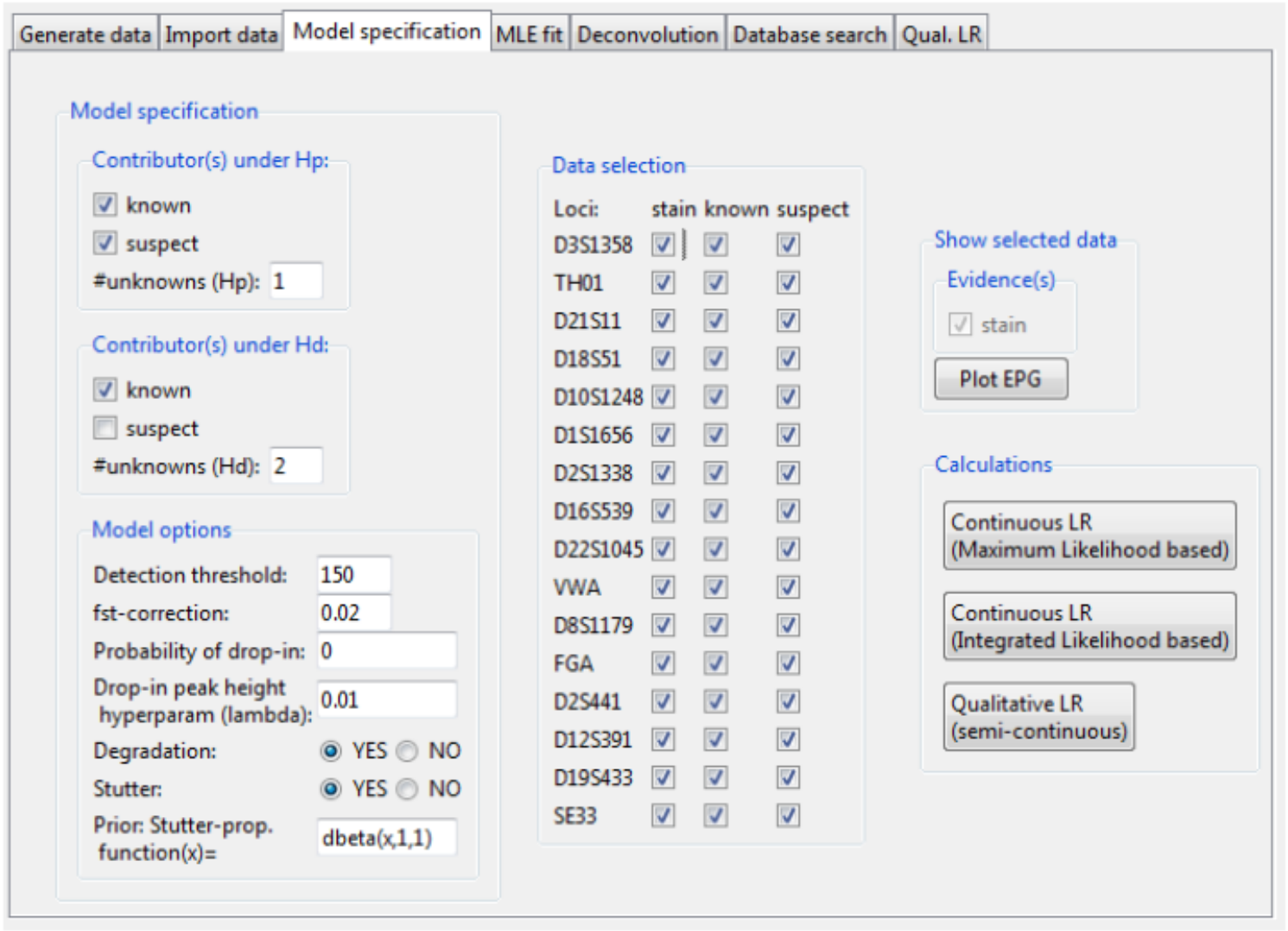
= 0:02, and this was used as a plug-in value to model the drop-in peak height in EuroForMix

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme



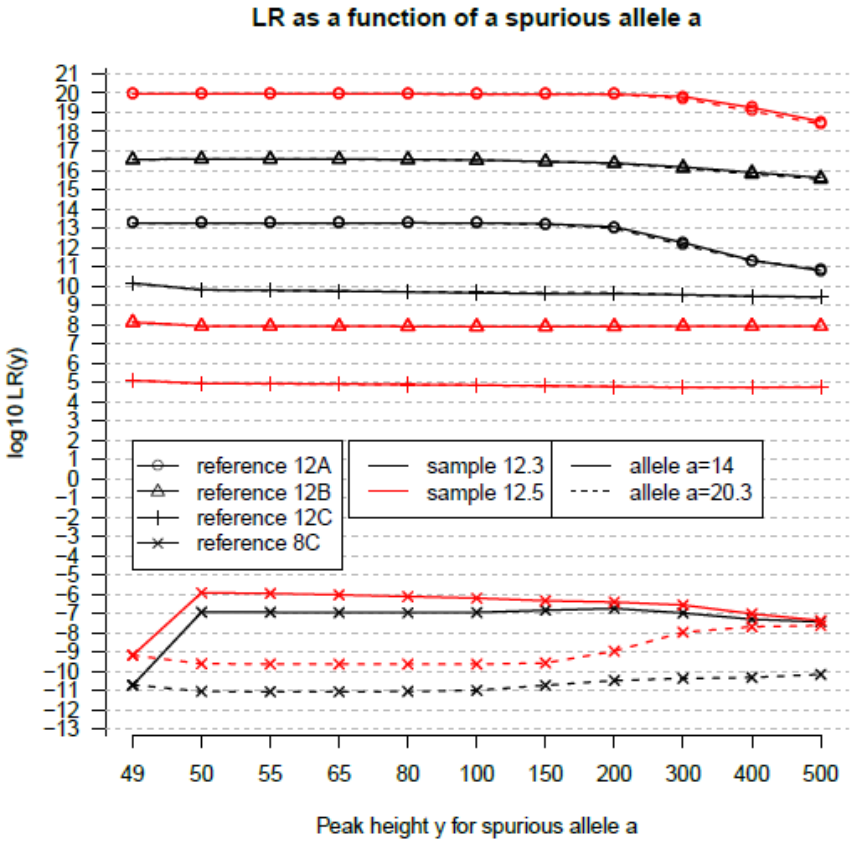
EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Model specification



EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

Effect of a single drop in event on the LR vs peak height (mixture of 12A,12B,12C)



DNA commission on the evaluation of evidence

Why?

- There have been two commissions on mixture analysis
- Over recent years, there has been development of theory, along with some examples of evaluation of evidence that goes beyond the DNA profile itself – to incorporate the activity of the transfer

Part I

- Where are we with forensic genetics?
- Do we adhere to the principle of ‘conservativeness’ in everything we do?
- Several methods of complex mixture analysis are in use (see other commissions). We don’t intend to advise on which models to use. But we will advise on general usage.
- Numbers of contributors
- Contamination
- Secondary transfer
- How to prepare propositions

What about haploid?

- Although there have been commissions on Y chromosome and mtDNA, we havent formulated clear guidelines on newer methods of analysis, other than the counting method
- Method of Charles Brenner
- Method of Andersen (discrete Laplace method)
- Do we accept these as useful alternatives since we currently greatly underestimate strength of evidence of singletons.

The likelihood ratio

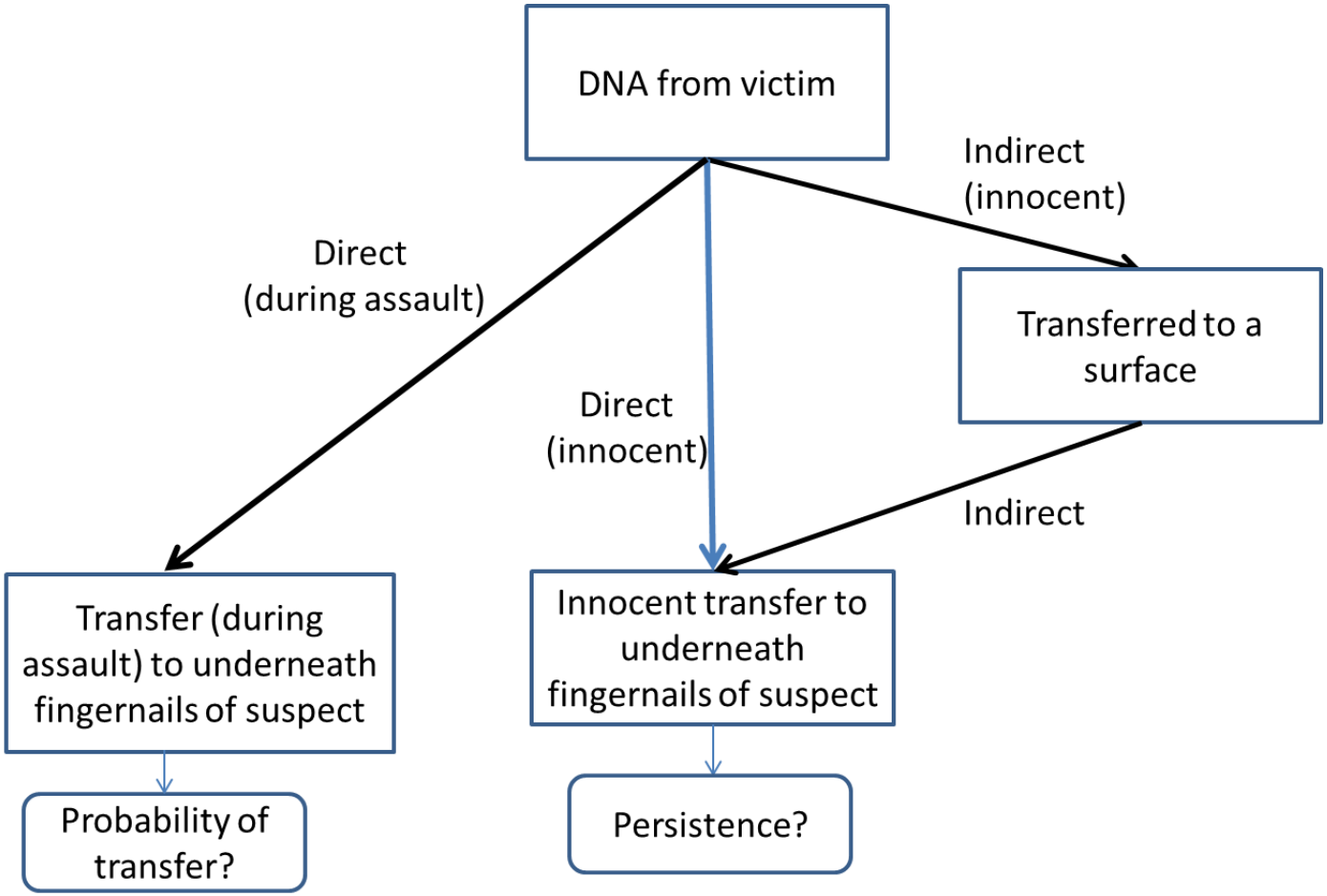
- Discussed in earlier commissions – should be standard by now as it is impossible to use any other method for most modern tests

Part II – beyond the DNA profile

The hierarchy of propositions

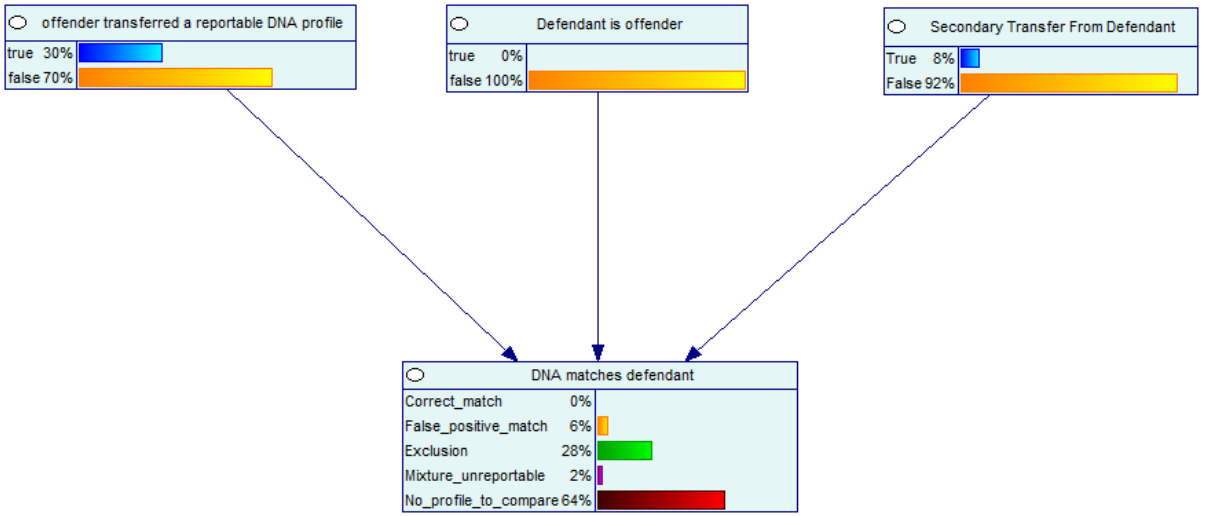
- A standard? way to think about evidence
  - Sub-source (the fact of the DNA profile)
  - Source (did the DNA come from a defined body fluid
  - Activity (how did the transfer occur)
  - Ultimate issue (Guilt/innocence)
- The probability (LR) associated with the sub-source cannot be automatically be transposed to the source – or the activity level

Example provided by DNA underneath fingernails



Likelihood ratio of interest at activity level

- $LR = \dfrac{Pr\ of\ direct\ transfer\ matching\ defendant\ and\ no\ secondary\ transfer}{\Pr\ of no\ direct\ transfer\ and\ secondary\ transfer\ matching\ defendant.\ Or\ direct\ transfer\ matching\ defendant\ from\ an\ unknown\ unrelated\ assailant}$



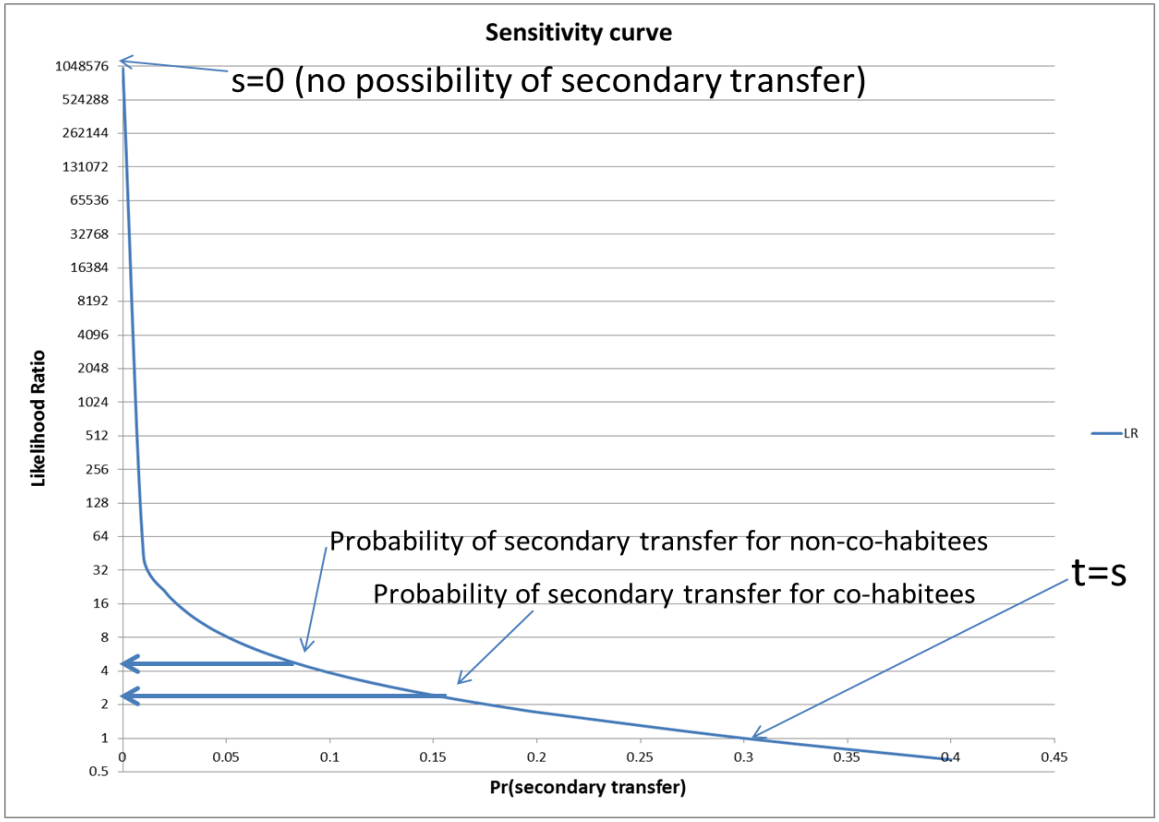
Consider all possibilities

- Offender transferred a reportable DNA profile: either true or false
- Defendant is the offender: either true or false
- There was secondary transfer from the defendant: either true or false
- The DNA matches the defendant: either a correct match, a false positive match, an exclusion, a mixture or no profile to compare

Data

- Data are needed of secondary transfer and direct transfer
- Experimental design needs to accommodate both
- Background experiments
- Scratching experiments

LR for Pr(direct transfer)=0.3



Limitations

- The probabilities used to inform probabilistic models are dependent upon the experimental design.
- Different papers in the literature can reveal markedly different results. This is because of different experimental designs in use.

Part III

- Subjective vs Objective inference
- Forensic genetics is based upon very sound theory
- We must ensure that we maintain an equivalent standard for all other aspects relating to interpretation of evidence ‘beyond the DNA profile’

Knowledge bases

- Probabilities depend upon a knowledge base (which might include ‘expert opinion’)
- However if two experts use two different knowledge bases then we get two different answers.

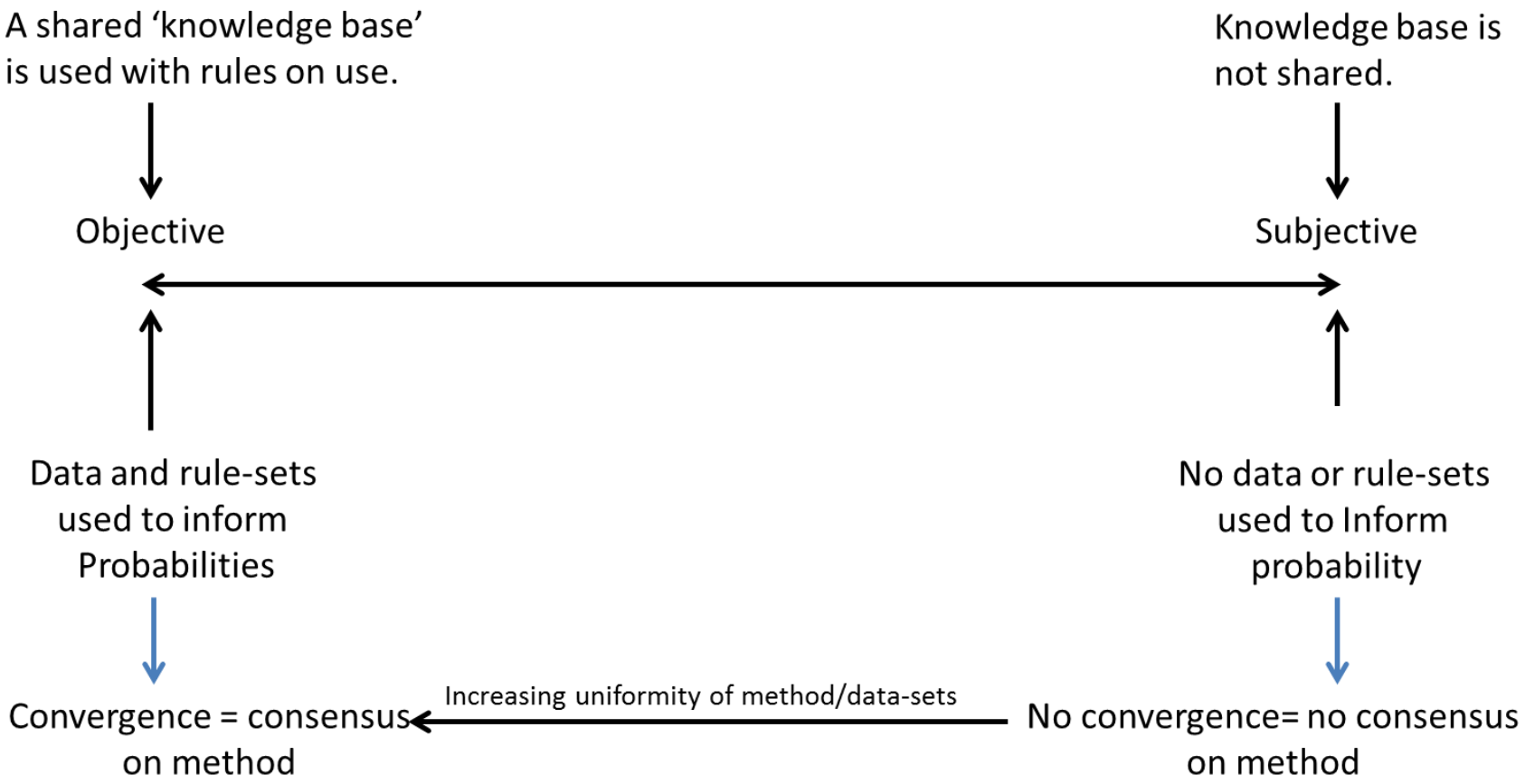
Expert opinion and cognitive bias

- If opinion is not based on a coherent dataset then there is an increased danger of cognitive bias – or ‘belief’
- Belief is a state of mind when someone thinks something is reality, true, when they have no absolute verified foundation for their certainty of the truth or realness of something.

The scientific method

- "a method or procedure that has characterized natural science since the 17th century, consisting of systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses."

Subjectivity vs objectivity



Conclusion

- As the community carries out more research into non-forensic-genetics, we need to be sure that:
  - Experimental designs are sound
  - We develop standard ways to interpret the evidence that are open-source /freely available
  - We adopt the principles of peer review
  - We adopt the same level of quality/ standards that are used for forensic genetics

EUROFORGEN / EDNAP mRNA NGS exercise 1

Assay for body fluid/tissue identification

Cordula Haas / Sabrina Ingold /
Erin Hanson / Jack Ballantyne

26. April 2016, Warsaw

Association of a Body Fluid with a DNA Profile by Targeted RNA/DNA Deep Sequencing

Cordula Haas*, Sabrina Ingold*, Erin Hanson°, Jack Ballantyne°
*University of Zurich, °University of Central Florida

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Objectives

1. set up a targeted mRNA/miRNA NGS approach for body fluid/tissue identification
   → establish a probabilistic approach to call/predict the presence of a body fluid

2. select a set of SNPs for each body fluid/tissue, that discriminates individuals the most
   → assign a body fluid to a specific individual

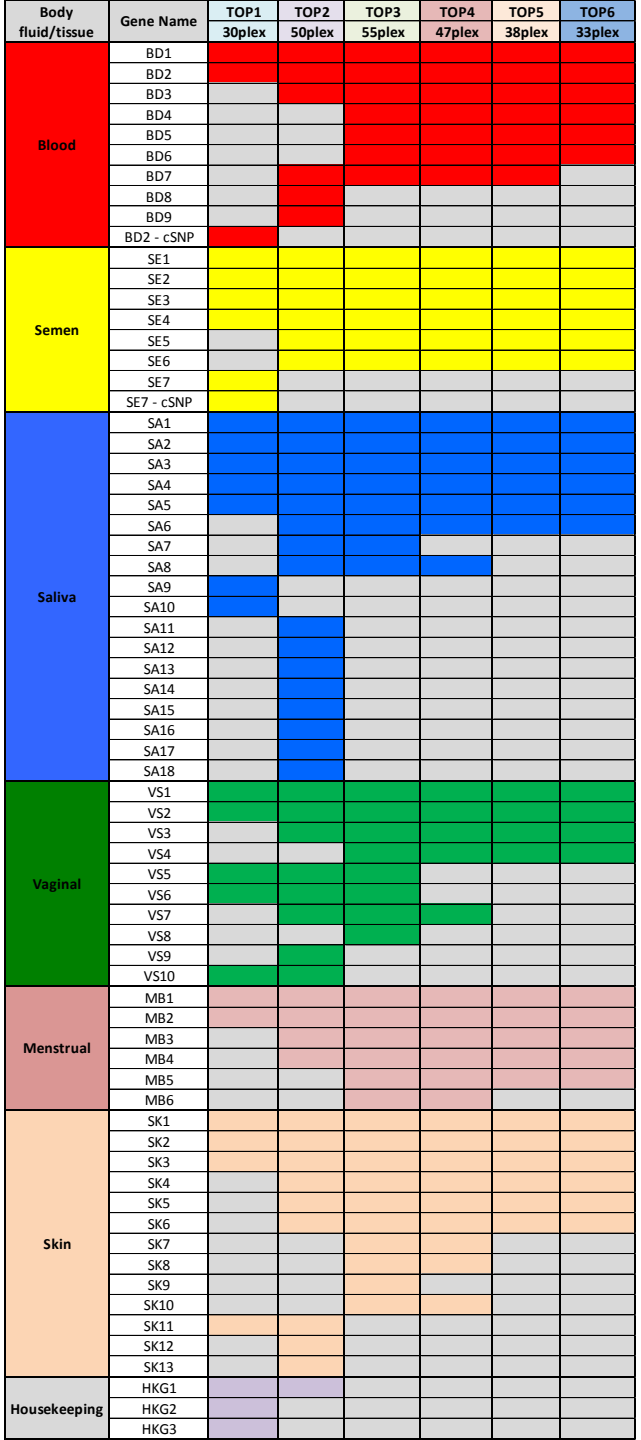
3. combine the RNA analysis with gDNA STR sequencing, allowing simultaneous human individual identification and forensic tissue identification

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

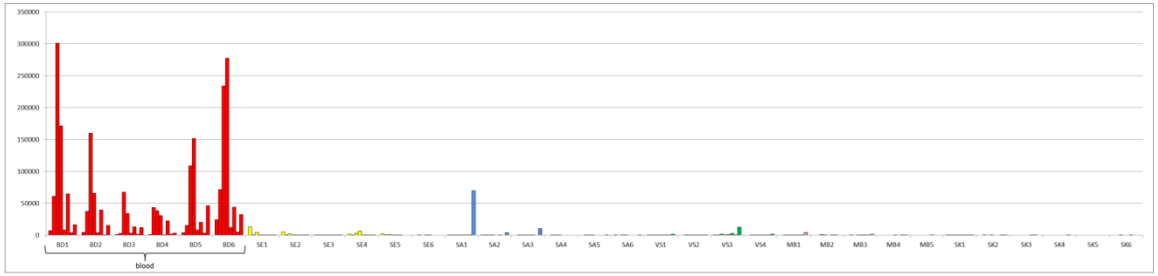
1A. targeted mRNA NGS approach for body fluid/tissue identification (MiSeq)

- Illumina DesignStudio
- TruSeq Targeted RNA Custom Panel
- TruSeq Targeted RNA Index Kit
- Illumina MiSeq
- Bioinformatics pipeline

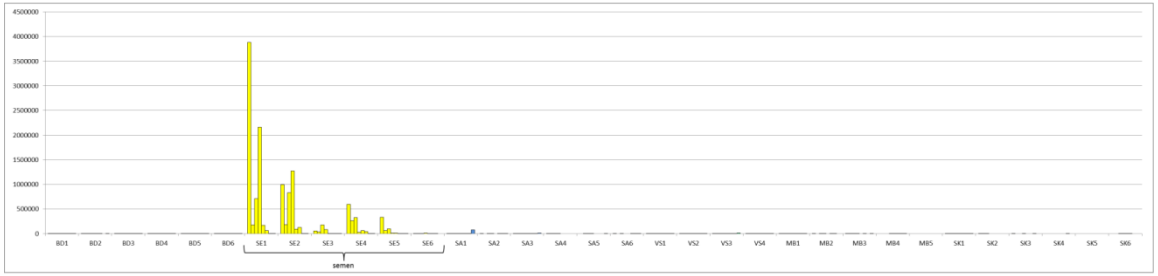
- blood, semen, saliva, vaginal secretions, menstrual blood, skin
- 66 mRNA biomarkers evaluated
- TOP6: 33 biomarkers



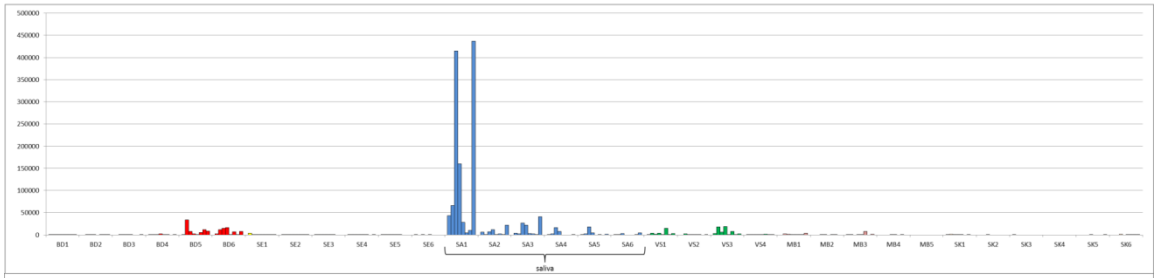
Blood
(n=8)



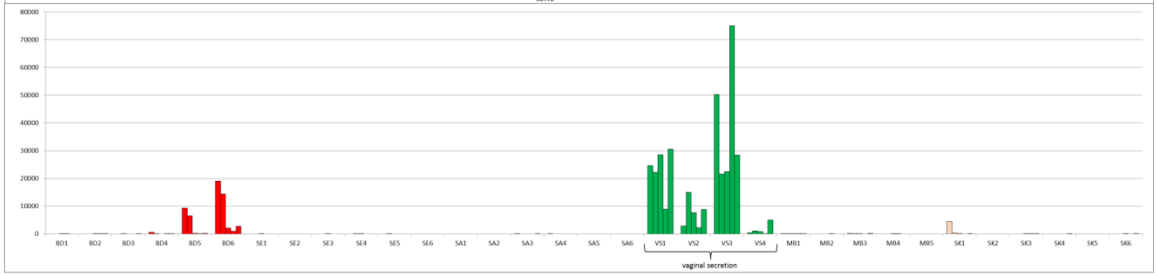
Semen
(n=8)



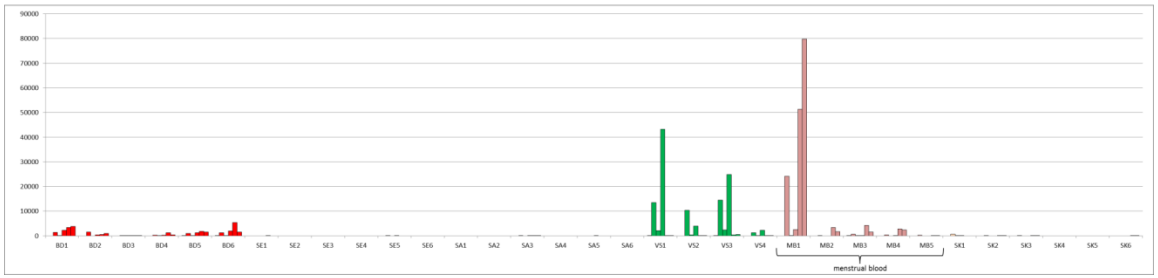
Saliva
(n=8)

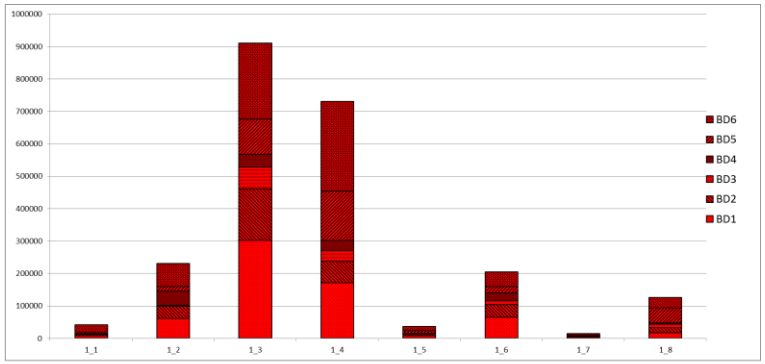


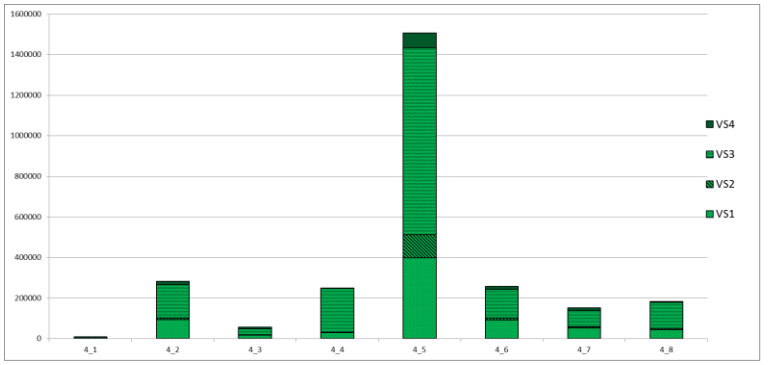
Vaginal
secretion
(n=8)

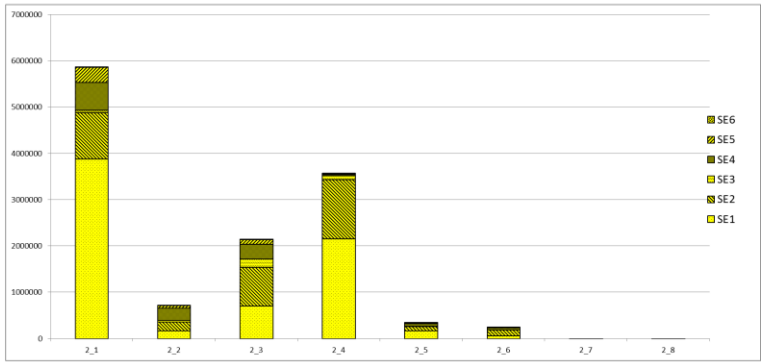


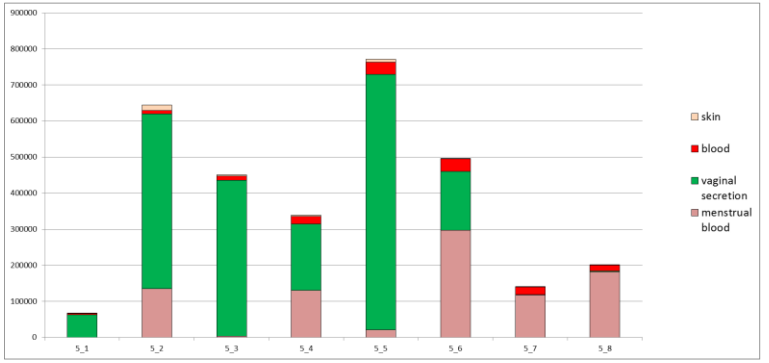
Menstrual
blood
(n=8)

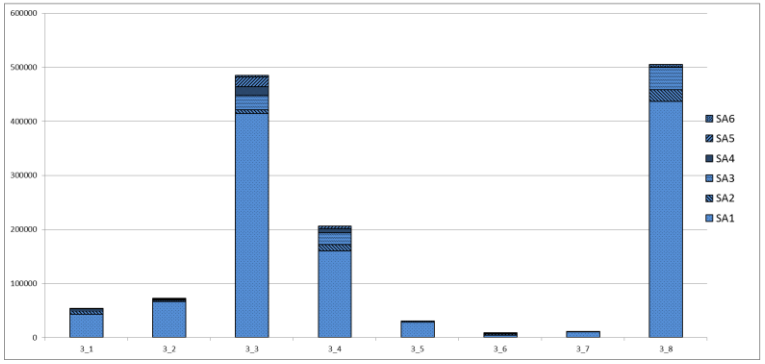


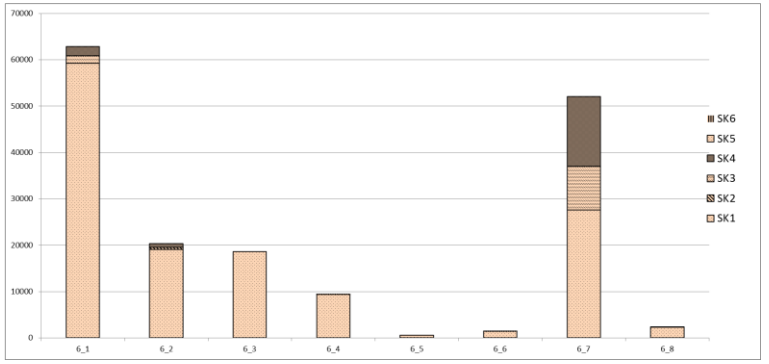

1000000
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
1_1
1_2
1_3
1_4
1_5
1_6
1_7
1_8
BD6
BD5
BD4
BD3
BD2
BD1


1600000
1400000
1200000
1000000
800000
600000
400000
200000
0
4_1
4_2
4_3
4_4
4_5
4_6
4_7
4_8
VS4
VS3
VS2
VS1


7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
2_1
2_2
2_3
2_4
2_5
2_6
2_7
2_8
SE6
SE5
SE4
SE3
SE2
SE1


900000
800000
700000
600000
500000
400000
300000
200000
100000
0
5_1
5_2
5_3
5_4
5_5
5_6
5_7
5_8
skin
blood
vaginal secretion
menstrual blood

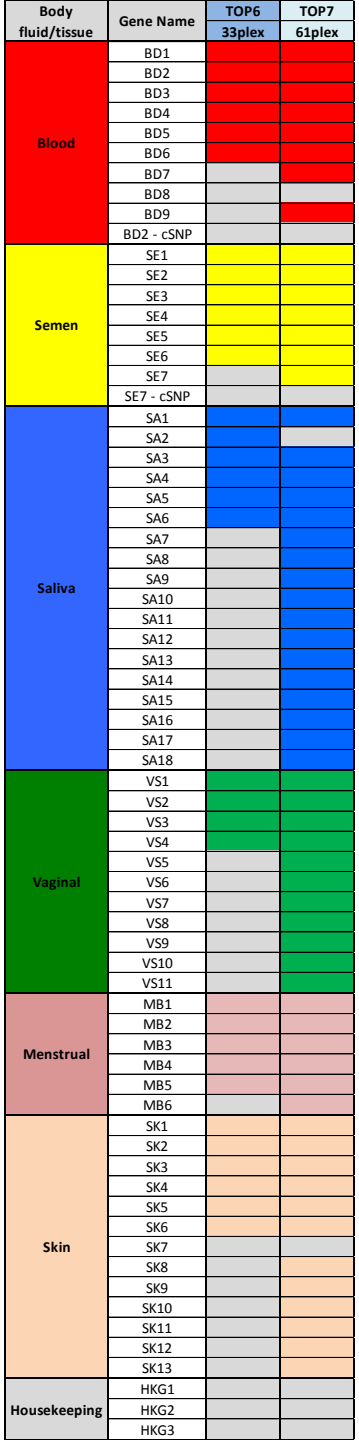

600000
500000
400000
300000
200000
100000
0
3_1
3_2
3_3
3_4
3_5
3_6
3_7
3_8
SA6
SA5
SA4
SA3
SA2
SA1


70000
60000
50000
40000
30000
20000
10000
0
6_1
6_2
6_3
6_4
6_5
6_6
6_7
6_8
SK6
SK5
SK4
SK3
SK2
SK1

1B. targeted mRNA NGS approach for body fluid/tissue identification (PGM)

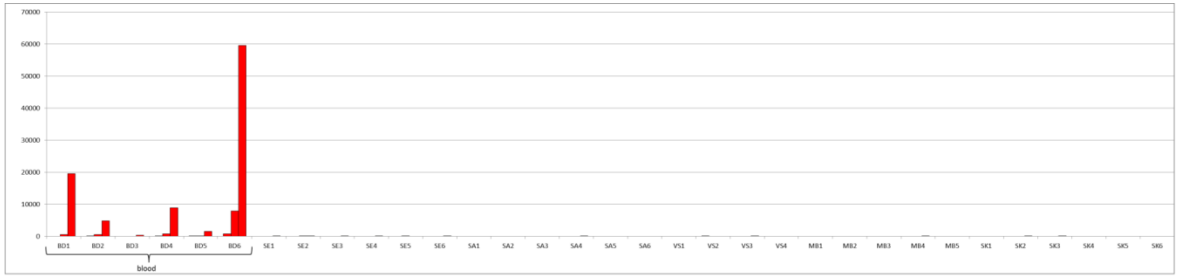
- Ion AmpliSeq Designer
- AmpliSeq RNA library preparation kits
- IonTorrent PGM
- Bioinformatics pipeline

- same 33 mRNA biomarkers

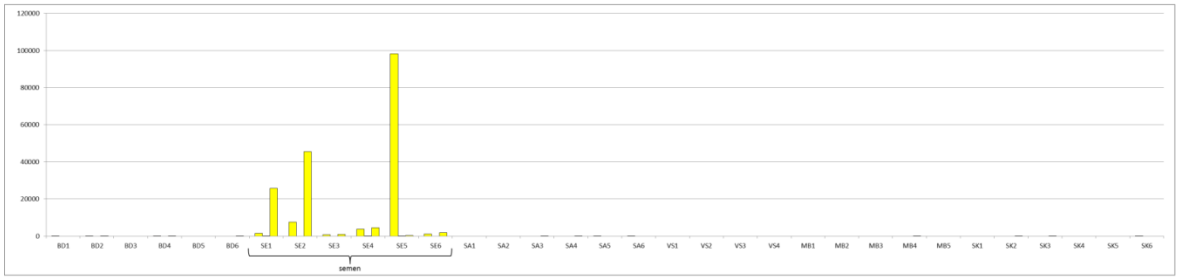


Body fluid/tissue	Gene Name	TOP6 33plex	TOP7 61plex
Blood	BD1		
	BD2		
	BD3		
	BD4		
	BD5		
	BD6		
	BD7		
	BD8		
	BD9		
	BD2 - cSNP		
Semen	SE1		
	SE2		
	SE3		
	SE4		
	SE5		
	SE6		
	SE7		
	SE7 - cSNP		
Saliva	SA1		
	SA2		
	SA3		
	SA4		
	SA5		
	SA6		
	SA7		
	SA8		
	SA9		
	SA10		
	SA11		
	SA12		
	SA13		
	SA14		
	SA15		
	SA16		
	SA17		
	SA18		
Vaginal	VS1		
	VS2		
	VS3		
	VS4		
	VS5		
	VS6		
	VS7		
	VS8		
	VS9		
	VS10		
	VS11		
Menstrual	MB1		
	MB2		
	MB3		
	MB4		
	MB5		
	MB6		
Skin	SK1		
	SK2		
	SK3		
	SK4		
	SK5		
	SK6		
	SK7		
	SK8		
	SK9		
	SK10		
	SK11		
	SK12		
	SK13		
Housekeeping	HKG1		
	HKG2		
	HKG3		

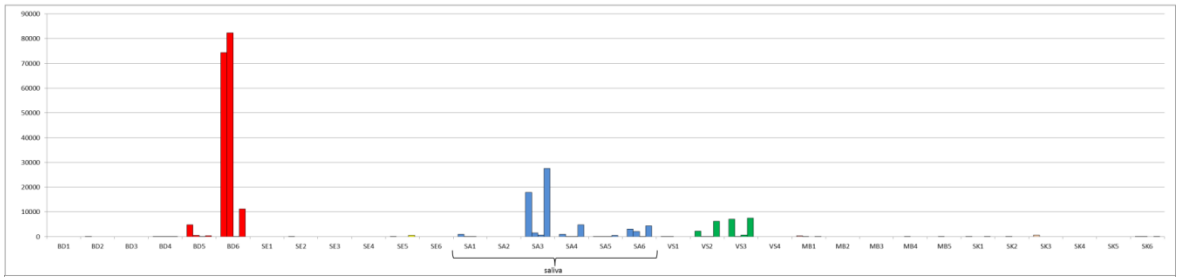
Blood
(n=3)



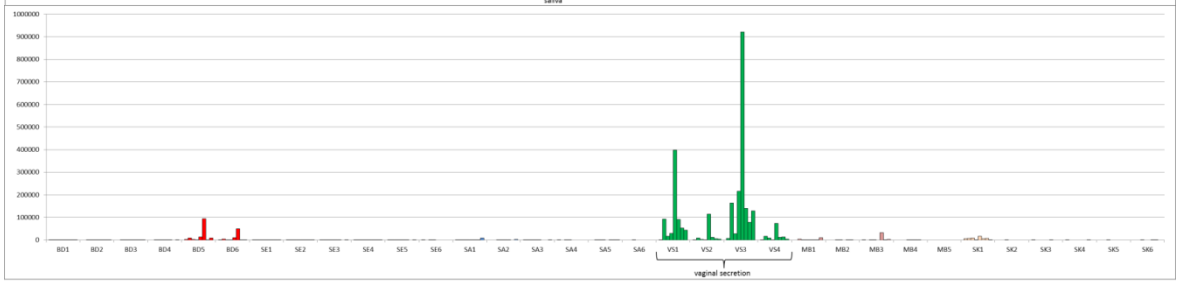
Semen
(n=3)



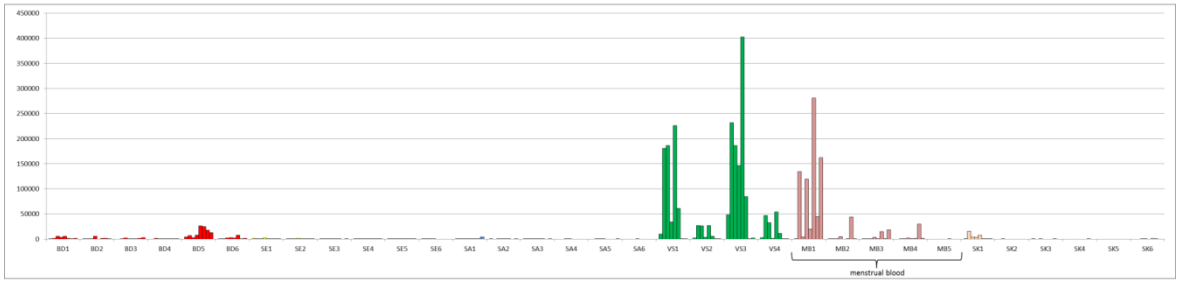
Saliva
(n=4)



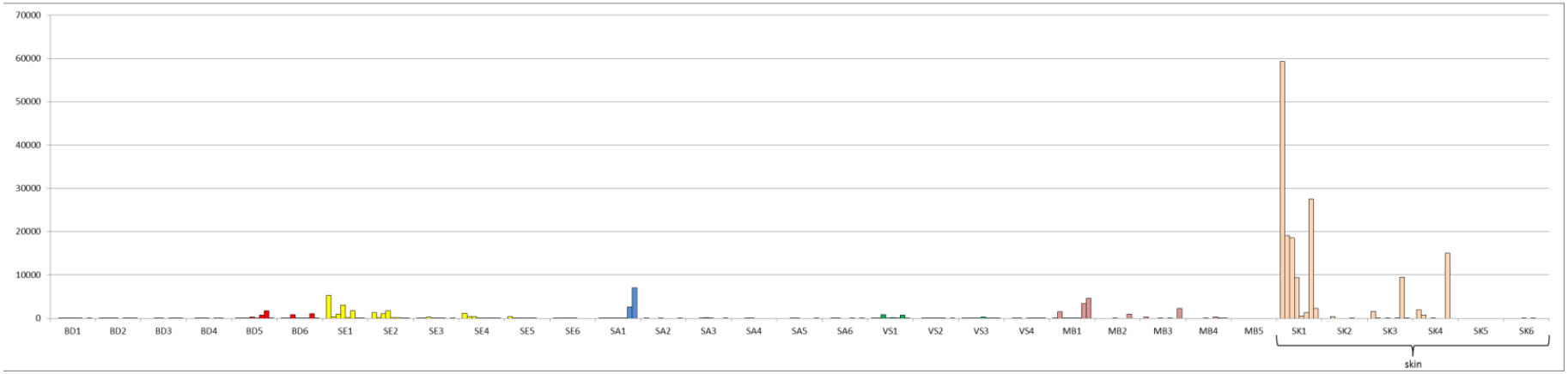
Vaginal
secretion
(n=5)



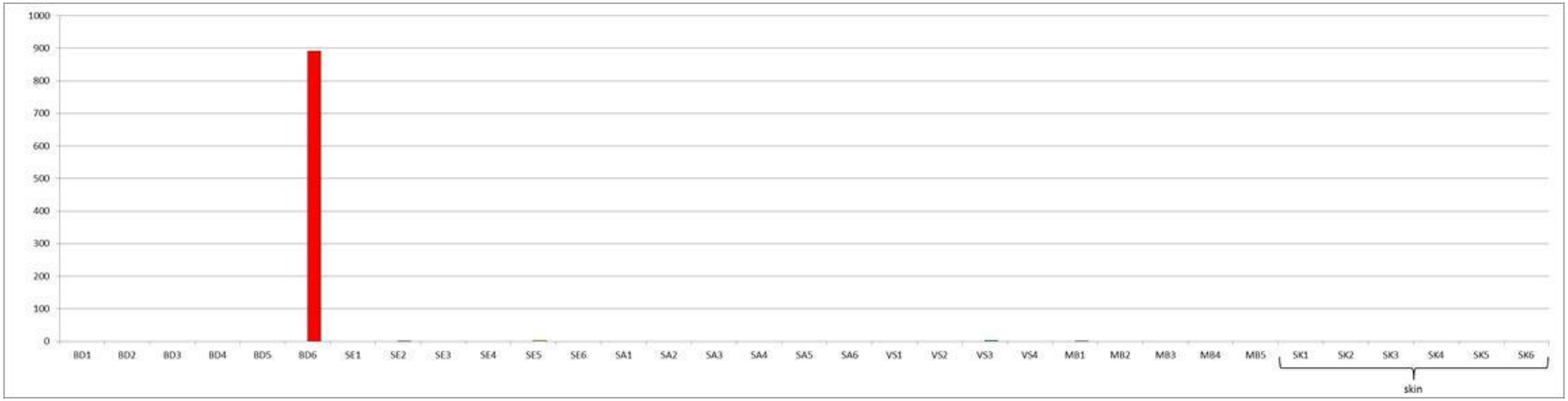
Menstrual
blood
(n=6)



Illumina MiSeq
Skin (n=8)



PGM
Skin (n=2)



1C. miRNA NGS approach for body fluid/tissue identification

- whole miRNome
- miRNeasy mini kit
- DNase treatment
- total RNA quantitation
- NEBNext multiplex small RNA library Prep Set 1
- Illumina HiSeq
- Bioinformatics pipeline

2 Experiments:

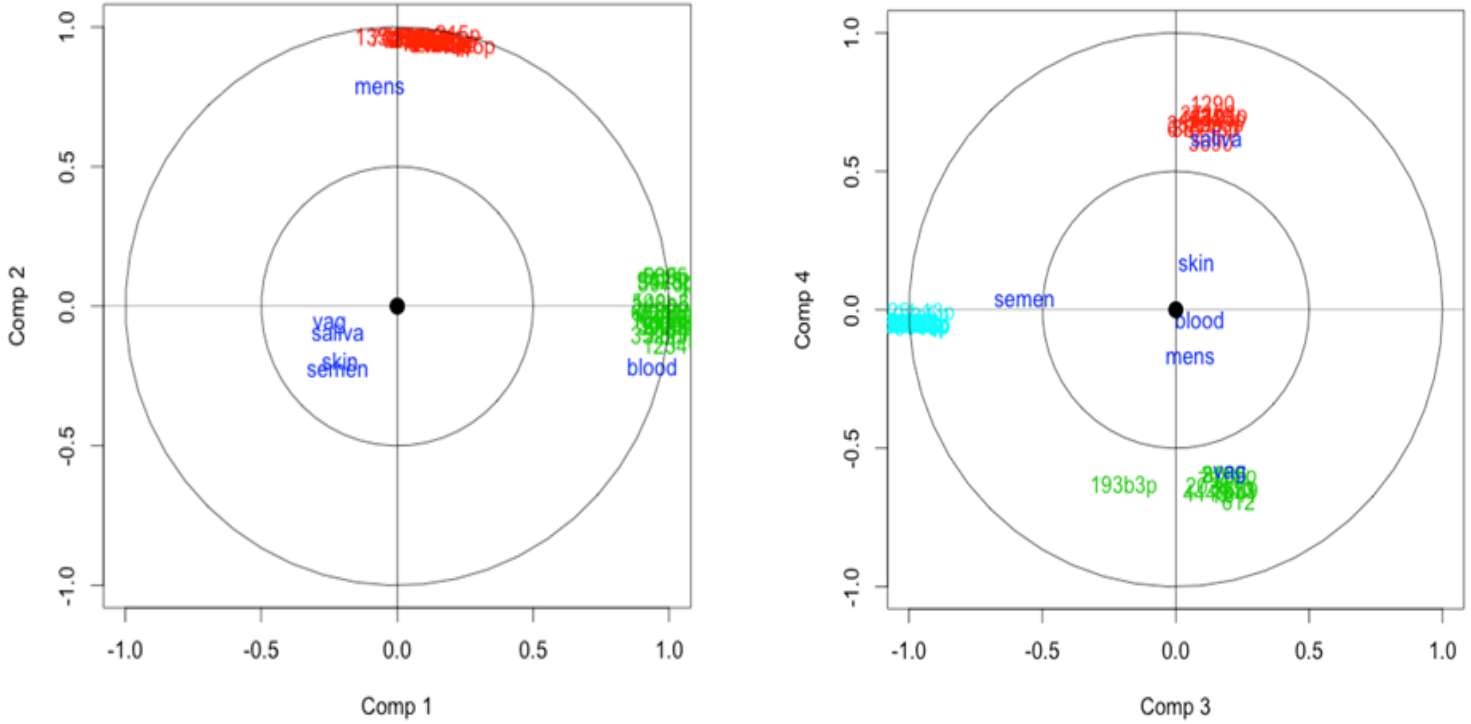
- 2 samples per body fluid:
  fresh samples (directly into lysis buffer)
  vs. 'aged' samples
  (1-4 weeks on swab at RT)
- pools of 6 samples per body fluid

PGM assay prone to adapter-dimers?

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

1C. miRNA NGS approach for body fluid/tissue identification

Correlation loading plots from miRNA results (8 samples per body fluid)



G. Dorum (NIHP)

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

2. Body fluid/tissue specific SNPs

→ associate specific mRNA transcripts to an individual (on mRNA)

- in preparation
- IonTorrent PGM and Illumina MiSeq

→ estimate RNA-SNP allele frequency by testing of population samples (on DNA)

- 100 cSNPs
- Ion AmpliSeq Designer
- Ion AmpliSeq Library preparation
- Ion PGM OT2 400 kit
- Ion PGM HiQ Sequencing kit
- IonTorrent PGM
- Bioinformatics pipeline

EUROFORGEN-NoE is funded by the European Commission within the 7th Framework Programme

2. Body fluid/tissue specific SNPs

SNaPshot results

Saliva

HTN3_2	HTN3_3	MUC7_1	MUC7_3	Incidence (No)	(%)	1/frequency
CT	GA	GG	CC	5/22	22.7	25.4
CC	GA	GA	GC	4/22	18.2	31.6
CC	AA	GG	CC	3/22	13.6	8.2
CC	GA	GG	CC	2/22	9.1	5.6
CC	GA	GG	GC	2/22	9.1	9.5
TT	AA	GG	CC	1/22	4.5	662.4
CT	GG	GG	CC	1/22	4.5	70.2
CC	AA	GG	GC	1/22	4.5	13.7
CT	GG	GG	GC	1/22	4.5	117.5
CC	GG	GG	CC	1/22	4.5	15.6
CC	AA	GA	GC	1/22	4.5	45.8

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

2. Body fluid/tissue specific SNPs

NGS results

cSNPs: 25 blood, 6 saliva, 15 semen, 9 vag, 18 mens, 25 skin, 2 nasal

Vaginal secretion

Gene	chr	65		66		67		68		69		70		71	
V1	chr19	G	G	C	G	C	G	C	G	C	C	C	G	C	G
V1	chr19	C	G	C	G	G	G	0	0	G	G	G	G	C	G
V1	chr19	C	T	C	C	C	T	C	C	C	C	C	T	C	C
V2	chr19	C	T	C	T	C	T	C	T	C	T	T	T	C	T
V2	chr19	C	C	C	C	C	C	C	T	C	T	C	C	C	T
V3	chr20	G	G	G	G	C	G	G	G	C	C	C	G	C	G
V3	chr20	A	A	A	A	A	G	A	A	A	G	A	G	A	A
V3	chr20	G	G	G	G	A	G	G	G	A	A	A	G	A	G
V4	chr8	G	G	G	G	G	G	G	G	G	G	G	G	G	G

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Collaborative exercise mRNA NGS part 1

targeted mRNA NGS approach for the identification of blood, saliva, semen, vaginal secretion, menstrual blood, skin

RNA extraction manual or kit, DNase treatment, quantification

Protocols for PGM and MiSeq will be provided

- PGM (primerpool provided)
- MiSeq (primerpool has to be ordered by the laboratories)

Laboratories will analyse 16 samples provided by UZH and about 30 own body fluid samples

Results (FASTQ files) will be collected and evaluated by UZH

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Timeline

04/2016 Presentation of a collaborative exercise on the developed mRNA NGS method, part 1

06/2016 Shipment of samples, primers, protocols

10/2016 Submission of results

11/2016 - Presentation of results at next EDNAP meeting
- Suggestion for Collaborative exercise, part 2 (cSNPs)

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme

Who wants to participate?

EUROFORGEN-NoE is funded by the European Commission
within the 7th Framework Programme