AGENDA FOR THE EDNAP MEETING

COPENHAGEN, DENMARK

29 MAY 2024

The Maersk Tower, Blegdamsvej 3B, 2200 Copenhagen N, 15th floor, Room 07.15.92

Expected duration: 09.00 - 17.00

Coffee: 10.30-11.00 – Lunch: 12.30-14.00 – Coffee: 15.30-15.45

Host: Bo Simonsen
Chairman: Niels Morling

Welcome	Bo Simonsen
EDNAP's new structure	
Discussion of the draft of the EDNAP Statutes	Niels Morling and
Discussion of the draft of the EDNAP Terms of Reference	the editing group
Election of the EDNAP board members	
Chairman	
Deputy Chairman	
Secretary	
Treasurer	
Future activities	
Paper Exercise on Estimating Biogeographic Ancestry from DNA	M Diepenbroek, C Phillips & W Parson
Update on activities	
mtDNA heteroplasmy exercise	Walther Parson
Methylated DNA and age exercise	Denise S. Court
Exercise no. four on bcSNPs (vaginal secretion, menstrual blood, and skin)	Cordula Haas
The series of exercises relating to DNA transfer	Rooland van Oorshot (written report)
Updates from other groups	
The ENFSI React project	Peter Gill
Presentations	
A comparative study of MPS vs CE using eDNA samples: a discussion on current limitations of MPS in casework and recommendations for future research	Peter Gill
Next EDNAP meeting	Niels Morling
Barcelona on 12 November 2024 followed by the ENFSI Expert DNA Working Group meeting on 13 – 15 November 2024 (to be confirmed)	
Any other business	Niels Morling

Dr. Jeppe Dyrberg Andersen
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V´s Vej 11
DK-2100 Copenhagen
Denmark
Tel:
E-mail: jeppe.dyrberg.andersen@sund.ku.dk

Dr. Ricky Ansell
National Forensic Centre
S-58194 Linköping
Sweden
Tel:
E-mail: ricky.ansell@polisen.se

Dr. Anders Buchard
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V´s Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 27319925
E-mail: Anders.buchard@sund.ku.dk

Dr. Claus Børsting
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V's Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 35 32 62 25
E-mail: claus.boersting@sund.u.dk

Dr. Heloiisa Alonso Costa
National Institute of Legal Medicine and Forensic Sciences
Portugal
Tel: +351 962 970 525
E-mail: heloisa.a.costa@immlaf.mj.pt

Professor Denise Syndercombe Court
King's Forensics
Institute of Pharmaceutical Science
King's College London
Franklin Wilkins Building
Strand
WC2R 2LS London
UK
Tel: +44 20 7848 4155
E-mail: denise.syndercombe-court@kcl.ac.uk

Prof. Dr. Cornelius Courts
Institut für Rechtsmedizin
Melatengürtel 60/62
D-50823 Köln
Germany
Tel: +49 177 7251 873
E-mail: cornelius.courts@uk-koeln.de

Dr. Marta Diepenbroek
LMU Munich
Germany
Tel:
E-mail: marta.diepenbroek@med.uni-muenchen.de

Dr. Ane Elida Fonneløp
Department of Forensic Biology
Oslo University Hopspital
P.O.Box 4950
Nydalen
N-0424 Oslo
Norway
Tel: +47 99 55 95 11
E-mail: rmanfo@ous-hf.no

Dr. Christina Forsberg
National Swedish Forensic Science
S-58194 Linköping
Sweden
Tel: +46 765255154
E-mail: christina-u.forsberg@polisen.se

Dr. Mathieu Gabut
Laboratoire de Police Scientifique de Lyon
31, avenue Franklin Roosevelt
69134 Ecully
France
Tel: +33 640 70 75 83
E-mail: mathieu.gabut@interieur.gouv.fr

Professor Peter Gill
Department of Forensic Biology
Oslo University Hospital
PO Box 4404
Nydalen
N-0403 Oslo
Norway
Tel: +44 7786 126 571
E-mail: peterd.gill@gmail.com

Dr. Cordula Haas
Institut für Rechtsmedizin Zurich
Winterthurerstr. 190
CH-8057 Zurich
Switzerland
Tel: +41 44 635 5656
E-mail: cordula.haas@irm.uzh.ch

Dr. Arnoud Kal
Department of Human Biological Traces
Netherlands Forensic Institute
Laan van Ypenburg 6
24 97 GB The Haque
The Netherlands
Tel: +31 6 4813 1812
E-mail: a.kal@nfi.nl

Dr. Helle Smidt Mogensen
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V's Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 3532 6212
E-mail: helle.smidt@sund.ku.dk

Professor, dr.med. Niels Morling
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V´s Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 3532 6194
E-mail: niels.morling@sund.ku.dk

Dr.Fabrice Noël
National Institute of Forensic Science
98-100 Chaussée de Vilvorde
B-1120 Bruxelles
Belgium
Tel: +32 2243 4604
E-mail: fabrice.noel@just.fgov.be

Dr. Leira Palencia Madrid
Forensic Scientific Laboratory
Basque Police
Spain
Tel: +34 652 71 60 83
E-mail: leire.palencia@seg.euskadi.eus

Prof. Dr. Walther Parson
Institute of Legal Medicine
Medical University of Innsbruck
Müllerstrasse 44
A-6020 Innsbruck
Austria
Tel: +43 512900370640
E-mail: walther.parson@i-med.ac.at

Dr. Vania Pereira
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V´s Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 35 32 60 22
E-mail: vania.pereira@sund.ku.dk

Dr. Chris Phillips
Forensic Genetic Unit
Department of Legal Medicine
University of Santiago de Compostela
San Francisco, s/n
E-15782 Santiago de Compostela
Spain
Tel: +34 98158 2327
E-mail: c.phillips@mac.com

Dr. Bo Simonsen
Section of Forensic Genetics
Department of Forensic Medicine
Faculty of Health Sciences
University of Copenhagen
Frederik V´s Vej 11
DK-2100 Copenhagen
Denmark
Tel: +45 3532 6136
E-mail: bo.simonsen@sund.ku.dk

Dr. Monika Stoljarova-Bibb
Estonia Forensic Science Institute
Estonia
Tel:
E-mail: monika.stoljarova-bibb@ekei.ee


Salonen

EUROPEAN DNA PROFILING GROUP (EDNAP) MEETING

Copenhagen, Denmark

29 May 2024

Host: Bo Simonsen
Chairman: Niels Morling

A list of participants is attached.

Welcome
Bo Simonsen welcomed members to Copenhagen.

The future of EDNAP Niels Morling
The editing group

Niels Morling summarised the suggested changes of EDNAP. The drafts of the Statutes and the EDNAP terms of references (ToR) suggested by the editing group consisting of Cordula Haas, Niels Morling, Geraldine O'Donnell, Walther Parson, Vince Pascali, Chris Phillips, and Bo Simonsen were circulated. The future organisation and work were discussed.

Election of the EDNAP board members

Chair: Cordula Haas
Deputy Chair: Bo Simonsen
Secretary and Treasurer: Walther Parson

All members were elected with acclamation.

The new Statutes and the board will take action on 1 June 2024.

Niels Morling was elected as honorary member of EDNAP.

Update on exercises

Paper Exercise on Estimating Biogeographic Ancestry from DNA M Diepenbroek, C Phillips & W Parson

Marta Diepenbroek, Chris Phillips, and Walther Parson reported on the current status of the joint exercise on biogeographic ancestry (BGA) estimation. Walther Parson gave an introduction and reviewed previous BGA work, which showed that it was not the wet lab part but the interpretation of results that was challenging for labs. In addition, he noted a lack of standardisation and harmonisation of data interpretation and BGA reporting. The first EDNAP BGA exercise was therefore tailored to reporting BGA results. Since privacy issues did not allow the sharing of DNA or genotypes, the exercise was based on the results of the BGA software available to study participants.
Chris Phillips presented the currently available solutions for BGA estimation and outlined their performance. He summarised the tools and analyses included in the exercise. Samples used in the study were typed using VISAGE Basic and Enhanced Tools and Phenotrivium. BGA software output was provided for autosomal, X, Y chromosomal, and mitochondrial DNA data. Chris Phillips presented the following analysis tools: SNIPPER (PCA), STRUCTURE, GenoGeographer, YHRD, and EMPOP.

Marta Diepenbroek presented the joint BGA exercise results from nine samples from three individuals of shared ancestry and six of mixed ancestry. Marta presented the results of admixture analysis with STRUCTURE and Converge, PCA, and GenoGeographer, followed by the p-values for phenotype and uniparental data obtained with the remaining markers. Finally, the true ancestry was revealed based on the paternal and maternal grandparents of the examined individuals.
Chris Phillips discussed the limitations of the tools and software used in the study and the caveats of the analyses. After a discussion among the EDNAP participants, it was decided to disseminate the results.
Marta Diepenbroek has submitted an abstract for the ISFG Congress 2024 that was accepted as a poster presentation. She agreed to prepare a draft manuscript of the exercise results before the end of 2024. A preliminary version will, hopefully, be ready for discussion at the next EDNAP meeting in Barcelona.
It was decided that further BGA experiments shall be conducted. Marta Diepenbroek will explore the possibility of providing individual genotypes for sharing among study participants. Such an exercise will contribute to harmonising BGA estimation in forensic genetics.

mtDNA heteroplasmy exercise Walther Parson
Walther Parson summarised the information on the mtDNA sequences used for the heteroplasmy exercise. Extracted DNA from five reference samples was distributed among the participants, who sequenced the mtDNA following their established protocols using Sanger, Ion Torrent, and Illumina sequencing. Walther reported agreement between the resulting mitotypes except for typographical errors, which were ignored for this study and heteroplasmy reporting. The analysis revealed that the dominant length variants were not consistent among participants. The deviations were due to the technologies and software used and individual settings and interpretations. To be able to compare data from different laboratories and platforms, agnostic software was developed in Innsbruck. The results indicate that revising the length heteroplasmy interpretation guidelines would be beneficial. A manuscript on the subject, intended for publication, will be disseminated for comments.

Second exercise on methylated DNA and age Denise Syndercombe Court
A two-part exercise was completed some years ago. The work has been presented at meetings. The results of the exercise remain relevant. The organisers commit to publishing the data.

Exercise no. 4 on mRNA typing with MPS Cordula Haas
Cordula Haas gave an update on mRNA MPS exercise 4, a collaborative exercise to test a targeted MPS assay for body fluid identification and the assignment to donors using cSNPs (attached). Cordula Haas presented new data from the sequencing platforms IonTorrent S5 and Illumina MiSeq and a comparison with an alternative cSNP panel developed by the Cologne laboratory. A manuscript draft is almost finished and will be shared with the participants soon. The results will also be presented as a poster at the ISFG conference 10 – 13 September 2024 in Santiago de Compostela.

The series of exercises relating to DNA transfer Roland van Oorschot
Roland van Oorshot had sent an update on the exercise (attached). Data on 1,427 tool handles and 1,357 glove samples have been submitted from 18 laboratories. The data are being analysed.

Updates from other groups
The ENFSI React project Peter Gill
Peter Gill presented the result of the ENFSI React project that aims at collecting data from common casework, making DNA data available as non-sensitive, open access data, and

developing open-source software to analyse the data using Bayesian Networks to analyse data and report LRs using bootstrap confidence intervals (attached). More than 2,700 data sets from 23+ labs have been collected, and the programme Shiny_React has been developed..

Presentations

A comparative study of MPS vs CE using eDNA samples: a discussion on the current limitations of MPS in casework and recommendations for future research — Peter Gill

Peter Gill presented a comparison of the results of MPS and CE examinations of surface DNA mixtures (attached). CE ultimately provided higher information content than MPS for compromised, low-quality samples.

Next meeting — Niels Morling

Barcelona on 12 November 2024, followed by the ENFSI Expert DNA Working Group meeting on 13 – 15 November 2024 (to be confirmed)

Any other business — Niels Morling

There was no other business.

Closing of the meeting — Niels Morling

The meeting closed with sincere thanks to Bo Simonsen, who organised the meeting.

The minutes and attachments are found on the EDNAP website:
http://www.isfg.org/EDNAP/Meetings, including:

- Agenda.
- List of participants.
- Group photo.
- Minutes.
- EDNAP Statutes.
- EDNAP Terms of Reference.
- Presentations.
  - Marta Diepenbroek: Estimating Biogeographic Ancestry.
  - Chris Phillips: Estimating Biogeographic Ancestry.
  - Walther Parson: mtDNA heteroplasmy exercise.
  - Denise Syndercombe Court: Methylated DNA and age exercise.
  - Cordula Haas: mRNA exercise no. 4 on bcSNPs.
  - Roland van Oorshot: Update on the series of exercises relating to DNA transfer.
  - Peter Gill: The ENFSI React project.
  - Peter Gill: A comparative study of MPS vs CE using eDNA samples.

STATUTES OF THE
EUROPEAN DNA PROFILING GROUP – EDNAP

Adapted at the EDNAP meeting 29 May 2024 in Copenhagen

§1 Name and Place of the Group

The European DNA Profiling Group (EDNAP) is a working party under the International Society for Forensic Genetics.

§2 Aims

(1) EDNAP aims to advance scientific knowledge in genetic markers used in forensic crime investigations and harmonization of DNA and other forensic genetic typing technologies.

(2) This is achieved by personal and Internet-based scientific meetings in the group, discussion groups, identification of new research areas, grant applications, scientific projects, collaborative exercises, collaboration with other genetic scientist groups and organizations, and scientific publications.

§3 Non-Profit Status

(1) EDNAP's purpose is exclusively scientific and non-profit; see §3 of the statutes of the International Society for Forensic Genetics.

(2) EDNAP does not pursue an economic purpose. EDNAP's funds may only be used for appropriate statutory purposes. The members do not receive any financial contributions from EDNAP's funds.

(3) Members of the EDNAP Executive Committee can reimburse relevant expenses after receiving an invoice. No individual may benefit from expenses inappropriate to EDNAP's purpose or receive excessive allowances.

§4 Membership

(1) European academic graduate or technical support staff members working scientifically with forensic genetics can apply for membership in EDNAP.

(2) Membership application must be sent to EDNAP's Executive Committee, and the General Assembly decides on membership with a simple majority.

(3) Membership is terminated by written cancellation, death, or disqualification. Notice of withdrawal must be submitted in writing to the Executive Committee and will take effect at the end of each calendar year.

(4) Membership is considered terminated if the annual EDNAP or ISFG membership fee, according to §9, has not been paid for at least two years.

(5) The Executive Committee can nominate scientists whose contributions are of great importance to forensic genetics and any other individuals who deserve acknowledgment by EDNAP as honorary members. The appointment is subject to a decision from EDNAP's General Assembly.

(7) Membership may not be used for advertising purposes. The Executive Committee can terminate the membership if this regulation is violated despite a warning.

§5 Executive Committee

(1) The Executive Committee comprises a Chairman, a Deputy Chairman, a Secretary, and a Treasurer. The Secretary and the Treasurer can be the same person.

(2) The General Assembly elects the Executive Committee by a simple majority vote among the attending members. Voting can be done by show of hands. A secret ballot must be taken if only one member so wishes. The Executive Committee members are elected for two consecutive years. Re-election is possible.

(3) The newly elected Executive Committee takes over its official function after the election.

(4) The Executive Committee is still authorized to pass resolutions, even if a member should withdraw prematurely. The required supplementary election will then take place at the next General Assembly. In particular situations, the Executive Committee can constitute itself with a substituting member.

(5) The Executive Committee can invite guests to the Executive Committee meetings.

§6 Authorization for Representation

Two members of the Executive Committee of the ISFG represent EDNAP in the sense of § 26 BGB (German Civil Code).

§7 General Assembly

(1) The General Assembly pays particular attention to the following:

- Receipt of the Executive Committee's annual report and the annual financial report.
- Election of members of the Executive Committee.
- Election of two auditors.
- Relieve of the members of the Executive Committee.
- Appointment of honorary members.
- Determination of the dates and places for scientific meetings.
- Determination of the membership fee.
- Decisions on procedural motions, if necessary.
- Decisions on changes to the statutes and society regulations.
- Decision on the dissolution of EDNAP.

(2) The General Assembly takes decisions by simple majority.

(3) The General Assembly takes place during EDNAP's yearly personal scientific meeting. The Executive Committee sends an invitation and the planned agenda via e-mail two weeks in advance.

(4) The Executive Committee can call an extraordinary General Assembly if at least 1/4 of the members submit a request to the Executive Committee.

(5) Following the Executive Committee's decision or a request of at least 1/4 of the ordinary members or 2/3 of the members attending the General Assembly, an e-mail vote can be taken regarding business matters.

§8 Scientific Meetings

(1) A scientific personal meeting is held at least once a year. If a personal meeting is impossible, the Executive Committee can arrange an Internet-based meeting. The Executive Committee can decide to arrange Internet-based participation at personal meetings.

(2) The Executive Committee can arrange additional personal and Internet-based meetings.

(3) Members can make proposals regarding the program. The Executive Committee should receive proposals at least one month in advance. The Executive Committee decides on the scientific topics for the presentations and chooses the speakers.

§9 Membership Subscriptions, Auditor, and Fiscal Year

(1) The annual membership fee is due at the beginning of the calendar year. The General Assembly determines the amount of the annual fee. A change to the membership fee takes effect at the beginning of the following calendar year.

(2) Honorary members and guests do not pay membership fees.

(3) After every fiscal year, the two auditors check EDNAP's accounts and report to the General Assembly.

(4) The fiscal year is identical to the calendar year.

§10 Working-, Project-, and Other Groups

(1) The General Assembly and the Executive Committee can decide to form working-, discussion-, project, and other groups.

§11 Changes of the Statutes

(1) Propositions for changes to the statutes, sent by at least 1/5 of the members and with reasons, must be received by the Executive Commission via e-mail or registered mail at least one month before the General Assembly.

(2) The Executive Committee can propose changes to the statutes to the General Assembly. The proposed changes must be notified to the members together with the notice of the General Assembly.

(3) The General Assembly can change the statutes or aims of EDNAP with a 3/4 majority of the attending members.

§12 Dissolution of the Society

(1) The dissolution of EDNAP can only be considered after the Executive Committee via e-mail or registered mail has received a proposal, duly signed by at least 2/3 of the ordinary members with reasons, at least three months before the General Assembly. A 3/4 majority of the attending ordinary members is required for a decision to dissolve EDNAP.

(2) If EDNAP is dissolved, its property shall be transferred to the International Society for Forensic Genetics, which shall use the property exclusively for non-profit purposes.

The EDNAP statutes were adopted at the General Assembly in Copenhagen on May 29, 2024. The statutes have been approved by the International Society for Forensic Genetics.

TERMS OF REFERENCES
THE EUROPEAN DNA PROFILING GROUP – EDNAP

Adapted at the EDNAP meeting 29 May 2024 in Copenhagen

History of EDNAP

EDNAP's founding members met on 15/16 October 1988 in Sunbury near London, UK, at a meeting arranged by Peter Martin, the Metropolitan Police Laboratory. Forensic genetic scientists from 11 European countries came together to find a way of harmonizing DNA technology for crime investigation. An integrated Europe with open borders led to an escalation of cross-border crimes and increased demand for exchanging intelligence data among European countries. Data obtained with the new DNA technology could only be exchanged efficiently among laboratories and across borders if the DNA methods used in the European countries were harmonized with common standards.

EDNAP was accepted as a working group of the International Society for Forensic Genetics (ISFG) during the 1991 congress in Mainz, Germany. It was intended that each European country should be represented by one laboratory with sufficient scientific expertise in forensic DNA technology. As a small group, EDNAP could make decisions solely based on scientific considerations and operate in the spirit of helping each other.

EDNAP's main objective was to harmonize DNA typing technology for crime investigations. This has been addressed by organizing collaborative intercomparison exercises and discussing the results at group meetings.

In two initial collaborative exercises, EDNAP addressed the application of single-locus DNA probes. The exercise demonstrated that the results obtained in the participating laboratories were sufficiently close to each other to be used to compare DNA typing results.

In further collaborative exercises, EDNAP evaluated the possibility of standardizing DNA typing of several PCR-amplified STR loci. Later, automated STR typing using fluorescent detection of PCR fragments was tested. This new technique dramatically increased sample throughput and computerized online recording of typing results. The reliability of STR typing was confirmed in a collaborative exercise including EDNAP and non-EDNAP laboratories and led to the selection of the "European standard set of loci" by ENFSI and the European Union. Further exercises have dealt with Y-chromosome STR systems and mitochondrial DNA polymorphism.

Between 1997 and 2000, a network project, Standardization of DNA Profiling in Europe – STADNAP, funded by the European Commission, was carried out. One of STADNAP's achievements was identifying the need for appropriate quality control in mtDNA databasing.

In 1999, the EDNAP Forensic mtDNA Population Database – EMPOP – was established by Walther Parson, Institute of Legal Medicine, Medical University of Innsbruck. New laboratory strategies for fail-safe mtDNA typing and mathematical data analysis models were developed and have been optimized continuously. EMPOP went online in October 2006. EMPOP serves as a repository of high-quality mtDNA data scrutinized

by EMPOP. Furthermore, EMPOP is a research centre with many collaborative research projects and advises other scientists in their mtDNA research.

The collaborative exercises in the last years have included the evaluation of single nucleotide polymorphism (SNP) typing, identification of body fluids in forensic stain samples using mRNA analysis, forensic ancestry analysis by SNP typing, forensic phenotyping of eye colour, and age estimation by measuring DNA methylation of informative CpG DNA nucleotides.

The results of the collaborative exercises have been published in the scientific literature; see Appendix 1.

Since 2004, the EDNAP Group and the DNA Expert Working Group of the European Network of Forensic Science Institutes (ENFSI) have coordinated their annual personal meetings. The ENFSI DNA Expert Working Group members are invited to attend the EDNAP meetings, and the EDNAP members are invited to attend the sessions of the DNA Analysis Methods & Interpretation Group of the ENFSI DNA Expert Working Group.

New Directions for EDNAP

Today, forensic genetic typing methods are well-established and harmonized in Europe and other parts of the World. The EDNAP laboratories perform accredited DNA analyses. Most EDNAP laboratories – but not all - are members of the DNA Expert Working Group of ENFSI, which serves as a platform for practical scientific collaboration among European forensic genetic laboratories. The ENFSI group competently addresses many of the issues initially dealt with by EDNAP. The need for EDNAP's role in harmonizing standard DNA typing methods no longer exists.

EDNAP's role should change into forensic genetic research. However, most of the founding laboratories of EDNAP have limited resources for research. Therefore, EDNAP's future structure must be reconsidered.

EDNAP must have statutes, and the new directions must be formulated, e.g., Terms of References (ToR), which can support the scientific activities of EDNAP as a working group under the International Society for Forensic Genetics.

Terms of Reference (ToR) means 'The purpose and structure of a project, committee, meeting, negotiation, or a similar collection of people who have agreed to work together to accomplish a shared goal.' The ToR of a dynamic group is updated continuously.

Suggestion for EDNAP Terms of Reference

Aims:

EDNAP

1. is a forum for experts to share information, explore new research areas, and drive the development of forensic genetics,
2. organizes collaborative exercises, workshops, in-depth discussions of research results and ideas, presentations, etc.
3. strives to act as an informal scientific environment and create an atmosphere where members help each other with research-related questions.

Members:

EDNAP invites

4. European academic and technical support staff members working scientifically within the field of forensic genetics and who are members of the ISFG to apply for membership in EDNAPand the ISFG,
5. European forensic genetic laboratories with high scientific expertise in forensic DNA technologies to be represented in EDNAP by qualified staff members (ISFG membership required), and
6. Scientific colleagues with particular competencies relating to forensic genetics to participate as guests (non-EDNAP members) in meetings, research projects, collaborative exercises, etc. ISFG membership is not required for guests.

Meeting forms:

EDNAP

7. organizes at least one annual personal scientific meeting. The meeting can take place (but it is not a requirement) in conjunction with the personal meeting of the ENFSI DNA Expert Working Group,
8. can, in exceptional circumstances, organize the annual meeting as an Internet-based meeting,
9. decides on additional personal, Internet-based, or combined meetings,
10. invites guests to participate in EDNAP's meetings and will continue to communicate through the EDNAP part of the ISFG website and explore the necessity of establishing other communicative web-based platforms to exchange information.

Projects:

EDNAP

11. explores new forensic genetic research areas suitable for research projects,
12. forms research groups, including EDNAP and non-EDNAP members,
13. explores the possibilities of obtaining funding, and
14. applies for funding of forensic genetic research projects.

Publication:

15. EDNAP supports the timely publication of results in high-impact journals.

Structure and organization of EDNAP's work:

EDNAP

16. is a working group with its own statutes under the International Society for Forensic Genetics,
17. is organized with an elected board with a Chairman, a Deputy Chairman, a Secretary, and a Treasurer according to EDNAP statutes; the Secretary and the Treasurer can be the same person,
18. coordinates its work with the board of the ISFG and the DNA Expert Working Group of the European Network of Forensic Science Institutes, and
19. has currently no membership fee in addition to the membership fee of the ISFG.

Appendix 1.

EDNAP Publications:

Schneider PM, Fimmers R, Woodroffe S, Werrett DJ, Bär W, Brinkmann B, Eriksen B, Jones S, Kloosterman AD, Mevag B, Pascali VL, Rittner C, Schmitter H, Thomson JA, Gill P. (1991) Report of a European collaborative exercise comparing DNA typing results using a single locus VNTR probe. Forensic Sci. Int. 49:1-15

Gill P, Woodroffe S, Bär W, Brinkmann B, Carracedo A, Eriksen B, Jones S, Kloosterman AD, Ludes B, Mevag B, Pascali VL, Rudler M, Schmitter H, Schneider PM, Thomson JA. (1992) A report of an international collaborative experiment to demonstrate the uniformity obtainable using DNA profiling techniques. Forensic Sci. Int. 53:29-43

Gill P, Kimpton C, D'Aloja E, Andersen JF, Bär W, Brinkmann B, Holgerssen S, Johnsson V, Kloosterman AD, Lareu MV, Nellemann L, Pfitzinger H, Phillips CP, Schmitter H, Schneider PM, Stenersen M. (1994) Report of the European DNA profiling group (EDNAP) - towards standardization of short tandem repeat (STR) loci. Forensic Sci. Intern. 65:51-59

Kimpton C, Gill P, D'Aloja E, Andersen JF, Bär W, Holgerssen S, Jacobsen S, Johnsson V, Kloosterman AD, Lareu MV, Nellemann L, Pfitzinger H, Phillips CP, Rand S, Schmitter H, Schneider PM, Stenersen M, Vide MC (1995) Report on the second EDNAP collaborative STR exercise. Forensic Sci. Int. 71:137-152

Andersen J, Martin P, Carracedo A, Dobosz M, Eriksen B, Johnsson V, Kimpton C, Kloosterman A, Konialis C, Kratzer A, Phillips P, Mevag B, Pfitzinger H, Rand S, Rosen B, Schmitter H, Schneider P, Vide M. (1996) Report on the third EDNAP collaborative STR exercise. Forensic Sci. Int. 78:83-93

Gill P, d'Ajola A, Andersen J, Dupuy B, Jangblad M, Johnsson V, Kloosterman AD, Kratzer A, Lareu MV, Meldegaard M, Philips C, Pfitzinger H, Rand S, Sabatier M, Scheithauer R, Schmitter H, Schneider PM, Vide MC. (1997) Report of the European DNA profiling group (EDNAP): an investigation of the complex STR loci D21S11 and HUMFIBRA (FGA). Forensic Sci Int 86:25-33

Gill P, Brinkmann B, d'Ajola E, Andersen J, Bär W, Carracedo A, Dupuy B, Eriksen B, Jangblad a, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider PM, Vide MC. (1997) Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. Forensic Sci Int. 87:185-192

Carracedo A, d'Aloja E, Dupuy B, Jangblad A, Karjalainen M, Lambert C, Parson W, Pfeiffer H, Pfitzinger H, Sabatier M, Syndercombe-Court D, Vide C. (1998) Reproducibility of mtDNA analysis between laboratories: a report of the European DNA Profiling group (EDNAP). Forensic Sci. Int. 97:155-164

Gill P, d'Aloja E, Dupuy B, Eriksen B, Jangblad A, Johnsson V, Kloosterman AD, Lareu MV, Mevag B, Morling N, Phillips C, Pfitzinger H, Rand S, Sabatier M, Scheithauer R, Schmitter H, Schneider PM, Skita I, Vide MC. (1998) Report of the European DNA Profiling group (EDNAP) - an investigation of the hypervariable loci ACTBP2, APOA11 and D11S554 and the compound loci D12S391 and D1S1656. Forensic Sci. Int. 98:193-200

Schneider PM, d'Aloja E, Dupuy BM, Eriksen B, Jangblad A, Kloosterman AD, Kratzer A, Lareu MV, Pfitzinger H, Rand S, Scheithauer R, Schmitter H, Skitsa I, Syndercombe-Court D, Vide MC (1999) Results of a collaborative study regarding the standardization of the Y-linked STR system DYS385 by the European DNA Profiling (EDNAP) group. Forensic Sci. Int. 102:159-165

Carracedo A, Beckmann A, Bengs A, Brinkmann B, Caglia A, Capelli C, Gill P, Gusmao L, Hagelberg C, Hohoff C, Hoste B, Kihlgren A, Kloosterman A, Myhre Dupuy B, Morling N, O'Donnell G, Parson W, Phillips C, Pouwels M, Scheithauer R, Schmitter H, Schneider PM, Schumm J, Skitsa I, Stradmann-Bellinghausen B, Stuart M, Syndercombe Court D, Vide C (2001) Results of a collaborative study of the EDNAP group regarding the reproducibility and robustness of the Y-chromosome STRs DYS19, DYS389 I and II, DYS390 and DYS393 in a PCR pentaplex format. Forensic Sci. Int. 119:28-41

Tully G, Bär W, Brinkmann B, Carracedo A, Gill P, Morling N, Parson W, Schneider PM (2001) Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. Forensic Sci. Int. 124:83-91

Schneider PM, Bender K, Mayr WR, Parson W, Hoste B, Decorte R, Cordonnier J, Vanek D, Morling M, Karjalainen M, Carlotti CMP, Sabatier M, Hohoff C, Schmitter H, Pflug W, Wenzel R, Patzelt D, Lessig R, Dobrowolski P, O'Donnell G, Garafano L, Dobosz M, de Knijff P, Mevag B, Pawlowski R, Gusmao L, Vide MC, Alonso A, Garcia Fernandez O, Pilar Sanz N, Kihlgreen A, Baer W, Meier V, Teyssier A, Coquoz R, Brandt C, Germann U, Gill P, Hallett J, Greenhalgh M (2004) STR analysis of artificially degraded DNA: Results of a collaborative European exercise. Forensic Sci. Int. 139:123-134

Parson W, Brandstätter A, Alonso A, Brandt N, Brinkmann B, Carracedo A, Corach D, Froment O, Furach I, Grzybowski T, Hedberg K, Keyser-Tracqui C, Kupiec T, Lutz-Bonengel S, Mevag B, Ploski R, Schmitter H, Schneider PM, Syndercombe-Court D, Sörensen E, Thew H, Tully G, Scheithauer R (2004) The EDNAP mitochondrial DNA population database (EMPOP) collaborative exercises: organisation, results and perspectives. Forensic Sci. Int. 139:215-226

Tully G, Barritt SM, Bender K, Brignon E, Capelli C, Dimo-Simonin N, Eichmann C, Ernst CM, Lambert C, Lareu MV, Ludes B, Mevag B, Parson W, Pfeiffer H, Salas A, Schneider PM, Staalstrom E (2004) Results of a collaborative study of the EDNAP group regarding mitochondrial DNA heteroplasmy and segregation in hair shafts. Forensic Sci. Int. 140:1-11

Brion M, Dupuy BM, Heinrich M, Hohoff C, Hoste B, Ludes B, Mevag B, Morling N, Niederstatter H, Parson W, Sanchez J, Bender K, Siebert N, Thacker C, Vide C, Carracedo A (2005) A collaborative study of the EDNAP group regarding Y-chromosome binary polymorphism analysis. Forensic Sci Int. 153:103-108

Dixon LA, Dobbins AE, Pulker HK, Butler JM, Vallone PM, Coble MD, Parson W, Berger B, Grubwieser P, Mogensen HS, Morling N, Nielsen K, Sanchez JJ, Petkovski E, Carracedo A, Sanchez-Diz P, Ramos-Luis E, Brion M, Irwin JA, Just RS, Loreille O, Parsons TJ, Syndercombe-Court D, Schmitter H, Stradmann-Bellinghausen B, Bender K, Gill P (2006) Analysis of artificially degraded DNA using STRs and SNPs -– results of a collaborative European (EDNAP) exercise. Forensic Sci Int. 164:33-44

Gill P, Fereday L, Morling N, Schneider PM (2006) The evolution of DNA databases -– recommendations for new European STR loci. Forensic Sci Int. 156:242-244.

Gill P, Fereday L, Morling N, Schneider PM (2006) New multiplexes for Europe -– amendments and clarification of strategic development. Forensic Sci Int. 163:155-157

Parson W, Fendt L, Ballard D, Børsting C, Brinkmann B, Carracedo A, Carvalho M, Coble MD, Corte Real F, Desmyter S, Dupuy BM, Harrison C, Hohoff C, Just R, Krämer T, Morling N, Salas A, Schmitter H, Schneider PM, Sonntag ML, Vallone PM, Brandstätter A (2008) Identification of West Eurasian mitochondrial haplogroups by mtDNA SNP screening: Results of the 2006–2007 EDNAP collaborative exercise. Forensic Sci. Int. Genet. 2:61-68. http://dx.doi.org/10.1016/j.fsigen.2007.08.007

Sanchez JJ, Børsting C, Balogh K, Berger B, Bogus M, Butler JM, Carracedo A, Syndercombe Court D, Dixon LA, Filipovic B, Fondevila M, Gill P, Harrison CD, Hohoff C, Huel R, Ludes B, Parson W, Parsons TJ, Petkovski E, Phillips C, Schmitter H, Schneider PM, Vallone PM, Morling N (2008) Forensic typing of autosomal SNPs with a 29 SNP-multiplex -— Results of a collaborative EDNAP exercise. Forensic Sci. Int. Genet. 2:176-183. http://dx.doi.org/10.1016/j.fsigen.2007.12.002

Haas C, Hanson E, Bär W, Banemann R, Bento AM, Berti A, Borges E, Bouakaze C, Carracedo A, Carvalho M, Choma A, Dotsch M, Duriancikova M, Hoff-Olsen P, Hohoff C, Johansen P, Lindenbergh PA, Loddenkötter B, Ludes B, Maronas O, Morling N, Niederstätter H, Parson W, Patel G, Popielarz C, Salata E, Schneider PM, Sijen T, Sviezena B, Zatkalikova L, Ballantyne J (2011) mRNA profiling for the identification of blood -– Results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet. 5:21-26, http://dx.doi.org/10.1016/j.fsigen.2010.01.003

Welch L, Gill P, Tucker VC, Schneider PM, Parson W, Smidt Mogensen H, Morling N (2011) A comparison of mini-STRs versus standard STRs -– Results of a collaborative European (EDNAP) exercise. Forensic Sci. Int. Genet. 5:257-258. http://dx.doi.org/10.1016/j.fsigen.2010.01.004

Tomas C, Axler-DiPerte G, Budimlija ZM, Borsting C, Coble MD, Decker AE, Eisenberg A, Fang R, Fondevila M, Frisk Fredslund S, Gonzalez S, Hansen AJ, Hoff-Olsen P, Haas C, Kohler P, Kriegel AK, Lindblom B, Manohar F, Maronas O, Mogensen HS, Neureuther K, Nilsson H, Scheible MK, Schneider PM, Sonntag ML, Stangegaard M, Syndercombe-Court D, Thacker CR, Vallone PM, Westen AA, Morling N (2011) Autosomal SNP typing of forensic samples with the GenPlex(TM) HID System: Results of a collaborative study. Forensic Sci. Int. Genet. 5:369-375, http://dx.doi.org/10.1016/j.fsigen.2010.06.007

Haas C, Hanson E, Anjos MJ, Bär W, Banemann R, Berti A, Borges E, Bouakaze C, Carracedo A, Carvalho M, Castella V, Choma A, De Cock G, Dotsch M, Hoff-Olsen P, Johansen P, Kohlmeier F, Lindenbergh PA, Ludes B, Maronas O, Moore D, Morerod ML, Morling N, Niederstätter H, Noel F, Parson W, Patel G, Popielarz C, Salata E, Schneider PM, Sijen T, Sviezena B, Turanska M, Zatkalikova L, Ballantyne J (2012) RNA/DNA co-analysis from blood stains -– Results of a second collaborative EDNAP exercise, Forensic Sci. Int. Genet. 6:70-80. http://dx.doi.org/10.1016/j.fsigen.2011.02.004

Haas C, Hanson E, Anjos MJ, Banemann R, Berti A, Borges E, Carracedo A, Carvalho M, Courts C, De Cock G, Dötsch M, Flynn S, Gomes I, Hollard C, Hjort B, Hoff-Olsen P, Hríbiková K, Lindenbergh A, Ludes B, Maroñas O, McCallum N, Moore D, Morling N, Niederstätter H, Noel F, Parson W, Popielarz C, Rapone C, Roeder AD, Ruiz Y, Sauer E, Schneider PM, Sijen T, Court DS, Sviežená B, Turanská M, Vidaki A, Zatkalíková L, Ballantyne J (2013) RNA/DNA co-analysis from human saliva and semen stains-–results of a third collaborative EDNAP exercise. Forensic Sci Int Genet. 7:230-239. http://dx.doi.org/10.1016/j.fsigen.2012.10.011

Haas C, Hanson E, Anjos MJ, Ballantyne KN, Banemann R, Bhoelai B, Borges E, Carvalho M, Courts C, De Cock G, Drobnic K, Dötsch M, Fleming R, Franchi C, Gomes I, Hadzic G, Harbison SA, Harteveld J, Hjort B, Hollard C, Hoff-Olsen P, Hüls C, Keyser C, Maroñas O, McCallum N, Moore D, Morling N, Niederstätter H, Noël F, Parson W, Phillips C, Popielarz C, Roeder AD, Salvaderi L, Sauer E, Schneider PM, Shanthan G, Syndercombe Court D, Turanská M, van Oorschot RAH, Vennemann M, Vidaki A, Zatkalíková L, Ballantyne J (2014) RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: Results of a fourth and fifth collaborative EDNAP exercise, Forensic Sci Int Genet 8:203-212. http://dx.doi.org/10.1016/j.fsigen.2013.09.009

Haas C, Hanson E, Banemann R, Bento AM, Berti AM, Carracedo A, Courts C, De Cock G, Drobnic K, Fleming R, Franchi C, Gomes I, Hadzic G, Harbison SA, Hjort B, Hollard C, Hoff-Olsen P, Keyser M, Kondili A, Maronas O, McCallum N, Miniati P, Morling N, Niederstätter H, Noel F, Parson W, Porto MJ, Roeder AD, Sauer E,

Schneider PM, Shantan G, Sijen T, Syndercombe Court D, Turanska M, van den Berge M, Vennemann M, Vidaki A, Zatkalikova L, Ballantyne J (2015) RNA/DNA co-analysis from human skin and contact traces - results of a sixth collaborative EDNAP exercise, Forensic Sci Int Genet 16: 139-47. http://dx.doi.org/10.1016/j.fsigen.2014.11.014

Chaitanya L, Walsh S, Andersen JD, Ansell R, Ballantyne K, Ballard D, Banemann R, Bauer CM, Bento AM, Brisighelli F, Capal T, ClarisseL, Groß T, Haas C, Hoff-Olsen P, Hollard C, Keyser C, Kiesler CM, Kohler P, Linacre A, Minawi A, Morling N, Nilsson H, Norén L, Ottens R, Parson W, Pascali VL, Phillips C, Porto MJ, Sajantila A, Schneider P, Sijen T, Söchtig J, Syndercombe-Court D, Tilmar A, Turanska M, Vallone PM, Zatkalíková L, Zidkova A, Branicki W and Kayser M. Collaborative EDNAP Exercise on the IrisPlex system for DNA based prediction of human eye colour. Forensic Sci Int Genet 2014; 11: 241-51. http://dx.doi.org/10.1016/j.fsigen.2014.04.006

Santos C, Fondevila M, Ballard D, Baneman R, Bentod AM, Børsting C, Branicki W, Brisighelli F, Burrington M, Capal T, Chaitanya N, Daniel R, Decroyer V, England R, Gettings KB, Gross TE, Haas C, Harteveld PJ, Hoff-Oisen P, Hoffmann A, Kayseri M, Linacre A, Kohler P, Mayr-Eduardoffu M, McGovern C, Morling N, Noel F, O'Donnell G,Parson W, Pascali VL, Porto MJ, Roset A, SchneiderPM, Sijen T, Sten V, Syndercombe Court D, Templeton J, Turanska M, Vallone PM, van Oorschot PAV, Zatkalikova L, The EUROFORGEN-NoE Consortium, Carracedo A, Phillips C. Forensic ancestry analysis with two simple capillary electrophoresis AlMs panels: Results of a collaborative EDNAP exercise. Forensic Sci Int Genet 2015; 19: 56-67. http://dx.doi.org/10.1016/j.fsigen.2015.06.004

Weiler NE, Baca K, Ballard D, Balsa F, Bogus M, Børsting C, Brisighelli F, Cervenáková J, Chaitanya L, Coble M, Decroyer V, Desmyter S, van der Gaag KJ, Gettings K, Haas C, Heinrich J, João Porto M, Kal AJ, Kayser M, Kúdelová A, Morling N, Mosquera-Miguel A, Noel F, Parson W, Pereira V, Phillips C, Schneider PM, Syndercombe Court D, Turanska M, Vidaki A, Wolinski P, Zatkalíková L, Sijen T. A collaborative EDNAP exercise on SNaPshot™-based mtDNA control region typing. Forensic Sci Int Genet 2017; 26: 77-84. http://dx.doi.org/10.1016/j.fsigen.2016.10.014.

Ingold S, Dørum G, Hanson E, Berti A, Branicki W, Brito P, Elsmore P, Gettings KB, Giangasparo F, Gross TE, Hansen S, Hanssen EN, Kampmann ML, Kayser M, Laurent FX, Morling N, Mosquera-Miguel A, Parson W, Phillips C, Porto MJ, Pospiech E, Roeder AD, Schneider PM, Schulze Johann K, Steffen CR, Syndercombe-Court D, Trautmann M, van den Berge M, van der Gaag KJ, Vannier J, Verdoliva V, Vidaki A, Xavier C, Ballantyne J, Haas C. Body fluid identification using a targeted mRNA massively parallel sequencing approach - results of a EUROFORGEN/EDNAP collaborative exercise. Forensic Sci Int Genet. 2018; 34:105-115. http://dx.doi.org/10.1016/j.fsigen.2018.01.002.

Ingold S, Dørum G, Hanson E, Ballard D, Berti A, Gettings KB, Giangasparo F, Kampmann ML, Laurent FX, Morling N, Parson W, Steffen CR, Ulus A, van den Berge M, van der Gaag KJ, Verdoliva V, Xavier C, Ballantyne J, Haas C. Body fluid identification and assignment to donors using a targeted mRNA massively parallel sequencing approach - results of a second EUROFORGEN / EDNAP collaborative exercise. Forensic Sci Int Genet. 2020; 45:102208. http://dx.doi.org/10.1016/j.fsigen.2019.102208.

Quo vadis, BGA?

A collaborative EDNAP exercise on estimating biogeographic ancestry from the DNA of unknown samples

Organized by:

MEDIZINISCHE UNIVERSITÄT INNSBRUCK

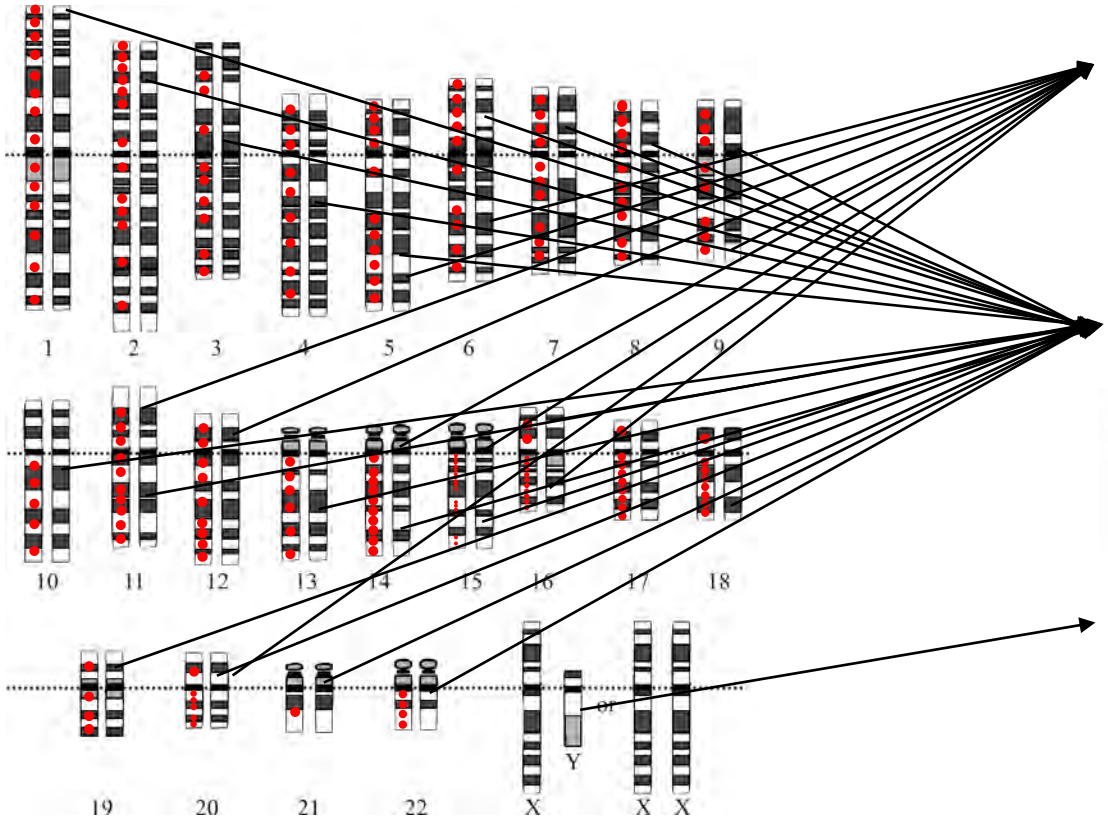
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Quo vadis, BGA?

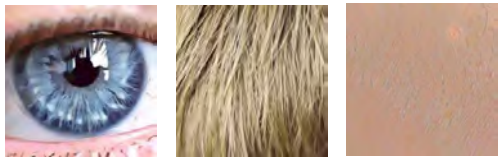
A collaborative EDNAP exercise on estimating biogeographic ancestry from the DNA of unknown samples

Sample collection



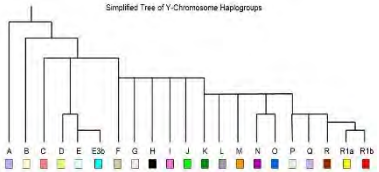
Source: Wikipedia

41 aSNPs - HirisPlex-S Panel



163 aSNPs - Applied Biosystems™ Precision ID Ancestry Panel

116 Y-SNPs



Sample collection

Institut für Rechtsmedizin München

Questionnaire

Implementation of Ancestry Informative Markers (AIMs) and Phenotype Informative SNPs (PI-SNPs) in forensic casework

Where are your roots? Please describe your "biogeographical ancestry" using the genealogical tree provided below. To describe the origin of your family members, please refer to continents and regions provided within the table and, if possible, describe any additional information known about your ancestry (please see an example):

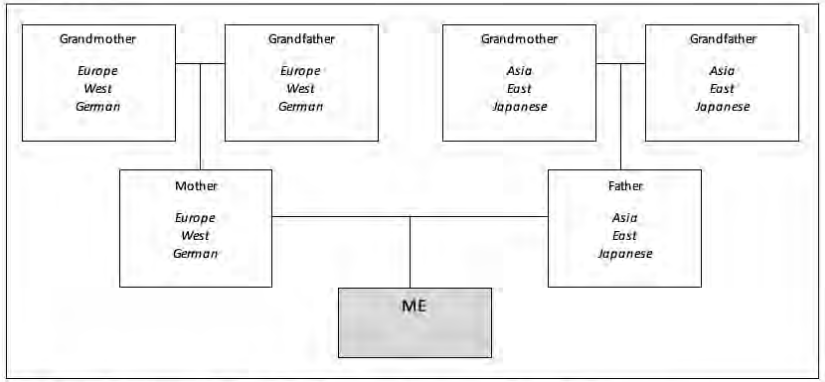
Continent	Region	Extra information
Europe	East Europe	Country/Ethnic group
	West Europe	
	South Europe	
	North Europe	
Asia	Middle East	
	East Asia	
	South Asia	
Africa	North Africa	
	East Africa	
	South Africa	
	Central Africa	
America	North America	
	Central America	
	South America	
Oceania	Australia	
	Oceania	

Institut für Rechtsmedizin München

Questionnaire

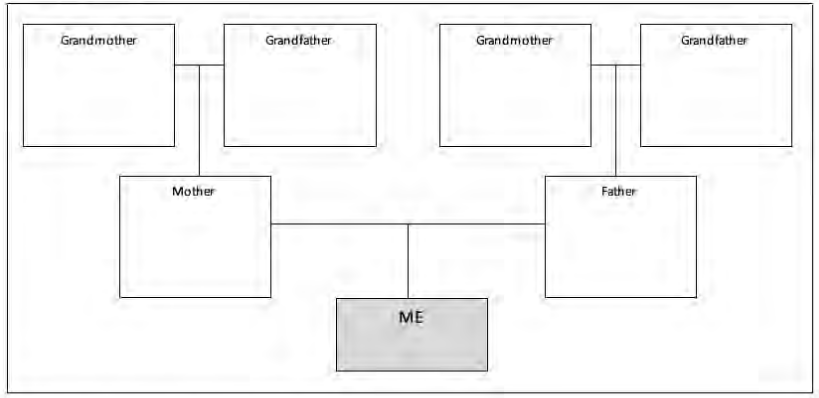
Implementation of Ancestry Informative Markers (AIMs) and Phenotype Informative SNPs (PI-SNPs) in forensic casework

Example:



EXTRA INFORMATION (if possible): *The relatives from mother side had Polish ancestry*

Your ancestry:



EXTRA INFORMATION (if possible):

Institut für Rechtsmedizin München

Questionnaire

Implementation of Ancestry Informative Markers (AIMs) and Phenotype Informative SNPs (PI-SNPs) in forensic casework

Please describe your physical appearance using tables below. Check a box corresponding with your eye, hair and skin color:

	blue	intermediate (green)	brown
Eye color			

	Color				Shade	
	blond	red	brown	black	light	dark
Hair color (natural)						

	very pale	pale	intermediate	dark	dark-black
Skin color					

Sample collection

Open Access Article

Evaluation of the Ion AmpliSeq™ PhenoTrivium Panel: MPS-Based Assay for Ancestry and Phenotype Predictions Challenged by Casework Samples

by Marta Diepenbroek [1,*], Birgit Bayer [1], Kristina Schwender [1], Roberta Schiller [1], Jessica Lim [2], Robert Lagacé [2] and Katja Anslinger [1]

[1] Department of Forensic Genetics, Institute of Legal Medicine, Ludwig Maximilian University of Munich, Nußbaumstraße 26, 80336 Munich, Bavaria, Germany

[2] Human Identification Group, Thermo Fisher Scientific, 180 Oyster Point Blvd, South San Francisco, CA 94080, USA

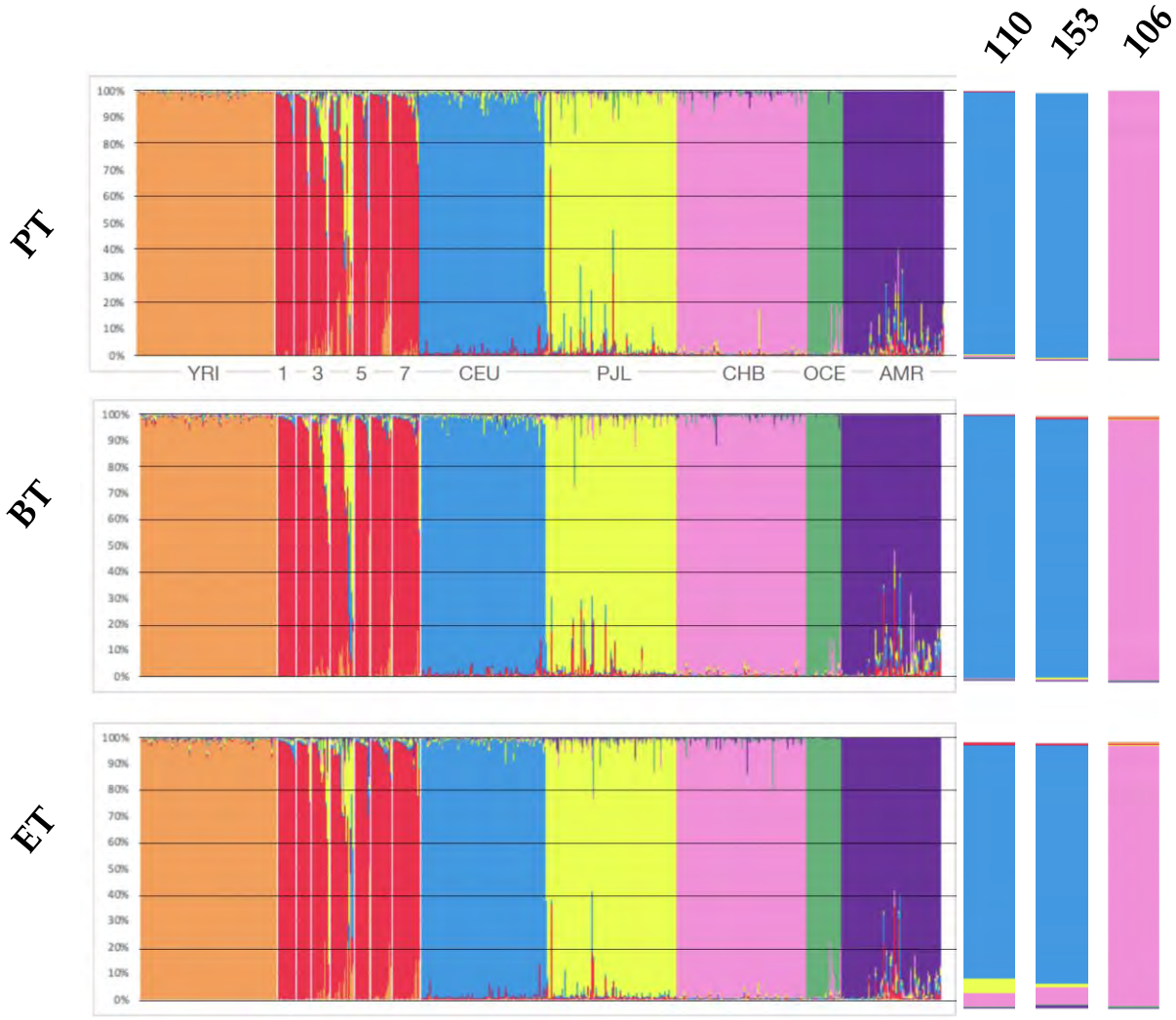
* Author to whom correspondence should be addressed.

Genes **2020**, *11*(12), 1398; https://doi.org/10.3390/genes11121398

140 samples published in the paper

9 samples selected for this study

KS110, KS153, KS106 – admixture analysis - STRUCTURE



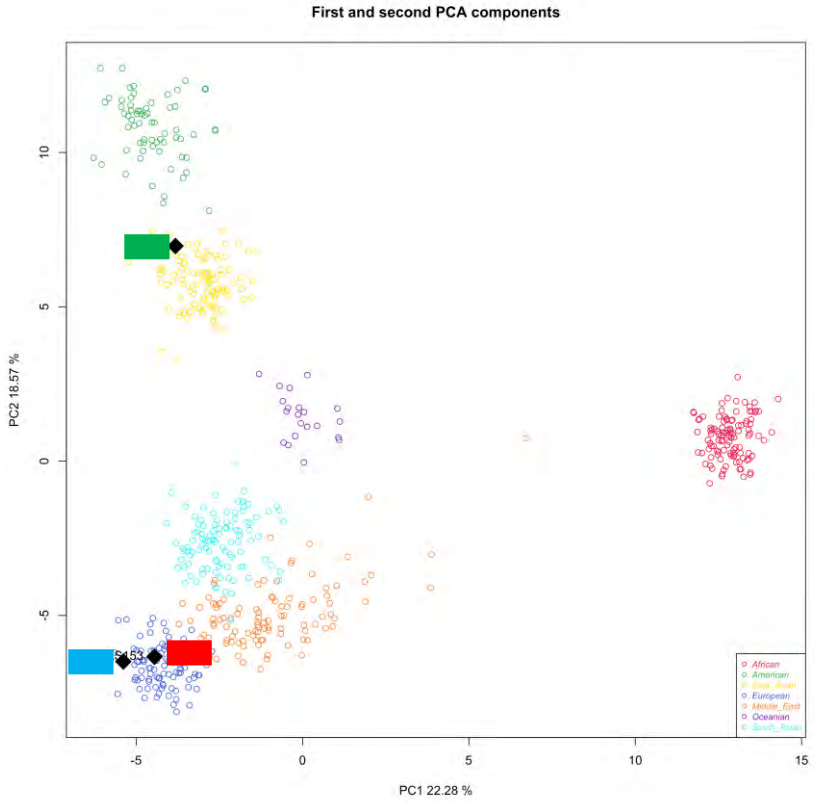
	110			153			106		
	PT	BT	ET	PT	BT	ET	PT	BT	ET
AFR	0.001	0.001	0.001	0.001	0.005	0.001	0.001	0.009	0.007
ME	0.006	0.005	0.009	0.003	0.007	0.008	0.001	0.002	0.003
EUR	0.977	0.987	0.880	0.988	0.975	0.900	0.001	0.001	0.001
SAS	0.005	0.002	0.056	0.004	0.007	0.013	0.001	0.002	0.002
EAS	0.006	0.004	0.051	0.001	0.002	0.066	0.991	0.980	0.982
OCE	0.002	0.001	0.002	0.002	0.002	0.003	0.003	0.003	0.005
AME	0.002	0.001	0.001	0.001	0.002	0.009	0.003	0.003	0.001

KS110, KS153, KS106 – admixture analysis - CONVERGE

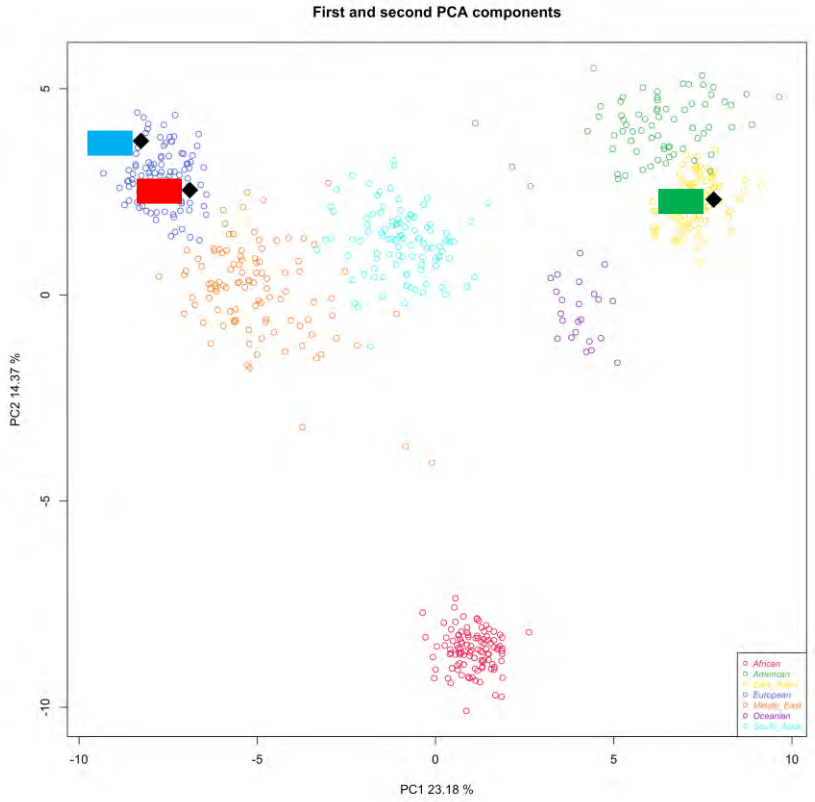
PT only

	110		153		106	
	TFS	MAC	TFS	MAC	TFS	MAC
AFR	0.000	0.000	0.000	0.000	0.000	0.000
ME	0.001	0.001	0.001	0.000	0.000	0.000
EUR	0.999	0.999	0.999	0.913	0.000	0.000
SAS	0.000	0.000	0.000	0.087	0.000	0.000
EAS	0.000	0.000	0.000	0.000	0.995	0.968
OCE	0.000	0.000	0.000	0.000	0.005	0.032
AME	0.000	0.000	0.000	0.000	0.000	0.000

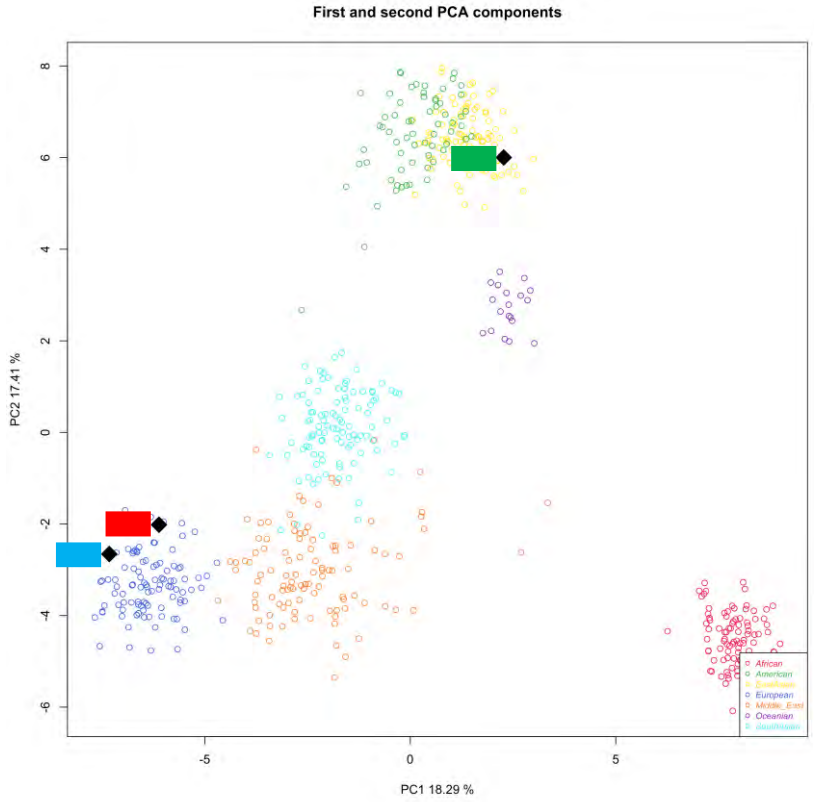
KS110, KS153, KS106 - PCA



PT



BT



ET

KS110, KS153, KS106 - GenoGeographer

110		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
Europe	-0.353	0.638
BT		
population	*z-score*	*p-value*
Europe	0.509	0.305
ET		
population	*z-score*	*p-value*
All populations rejected		

153		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
Europe	-1.159	0.877
BT		
population	*z-score*	*p-value*
Europe	-0.283	0.611
ET		
population	*z-score*	*p-value*
Europe	1.633	0.051

106		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
All populations are rejected		
BT		
population	*z-score*	*p-value*
East Asia	-1.33	0.908
ET		
population	*z-score*	*p-value*
East Asia	0.736	0.231

KS110, KS153, KS106 – extra markers

			110	153	106
			p-value		
eye colour		*blue*	0,099		
		intermediate	0,134		
		brown	**0,767**		
hair	colour	*blond*	**0,362**		
		brown	**0,521**		
		red	0,002		
		black	0,116		
	shade	*light*	**0,742**		
		dark	0,258		
skin colour		*very pale*	0,008		
		pale	**0,304**		
		intermediate	**0,681**		
		dark	0,007		
		dark to black	0,000		

	110	153	106
mtDNA	H2a1c		
87 ET Y-SNPs	♀		
116 PT Y-SNPs	♀		
16 ET X-SNPs	European specific X-chromosomes		

KS110, KS153, KS106 – extra markers

			110	153	106
			p-value		
eye colour		*blue*	0,099	**0,903**	
		intermediate	0,134	0,074	
		brown	**0,767**	0,023	
hair	colour	*blond*	**0,362**	**0,319**	
		brown	**0,521**	**0,605**	
		red	0,002	0,007	
		black	0,116	0,068	
	shade	*light*	**0,742**	**0,812**	
		dark	0,258	0,188	
skin colour		*very pale*	0,008	0,078	
		pale	**0,304**	**0,563**	
		intermediate	**0,681**	**0,369**	
		dark	0,007	0,000	
		dark to black	0,000	0,000	

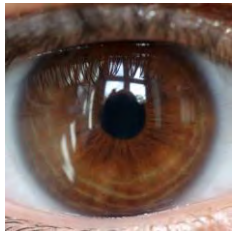
	110	153	106
mtDNA	H2a1c	U4c1	
87 ET Y-SNPs	♀	R-P312	
116 PT Y-SNPs	♀	R-U152	
16 ET X-SNPs	European specific X-chromosomes	European specific X-chromosome	

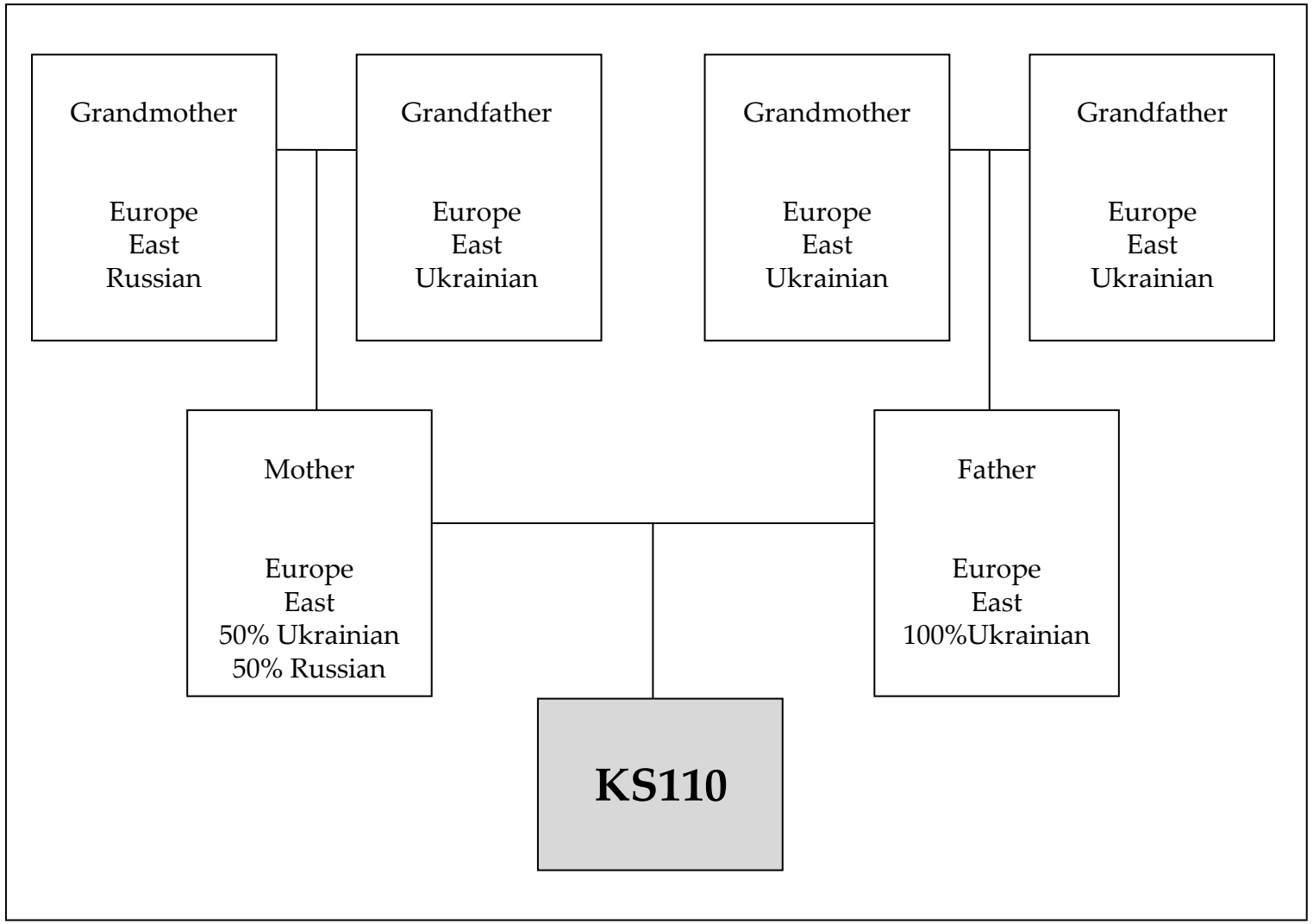
KS110, KS153, KS106 – extra markers

			110	153	106
			p-value		
eye colour		*blue*	0,099	**0,903**	0,000
		intermediate	0,134	0,074	0,002
		brown	**0,767**	0,023	**0,998**
hair	colour	*blond*	**0,362**	**0,319**	0,001
		brown	**0,521**	**0,605**	0,122
		red	0,002	0,007	0,000
		black	0,116	0,068	**0,877**
	shade	*light*	**0,742**	**0,812**	0,002
		dark	0,258	0,188	**0,998**
skin colour		*very pale*	0,008	0,078	0,000
		pale	**0,304**	**0,563**	0,000
		intermediate	**0,681**	**0,369**	**0,965**
		dark	0,007	0,000	0,035
		dark to black	0,000	0,000	0,000

	110	153	106
mtDNA	H2a1c	U4c1	F1e3
87 ET Y-SNPs	♀	R-P312	O-M119
116 PT Y-SNPs	♀	R-U152	O-M119
16 ET X-SNPs	European specific X-chromosomes	European specific X-chromosome	East Asian specific X-chromosome

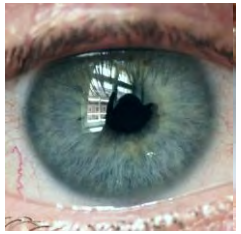
KS110 - reveal!

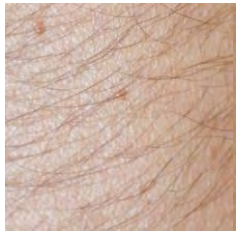


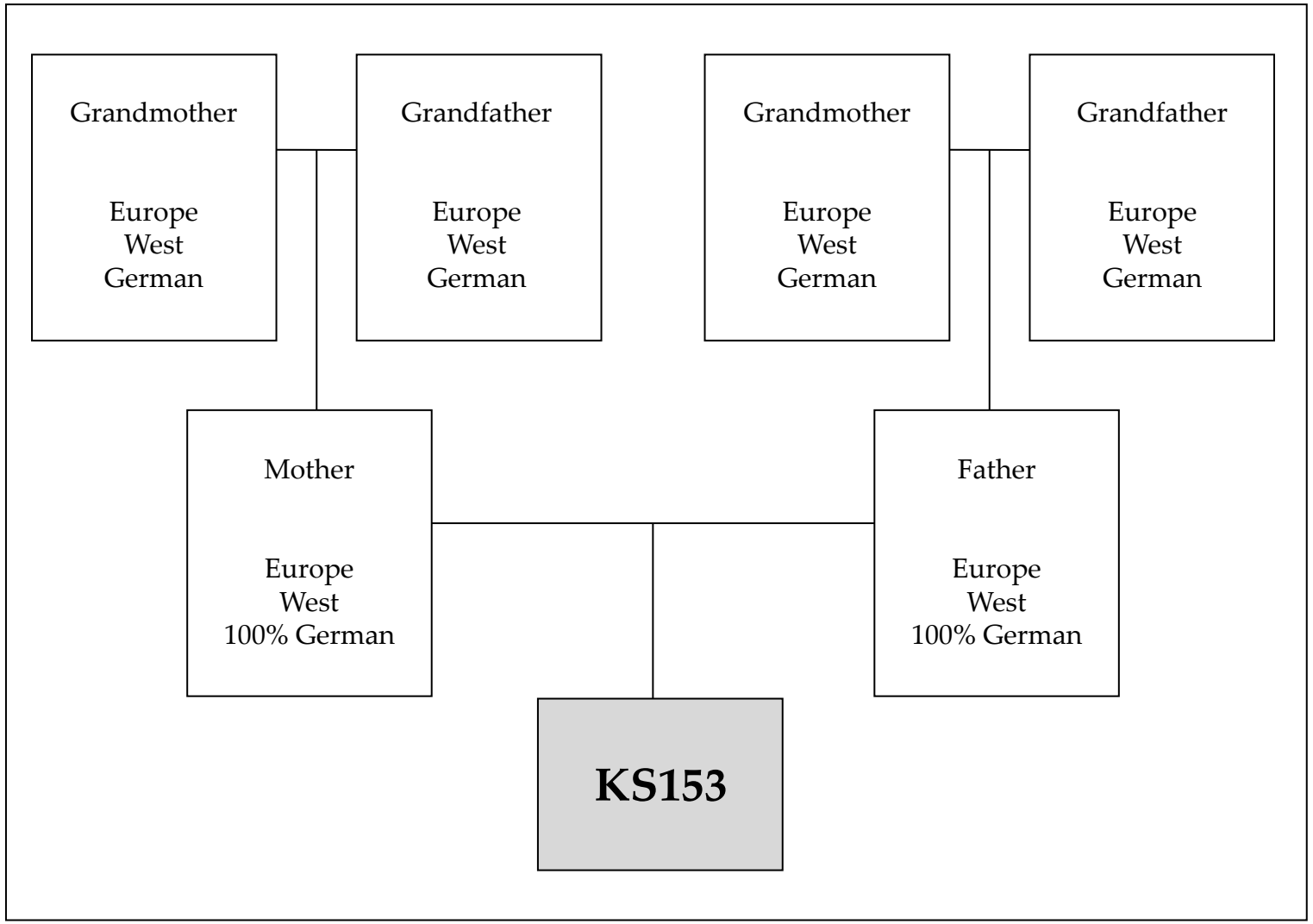


	PT	BT	ET
AFR	0.001	0.001	0.001
ME	0.006	0.005	0.009
EUR	0.977	0.987	0.880
SAS	0.005	0.002	0.056
EAS	0.006	0.004	0.051
OCE	0.002	0.001	0.002
AME	0.002	0.001	0.001

KS153 - reveal!

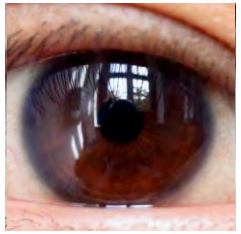


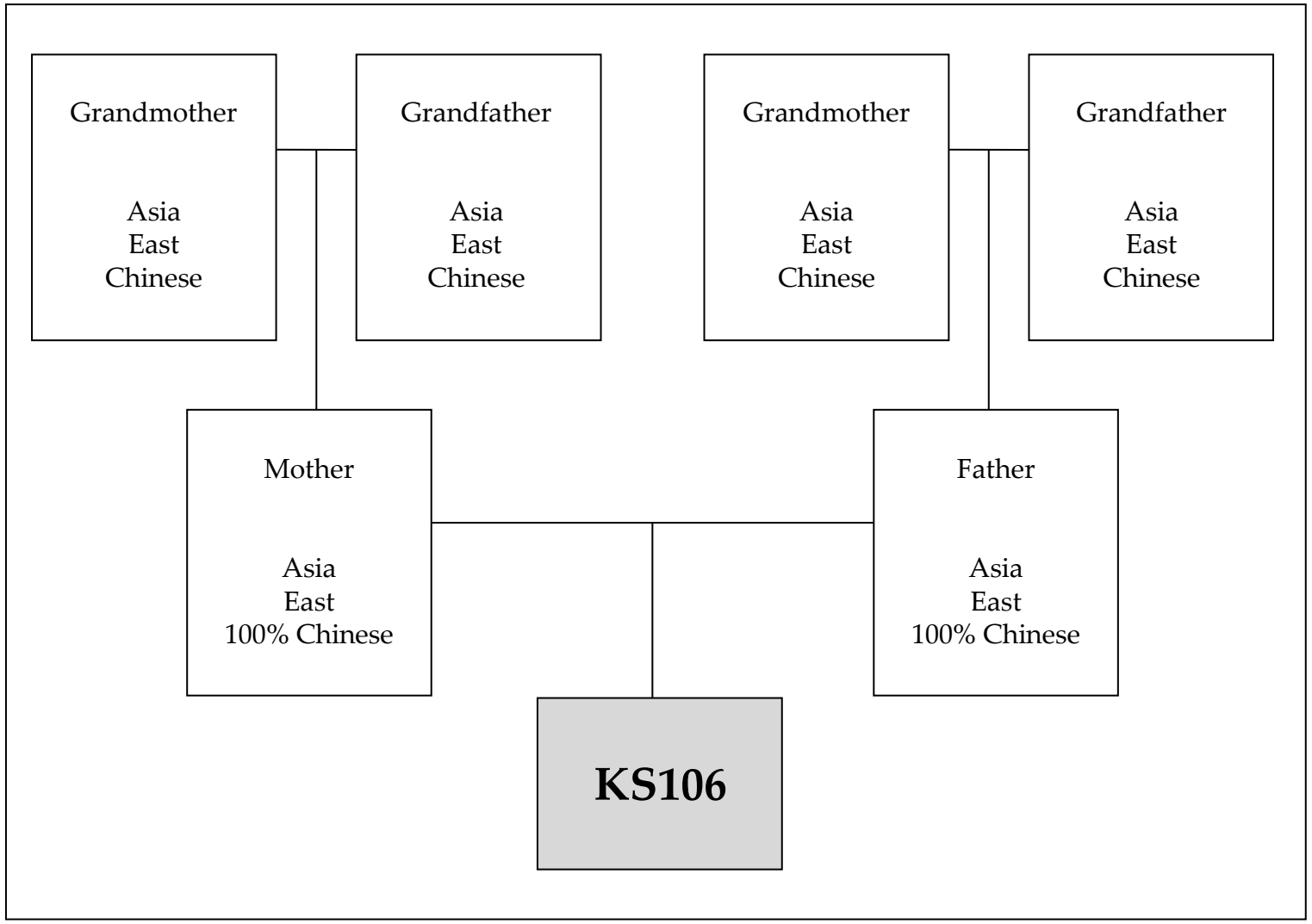




	PT	BT	ET
AFR	0.001	0.005	0.001
ME	0.003	0.007	0.008
EUR	0.988	0.975	0.900
SAS	0.004	0.007	0.013
EAS	0.001	0.002	0.066
OCE	0.002	0.002	0.003
AME	0.001	0.002	0.009

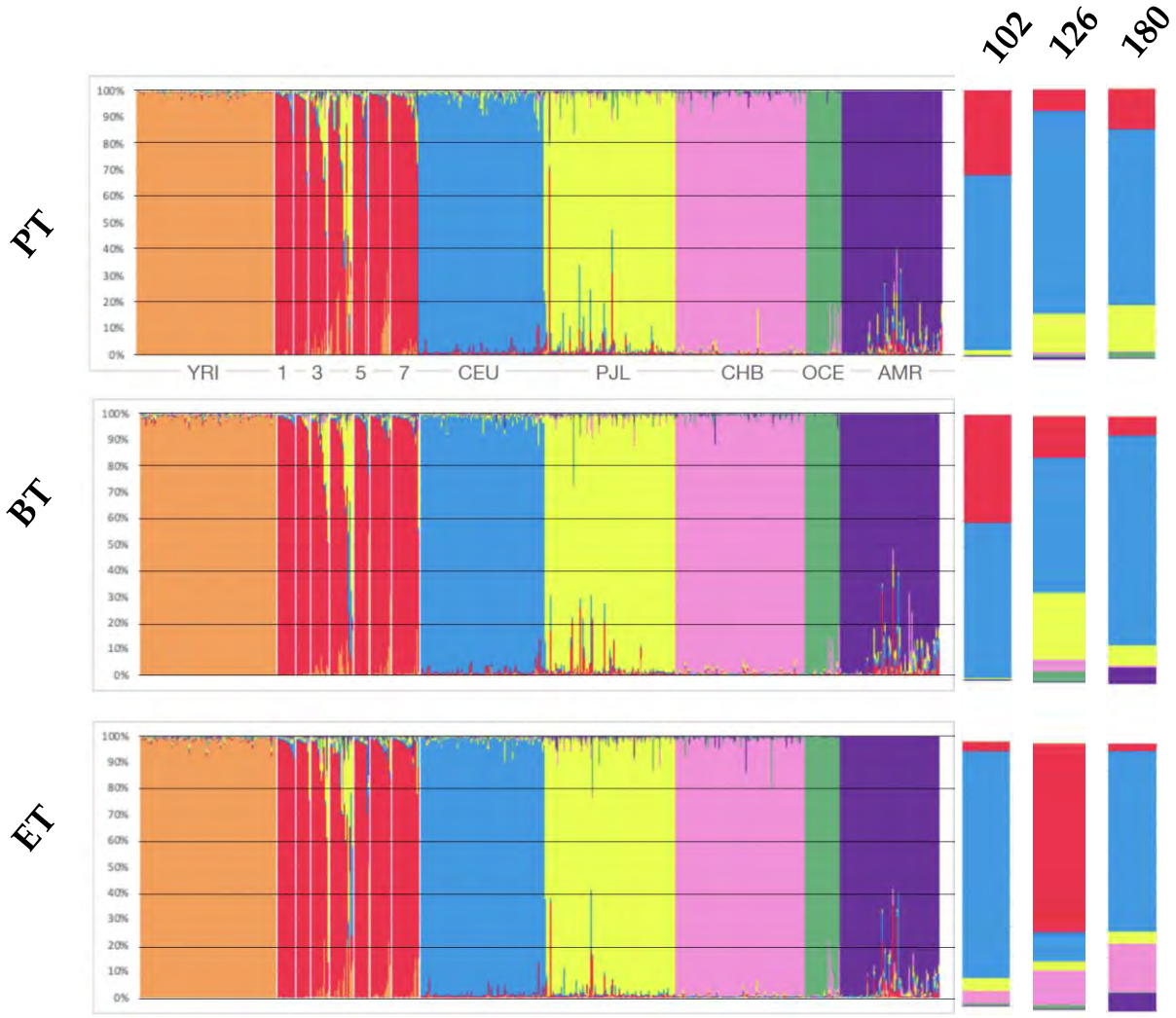
KS106 - reveal!





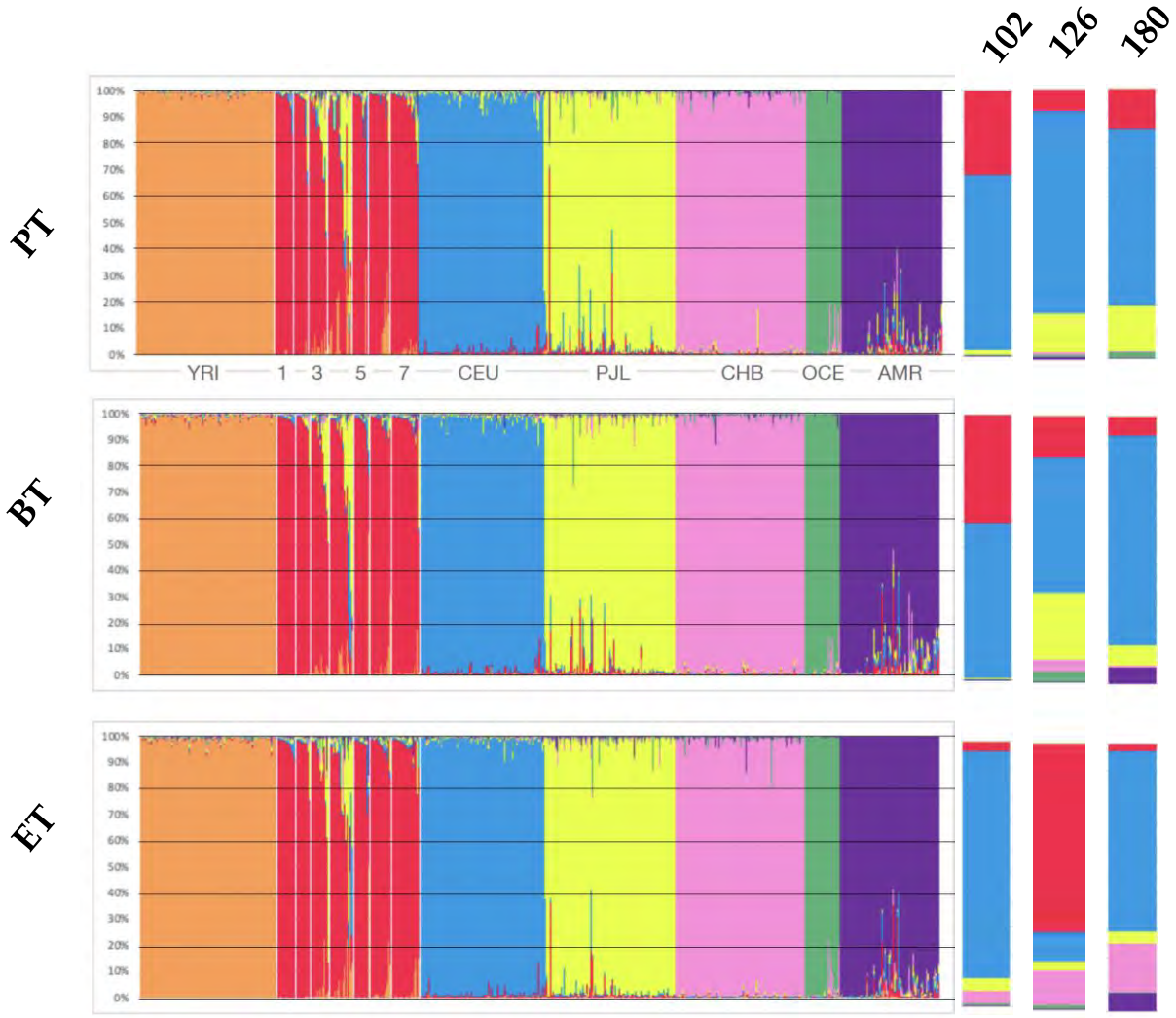
	PT	BT	ET
AFR	0.001	0.009	0.007
ME	0.001	0.002	0.003
EUR	0.001	0.001	0.001
SAS	0.001	0.002	0.002
EAS	0.991	0.980	0.982
OCE	0.003	0.003	0.005
AME	0.003	0.003	0.001

KS102, KS126, KS180 – admixture analysis



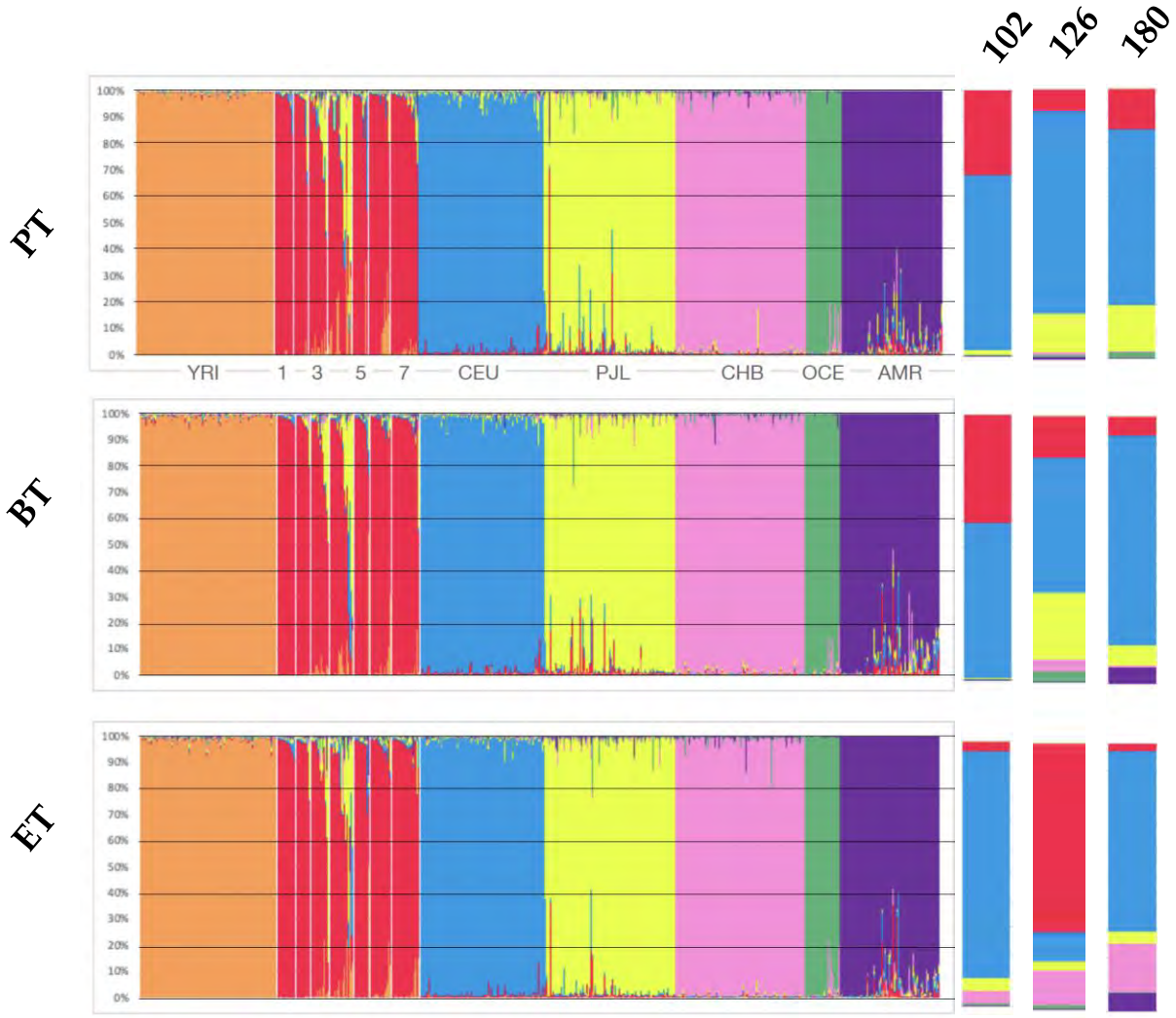
	102		
	PT	BT	ET
AFR	0.003	0.002	0.003
ME	0.317	0.406	0.036
EUR	0.658	0.585	0.855
SAS	0.017	0.004	0.048
EAS	0.002	0.001	0.048
OCE	0.004	0.002	0.004
AME	0.001	0.001	0.005

KS102, KS126, KS180 – admixture analysis



	102			126		
	PT	BT	ET	PT	BT	ET
AFR	0.003	0.002	0.003	0.003	0.003	0.004
ME	0.317	0.406	0.036	0.079	0.156	0.704
EUR	0.658	0.585	0.855	0.751	0.505	0.112
SAS	0.017	0.004	0.048	0.141	0.250	0.035
EAS	0.002	0.001	0.048	0.011	0.046	0.130
OCE	0.004	0.002	0.004	0.007	0.036	0.012
AME	0.001	0.001	0.005	0.008	0.005	0.003

KS102, KS126, KS180 – admixture analysis



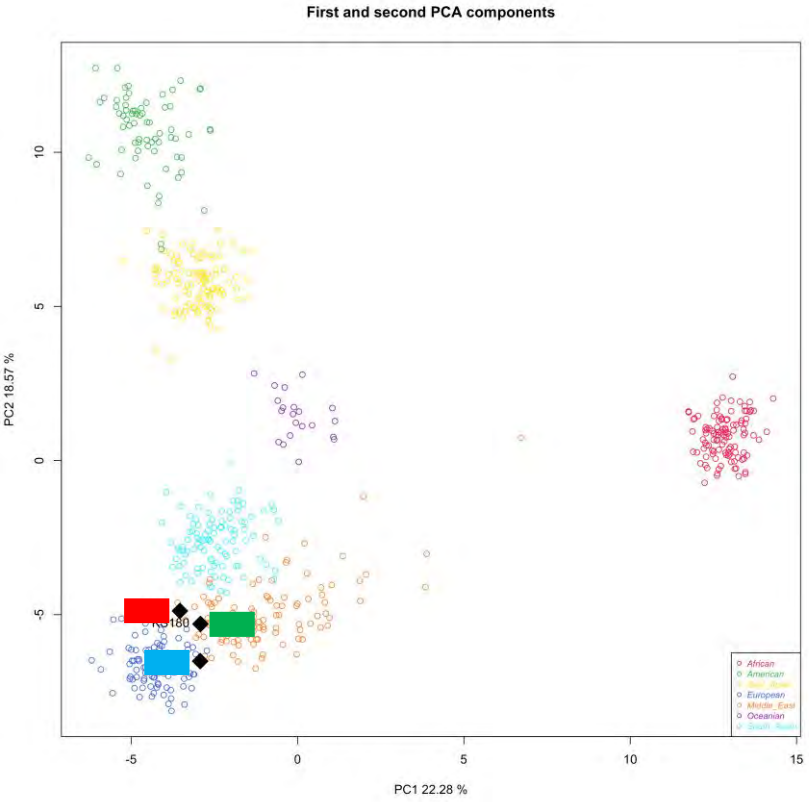
	102			126			180		
	PT	BT	ET	PT	BT	ET	PT	BT	ET
AFR	0.003	0.002	0.003	0.003	0.003	0.004	0.005	0.003	0.002
ME	0.317	0.406	0.036	0.079	0.156	0.704	0.149	0.069	0.025
EUR	0.658	0.585	0.855	0.751	0.505	0.112	0.648	0.786	0.674
SAS	0.017	0.004	0.048	0.141	0.250	0.035	0.171	0.074	0.047
EAS	0.002	0.001	0.048	0.011	0.046	0.130	0.005	0.007	0.179
OCE	0.004	0.002	0.004	0.007	0.036	0.012	0.019	0.001	0.006
AME	0.001	0.001	0.005	0.008	0.005	0.003	0.003	0.061	0.067

KS102, KS126, KS180 – admixture analysis - CONVERGE

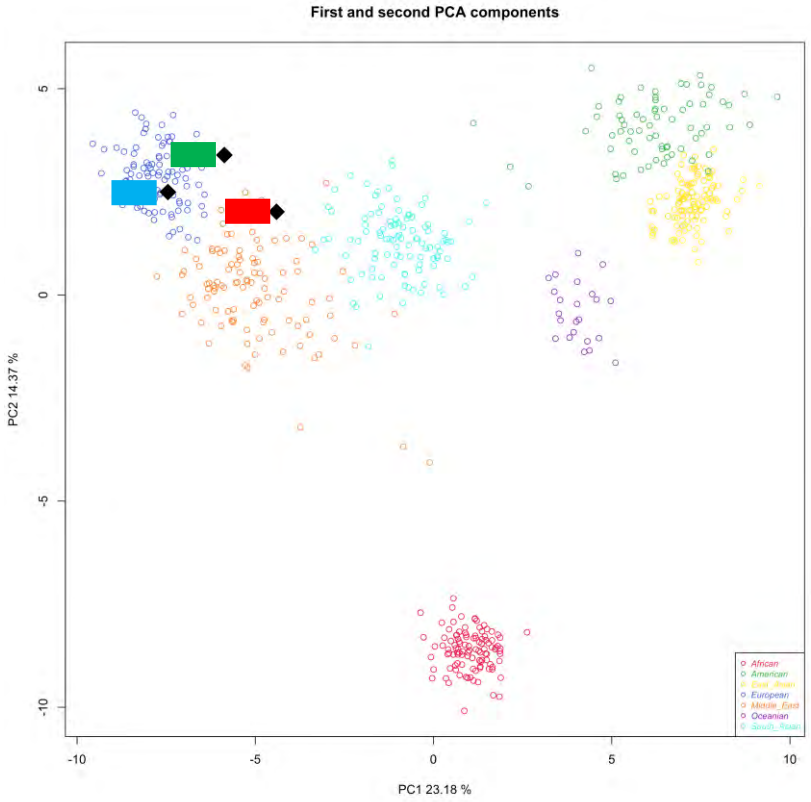
PT only

	102		126		180	
	TFS	MAC	TFS	MAC	TFS	MAC
AFR	0.000	0.000	0.000	0.000	0.007	0.005
ME	0.445	0.439	0.445	0.439	0.030	0.031
EUR	0.549	0.525	0.549	0.525	0.657	0.633
SAS	0.003	0.036	0.003	0.036	0.306	0.331
EAS	0.001	0.000	0.001	0.000	0.000	0.000
OCE	0.001	0.000	0.001	0.000	0.000	0.000
AME	0.001	0.000	0.001	0.000	0.000	0.000

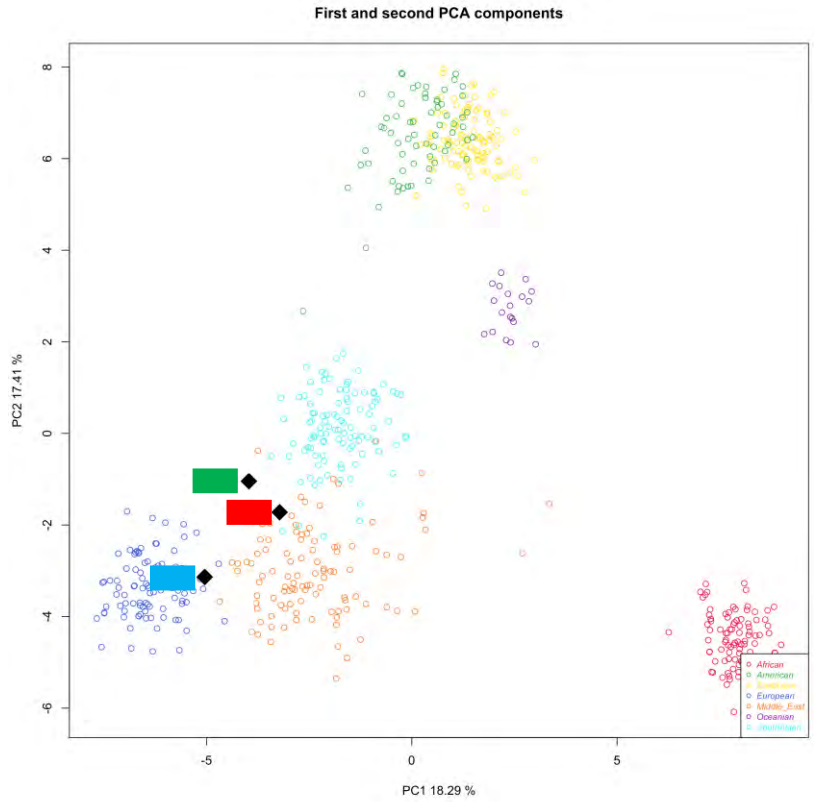
KS102, KS126, KS180 - PCA



PT



BT



ET

KS102, KS126, KS180 - GenoGeographer

102		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
Middle East	-0.956	0.830
BT		
population	*z-score*	*p-value*
Europe & ME	-1.546	0.939
ET		
population	*z-score*	*p-value*
Europe & ME	1.356	0.088

126		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
ME & S. Asia	0.306	0.380
BT		
population	*z-score*	*p-value*
Europe & S. Asia	1.525	0.064
ET		
population	*z-score*	*p-value*
Middle East	0.347	0.364

180		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
Europe & S. Asia	-0.865	0.806
BT		
population	*z-score*	*p-value*
Europe & S. Asia	0.257	0.398
ET		
population	*z-score*	*p-value*
Middle East	-0.087	0.535

KS102, KS126, KS180 – extra markers

			102	126	180
			p-value		
eye colour		*blue*	0,001		
		intermediate	0,032		
		brown	**0,967**		
hair	colour	*blond*	0,027		
		brown	**0,604**		
		red	0,000		
		black	**0,369**		
	shade	*light*	0,060		
		dark	**0,940**		
skin colour		*very pale*	0,006		
		pale	**0,159**		
		intermediate	**0,807**		
		dark	0,026		
		dark to black	0,002		

	102	126	180
mtDNA	W1		
87 ET Y-SNPs	R-CTS1078		
116 PT Y-SNPs	R-M269		
16 ET X-SNPs	European specific X-chromosome		

KS102, KS126, KS180 – extra markers

			102	126	180
			p-value		
eye colour		*blue*	0,001	0,000	
		intermediate	0,032	0,013	
		brown	**0,967**	**0,986**	
hair	colour	*blond*	0,027	0,101	
		brown	**0,604**	**0,657**	
		red	0,000	0,000	
		black	**0,369**	**0,241**	
	shade	*light*	0,060	0,246	
		dark	**0,940**	**0,754**	
skin colour		*very pale*	0,006	0,004	
		pale	**0,159**	0,103	
		intermediate	**0,807**	**0,755**	
		dark	0,026	**0,136**	
		dark to black	0,002	0,002	

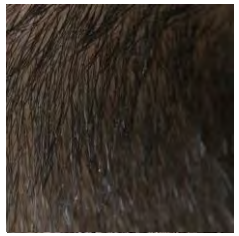
	102	126	180
mtDNA	W1	U7a	
87 ET Y-SNPs	R-CTS1078	R-CTS1078	
116 PT Y-SNPs	R-M269	R-M269	
16 ET X-SNPs	European specific X-chromosome	European specific X-chromosome	

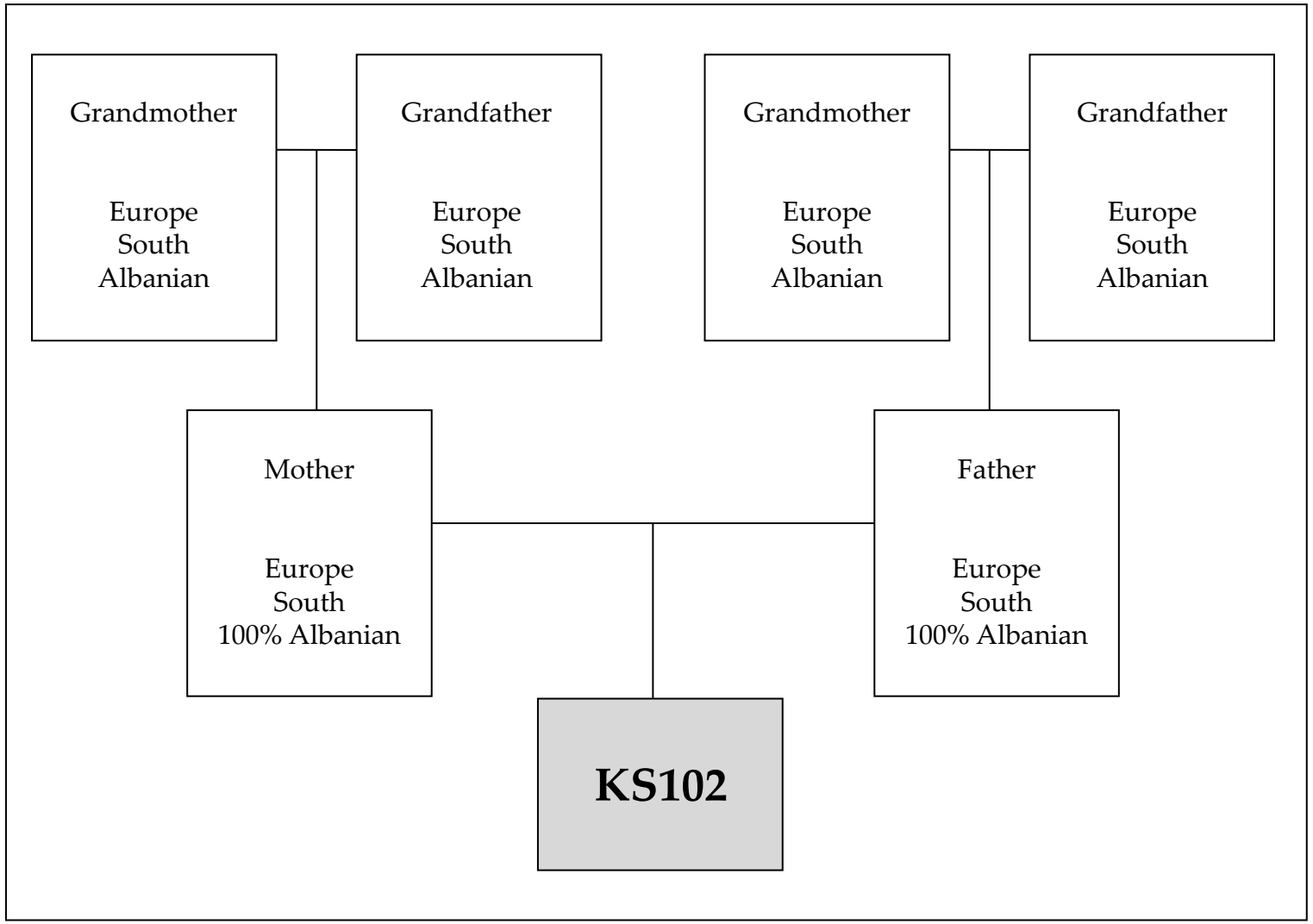
KS102, KS126, KS180 – extra markers

			102	126	180
			p-value		
eye colour		*blue*	0,001	0,000	0,082
		intermediate	0,032	0,013	0,142
		brown	**0,967**	**0,986**	**0,776**
hair	colour	*blond*	0,027	0,101	0,256
		brown	**0,604**	**0,657**	**0,617**
		red	0,000	0,000	0,058
		black	**0,369**	**0,241**	0,069
	shade	*light*	0,060	0,246	**0,864**
		dark	**0,940**	**0,754**	0,136
skin colour		*very pale*	0,006	0,004	0,138
		pale	**0,159**	0,103	**0,458**
		intermediate	**0,807**	**0,755**	**0,383**
		dark	0,026	**0,136**	0,021
		dark to black	0,002	0,002	0,000

	102	126	180
mtDNA	W1	U7a	U2e2a1a
87 ET Y-SNPs	R-CTS1078	R-CTS1078	J-M172
116 PT Y-SNPs	R-M269	R-M269	J-M67
16 ET X-SNPs	European specific X-chromosome	European specific X-chromosome	European specific X-chromosome

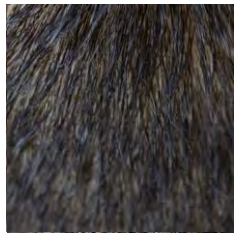
KS102 - reveal!

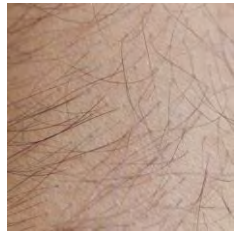


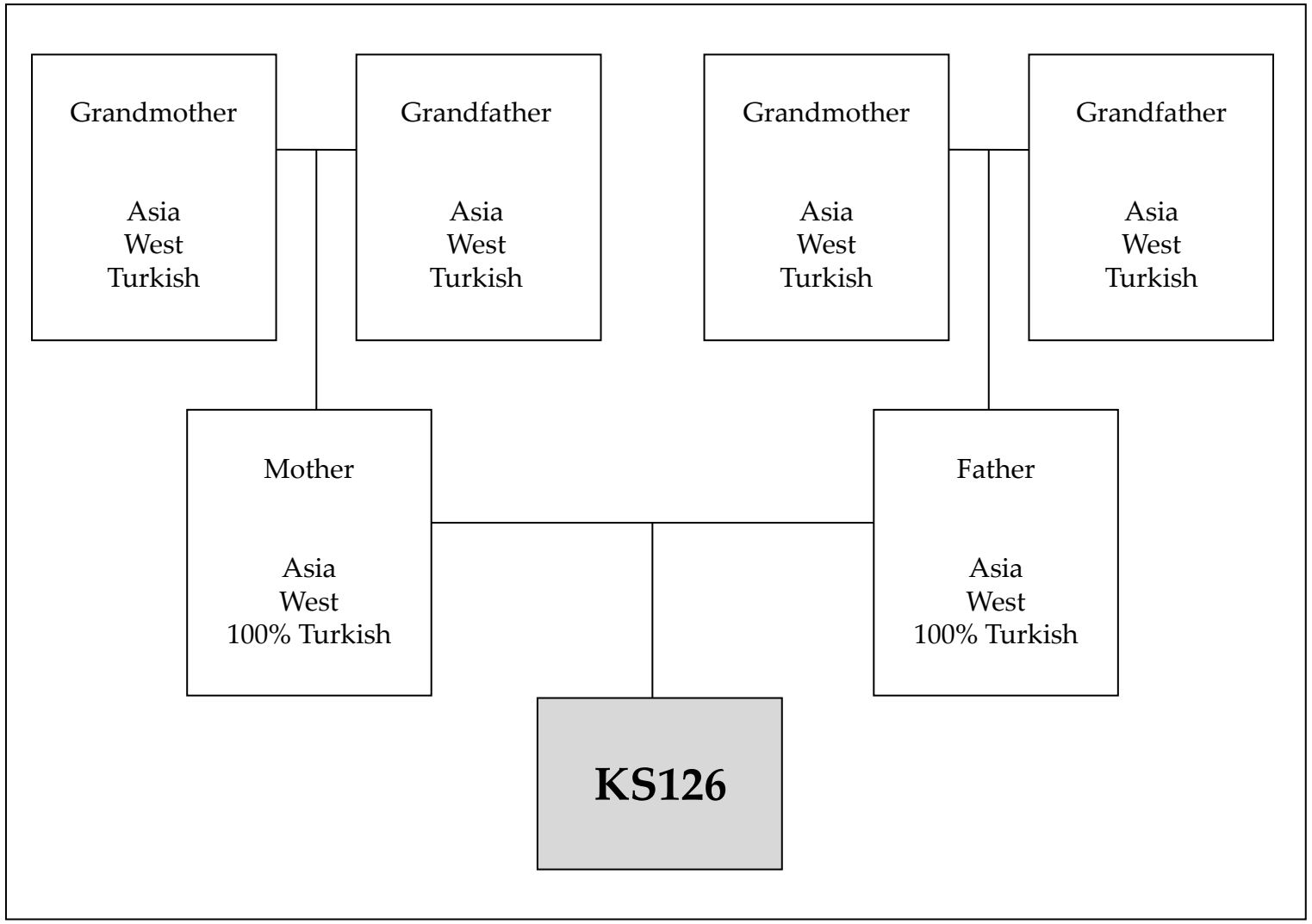


	PT	BT	ET
AFR	0.003	0.002	0.003
ME	0.317	0.406	0.036
EUR	0.658	0.585	0.855
SAS	0.017	0.004	0.048
EAS	0.002	0.001	0.048
OCE	0.004	0.002	0.004
AME	0.001	0.001	0.005

KS126 - reveal!

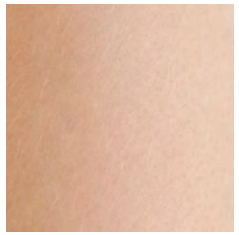


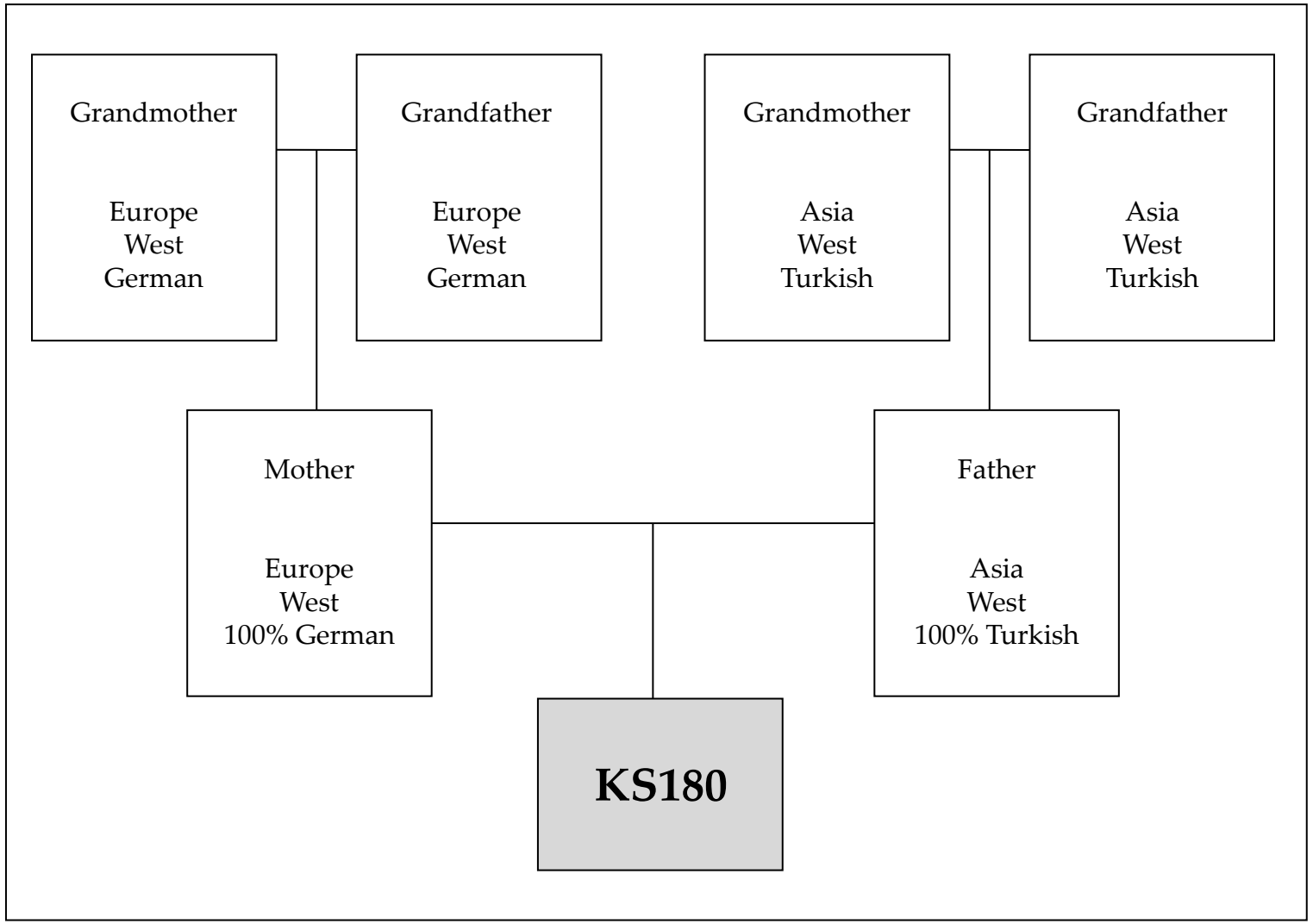




	PT	BT	ET
AFR	0.003	0.003	0.004
ME	0.079	0.156	0.704
EUR	0.751	0.505	0.112
SAS	0.141	0.250	0.035
EAS	0.011	0.046	0.130
OCE	0.007	0.036	0.012
AME	0.008	0.005	0.003

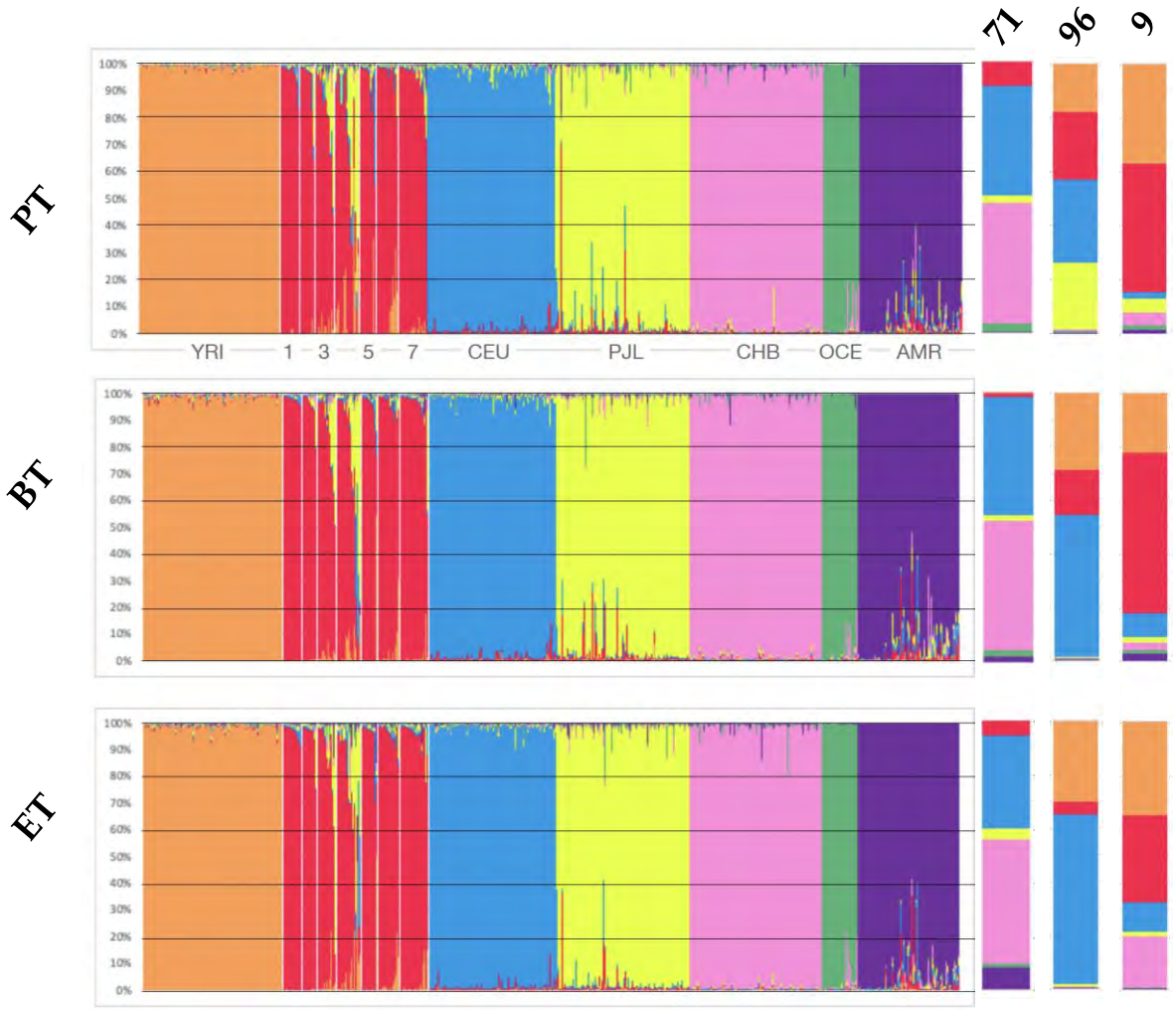
KS180 - reveal!





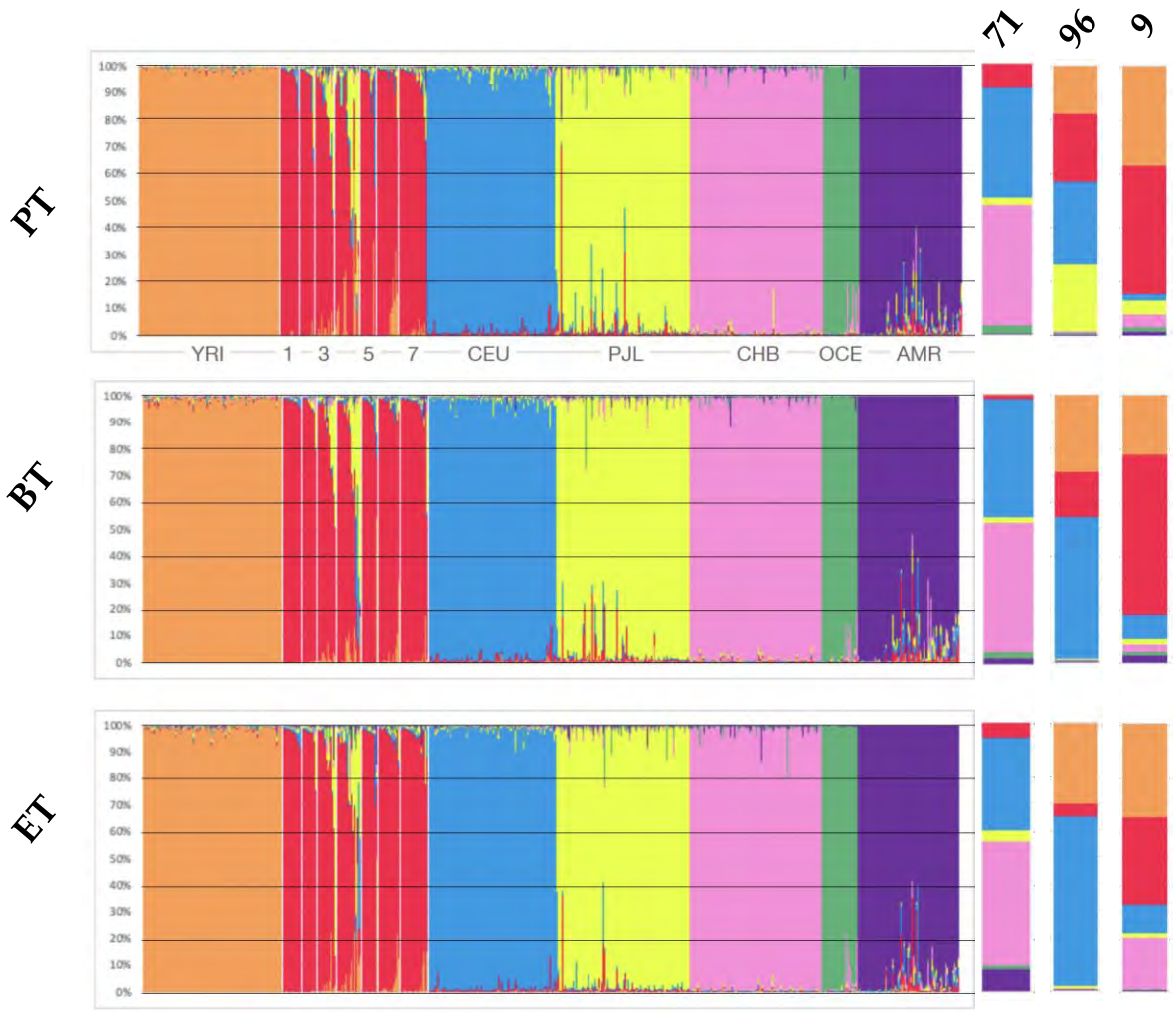
	PT	BT	ET
AFR	0.005	0.003	0.002
ME	0.149	0.069	0.025
EUR	0.648	0.786	0.674
SAS	0.171	0.074	0.047
EAS	0.005	0.007	0.179
OCE	0.019	0.001	0.006
AME	0.003	0.061	0.067

KS71, KS96, KS9 – admixture analysis



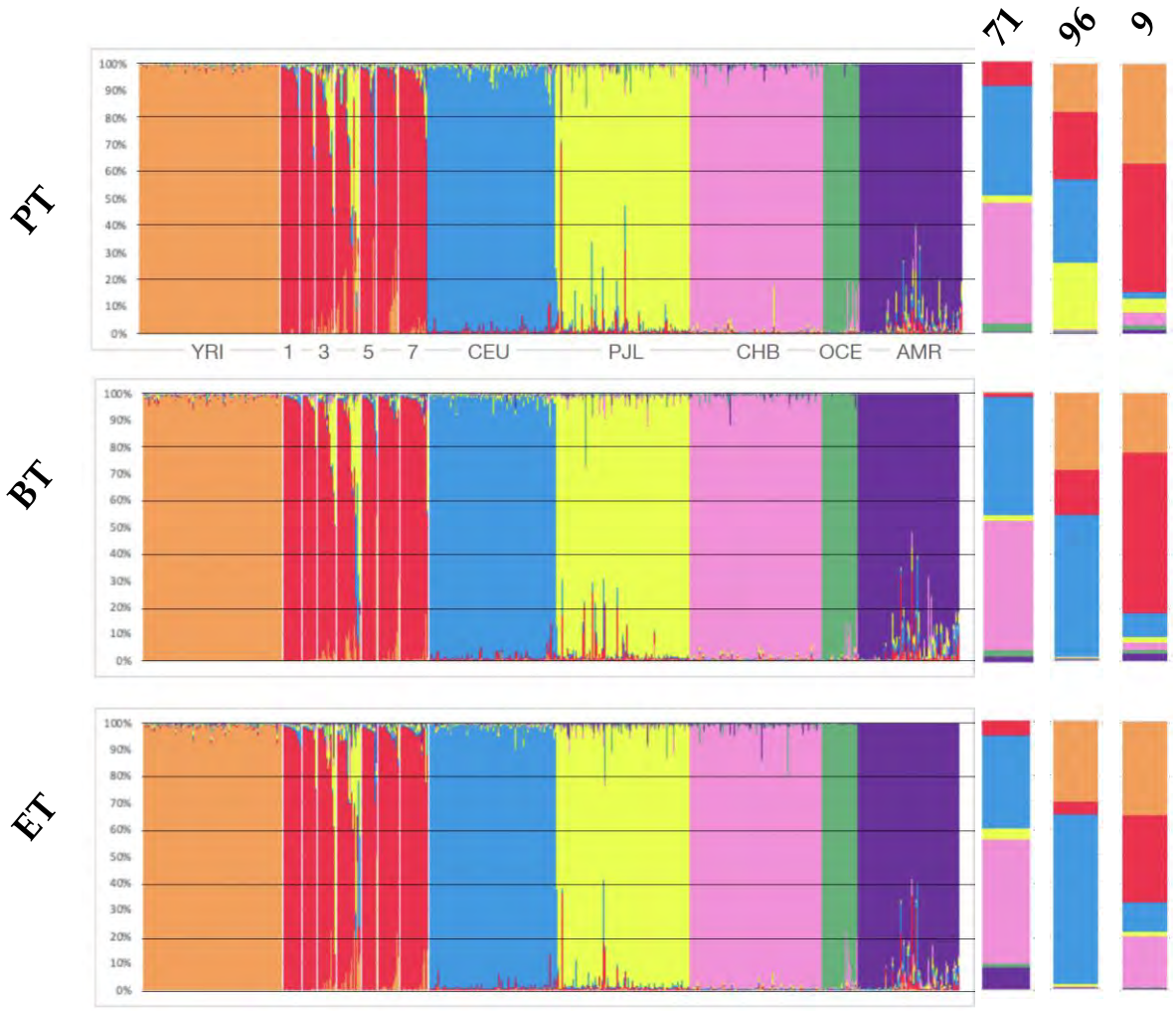
	71		
	PT	BT	ET
AFR	0.002	0.003	0.002
ME	0.089	0.014	0.054
EUR	0.405	0.438	0.344
SAS	0.026	0.021	0.042
EAS	0.447	0.479	0.463
OCE	0.029	0.025	0.012
AME	0.003	0.020	0.083

KS71, KS96, KS9 – admixture analysis



	71			96		
	PT	BT	ET	PT	BT	ET
AFR	0.002	0.003	0.002	0.180	0.287	0.302
ME	0.089	0.014	0.054	0.249	0.170	0.048
EUR	0.405	0.438	0.344	0.308	0.531	0.633
SAS	0.026	0.021	0.042	0.247	0.004	0.011
EAS	0.447	0.479	0.463	0.006	0.002	0.002
OCE	0.029	0.025	0.012	0.007	0.003	0.002
AME	0.003	0.020	0.083	0.003	0.003	0.002

KS71, KS96, KS9 – admixture analysis



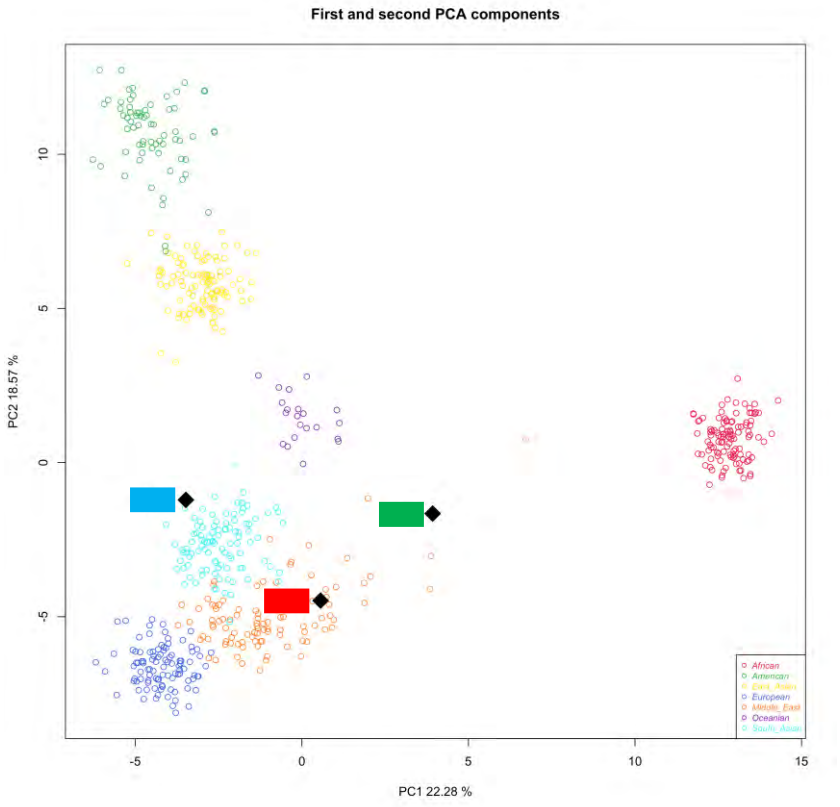
	71			96			9		
	PT	BT	ET	PT	BT	ET	PT	BT	ET
AFR	0.002	0.003	0.002	0.180	0.287	0.302	0.372	0.222	0.349
ME	0.089	0.014	0.054	0.249	0.170	0.048	0.475	0.600	0.326
EUR	0.405	0.438	0.344	0.308	0.531	0.633	0.026	0.090	0.111
SAS	0.026	0.021	0.042	0.247	0.004	0.011	0.049	0.019	0.016
EAS	0.447	0.479	0.463	0.006	0.002	0.002	0.047	0.029	0.192
OCE	0.029	0.025	0.012	0.007	0.003	0.002	0.018	0.011	0.003
AME	0.003	0.020	0.083	0.003	0.003	0.002	0.012	0.029	0.004

KS71, KS96, KS9 – admixture analysis - CONVERGE

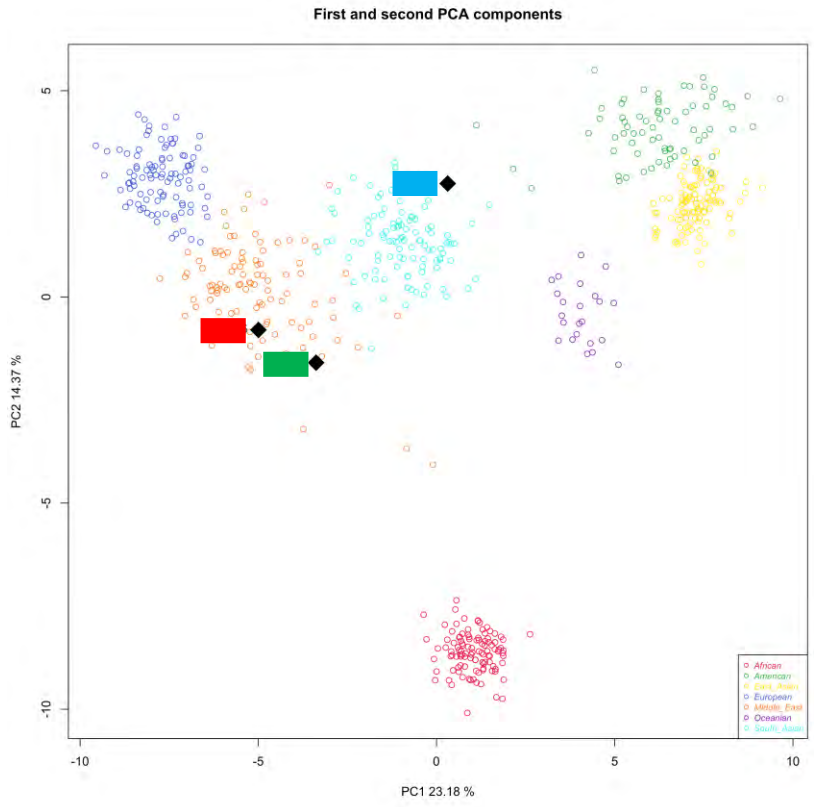
PT only

	71		96		9	
	TFS	MAC	TFS	MAC	TFS	MAC
AFR	0.000	0.000	0.283	0.250	0.463	0.115
ME	0.012	0.242	0.115	0.137	0.236	0.732
EUR	0.479	0.291	0.582	0.551	0.007	0.000
SAS	0.033	0.000	0.020	0.062	0.294	0.002
EAS	0.471	0.467	0.000	0.000	0.000	0.151
OCE	0.005	0.000	0.000	0.000	0.000	0.000
AME	0.000	0.000	0.000	0.000	0.000	0.000

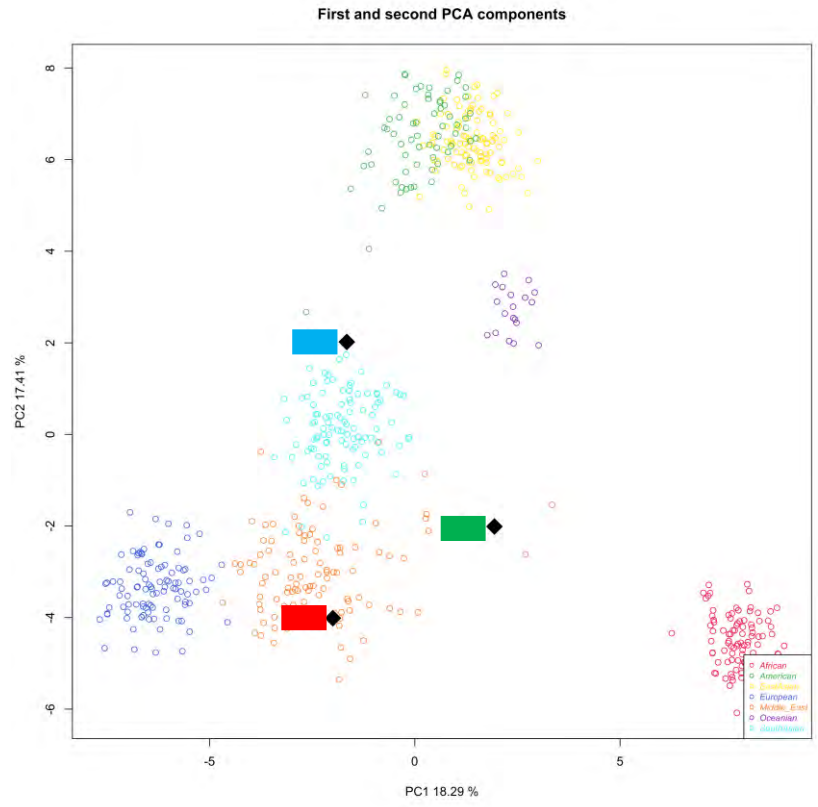
KS71, KS96, KS9 - PCA



PT



BT



ET

KS71, KS96, KS9 - GenoGeographer

71		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
E. Asia & ME	-0.153	0.561
BT		
population	*z-score*	*p-value*
South Asia	1.241	0.107
ET		
population	*z-score*	*p-value*
All populations are rejected		

96		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
Middle East	0.823	0.205
BT		
population	*z-score*	*p-value*
All populations are rejected		
ET		
population	*z-score*	*p-value*
Middle East	0.957	0.169

9		
z-score ≤ 1.64; P ≥ 0.05		
PT		
population	*z-score*	*p-value*
All populations are rejected		
BT		
population	*z-score*	*p-value*
All populations are rejected		
ET		
population	*z-score*	*p-value*
All populations are rejected		

KS71, KS96, KS9 – extra markers

			71	96	9
			p-value		
eye colour		*blue*	0,000		
		intermediate	0,007		
		brown	**0,993**		
hair	colour	*blond*	0,004		
		brown	**0,347**		
		red	0,000		
		black	**0,649**		
	shade	*light*	0,005		
		dark	**0,995**		
skin colour		*very pale*	0,000		
		pale	0,000		
		intermediate	**0,705**		
		dark	**0,264**		
		dark to black	0,030		

	71	96	9
mtDNA	M7b1a1a1		
87 ET Y-SNPs	♀		
116 PT Y-SNPs	♀		
16 ET X-SNPs	European and E.Asian specific X-chromosomes		

KS71, KS96, KS9 – extra markers

			71	96	9
			p-value		
eye colour		*blue*	0,000	0,003	
		intermediate	0,007	0,020	
		brown	**0,993**	**0,977**	
hair	colour	*blond*	0,004	0,038	
		brown	**0,347**	**0,725**	
		red	0,000	0,000	
		black	**0,649**	**0,236**	
	shade	*light*	0,005	0,103	
		dark	**0,995**	**0,897**	
skin colour		*very pale*	0,000	0,001	
		pale	0,000	0,028	
		intermediate	**0,705**	**0,262**	
		dark	**0,264**	**0,685**	
		dark to black	0,030	0,024	

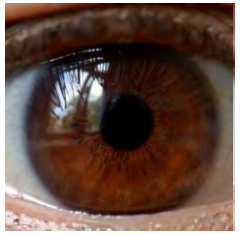
	71	96	9
mtDNA	M7b1a1a1	L3e2a1a	
87 ET Y-SNPs	♀	R-M343	
116 PT Y-SNPs	♀	R-M343	
16 ET X-SNPs	European and E.Asian specific X-chromosomes	African specific X-chromosome	

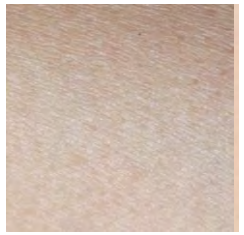
KS71, KS96, KS9 – extra markers

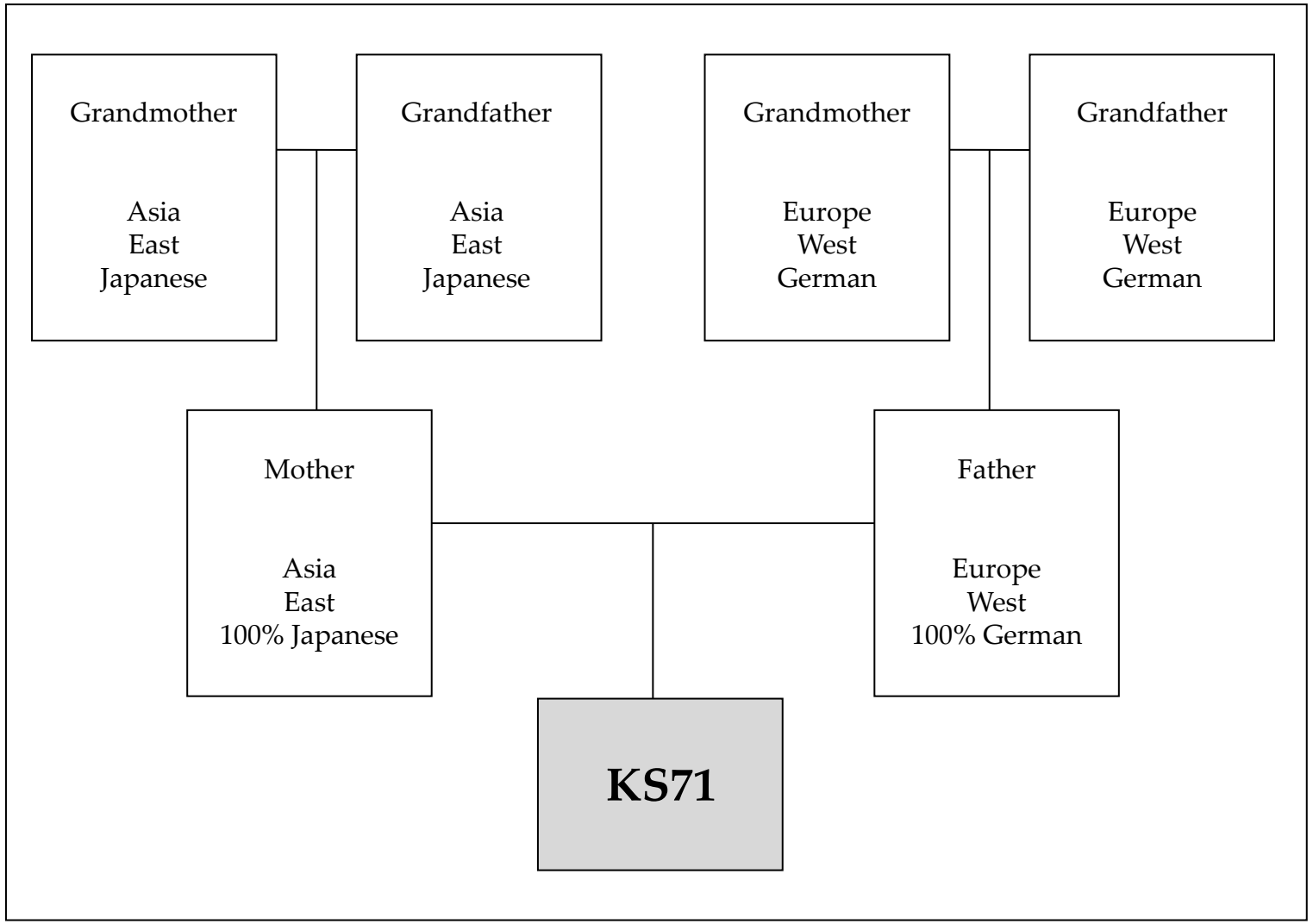
			71	96	9
			p-value		
eye colour		*blue*	0,000	0,003	0,000
		intermediate	0,007	0,020	0,002
		brown	**0,993**	**0,977**	**0,998**
hair	colour	*blond*	0,004	0,038	0,003
		brown	**0,347**	**0,725**	**0,474**
		red	0,000	0,000	0,000
		black	**0,649**	**0,236**	**0,523**
	shade	*light*	0,005	0,103	0,005
		dark	**0,995**	**0,897**	**0,995**
skin colour		*very pale*	0,000	0,001	0,000
		pale	0,000	0,028	0,000
		intermediate	**0,705**	**0,262**	0,000
		dark	**0,264**	**0,685**	0,009
		dark to black	0,030	0,024	**0,991**

	71	96	9
mtDNA	M7b1a1a1	L3e2a1a	E1a1a1
87 ET Y-SNPs	♀	R-M343	♀
116 PT Y-SNPs	♀	R-M343	♀
16 ET X-SNPs	European and E.Asian specific X-chromosomes	African specific X-chromosome	African specific X-chromosomes

KS71 - reveal!

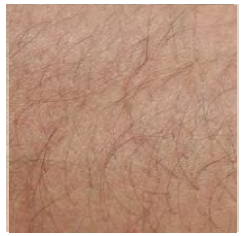


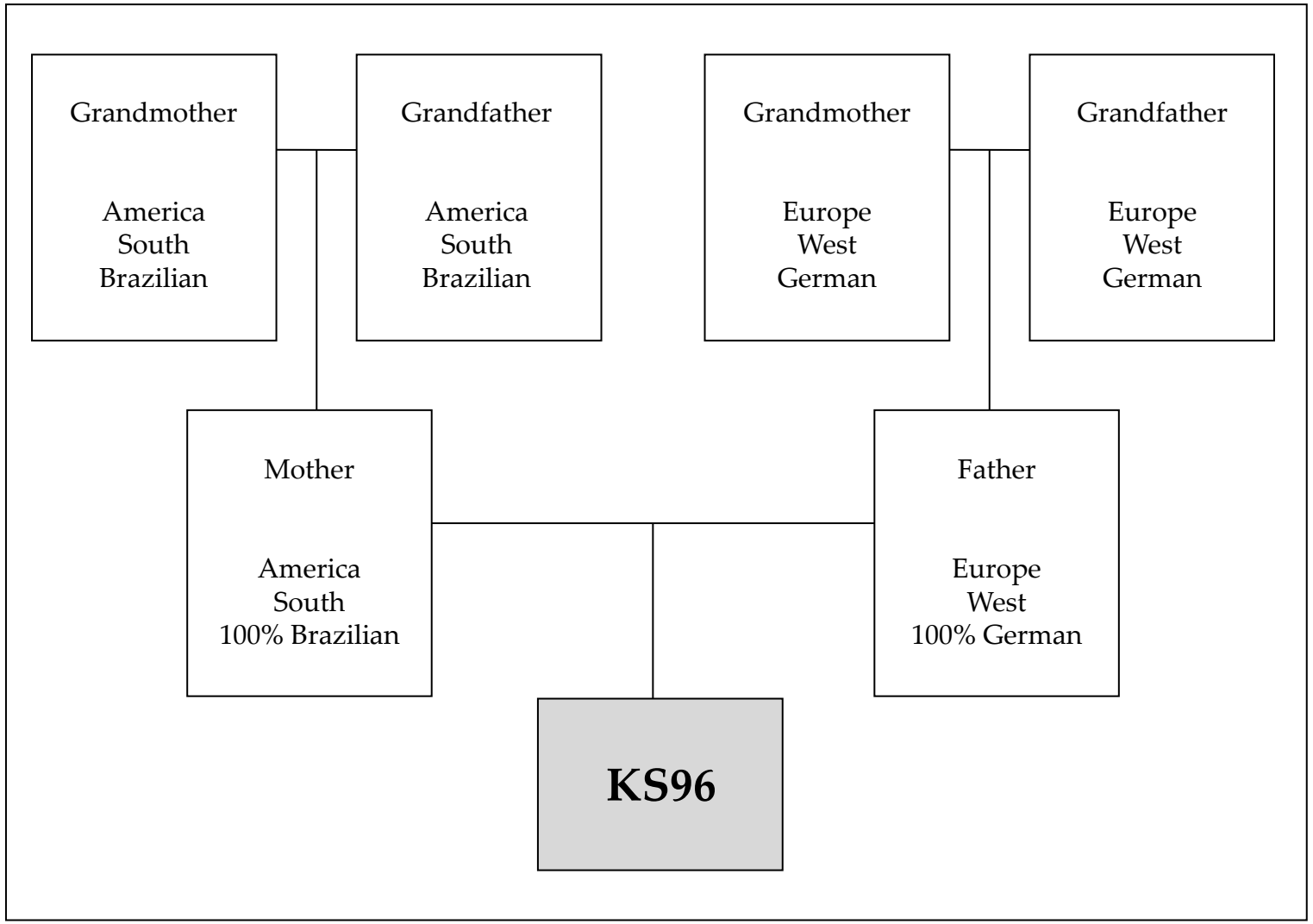




	PT	BT	ET
AFR	0.002	0.003	0.002
ME	0.089	0.014	0.054
EUR	0.405	0.438	0.344
SAS	0.026	0.021	0.042
EAS	0.447	0.479	0.463
OCE	0.029	0.025	0.012
AME	0.003	0.020	0.083

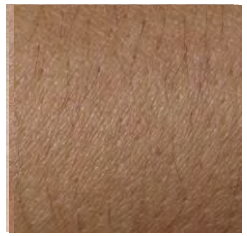
KS96 - reveal!

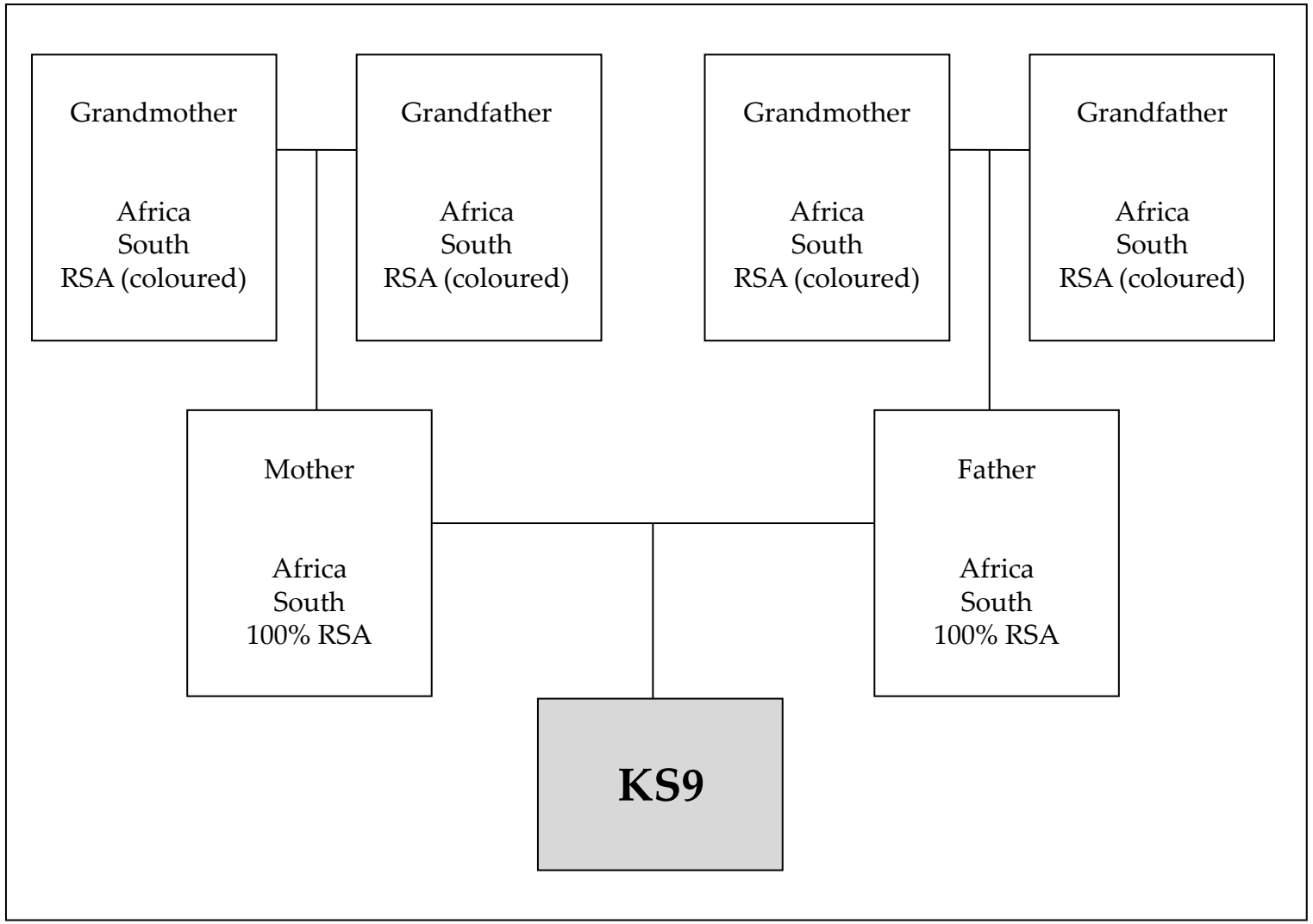




	PT	BT	ET
AFR	0.180	0.287	0.302
ME	0.249	0.170	0.048
EUR	0.308	0.531	0.633
SAS	0.247	0.004	0.011
EAS	0.006	0.002	0.002
OCE	0.007	0.003	0.002
AME	0.003	0.003	0.002

KS9 - reveal!





	PT	BT	ET
AFR	0.372	0.222	0.349
ME	0.475	0.600	0.326
EUR	0.026	0.090	0.111
SAS	0.049	0.019	0.016
EAS	0.047	0.029	0.192
OCE	0.018	0.011	0.003
AME	0.012	0.029	0.004

Your feedback

1. Did you find any of the data provided confusing and/or difficult to interpret?

PCA plots, especially for admixed

GenoGeographer – not familar with

Haploid markers – no clear reference data

Your feedback

2. Which individual's data did you find the most difficult to analyse?

KS9

KS126

KS96 / KS180

Your feedback

3. Which method of biparental SNP analysis did you find the most informative?

Structure

Converge / GenoGeographer

PCA plots

Your feedback

4. Which additional markers besides the biparental SNPs were most useful for interpreting the data?

Haploid markers

X-SNPs

Didn‘t use any

Your feedback

5. Did the provided phenotype predictions have an impact on your conclusions regarding the biogeographic origin of the studied individuals?

Yes, to confirm BGA

No

Didn‘t use

Your feedback

6. Did you refer to the HPS authors' guidelines or did you interpret the provided p-values using your own criteria?

Yes

Your feedback

7. To interpret the maternal and paternal lineages provided, what source of information did you use?

EMPOP

YHRD

Literature

Your feedback

8. Would you include any additional analyses to improve interpretation of the data?

No

New /different frequency databases

The 2024 EDNAP ancestry data interpretation exercise

- The forensic practice framework of the population analysis tests and reference data we applied in this exercise
- Join with Marta on discussion of the feedback sent to us from the questionnaire
- Group discussion about next steps

12 Labs completed data interpretations of 9 ‘real world’ samples

3 unadmixed : 3 unadmixed European or ‘peri-European’: 3 admixed*
*one parental co-ancestry, two complex co-ancestries

The main constraint was inability to share genotypes of donors

So, MPS SNP genotyping / mtDNA plus the applied population analysis algorithms were fixed and ‘universal’ reference data used

DNA Signature Kit AIM set and NFI ForAPP algorithm were excluded

The questionnaire replies helped understand which difficulties might arise with BGA interpretation when made with no prior data

Overall, findings could help guide BGA workshop scope, creating guidelines, and gauging error and interpretative pitfalls of algorithms

12 Labs completed data interpretations of 9 ‘real world’ samples

3 unadmixed : 3 unadmixed European or ‘peri-European’: 3 admixed*
*one parental co-ancestry, two complex co-ancestries

The main constraint was inability to share genotypes of donors

So, MPS SNP genotyping / mtDNA plus the applied population analysis algorithms were fixed and ‘universal’ reference data used

DNA Signature Kit AIM set and NFI ForAPP algorithm were excluded

The questionnaire replies helped understand which difficulties might arise with BGA interpretation when made with no prior data

Overall, findings could help guide BGA workshop scope, creating guidelines, and gauging error and interpretative pitfalls of algorithms

12 Labs completed data interpretations of 9 ‘real world’ samples

3 unadmixed : 3 unadmixed European or ‘peri-European’: 3 admixed*
*one parental co-ancestry, two complex co-ancestries

The main constraint was inability to share genotypes of donors

So, MPS SNP genotyping / mtDNA plus the applied population analysis algorithms were fixed and ‘universal’ reference data used

DNA Signature Kit AIM set and NFI ForAPP algorithm were excluded

The questionnaire replies helped understand which difficulties might arise with BGA interpretation when made with no prior data

Overall, findings could help guide BGA workshop scope, creating guidelines, and gauging error and interpretative pitfalls of algorithms

Evolution of current forensic BGA population analysis tests

	Snipper	*ForAPP*	*FROGkb*
Bayes LR tests	Fixed training sets (34-plex, Eurasiaplex, etc.)		Fixed training sets (34-plex, 56-plex, etc.)
	User-defined training sets		
PCA	Linked to LR tests		
STRUCTURE	Input file now generated in *Snipper* from user's data		
GenoGeographer Outlier test	In-house data + 1000 Genomes (allele frequencies)		
	Fixed training set for PT		
Converge Ancestry plug-in	ALFRED allele frequencies		
	Fixed training set for PT		

Evolution of current forensic BGA population analysis tests

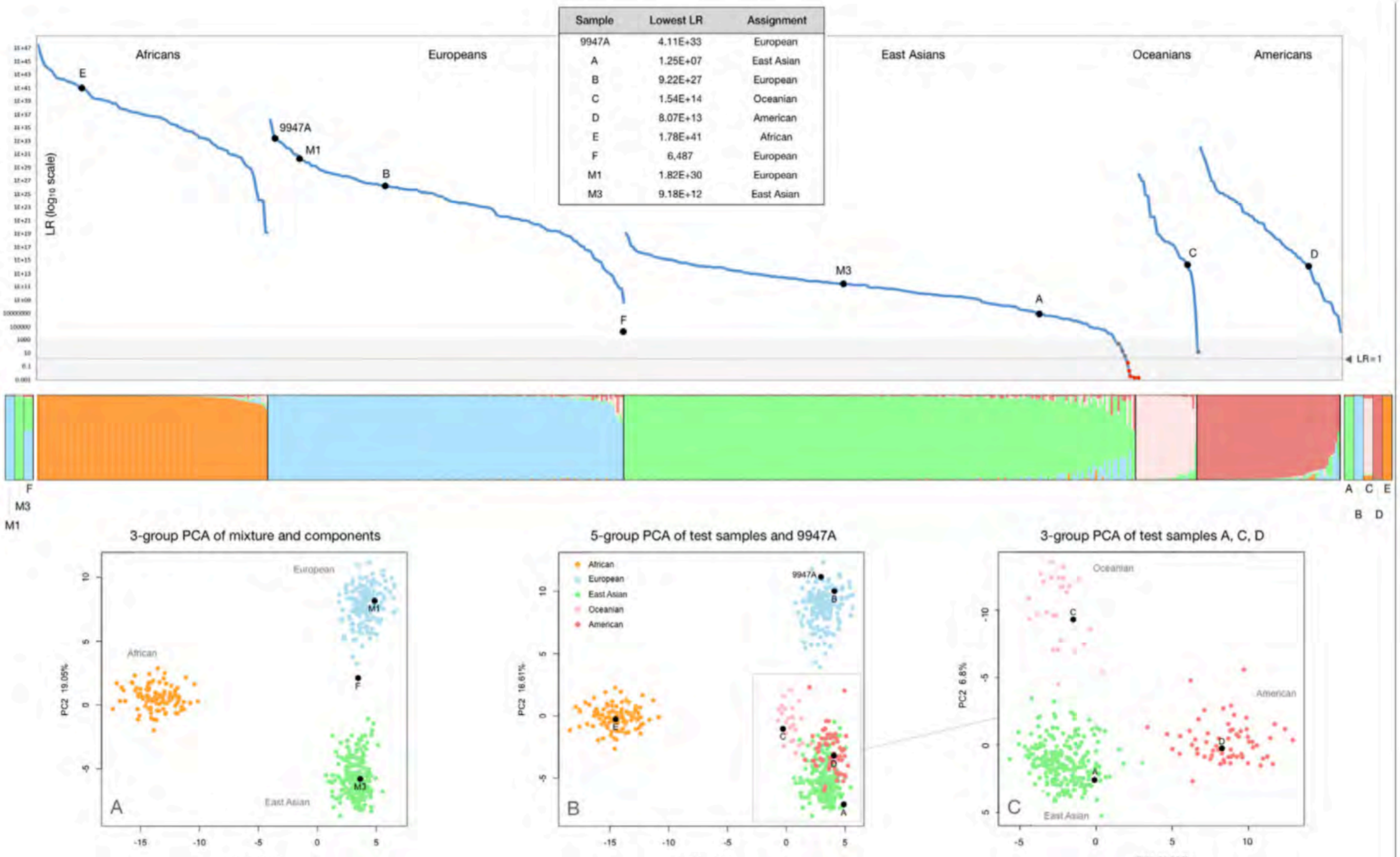
	Snipper	*ForAPP*	*FROGkb*	
Bayes LR tests	Fixed training sets (34-plex, Eurasiaplex, etc.)		Fixed training sets (34-plex, 56-plex, etc.)	CE
	User-defined training sets			
PCA	Linked to LR tests			
STRUCTURE	Input file now generated in *Snipper* from user's data			

GenoGeographer Outlier test	In-house data + 1000 Genomes (allele frequencies) Fixed training set for PT
Converge Ancestry plug-in	ALFRED allele frequencies Fixed training set for PT

Evolution of current forensic BGA population analysis tests

	Snipper	*ForAPP*	*FROGkb*
Bayes LR tests	Fixed training sets (34-plex, Eurasiaplex, etc.)		Fixed training sets (34-plex, 56-plex, etc.)
	User-defined training sets		
PCA			
STRUCTURE			

CE



Sample	Lowest LR	Assignment
9947A	4.11E+33	European
A	1.25E+07	East Asian
B	9.22E+27	European
C	1.54E+14	Oceanian
D	8.07E+13	American
E	1.78E+41	African
F	6,487	European
M1	1.82E+30	European
M3	9.18E+12	East Asian

Emergence of MPS has led to increasing use of STRUCTURE/GG

Snipper

Bayes LR tests	Fixed training sets (34-plex, Eurasiaplex, etc.) **User-defined training sets**	CE
PCA	Linked to LR tests Standard Reference Grid	
STRUCTURE	Input file now generated in *Snipper* from user's data	
GenoGeographer Outlier test	In-house data + 1000 Genomes (allele frequencies) Fixed training set for PT	MPS
Converge Ancestry plug-in	ALFRED allele frequencies Fixed training set for PT	

STRUCTURE cluster data has changed little

Snipper

Bayes LR tests

Fixed training sets
(34-plex, Eurasiaplex, etc.)

User-defined training sets

PCA

Linked to LR tests

Standard Reference Grid

STRUCTURE

Input file now generated in *Snipper* from user's data

GenoGeographer
Outlier test

In-house data + 1000 Genomes (allele frequencies)

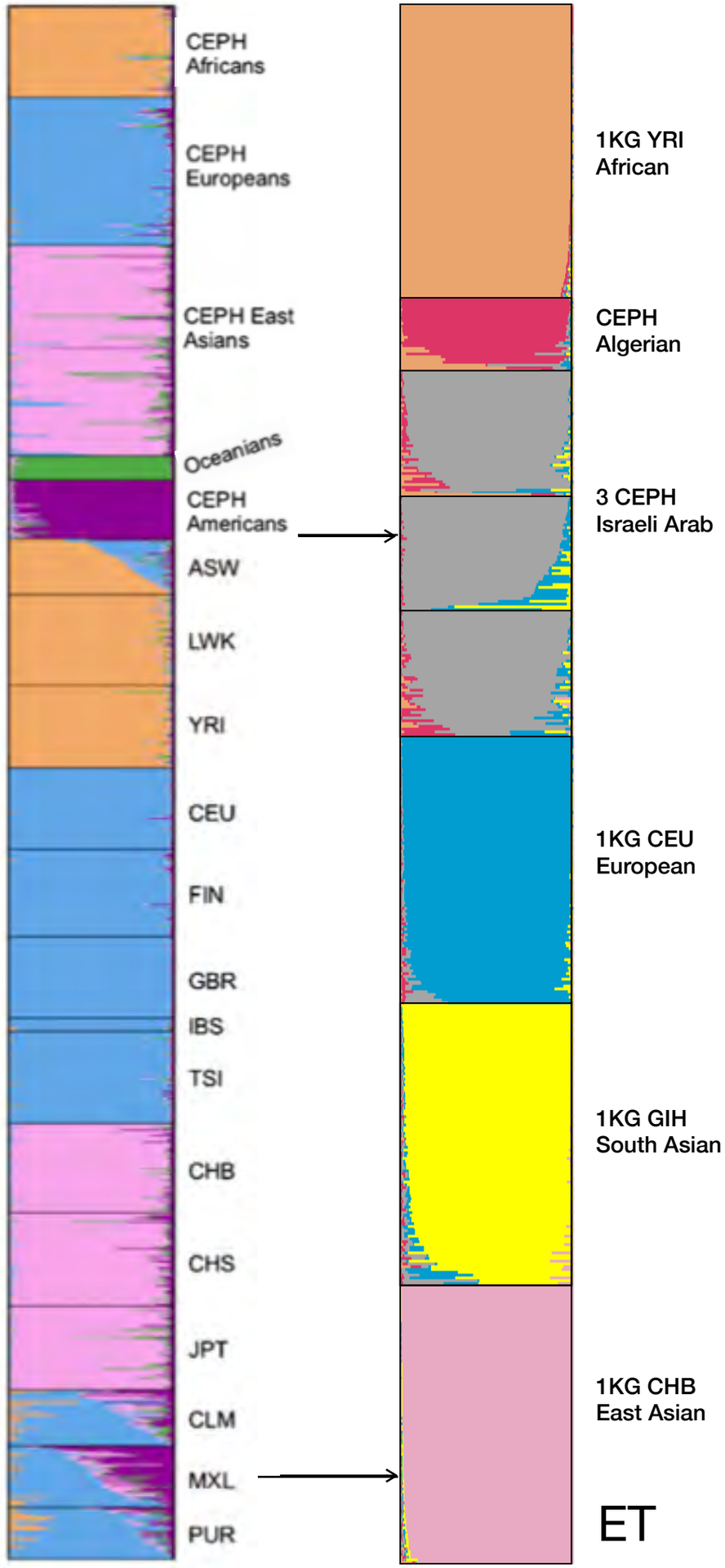
Fixed training set for PT

Converge
Ancestry plug-in

ALFRED allele frequencies

Fixed training set for PT

34plex



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	626	115	7	rs12142199	rs434504	rs776912	rs12130799	rs2139931	rs3737576	rs2814778	rs4657449	rs2227203	rs7531501	rs1834619	rs1796048
2			Chr:	1	1	1	1	1	1	1	1	1	1	2	2
3			Position GRCh37:	1249187	4815477	10847784	55663372	84590527	101709563	159174683	165465281	172879023	234338303	17901485	97643576
4			Position GRCh38:	1313807	4755417	10787727	55197699	84124844	101244007	159204893	165496044	172909883	234202557	17720218	96977839
5				1	2	3	4	5	6	7	8	9	10	11	12
6	1	African	NA18486	GG	AA	CC	AA	AA	TT	CC	GG	CC	AA	GG	TT
7	2	African	NA18488	AG	AA	CC	AA	AA	TT	CC	AG	CC	AA	GG	TT

Snipper Standard Reference Grid | 1KG Unadmixed Populations | 1KG Admixed Populations | HGDP-CEPH Populations | SGDP Test Samples | EGDP Test Samples

It is possible to download the Reference Population Set of choice: the VISAGE Basic Tool ancestry panel grid; Verogen ForenSeq DNA Signature Kit 56-AIMs grid; FORCE AIMs grid; and TFS Precision ID Ancestry Panel grid. Then paste in your profiles as extra rows ending in 0 (final column), and modify the total number of individuals in first leftmost cell accordingly.

PCA — Linked to LR tests

Standard Reference Grid

STRUCTURE — Input file now generated in *Snipper* from user's data

GenoGeographer Outlier test — In-house data + 1000 Genomes (allele frequencies)

Fixed training set for PT

Converge Ancestry plug-in — ALFRED allele frequencies

Fixed training set for PT

Metapopulations

Copy | CSV | Excel | PDF | Print

metapopulation
Europe
South / Central Asia
East Asia
Middle East
East Greenland
Horn of Africa
Sub-Saharan Africa
North Africa
West Greenland

Genogeographer

population	n
Albania	94
Bengali, Bangladesh	86
Chinese Dai in Xishuangbanna, China	93
Utah Residents (CEPH), N&W Europe ancestry	99
Han Chinese in Bejing, China	103
Danes, RGA	140
Druze, Israel	102
European Americans	89
Greenlanders East	112
Eritrea	88
Esan, Nigeria	99
Finns, Finland	99
British from England and Scotland	91
Gujarati (India), Houston, Texas	103
Greece, RGA	79
Gambians, Western Division,The Gambia	113
Southern Han Chinese	105
Hungarians	89
Iberians, Spain	107
Irish	113
Iranian, RGA	93
Iraqi, RGA	94
Telugu (India), collected in United Kingdom	102
Japanese, Tokyo, Japan	104
Kinh (Vietnamese), Ho Chi Minh City, Vietnam	99
Laotians	118
Luhya, Webuye, Kenya	99
Morroco, RGA	75
Mende, Sierra Leone	85
Nigeria	87
Punjabi, Lahore, Pakistan	96
Russians	80
Slovenia	96
Somali, RGA	75
Tamil (Sri Lanka), collected in United Kingdom	102
Toscani, Italy	107
Tunisia	208
Turkish, RGA	88
Greenlanders West	204
Yoruba, Benin City, Nigeria	77
Yoruba, Ibadan, Nigeria	108

Applied a universal reference set with balanced population numbers to reduce bias in STRUCTURE cluster analysis

Bayes LR tests — Provided for 16 X-SNPs

PCA — Linked to LR tests

Standard Reference Grid

STRUCTURE — Input file now generated in *Snipper* from user's data

Standard Reference Grid

GenoGeographer Outlier test — In-house data + 1000 Genomes (allele frequencies)

Standard Reference Grid

Converge Ancestry plug-in — ALFRED allele frequencies

Fixed training set for PT

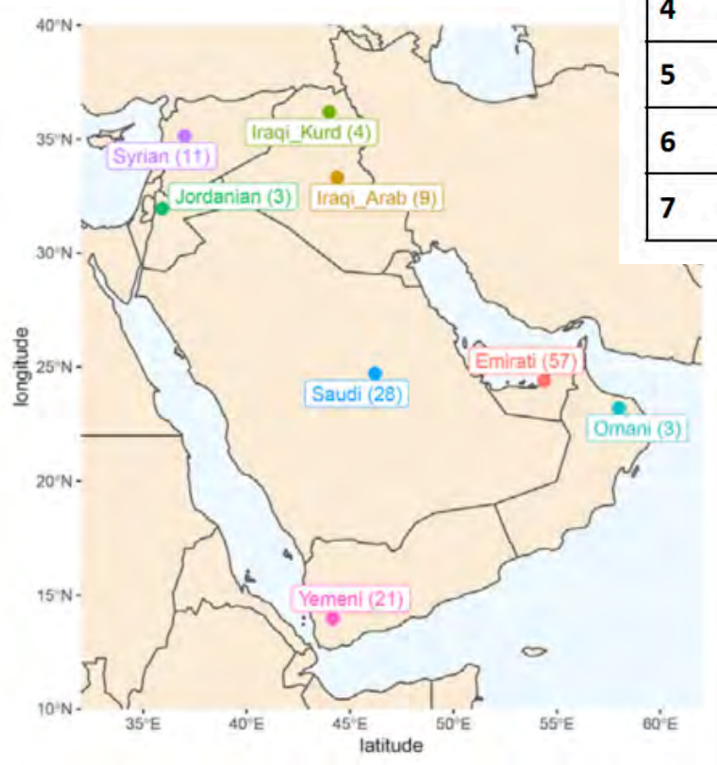
1000 Genomes genotype sets

YRI	108 Yoruba from Nigeria
CEU	99 NW Europeans from Utah
GIH	103 Gujarati from Houston
CHB	103 Han from Beijing

HGDP-CEPH genotype sets

OCE	28 Oceanians
AMR	61 Native Americans + 18 PEL

1	14 Emirati A
2	11 Emirati B
3	14 Emirati C
4	18 Emirati D
5	12 Saudi A
6	16 Saudi B
7	21 Yemeni



106 Almarri et al. Middle East genotype sets

Feedback on most informative data analysis regimes and which data was confusing/difficult to interpret

Bayes LR tests — Provided for 16 X-SNPs

PCA — Linked to LR tests

Standard Reference Grid

STRUCTURE — Input file now generated in *Snipper* from user's data

Standard Reference Grid

GenoGeographer
Outlier test — In-house data + 1000 Genomes (allele frequencies)

Standard Reference Grid

Converge
Ancestry plug-in — ALFRED allele frequencies

Fixed training set for PT

Which method of bi-parental SNP data analysis did you find the most informative?

PCA	3
STRUCTURE	6
GenoGeographer	3
Converge	2

Did you find any of the data provided confusing and/or difficult to interpret?

STRUCTURE rather hard to interpret

No information on STRUCTURE limitations

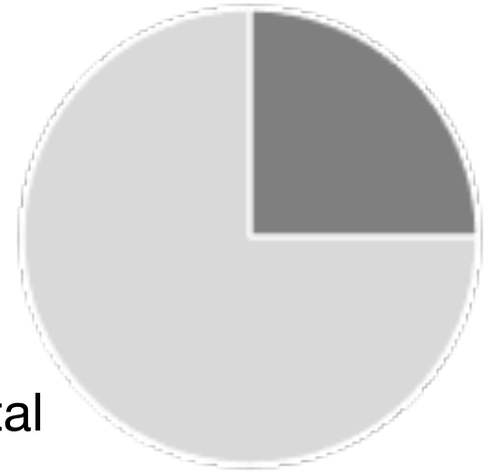
Lack of clarity with STRUCTURE reference data and guidelines

GenoGeographer often had contradictory z-score and likelihoods GenoGeographer

All tests lacked easy operational guidelines

Participants use of uniparental data

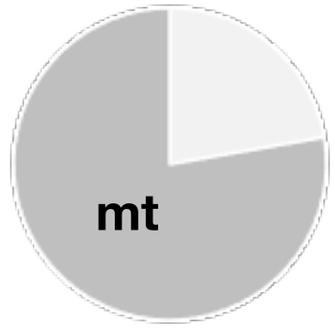
Which additional markers besides the bi-parental SNPs were most useful for interpreting data?



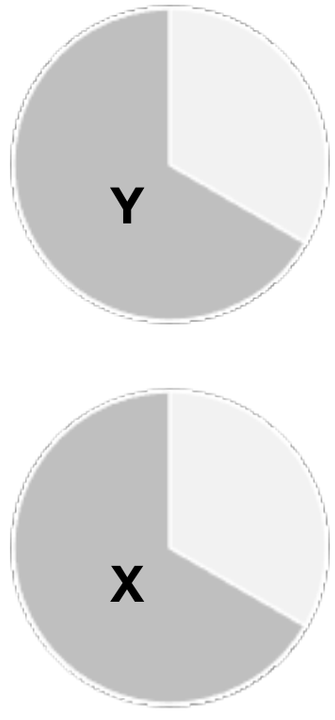
9 Labs used uniparental data

3 Labs did not use uniparental data

7 of 9 Labs



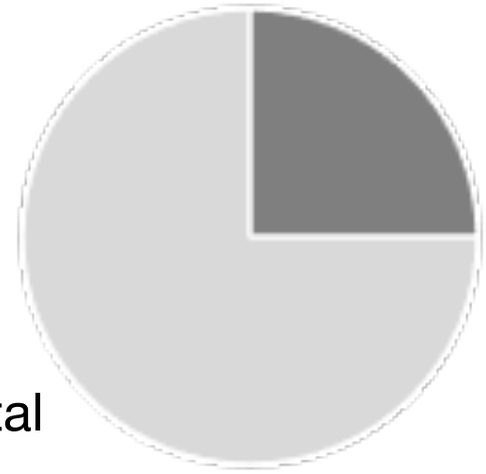
6 of 9 Labs



6 of 9 Labs

Participants use of uniparental data

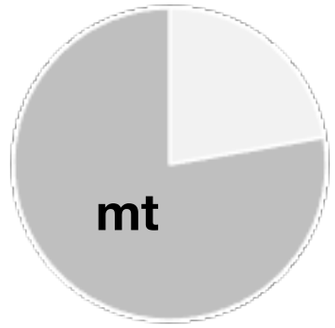
Which additional markers besides the bi-parental SNPs were most useful for interpreting data?



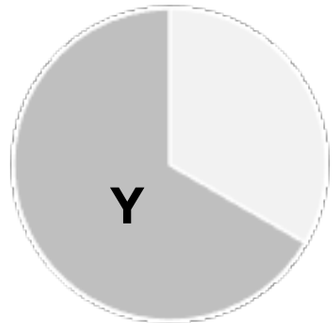
9 Labs used uniparental data

3 Labs did not use uniparental data

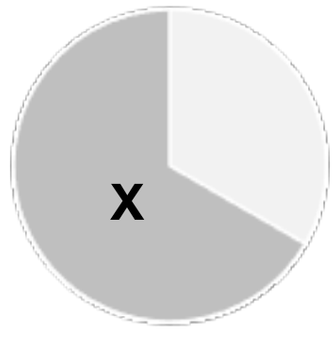
7 of 9 Labs



6 of 9 Labs



6 of 9 Labs



3 of the 9 samples = ♀

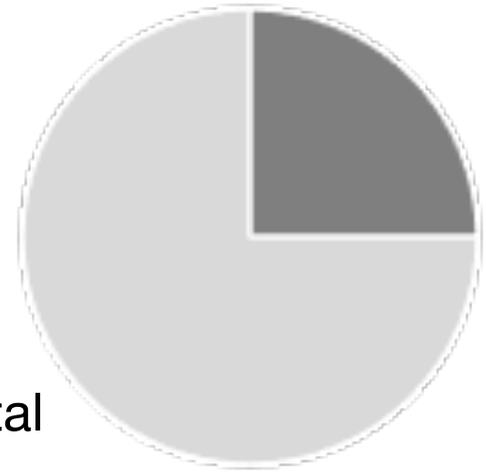
KS9 AGTTT African-specific / AGTCT Non-specific

KS71 AATCT East Asian-specific / AATTC European-specific
AATCC European-specific / AATTT Non-specific

KS110 AATCC European-specific / CATTC European-specific
AATTC European-specific / CATCC Non-specific

Participants use of uniparental data

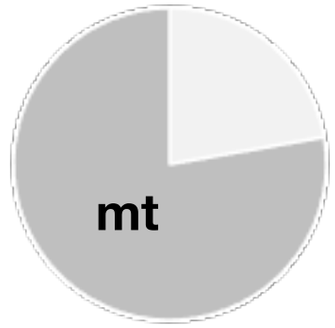
Which additional markers besides the bi-parental SNPs were most useful for interpreting data?



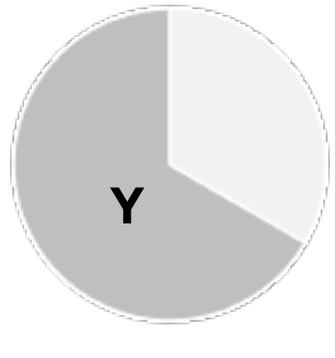
9 Labs used uniparental data

3 Labs did not use uniparental data

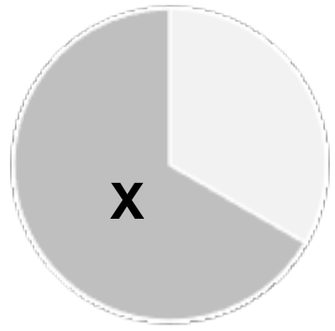
7 of 9 Labs



6 of 9 Labs



6 of 9 Labs



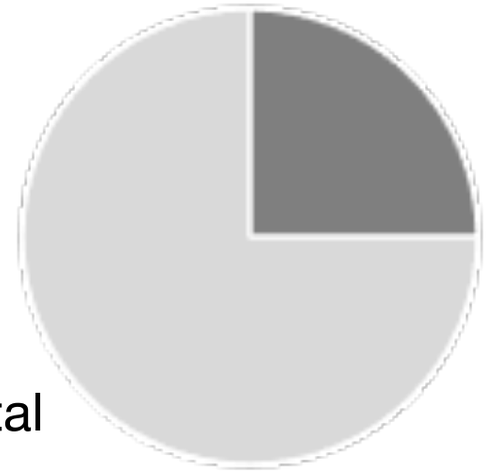
To interpret the maternal and paternal lineages provided, what source of reference information did you use?

Literature	3
EMPOP	8
YHRD	4
Online resources	5
one lab used X-SNPs to co-analyse male Y patterns	
one lab used the VISAGE Y-SNP mapping software	

Ian Logan (mt)
Eupedia
ISOGG Y-Tree

Participants use of uniparental data - Walther comments on mito

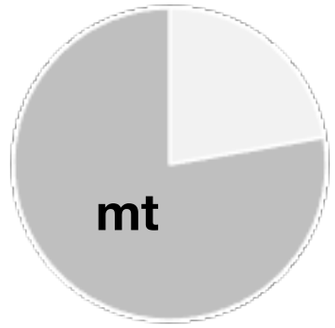
Which additional markers besides the bi-parental SNPs were most useful for interpreting data?



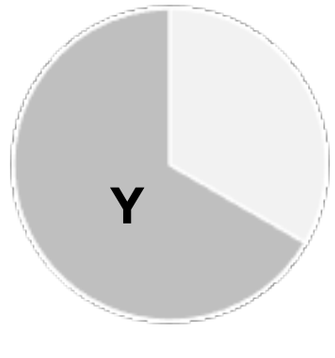
9 Labs used uniparental data

3 Labs did not use uniparental data

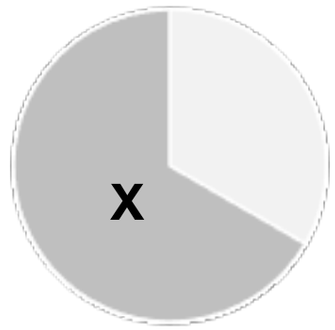
7 of 9 Labs



6 of 9 Labs



6 of 9 Labs



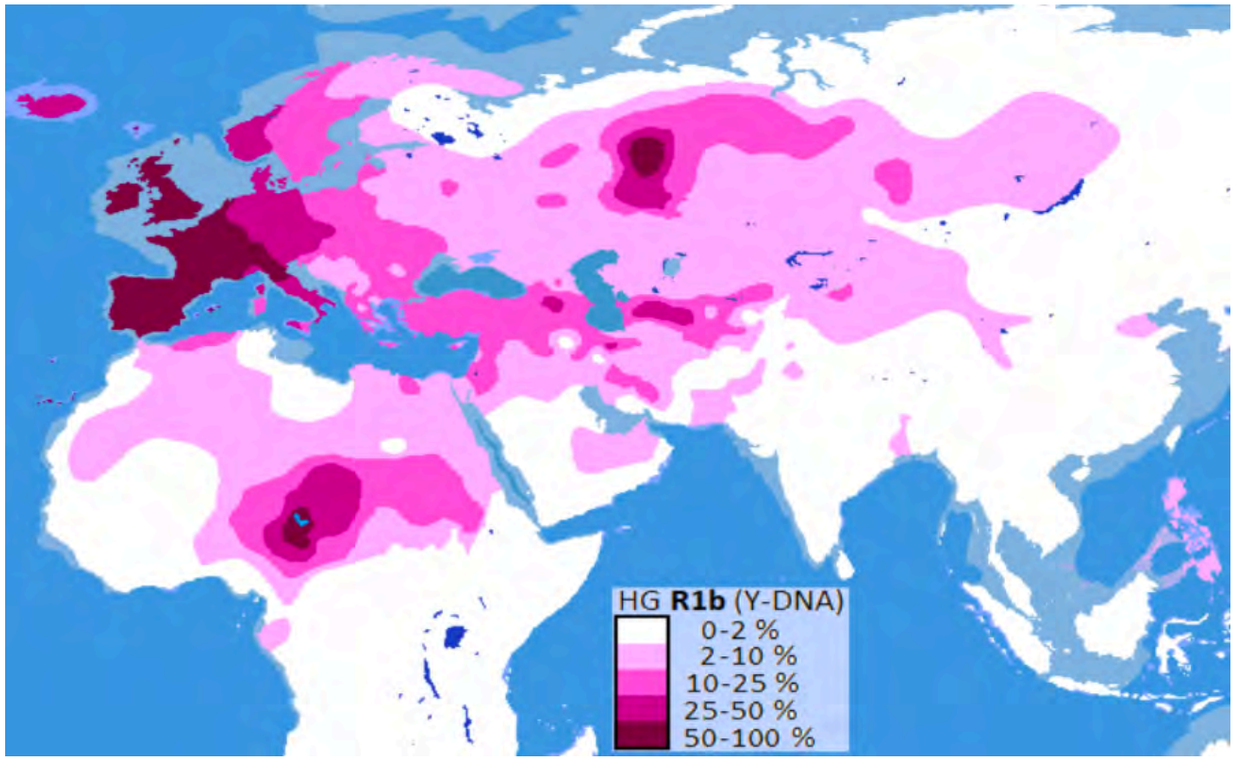
To interpret the maternal and paternal lineages provided, what source of reference information did you use?

Literature	3
EMPOP	8
YHRD	4
Online resources	5
one lab used X-SNPs to co-analyse male Y patterns	
one lab used the VISAGE Y-SNP mapping software	

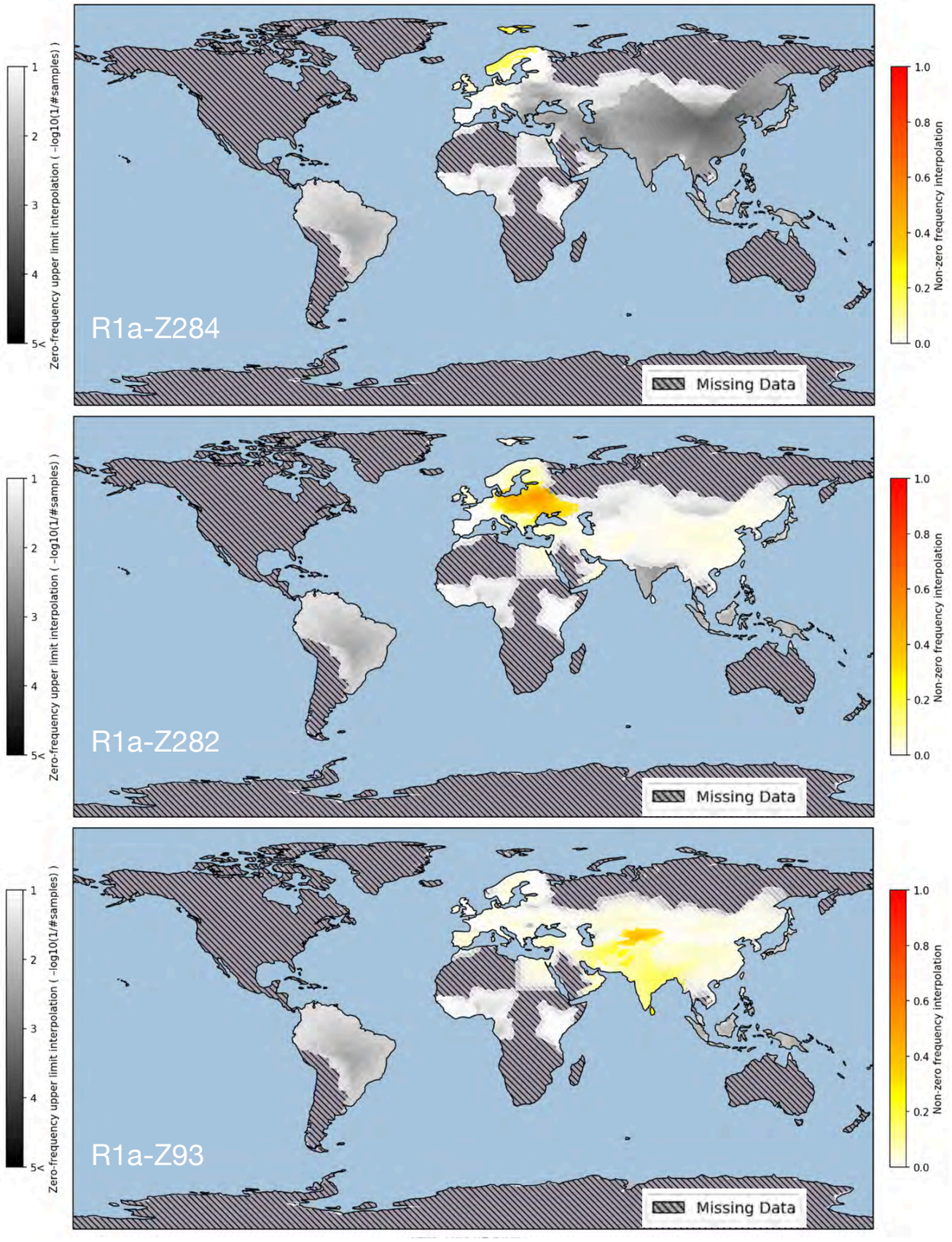
Ian Logan (mt)
Eupedia
ISOGG Y-Tree

Y-SNP data

ISOGG



VISAGE Maps

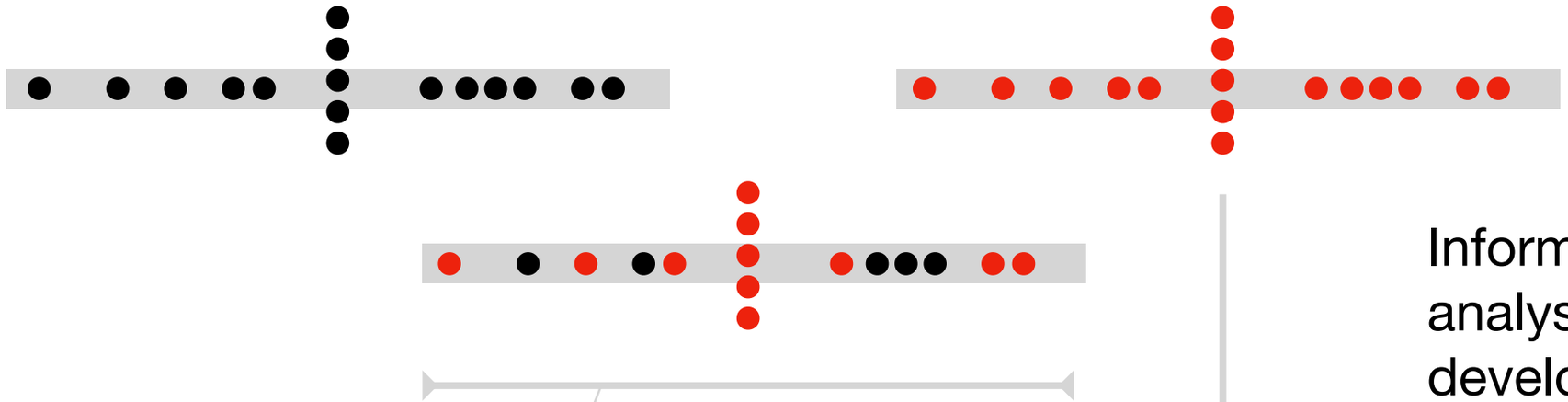


To interpret the maternal and paternal lineages provided, what source of reference information did you use?

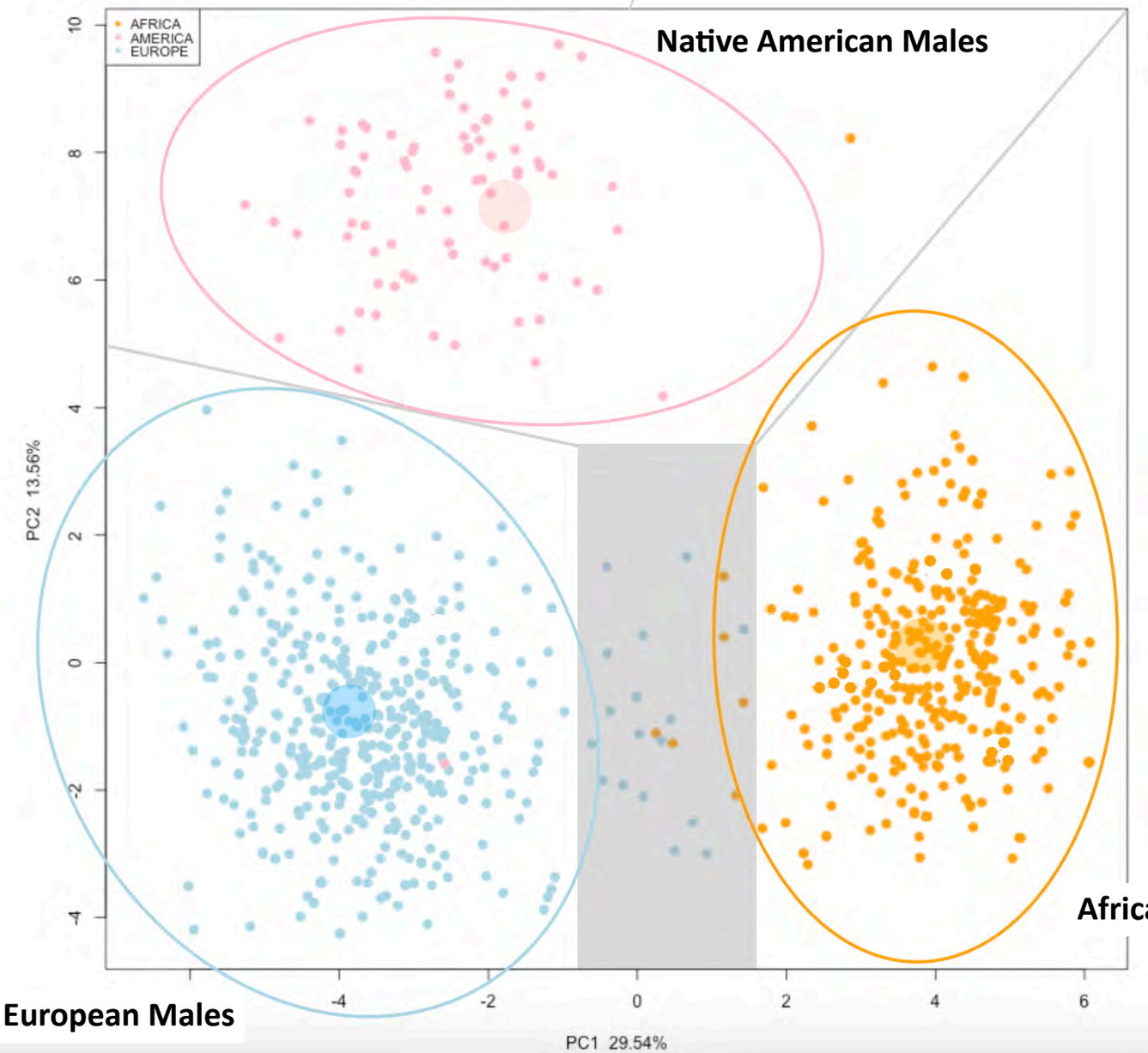
Literature	3
EMPOP	8
YHRD	4
Online resources	5
one lab used X-SNPs to co-analyse male Y patterns	
one lab used the VISAGE Y-SNP mapping software	

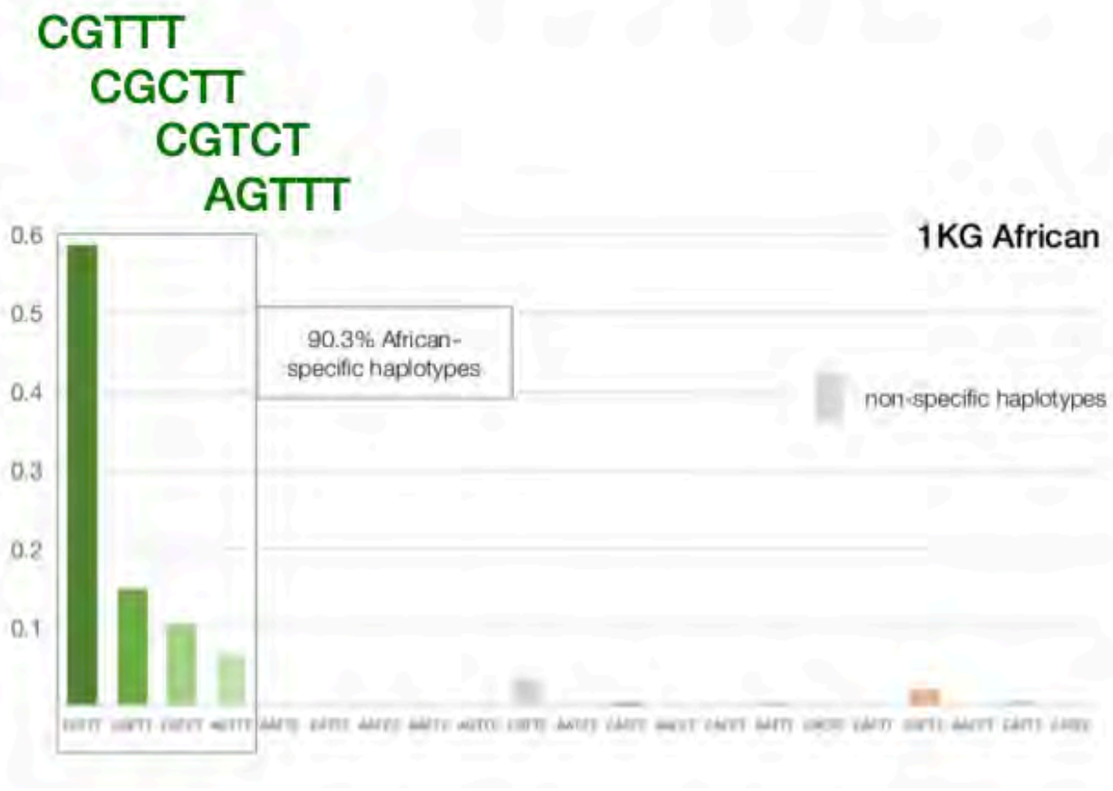
Ian Logan (mt)
Eupedia (Y German)
ISOGG Y-Tree

X-SNP data



Informative in admixed males - analysis systems are being developed, so we decided to provide interpretations direct





KS9 AGTTT African-specific / AGTCT Non-specific

STRUCTURE Guidance and avoiding risk of over-interpretation

genes

MDPI

Review

Review of the Forensic Applicability of Biostatistical Methods for Inferring Ancestry from Autosomal Genetic Markers

Torben Tvedebrink [1,2]

[1] Department of Mathematical Sciences, Aalborg University, DK-9220 Aalborg, Denmark; tvede@math.aau.dk
[2] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, DK-1165 Copenhagen, Denmark

3. Discussion

STRUCTURE remains to be a valuable methodology and approach for analysing population structure in forensic genetics. There exist several guides and tutorials for how to prepare the population data and choosing parameter settings for STRUCTURE (e.g., [60,61], with some emphasis on forensic applications). However, as warned by [44], the resulting barplots should not be overinterpreted. The badMIXTURE approach [45] provides the means to reduce the risk of being mislead by the summaries provided by STRUCTURE and supporting software (e.g., [7,48]). CLUMP

However, STRUCTURE's popularity in the field of population genetics to study population structure is well deserved. It complemented PCA with a quantitative method that assigns sample specific weights to each of the K populations specified in the study. This is a powerful way to summarise the data in terms of cluster membership probabilities and the estimated allele frequencies for the identified populations. Both PCA and STRUCTURE are valuable for exploratory analysis, where encoding errors (e.g., of missing data) and warnings about misspecification of origin may be detected by visual inspections of the results. From a forensic point of view, the results from both PCA and STRUCTURE are hard to report in terms of a weight of evidence calculation. The similarity (or dissimilarity) between the sample and reference materials can be reported but typically only in terms of their visual appearance.

frontiers in GENETICS

REVIEW ARTICLE
published: 29 May 2013
doi: 10.3389/fgene.2013.00098

60

An overview of *STRUCTURE*: applications, parameter settings, and supporting software

Liliana Porras-Hurtado[1,2†], Yarimar Ruiz[2†], Carla Santos[2†], Christopher Phillips[2], Ángel Carracedo[2,3] and Maria V. Lareu[2]*

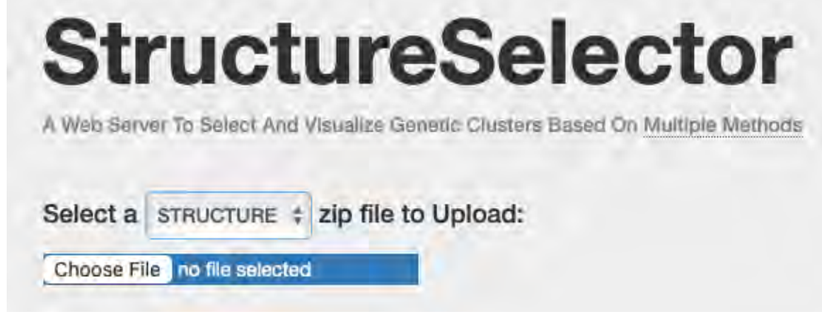
61

Inference of Ancestry in Forensic Analysis II: Analysis of Genetic Data

C. Santos, C. Phillips, A. Gomez-Tato, J. Alvarez-Dios, Á. Carracedo, and M.V. Lareu

doi:10.1007/978-1-4939-3597-0_19

Best current CLUMP web tool



PCA lacked quantitative metrics

PCA charts with admixed samples

STRUCTURE rather hard to interpret

No information on STRUCTURE limitations

Lack of clarity with STRUCTURE reference data and guidelines

GenoGeographer often had contradictory z-score and likelihoods

GenoGeographer

All tests lacked easy operational guidelines

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-05257-7 OPEN

44

A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots

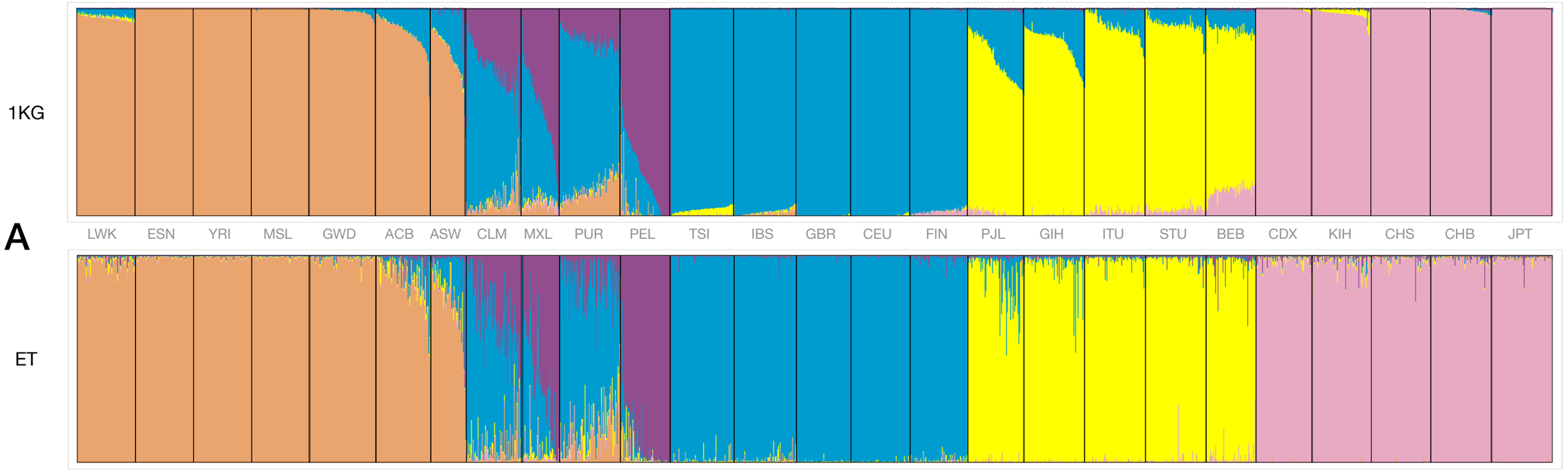
Daniel J. Lawson[1], Lucy van Dorp[2,3] & Daniel Falush[4]

45

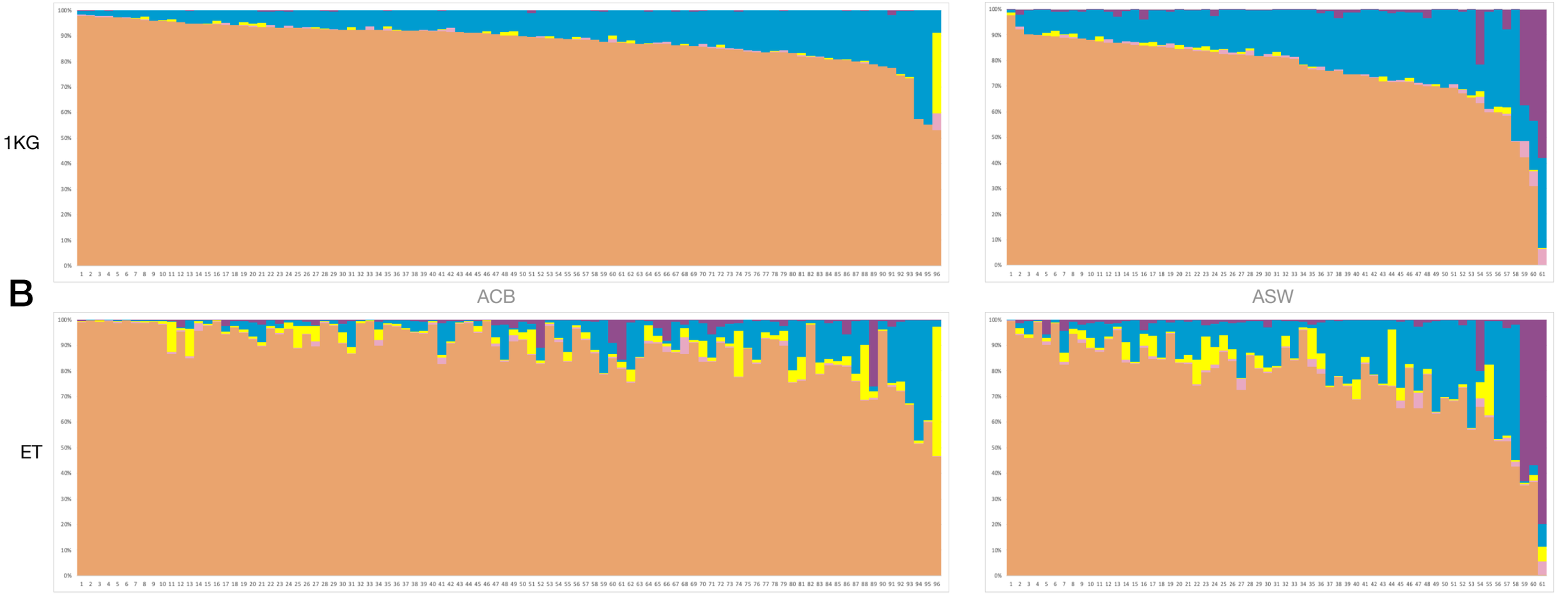
Lawson, D. badMIXTURE: Validating Structure With Chromosome Painting; R Package Version 0.0.0.9000; 2018. Available online: https://github.com/danjlawson/badMIXTURE

Previously compared BT and ET with 1000 Genomes WGS data

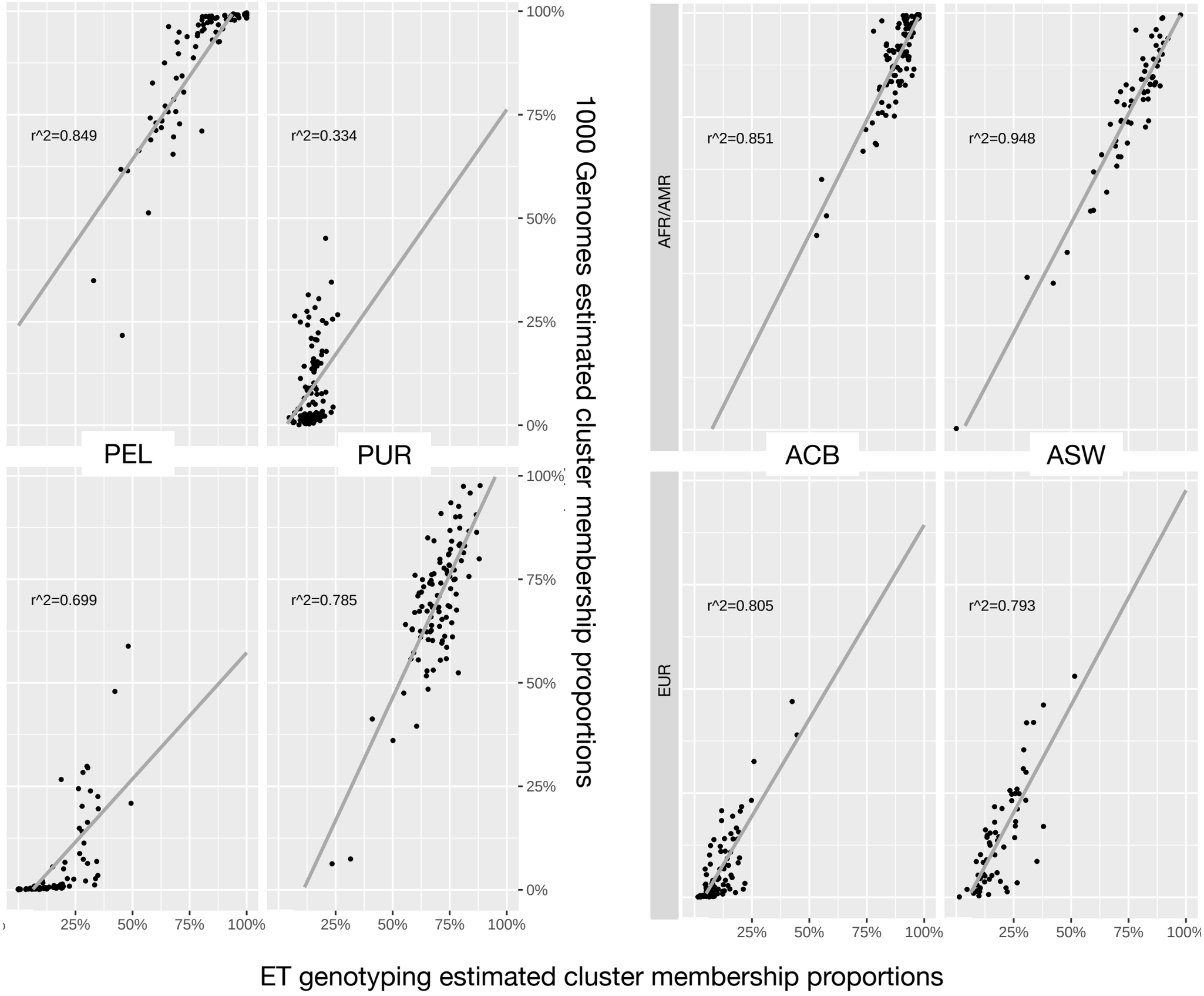
A



B

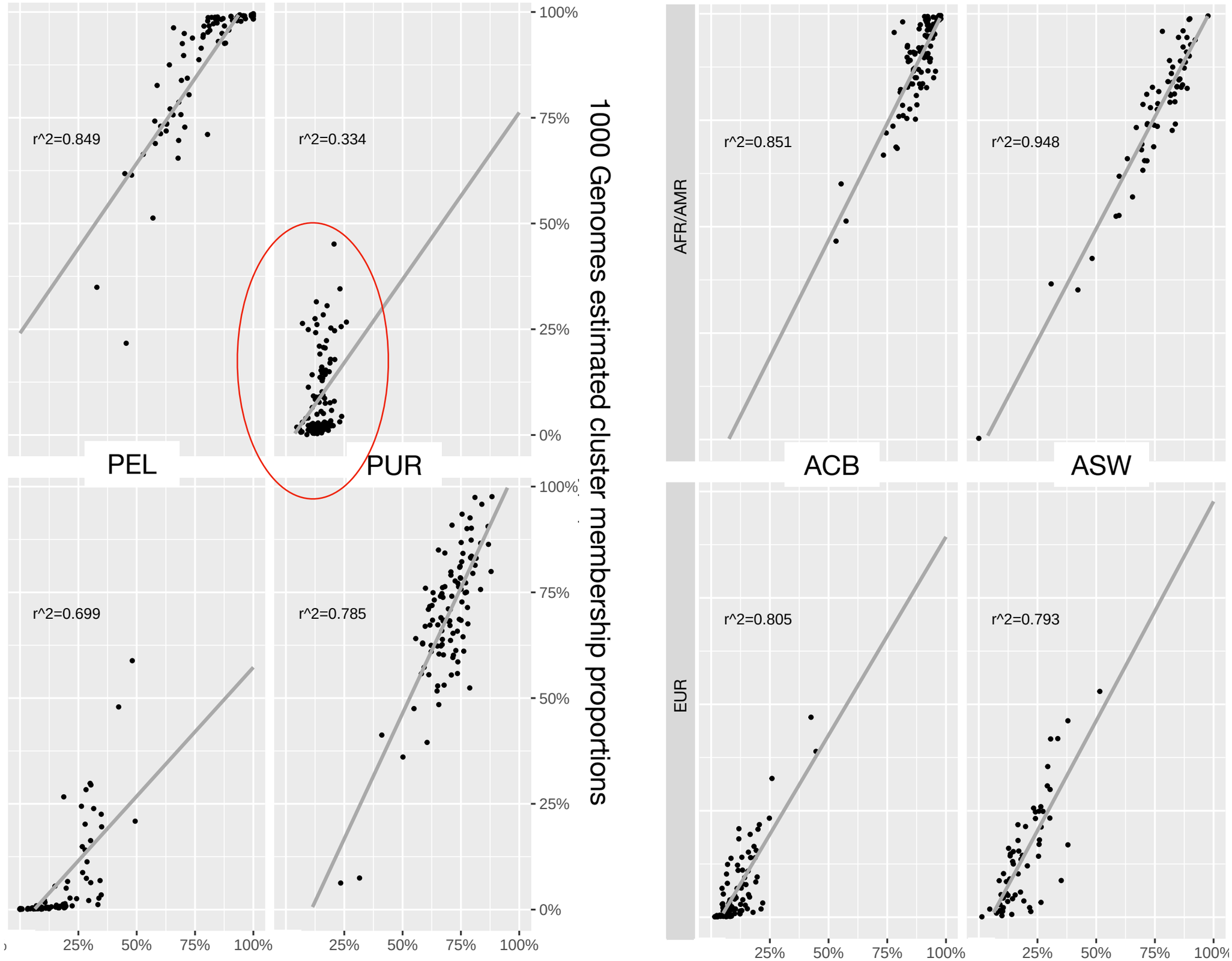


Previously compared BT and ET with 1000 Genomes WGS data



Previously compared BT and ET with 1000 Genomes WGS data

STRUCTURE analysis is a lot less informative when there is three-way admixture with two minor co-ancestries and one dominant co-ancestry - shown by Puerto Rican data here



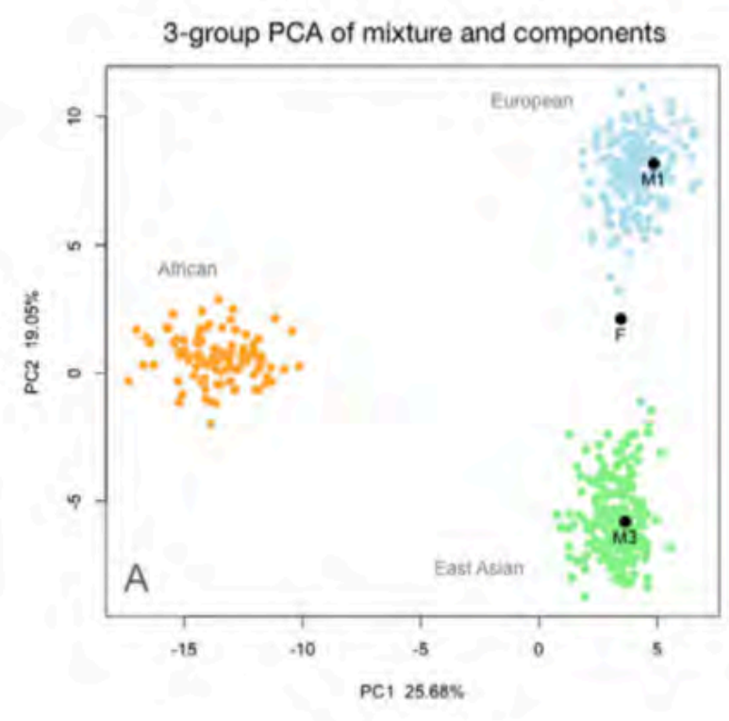
Autosomal SNP data - caveats

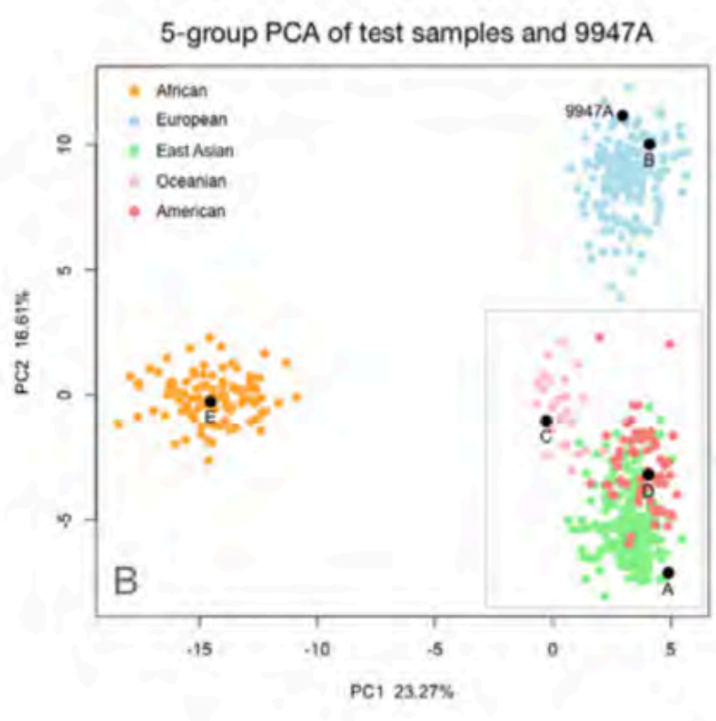
Likelihood ratio analysis does not work with large-scale datasets so was not provided outside of X-SNP analyses. VISAGE software used logistic regression with a fixed model so cannot adjust for missing data

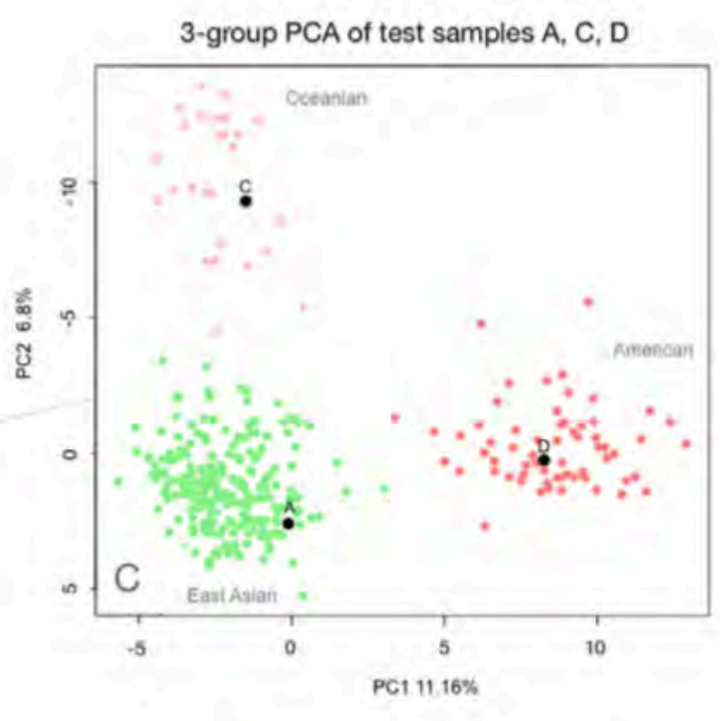
Autosomal SNP data - caveats

Likelihood ratio analysis does not work with large-scale datasets so was not provided outside of X-SNP analyses. VISAGE software used logistic regression with a fixed model so cannot adjust for missing data

PCA analyses a portion of the data, so less well differentiated populations do not contribute as much to distributions of clusters: nested PCA works well







Autosomal SNP data - caveats

Likelihood ratio analysis does not work with large-scale datasets so was not provided outside of X-SNP analyses. VISAGE software used logistic regression with a fixed model so cannot adjust for missing data

PCA analyses a portion of the data, so less well differentiated populations do not contribute as much to distributions of clusters: nested PCA works well

STRUCTURE results are always summaries of multiple runs - cluster membership proportions are averaged by CLUMP

STRUCTURE is challenged to summarise individuals with complex co-ancestry patterns and when poorly differentiated populations are analysed at the same time.

Individuals with three-way admixture can have cluster proportions too small to be reliably detected - we apply a 15% cut-off to exclude ‘noise’

Recommended to ‘calibrate’ STRUCTURE with GG and vice versa

Concluding remarks / discussion

There was no error or lack of precision in the reports provided, just differences in detail. Difficulties with interpretation of STRUCTURE and GG

We intend to write a paper with a summary of results, but more importantly sources of confusion, ambiguity and unclear statistical data from these tests

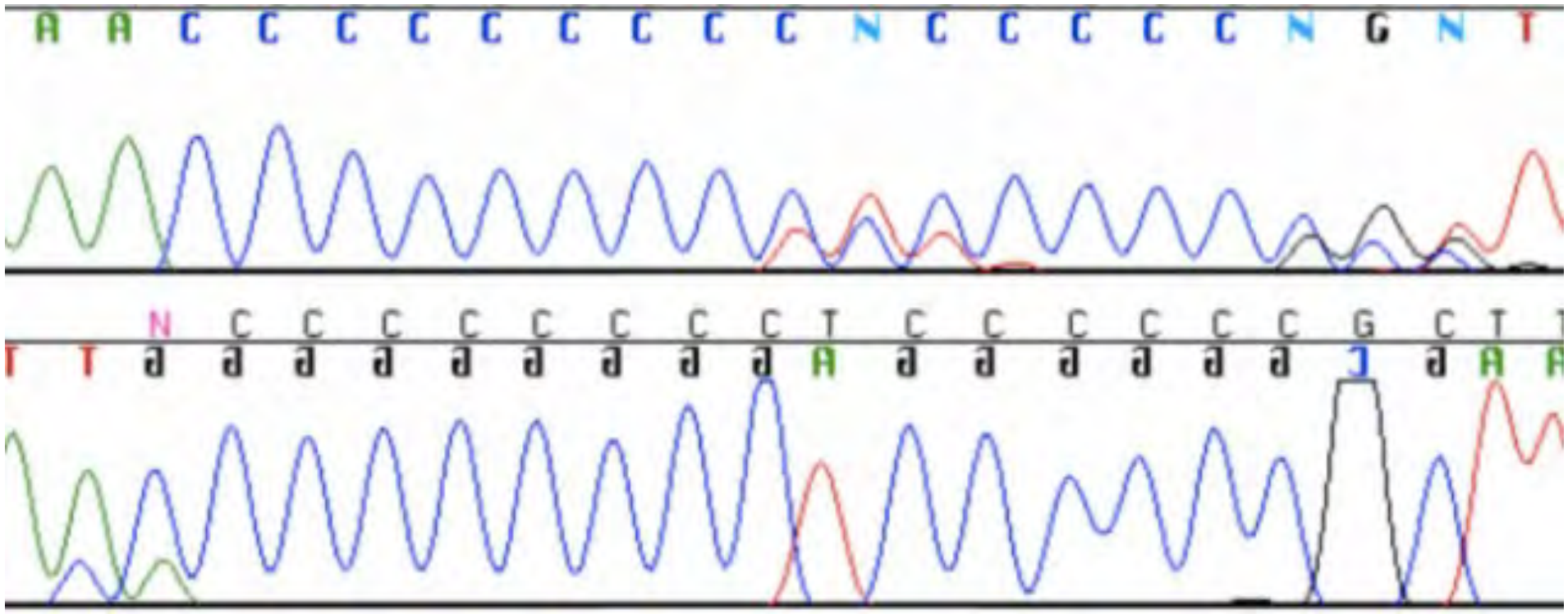
Do we have consensus about this step and agreement to do it?

Concluding remarks / discussion

We propose a follow-up paper exercise with full genotype data with free-form analyses made by participants

Would this be a welcome suggestion for the community interested or engaged in forensic BGA-FDP analysis?

EDNAP Meeting, Copenhagen, Denmark, May 29 2024



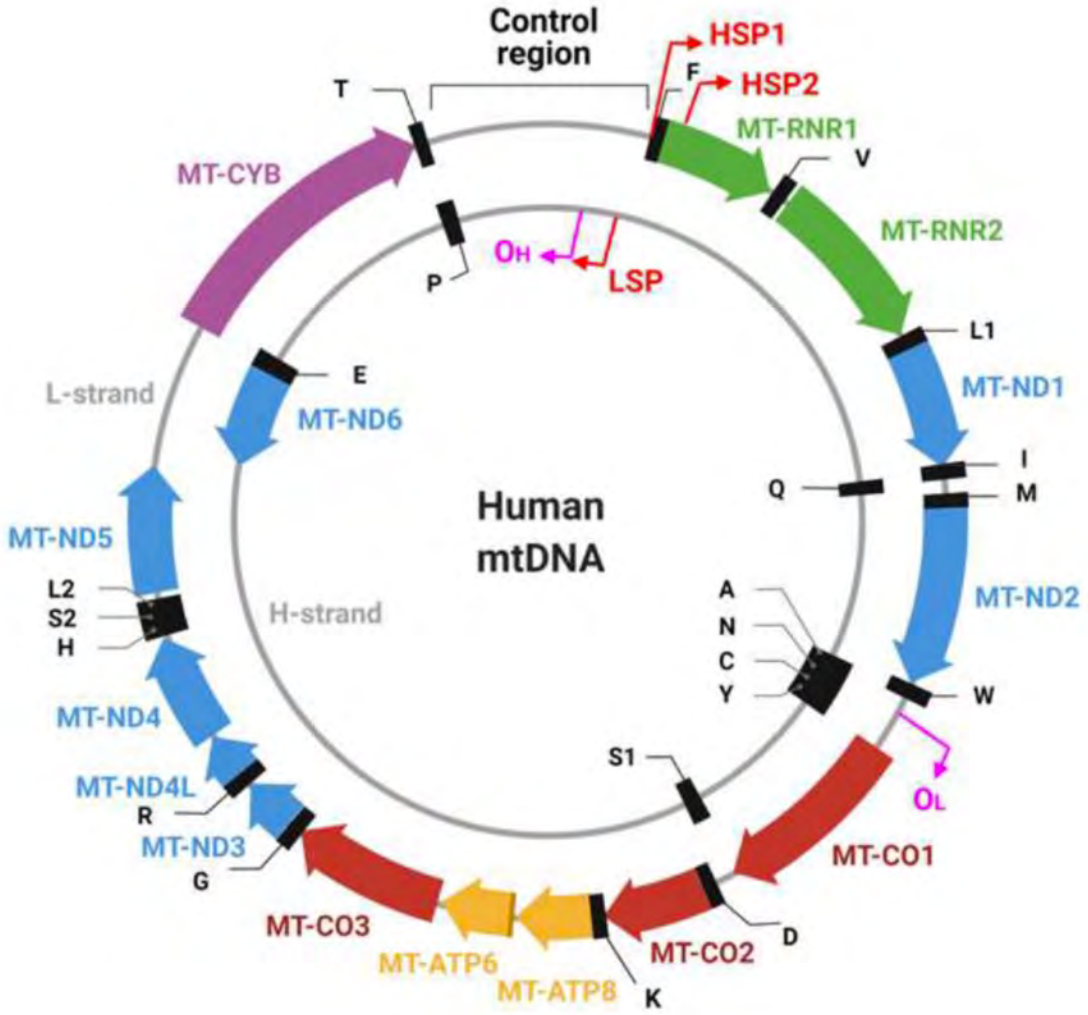
mtDNA heteroplasmy exercise

Walther Parson[1,2], Lena Ewers[1], Gabriela Huber[1], Nicole Huber[1], Claudia Wöss[1], Arne Dür[3]
[1] Institute of Legal Medicine, Medical University of Innsbruck, Austria
[2] Forensic Science Program, Penn State University, PA, USA
[3] Institute of Mathematics, University of Innsbruck, Austria

GMI DNA EMPOP

Mitochondrial DNA (mtDNA)

circular double-stranded molecule
16.5 kbp in size
compact and reduced
coding region (15 kb)
  13 OSPHOX proteins
  22 tRNAs
    2 rRNAs
control region (1.1 kb) includes d-loop
  non-coding, regulatory functions
evolutionary rate ~10x of nDNA



Shokolenko and Alexeyev 2022

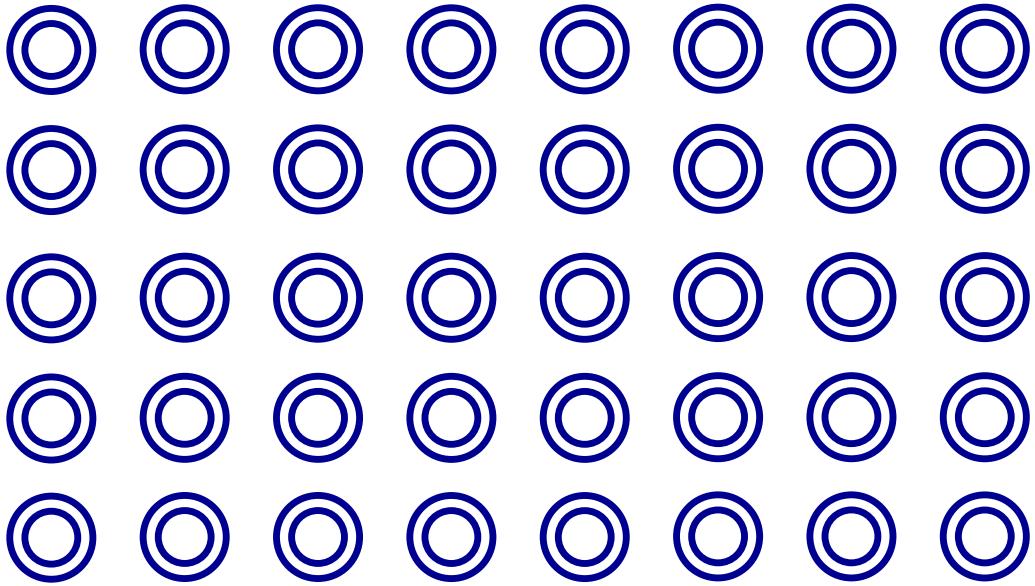
Mitochondrial DNA size and copy number

nuclear DNA (nDNA)

46 chromosomes, 3.4 x 10^9 bp

diploid

mitochondrial DNA (mtDNA)



100(0)s per cell, 16,568 bp*

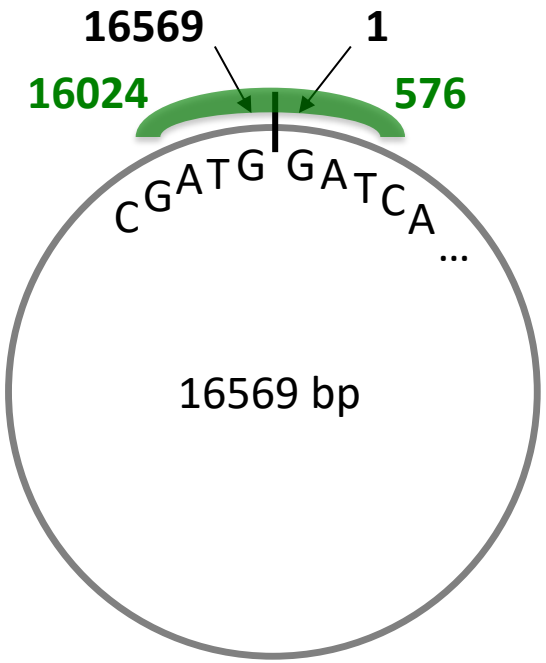
haploid

EMPOP

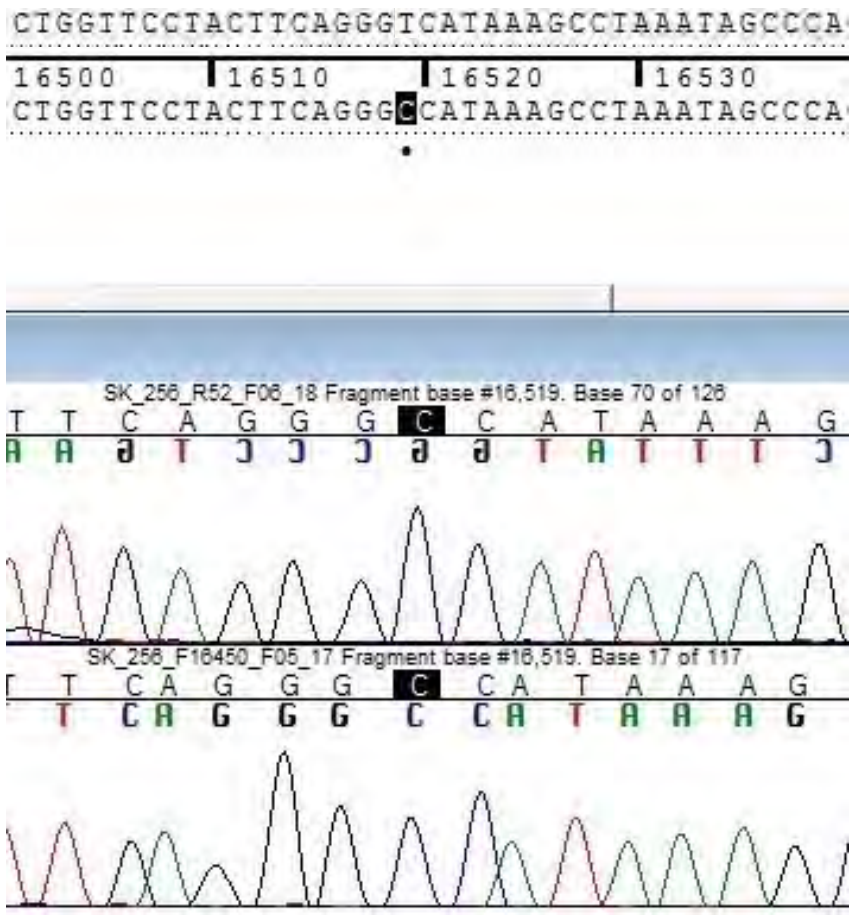
*revised Cambridge Reference Sequence (Andrews et al 1999)

MtDNA reference sequence

Control Region



rCRS



16500-16540: 16519C

first version of CRS (Cambridge Reference Sequence) published in 1981 (Anderson et al 1981)
revised version (**rCRS**; Andrews et al 1999)
GenBank >NC012920

Human mtDNA is reported relative to the **rCRS**

EMPOP

mtDNA nomenclature

prefix
rCRS variant

suffix
difference to rCRS

A73G

position
relative to rCRS

Prefixes are common in medical genetics
Forensic genetics typically uses position and suffix (73G)
Population genetics uses positions only (except transversions e.g. 73, 150A)
The IUPAC Code **lacks combinations of bases and deletions**
Introduced in forensics by using **lower case letters**
e.g. C150del/T -> **150t**, T152del/Y -> **152y,** C309del -> **309c**

IUPAC Code

Code	Represents	Complement
A	Adenine	T
G	Guanine	C
C	Cytosine	G
T	Thymine	A
Y	Pyrimidine (C or T)	R
R	Purine (A or G)	Y
W	weak (A or T)	W
S	strong (G or C)	S
K	keto (T or G)	M
M	amino (C or A)	K
D	A, G, T (not C)	H
V	A, C, G (not T)	B
H	A, C, T (not G)	D
B	C, G, T (not A)	V
X/N	any base	X/N
-	Gap	-

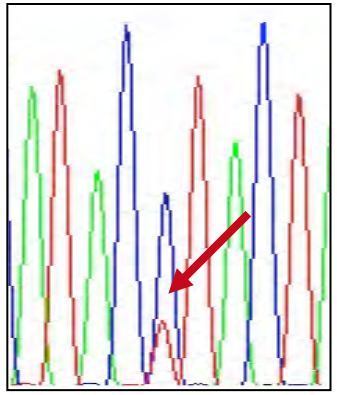
EMPOP

Parson et al 2014

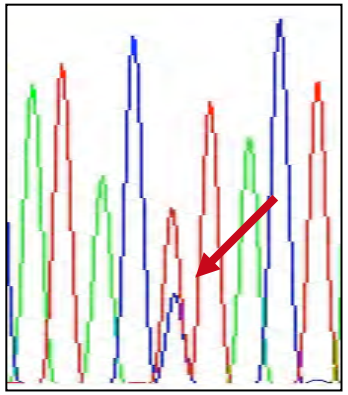
Heteroplasmy

Heteroplasmy describes the co-existence of (very) similar mtDNA molecules within an individual/tissue

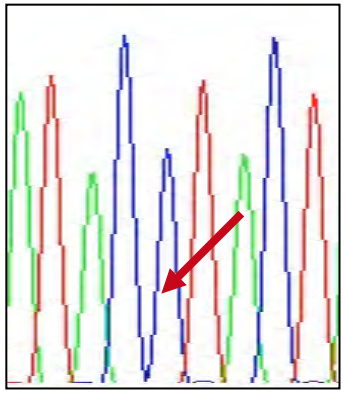
Blood



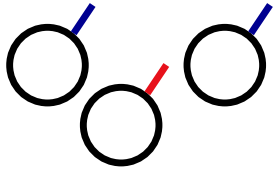
Hair 1

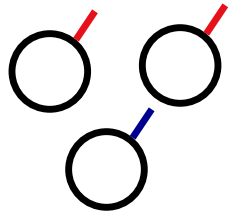


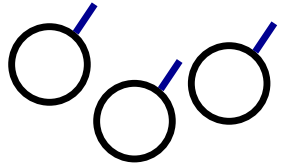
Hair 2

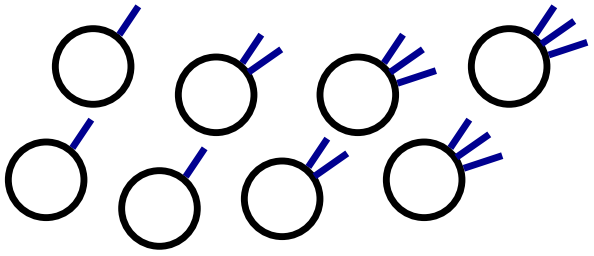


Point heteroplasmy
e.g. 152C/T = 152Y

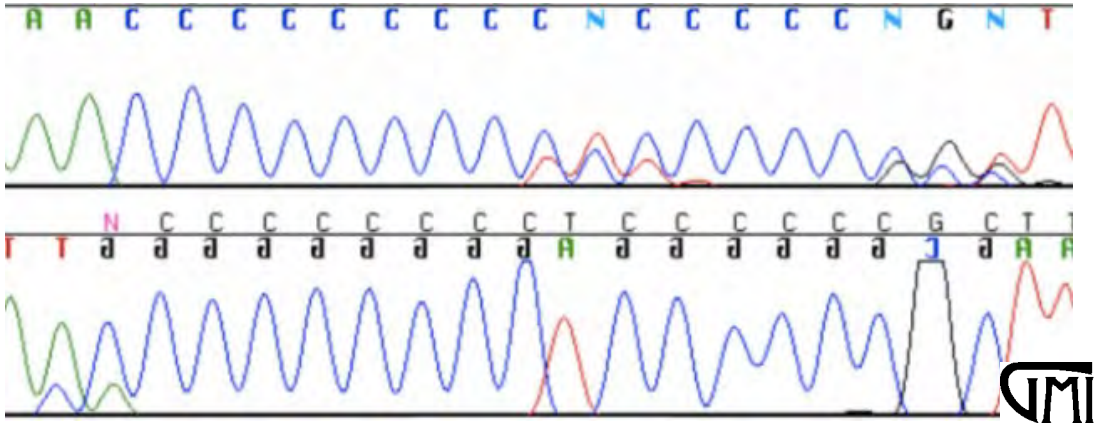




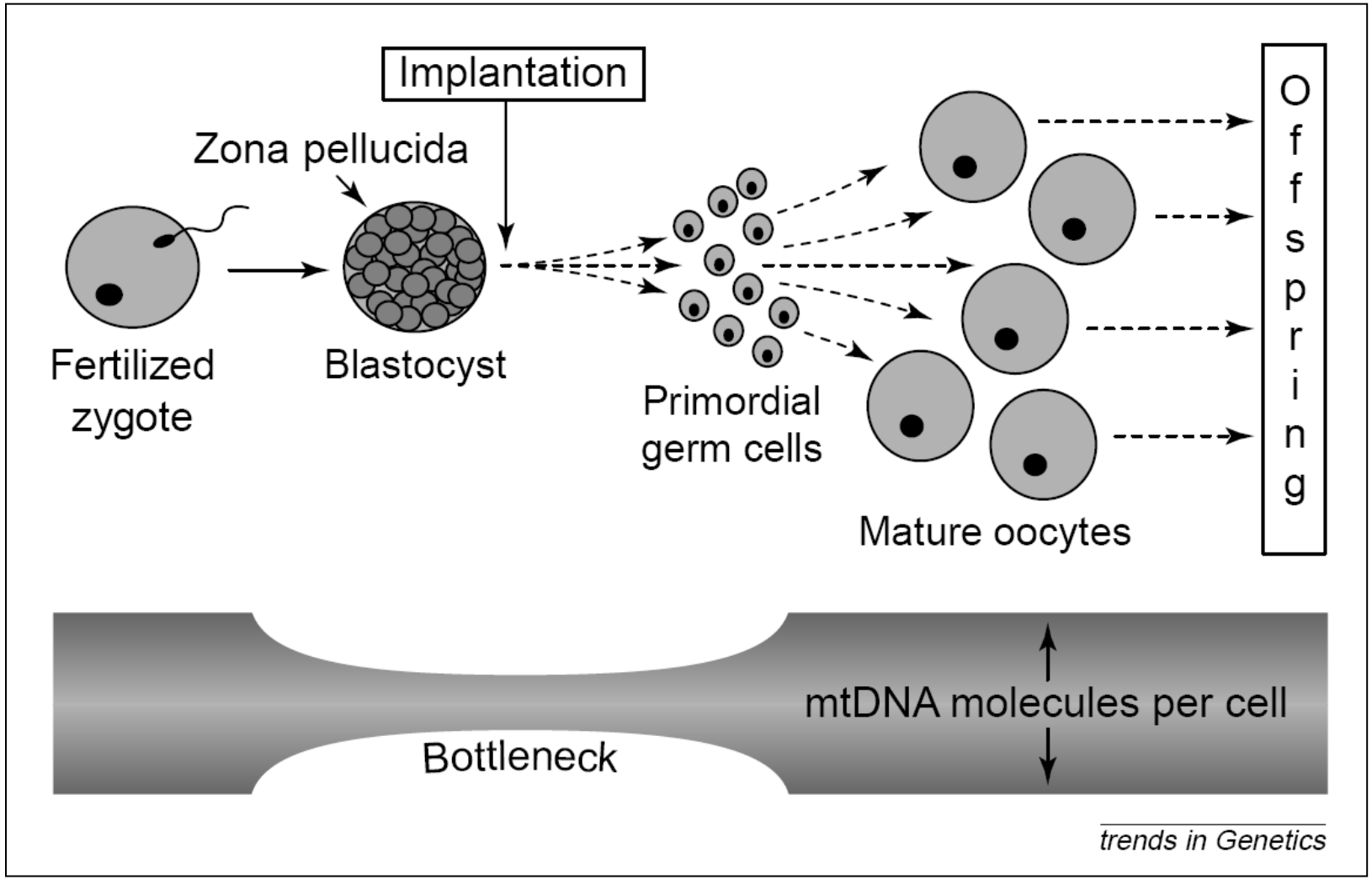




Length Heteroplasmy
e.g. 309+CCC



Bottleneck-Effect



Chinnery et al (2000)

Reproducibility of heteroplasmy detection

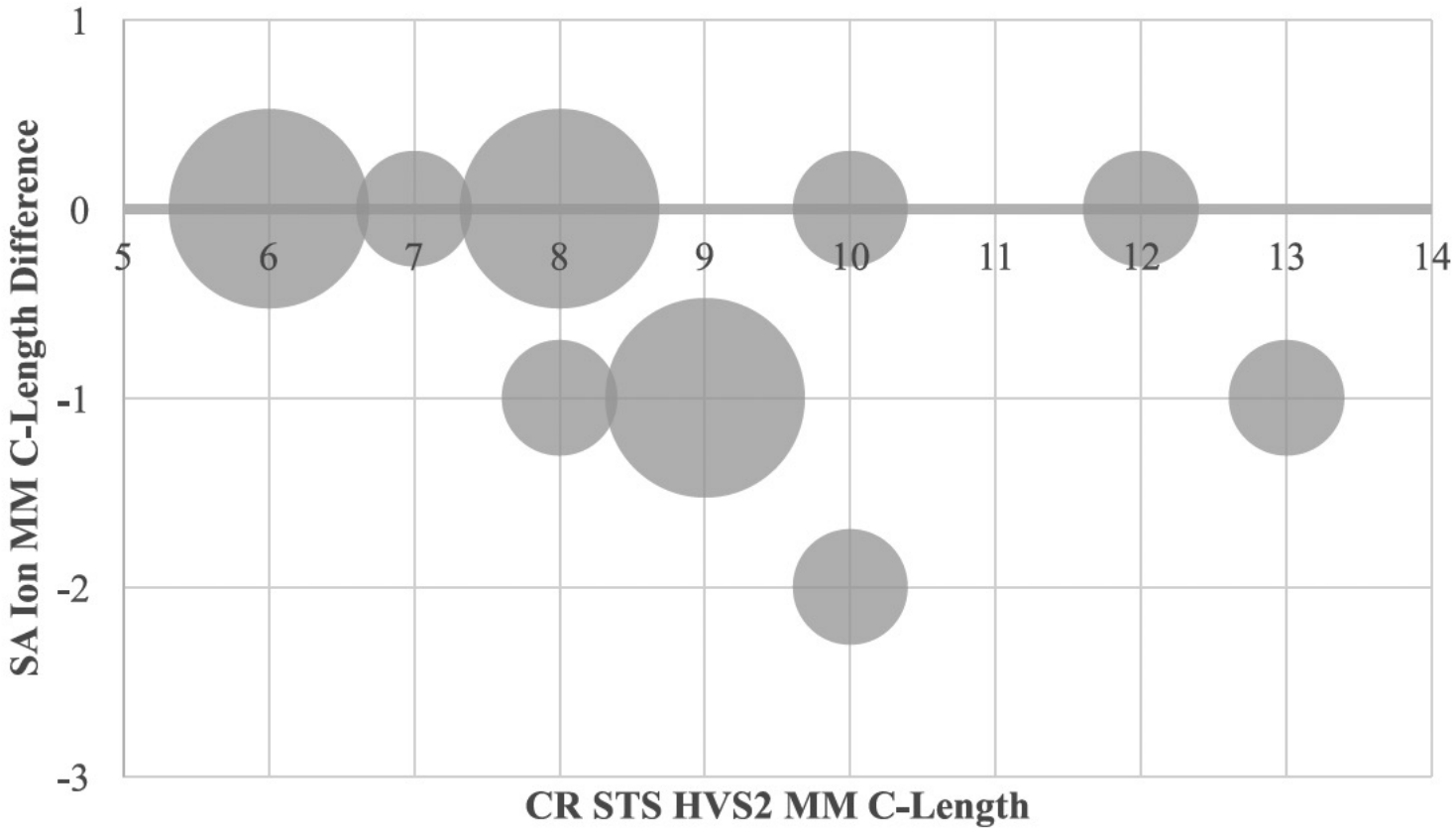
Sanger-type sequencing data showed little variation based on different primer and sequencing chemistries; results were by and large comparable

Massively Parallel Sequencing (MPS) technologies showed variation in heteroplasmy detection relative to Sanger-based results

First trend that MPS-based data yielded underestimation of length variation

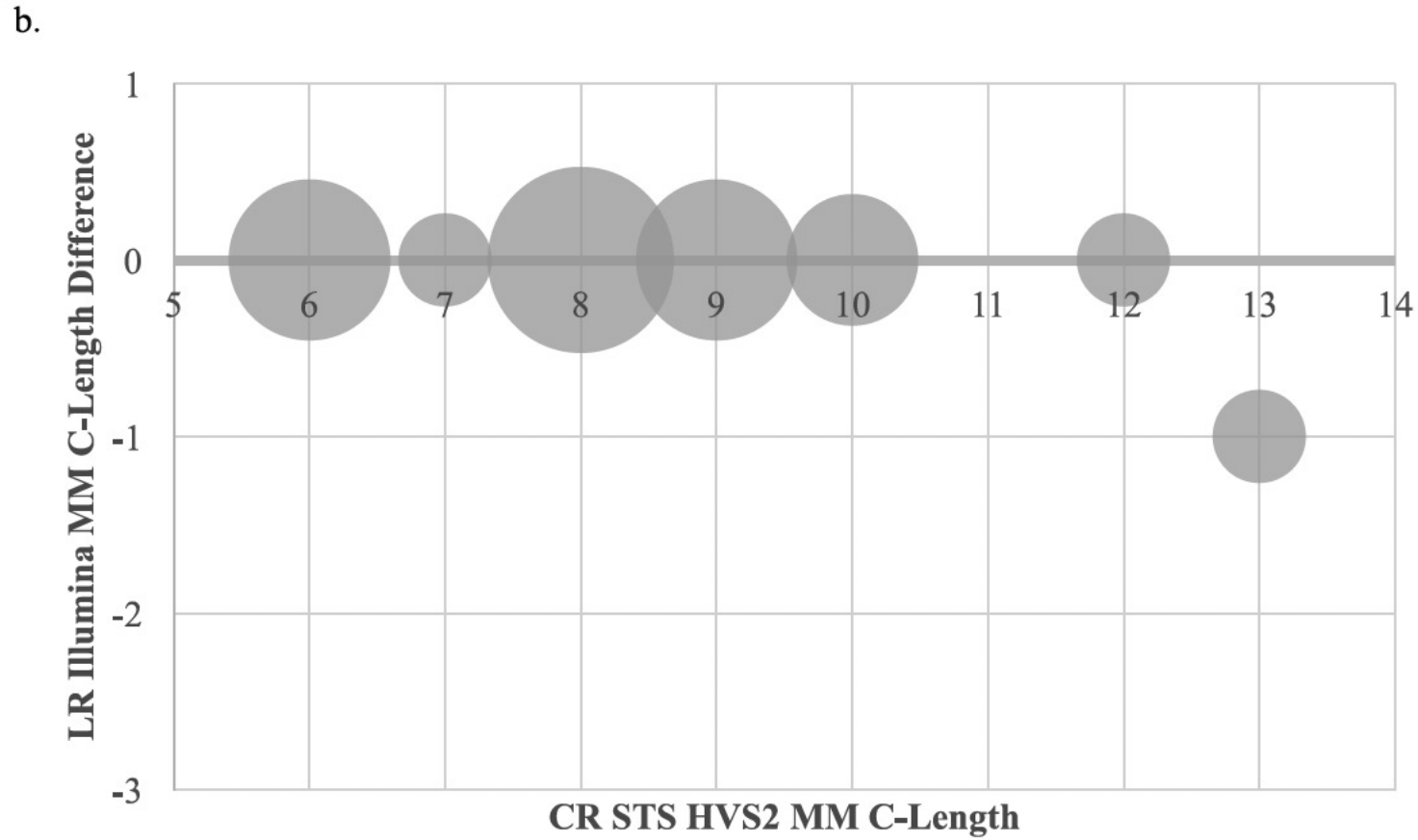
Comparing LHP DV between Sanger - Ion Torrent

b.



Sturk-Andreaggi et al ***FSIG*** (2020)

Comparing LHP DV between Sanger - Illumina



Sturk-Andreaggi et al ***FSIG*** (2020)

Suggestion for an EDNAP Heteroplasmy exercise 2019

Evaluate Point Heteroplasmy (PHP) and Length Heteroplasmy (LHP) between forensically relevant techniques, i.e. Sanger (CE) and Illumina/Ion Torrent (MPS)

First part

Evaluate the effect of different sequencing techniques on heteroplasmy detection, we shared DNA extracts prepared in Innsbruck from 5 reference samples (007, 9947A, 3 volunteers)

(**Second part**

To evaluate the somatic mutation rate in hair we shared hair shafts and DNA extracted from hair shafts) – will be presented later

Mitotypes from DNA extracts provided by Innsbruck

CO1: 007 (TFS 1707017)

73G 152C 199C 204C 207A 250C 263G 315.1C 460C **573.XC** 750G 1438G 1719A **2413Y** 2706G 4529T 4769G 6293C 7028T 8251A 8860G 9438A 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15758G 15924G 16129A 16223T 16391A 16519C

CO2: 9947A (Promega 18961603)

93G 195C 214G 263G **309.XC** 315.1C 750G 1438G 4135C 4769G 7645C **7861Y** 8448C 8860G 9315C 13572C 13759A 15326G 16311C 16519C

CO3: reference sample volunteer 1

73G 152C 195C 263G **309.XC** 315.1C **573.XC** 750G 1438G 2706G 3480G 4769G 5165T 7028T 8860G 9055A 9698C 11467G 11719A 12308G 12372A 14053G 14167T 14766T 15326G 15924G 16183C 16189C **16193.XC** 16234T 16324C 16519C

CO4: reference sample volunteer 2

16093Y 16311C 16519C 73G 263G 315.1C 750G 1438G 4769G 8860G **12483Y** 15326G

CO5: reference sample volunteer 3

16189C **16193.XC** 16356C 16362C 16519C **234R** 263G 315.1C 523del 524del **573.XC** 750G 1438G 3010A 3796G 4769G 8860G 15326G

Preliminary results – DNA extracts

Overview

We have received results and raw data for the 5 reference DNAs from the following technologies

Sanger (7-8 labs*)

Ion Torrent (12-13 labs*)

and Illumina technologies (6-7 labs*)

* not all analyses were successful

Preliminary results – DNA extracts

Concordance

Reported mitotypes were concordant between technologies, except for

clerical errors (e.g. 152T, lack of 16623T, lack of 460C …)

heteroplasmy reporting

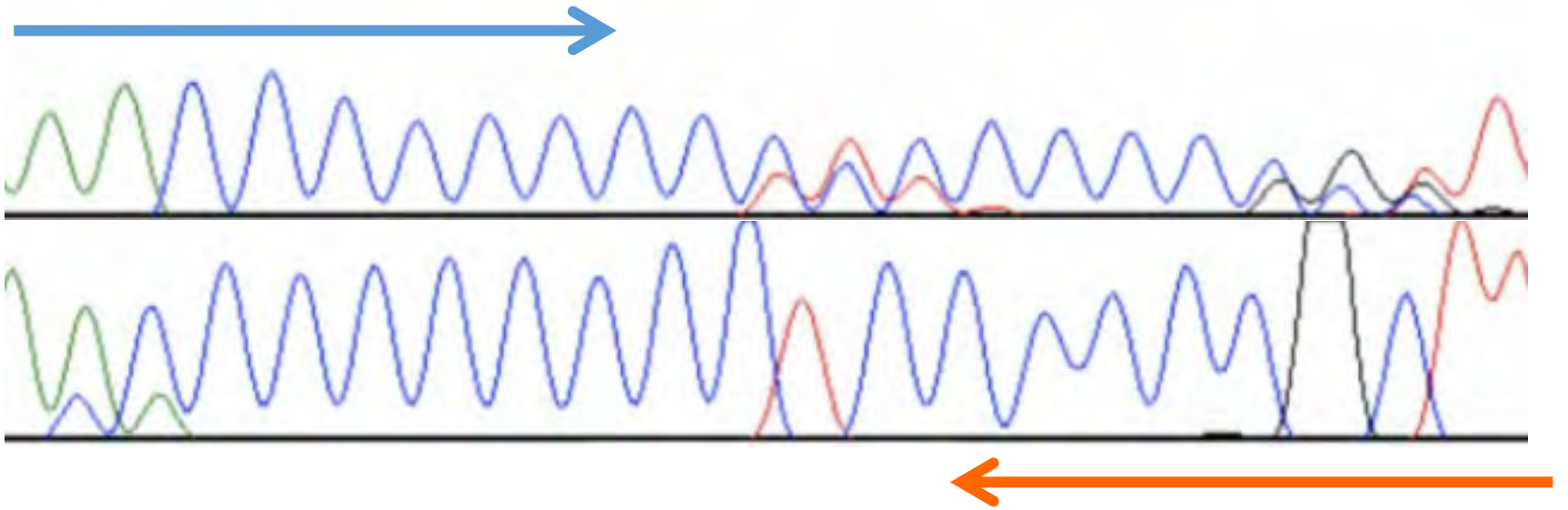
Preliminary results – DNA extracts

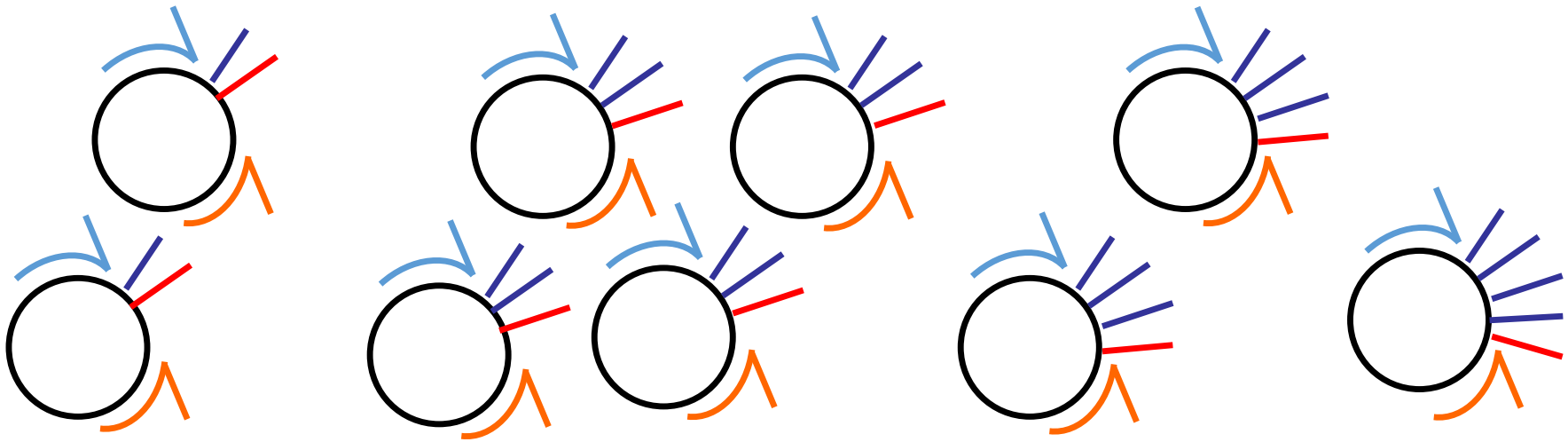
Length heteroplasmy (LHP)

Participants were asked to determine and report LHP according to their established guidelines

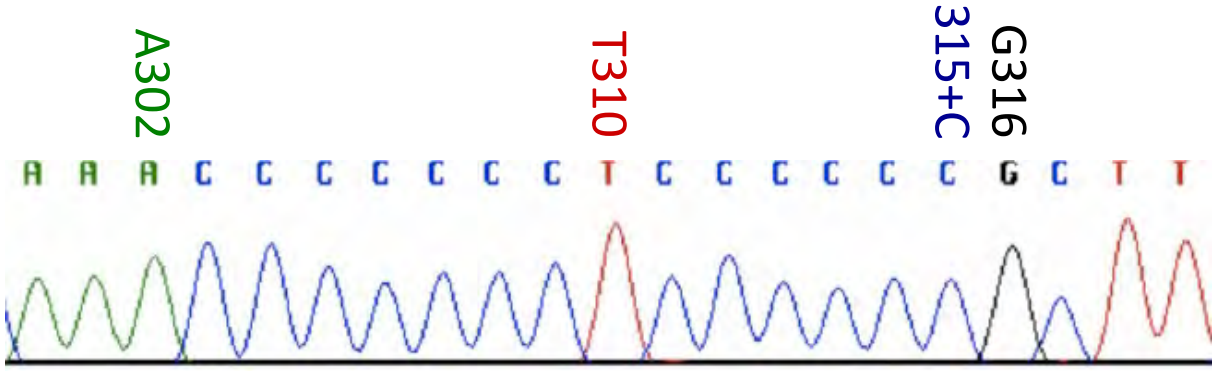
Most labs reported the **dominant variant** (= major molecule) as recommended by revised ISFG guidelines (Parson et al 2014)

Length heteroplasmy

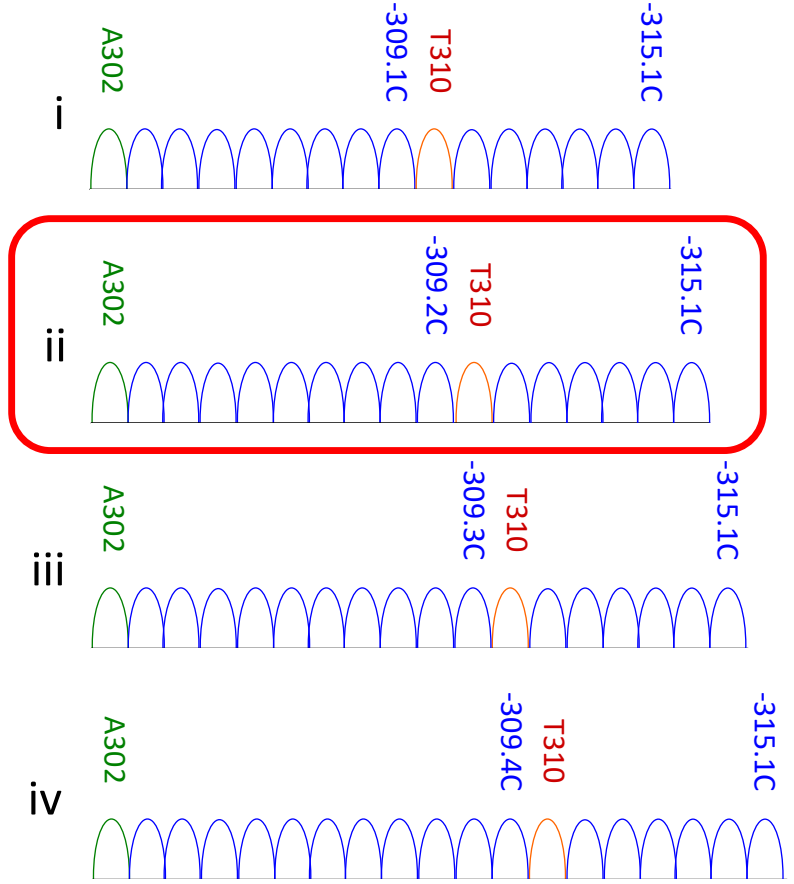




Length Heteroplasmy Dominant Variant

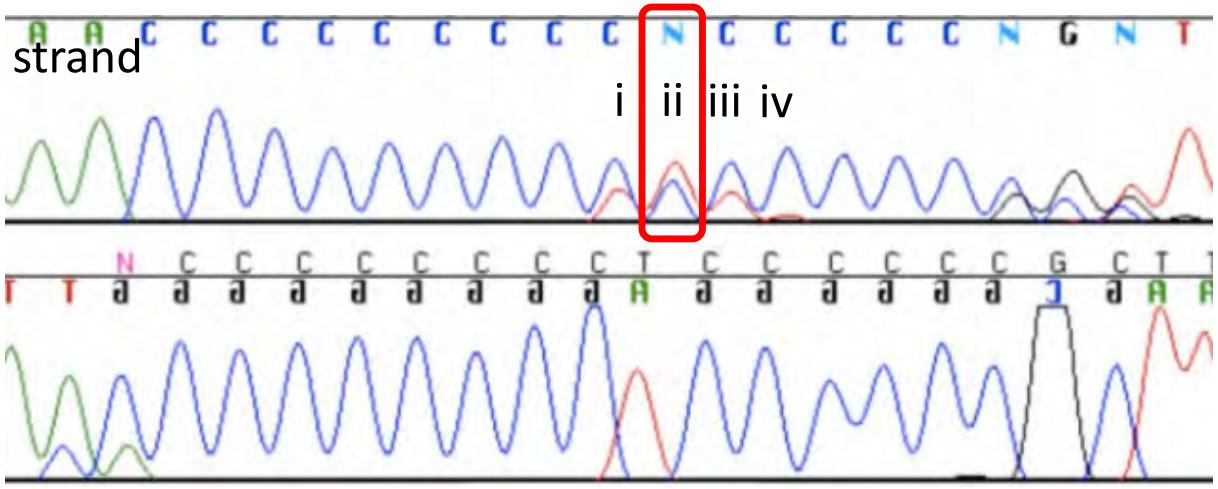


“dominant type/major molecule”



LHP example

forward strand



Parson et al (2014)

Preliminary results – DNA extracts

LHP CO5 – **Sanger** (reported)

CO5: reference sample volunteer 3

HVS-I motif: 16189C **16193.XC** 16356C 16362C 16519C

Lab	Software	Range	Mitotype					
Lab 0	Sequencher v5.1	16024-576	16189C	16193.1C	16193.2C	16356C	16362C	16519C
Lab 5	SeqScape v3	48-408 15997-16401	16189C			16356C	16362C	
Lab 6	Sequencher v5.1	15911-16396 39-395	16189C	16193.1C	16193.2C	16356C	16362C	
Lab 7	Sequencher v5.4.6	16024-576	16189C	16193.1C	16193.2C	16356C	16362C	16519C
Lab 21	SeqScape v3	16024-576	16189C	16193.1C	16193.2C	16356C	16362C	16519C
Lab 22	Sequencing Analysis v6.0	16020-576	16189C			16356C	16362C	16519C
Lab 23	BioEdit, MEGA6, Seq. analysis	65-431 15869-16502	16189C			16356C	16362C	
Lab 24	SeqScape 2.6	16024-576	16189C	16193.1C	16193.2C	16356C	16362C	16519C

Preliminary results – DNA extracts

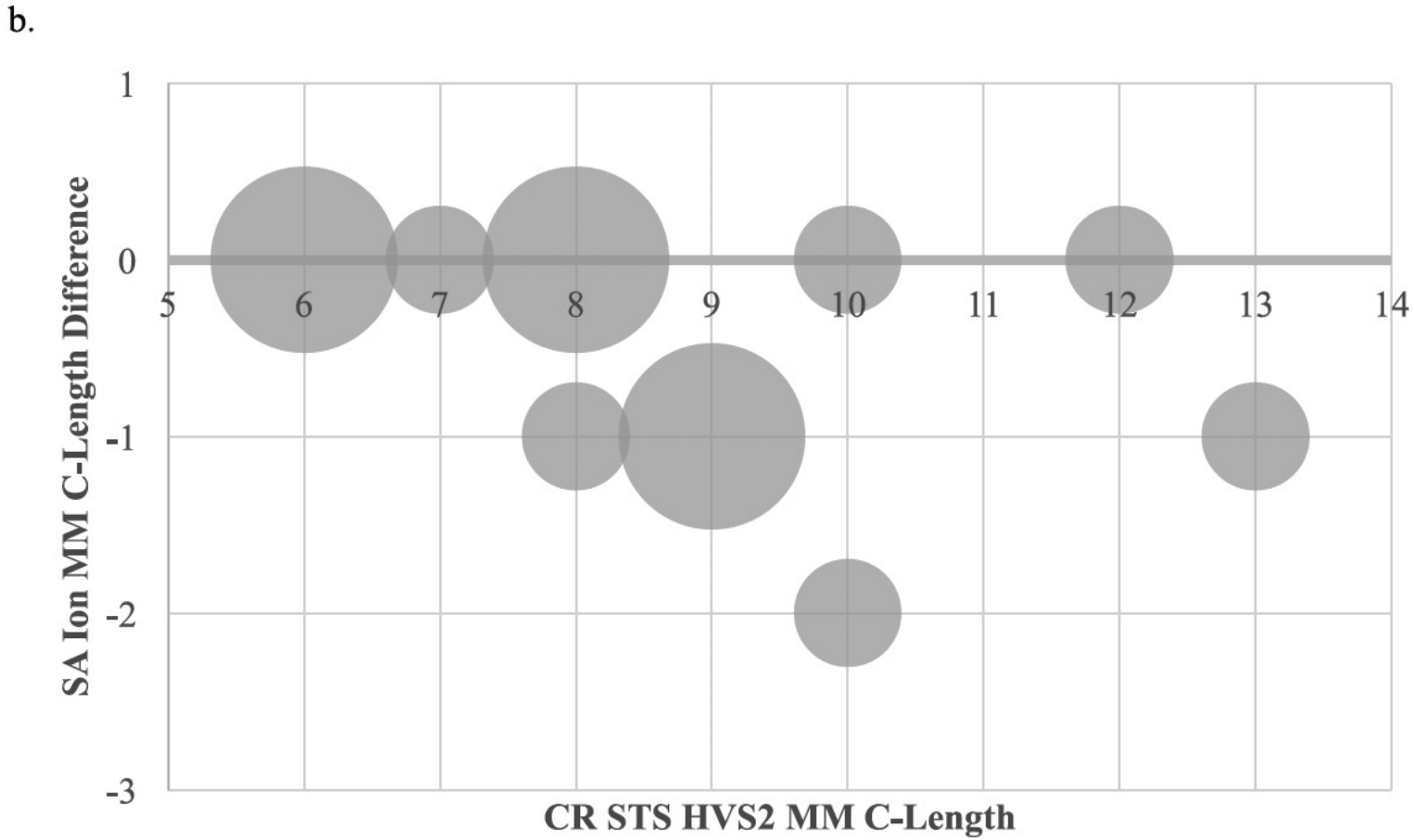
LHP CO5 – **Ion Torrent** (reported)

CO5: reference sample volunteer 3

HVS-I motif: 16189C **16193.XC** 16356C 16362C 16519C

Lab	Software	Range	LHP reported	
Lab 0	IGV	1-16569	16193.1C	
Lab 1	IGV+ mito IGV/ Converge v2.1	1-16569	rCRS	
Lab 3	IGV+Converge	1-16569	16193.1C	
Lab 4	Converge v2.2+IGV	1-16569	16193.1C	
Lab 5	IGV	16024-576	rCRS	
Lab 6	IGV v2.3.94	1-16569	rCRS	
Lab 10	Converge v2.1+mito IGV	1-16569	rCRS	
Lab 12	Converge v2.2	16024-576	16193.1C	
Lab 13	IGV v2.3.72+GeneMarker HTS	1-16569	16193.1C	16193.2C
Lab 14	TVC v5.10.1.19, IGV, Converge for LHP	1-16569	16193.1C	
Lab 15	Converge v2.2	1-16569	rCRS	
Lab 17	IGV	4-514 516-16569	rCRS	

Comparing LHP DV between Sanger - Ion Torrent



Sturk-Andreaggi et al ***FSIG*** (2020)

Preliminary results – DNA extracts

LHP CO5 – **Illumina** (reported)

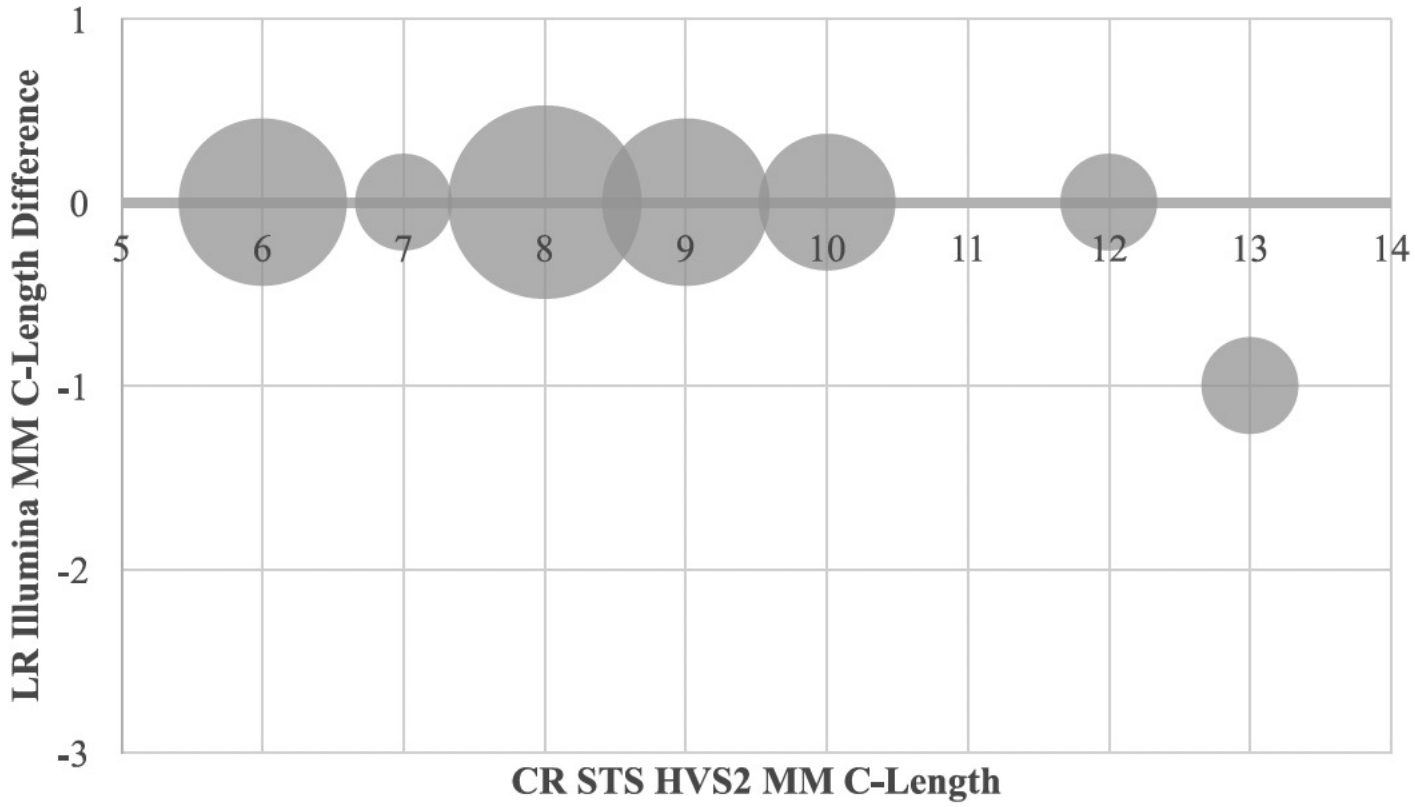
CO5: reference sample volunteer 3

HVS-I motif: 16189C **16193.XC** 16356C 16362C 16519C

Lab	**Software**	**Range**	**LHP reported**	
Lab 7	Base Space, IGV	16024-576	16193.1C	
Lab 8	IGV	1-16569		
Lab 9	CLC Gen. Workbench v12/AQME	1-16569	rCRS	
Lab 11	GeneMarker HTS	1-16569		
Lab 16	fdstools	16009-589	16193.1C	
Lab 18	IGV	15985-576	16193.1C	16193.2c
Lab 19	CLCBio v12.0	16024-576	16193.1C	16193.2C
Lab 20	GeneMarker HTS	1-16569	rCRS	

Comparing LHP DV between Sanger - Illumina

b.



Sturk-Andreaggi et al ***FSIG*** (2020)

Preliminary results – DNA extracts

First summary

Reporting of LHP in DNA extracts not concordant between technologies and labs

In total we discern three sources of variation: **technological, software** and **interpretational**

Limitation: we do not have access to all the different software versions used by labs

Lab	Software
Lab 0	IGV
Lab 1	IGV+ mito IGV/ Converge v2.1
Lab 3	IGV+Converge
Lab 4	Converge v2.2+IGV
Lab 5	IGV
Lab 6	IGV v2.3.94
Lab 10	Converge v2.1+mito IGV
Lab 12	Converge v2.2
Lab 13	IGV v2.3.72+GeneMarker HTS
Lab 14	TVC v5.10.1.19, IGV, Converge for LHP
Lab 15	Converge v2.2
Lab 17	IGV

Ion Torrent software versions

Lab	Software
Lab 7	Base Space, IGV
Lab 8	IGV
Lab 9	CLC Gen. Workbench v12/AQME
Lab 11	GeneMarker HTS
Lab 16	fdstools
Lab 18	IGV
Lab 19	CLCBio v12.0
Lab 20	GeneMarker HTS

Illumina software versions

Preliminary results – DNA extracts

Way forward

We would like to understand the variability caused by the sequencing technologies, ignoring differences in software (versions) and human interpretation

Decided to develop software that is **agnostic** to instrument/technology

"MPSaligner" is based on SAM2 (sequence alignment method; EMPOP; Dür et al 2022)

- reads and writes in sam formats
- SAM2 implements phylogenetic alignment from the estimated hg-motif
- SAM2 executes global alignment (no changes in the read lengths)
- Software converts sam to emp format, filters numts and assembles consensus

Agnostic software removes individual user settings and differences in software (versions) and highlights differences between technologies

(One reason for the delay)

Preliminary results – DNA extracts

LHP CO5 – **Ion Torrent** (reported vs. SAM2)

CO5: reference sample volunteer 3

HVS-I motif: 16189C **16193.XC** 16356C 16362C 16519C

Lab	Software	Range	LHP reported		LHP SAM2	
Lab 0	IGV	1-16569	16193.1C		rCRS	x
Lab 1	IGV+ mito IGV/ Converge v2.1	1-16569	rCRS		rCRS	
Lab 3	IGV+Converge	1-16569	16193.1C		rCRS	x
Lab 4	Converge v2.2+IGV	1-16569	16193.1C		16193.1C	
Lab 5	IGV	16024-576	rCRS		rCRS	
Lab 6	IGV v2.3.94	1-16569	rCRS		rCRS	
Lab 10	Converge v2.1+mito IGV	1-16569	rCRS		rCRS	
Lab 12	Converge v2.2	16024-576	16193.1C		rCRS	x
Lab 13	IGV v2.3.72+GeneMarker HTS	1-16569	16193.1C	16193.2C	16193.1C	x
Lab 14	TVC v5.10.1.19, IGV, Converge for LHP	1-16569	16193.1C		16193.1C	
Lab 15	Converge v2.2	1-16569	rCRS		rCRS	
Lab 17	IGV	4-514 516-16569	rCRS		rCRS	

Preliminary results – DNA extracts

LHP CO5 – **Illumina** (reported vs. SAM2)

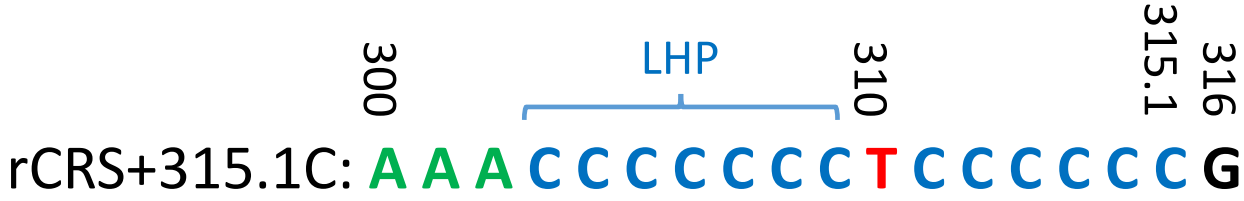
CO5: reference sample volunteer 3

HVS-I motif: 16189C **16193.XC** 16356C 16362C 16519C

Lab	Software	Range	LHP reported		LHP SAM2	
Lab 7	Base Space, IGV	16024-576	16193.1C		rCRS	x
Lab 8	IGV	1-16569				
Lab 9	CLC Gen. Workbench v12/AQME	1-16569	rCRS		16193.1C	x
Lab 11	GeneMarker HTS	1-16569			16193.1C	
Lab 16	fdstools	16009-589	16193.1C		16193.1C	
Lab 18	IGV	15985-576	16193.1C	16193.2c	16193.1C	x
Lab 19	CLCBio v12.0	16024-576	16193.1C	16193.2C	16193.1C	x
Lab 20	GeneMarker HTS	1-16569	rCRS		16193.1C	

Figure 1

LHP in HVS-II C-stretch (at 309) - preliminary



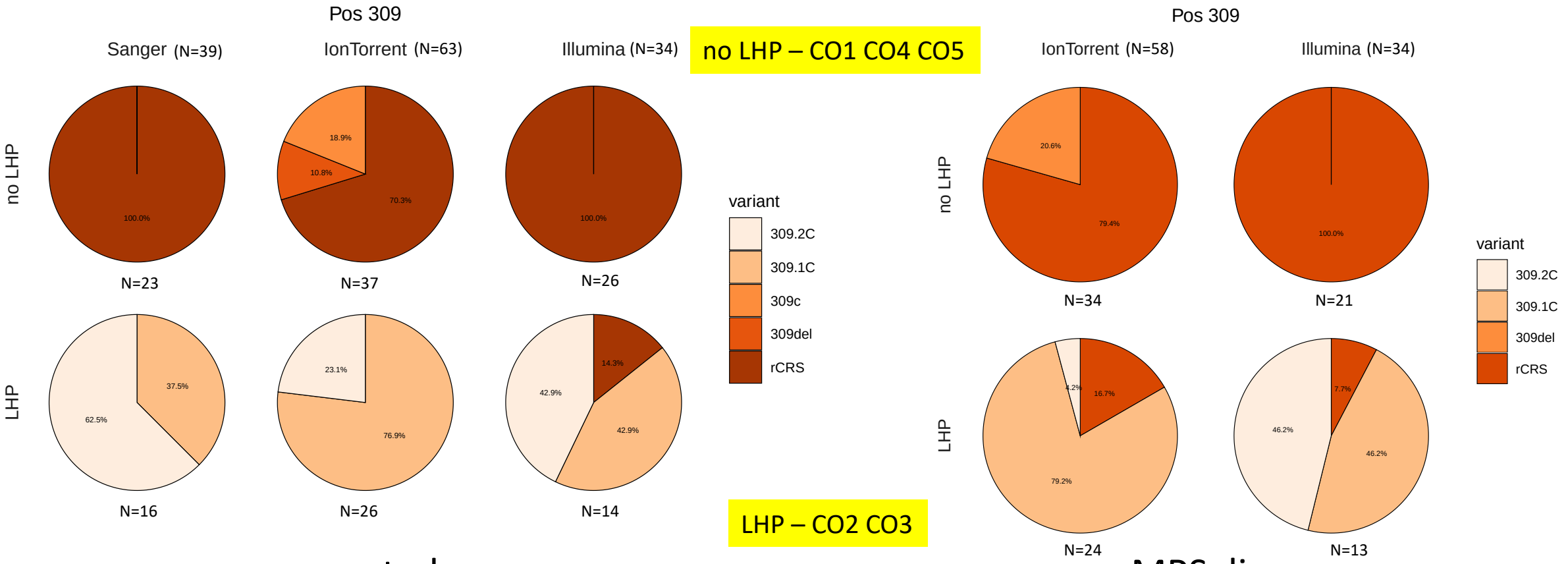


Figure 2

LHP in HVS-III C-stretch (at 573) - preliminary

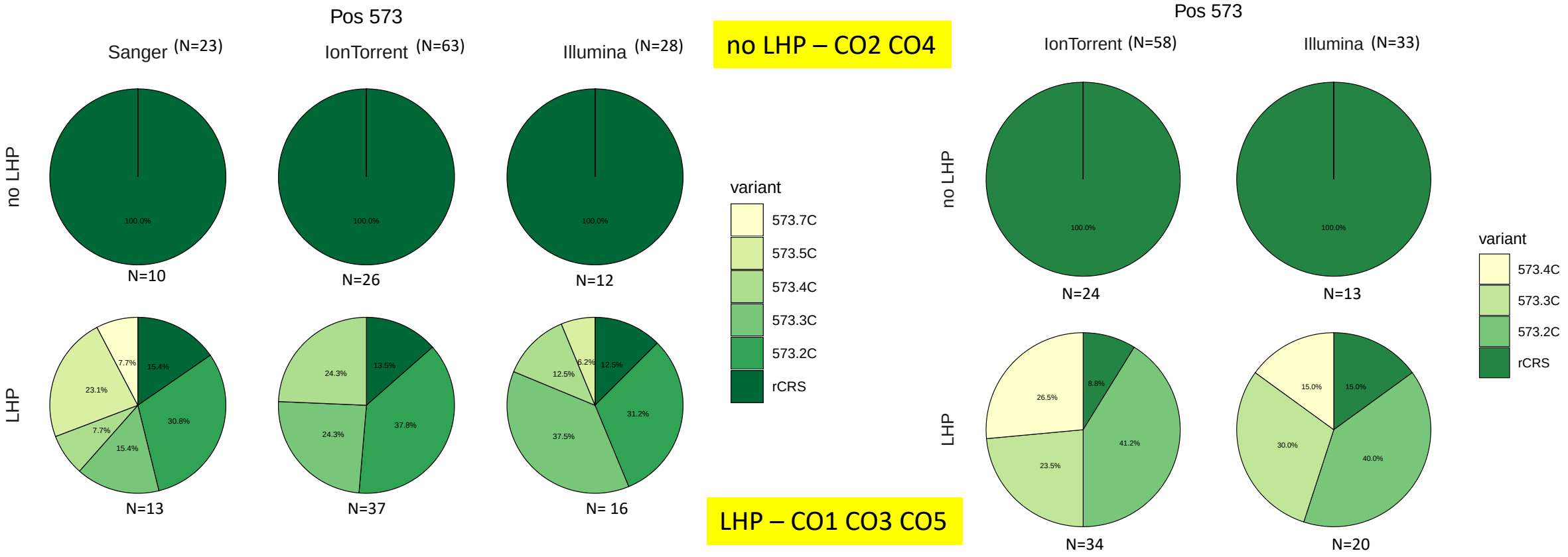
565 573

rCRS: A C A C C C C C A C A

LHP

no LHP – CO2 CO4

LHP – CO1 CO3 CO5

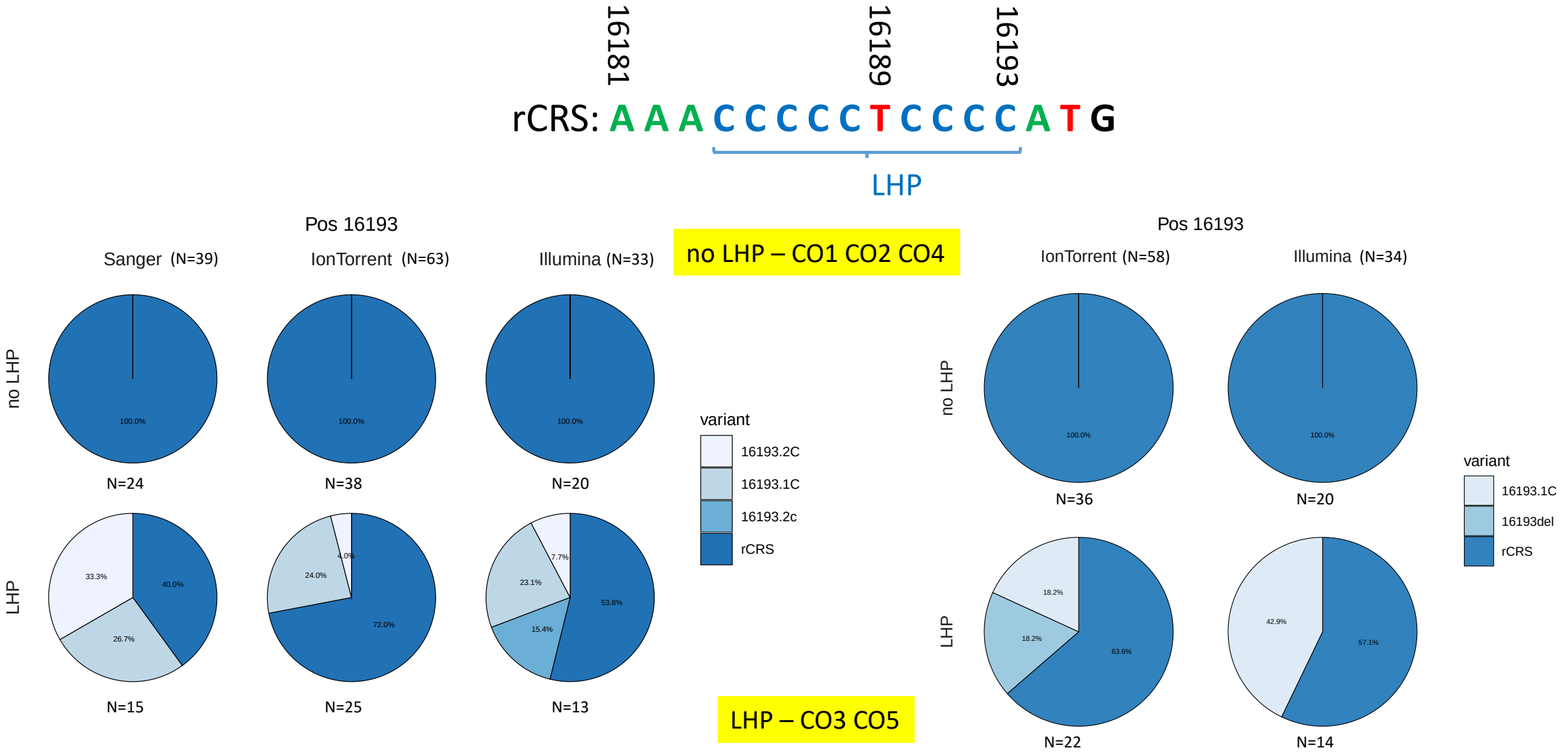


reported

MPSAligner

Figure 3

LHP in HVS-I C-stretch (at 16193) - preliminary



LHP in DNA extracts – preliminary results

C-stretch around **309** (Figure 1)

no LHP: Sanger and Illumina identical, Ion Torrent shows some variation

LHP reported: variation reported for Sanger (2), Ion Torrent (2) and Illumina (3)

LHP MPSAligner (MPS only): variation observed for Ion Torrent (4) and Illumina (3)

C-stretch around **573** (Figure 2)

no LHP: Sanger, Ion Torrent and Illumina identical

LHP reported: variation reported for Sanger (6), Ion Torrent (4) and Illumina (5)

LHP MPSAligner (MPS only): variation observed for Ion Torrent (4) and Illumina (4)

C-stretch around **16193** (Figure 3)

no LHP: Sanger, Ion Torrent and Illumina identical

LHP reported : variation reported for Sanger (3), Ion Torrent (4) and Illumina (4)

LHP MPSAligner (MPS only): variation observed for Ion Torrent (4) and Illumina (2)

LHP in DNA extracts – preliminary conclusions

We observed variation at different levels of analysis

1) technological,
2) software,
3) interpretation

Developed software to understand technological/software differences

Level of variation between technologies and software seems high

May not be able to call Dominant LHP variants consistently

(New technologies may add more variability)

Future guidelines may need to account for this variation by further relaxing recommendations on LHP reporting

EDNAP age exercise

MiSEQ / PGM comparison

Introduction

The development of methods that can accurately estimate an individual's chronological age from trace evidence is an ongoing quest in the field of forensic DNA intelligence. The retrieval of this information, as well as information regarding externally visible characteristics, like eye, hair or skin colour and hair morphology [1-4], from DNA samples recovered from crime scenes, can significantly aid police investigations, especially in cases lacking eye witness testimonies and/or intact human remains.

In order to investigate these issue further, this exercise focuses on the transferability of a previously described DNA methylation-based age prediction method originally developed on the MiSeq FGx platform [40]. The same protocol, with minor instrument-related alterations was performed in 15 different laboratories using different types of MPS technology including the MiSeq, MiSeq FGx, Ion PGM and Ion S5 systems and the results were compared both for standards of known methylation and real samples.

Methods

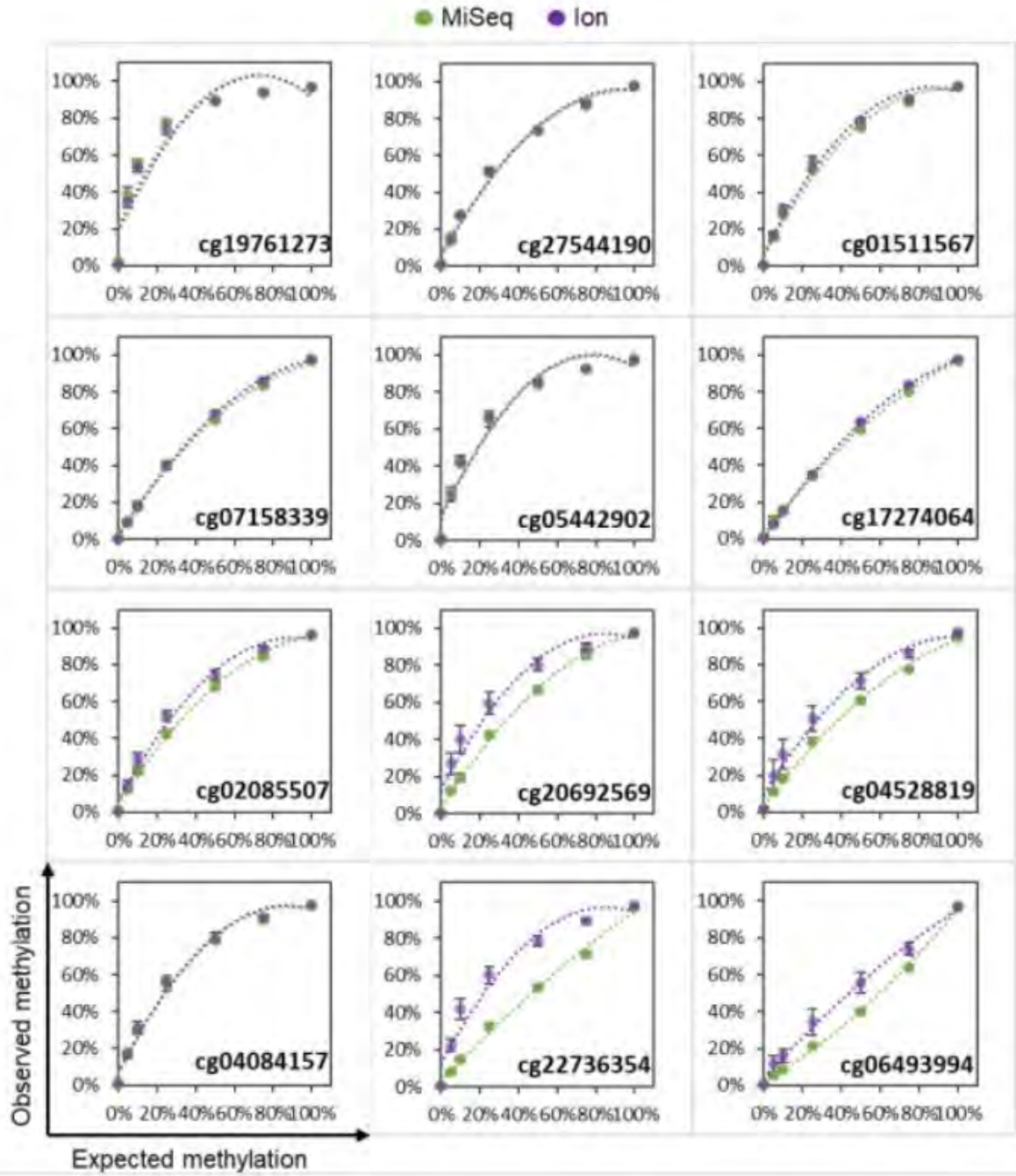
For the first part of this study 7 pre-mixed methylation standards ranging from 0% to 100% methylation (0%, 5%, 10%, 25%, 50%, 75% and 100%) were purchased from EpigenDx (Massachusetts, USA) at a concentration of 50 ng/µL. Standards were diluted and delivered to the participating laboratories at a final concentration of 2.5 ng/µL. Each laboratory then proceeded with bisulphite conversion, PCR, library preparation and methylation quantification for the set of 12 age-associated CpG markers, as set out below.

The second element of this study involved the participating laboratories each analysing the same set of 7 blood samples to generate methylation values for the 12 age-associated CpG markers, following which an age prediction was generated for each sample/laboratory combination. Principal sample collection for this aspect of the study was performed by King's College London under ethical approval granted by the Biomedical Sciences, Dentistry, Medicine and Natural & Mathematical Sciences Research Ethics Subcommittee (BDM/13/14-30). A total of 7 donors aged between 27.7 and 79.7 years were recruited for the collection of whole blood samples (samples A-G) via venepuncture following the acquisition of full informed consent. Samples were stored at 4 oC.

Sequencing

MiSeqFGx: The MiSeqFGx instruments were used in combination with MiSeq Reagent Kits version 2 (300 cycles set to 151bp reads each way) or version 3 (150bp or 200bp single read cycles) (Verogen, California, USA). Prior to sequencing, libraries were pooled together at 4nM and diluted to a final concentration of 6-10pM, while a PhiX library was spiked in at approximately 10%.

PGM: For sequencing with the Ion PGM instrumentation, the Ion PGM™ Hi-Q™ OT2 kit, Ion PGM™ Hi-Q™ Sequencing kit and Ion 316/318 Chip kit version 2 (Thermo Fisher Scientific, Massachusetts, USA) were used. Prior to sequencing, libraries were diluted to 100pM and pooled together to a final concentration of 1.3pM.


MiSeq
Ion
cg19761273
cg27544190
cg01511567
cg07158339
cg05442902
cg17274064
cg02085507
cg20692569
cg04528819
cg04084157
cg22736354
cg06493994
0% 20% 40% 60% 80% 100%
Observed methylation
Expected methylation

Summary

- PGM lab predictions significantly different
- Global model corrections applied to the PGM results but the between PGM lab variation was significantly higher
- Blind samples – to b be reanalysed with a new model

University of Zurich[UZH]

Zurich Institute of Forensic Medicine

EDNAP mRNA MPS collaborative exercise 4 IonTorrent S5 and Illumina MiSeq (BFID-cSNP-6F)

Cordula Haas, Nadescha Hänggi, Erin Hanson, Jack Ballantyne

EDNAP Meeting, 29. May 2024, Copenhagen

EUROFORGEN Network of Excellence

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

Association of Body Fluids with a Donor: cSNPs

DNA Profile:

- Mixture of 2 persons

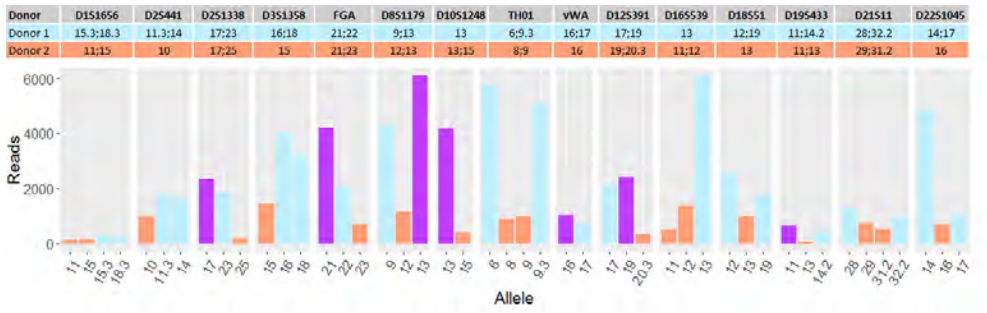
RNA Profile:

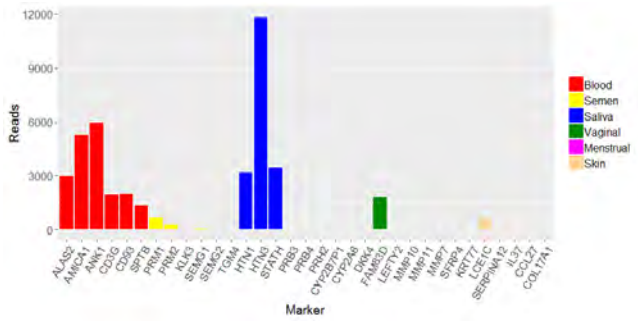
- Body fluid identification (BFID)

cSNPs:

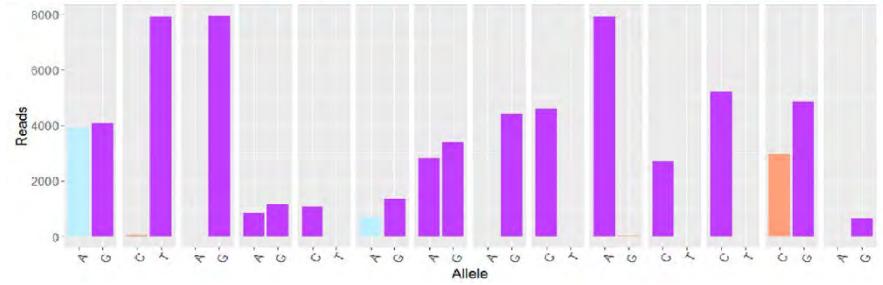
- Association to donors

Donor	D1S1656	D2S441	D2S1338	D3S1358	FGA	D8S1179	D10S1248	TH01	vWA	D12S391	D16S539	D18S51	D19S433	D21S11	D22S1045
Donor 1	15.3;18.3	11.3;14	17;23	16;18	21;22	9;13	13	6;9.3	16;17	17;19	13	12;19	11;14.2	28;32.2	14;17
Donor 2	11;15	10	17;25	15	21;23	12;13	13;15	8;9	16	19;20.3	11;12	13	11;13	29;31.2	16





	Blood markers											Saliva markers		
Donor	AMICA1_1	AMICA1_2	ANK1_1	ANK1_2	ANK1_3	ANK1_4	CD3G	CD93_1	CD93_2	CD93_3	SPTB	MUC7_1	MUC7_2	PRB3
Donor 1	AG	TT	GG	AG	CC	AG	AG	GG	CC	AA	CC	CC	CC	GG
Donor 2	GG	CT	AG	AG	CC	GG	AG	GG	CT	AG	CC	CT	CG	GG



Association of Body Fluids with a Donor: cSNPs

International Journal of Legal Medicine (2023) 137:13–32
https://doi.org/10.1007/s00414-022-02908-9

ORIGINAL ARTICLE

Check for updates

Targeted S5 RNA sequencing assay for the identification and direct association of common body fluids with DNA donors in mixtures

Erin Hanson[1,2] · Guro Dørum[3] · Manuel Zamborlin[3] · Shouyu Wang[3] · Mario Gysi[3] · Sabrina Ingold[3] · Robert Lagace[4] · Chantal Roth[4] · Cordula Haas[3] · Jack Ballantyne[1,2]

BFID-cSNP-BSS blood, semen, saliva

BFID-cSNP-6F 6 fluids/tissue

Recap EDNAP mRNA MPS Exercise 3

- BFID-cSNP-BSS RNA assay
- BFID-cSNP-BSS DNA assay

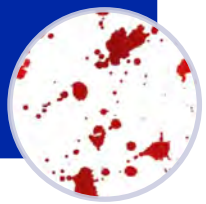
→ reference donor genotypes

- 6 participating laboratories on IonTorrent S5
- mRNA profiling + cSNP typing for 16 provided stains + 8 own stains incl. up to 8 DNA references

	BFID-cSNP-BSS	
BD	ANK1	2
	CD3G	1
	SPTB	4
SE	PRM1	1
	SEMG2	1
	KLK3	2
	TGM4	4
SA	HTN3	3
	PRB4	1
	PRH2	1
	MUC7	1
	STATH	
VS	CYP2A6	
	MUC22	
	CYP2B7P1	
MB	MMP10	
	MMP3	
	COL6A3	
	COL12A1	
	LEFTY2	
SK	LCE1C	
	COL17A1	
	IL37	
Total		19 genes (BFID) 23 cSNPs (11 genes) *cSNP microhaps:* *3 bi-local, 1 tri-local*

13/16 provided
21/32 own stains ✓

BFID

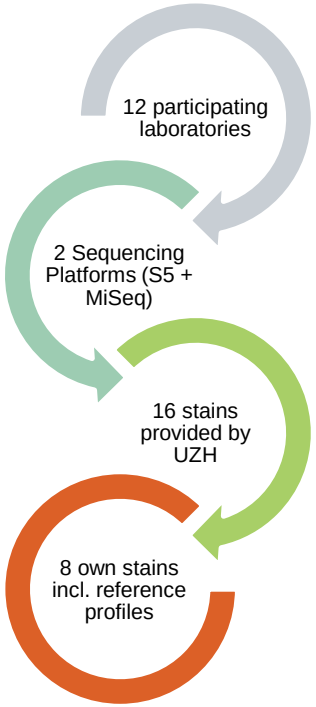


Performance dependent on how many reads detected

cSNPs

EDNAP mRNA MPS Exercise 4

	BFID-cSNP-6F	
BD	ANK1	2
	CD3G	1
	SPTB	4
SE	PRM1	1
	SEMG2	1
	KLK3	2
	TGM4	4
SA	HTN3	3
	PRB4	1
	PRH2	1
	MUC7	1
	STATH	
VS	CYP2A6	1
	MUC22	7
	CYP2B7P1	
MB	MMP10	2
	MMP3	1
	COL6A3	5
	COL12A1	3
	LEFTY2	
SK	LCE1C	3
	COL17A1	1
	IL37	2
Total		23 genes (BFID) 46 cSNPs (20 genes) *cSNP microhaps:* *8 bi-local, 3 tri-local*



Stain N°	BF/T	Amount	Stain Provided
1	SK	1 swab	1 swab
2	BL-MB	1 swab + 25ul	1/4 swab
3	SA-VAG	1 swab + 25ul	1/4 swab
4	SE-MB	1 swab + 25ul	1/4 swab
5	BL-SE	25ul + 25ul	part of T-Shirt
6	SE-SE	25ul + 25ul	1 swab
7	SA-MB	1 swab + 50ul	1/4 swab
8	SA-SK	1 swab + 25ul	1 swab
9	VAG	cotton part of a slip	a piece of it
10	MB	menstrual pad	a part of it
11	SE	50ul	part of a glove (latex)
12	BL	20ul	part of a T-Shirt
13	SA-SE	50ul + 10ul	artificial cotton
14	VAG-BL	1 swab + 25ul	1/4 swab
15	SA	50ul	stockings (nylon)
16	VAG-SE	1 swab + 25ul	1/4 swab

Light blue: single donor, low input
Dark blue: single donor, high input
Orange: mixtures

RNA/DNA Co-Extraction → RT → cDNA STR-Analysis → Library Preparation → Sequencing: S5/MiSeq → Data Analysis

Participating Laboratories

6x S5
3x MiSeq
2x both sequencing platforms

Netherlands Forensic Institute, Ministry of Justice and Security, Netherlands

National Forensic Center, Swedish Police Authority, Sweden

Department of Analytical, Environmental and Forensic Sciences, King's College London, UK

Institute of Forensic Medicine, University of Zurich, Switzerland

Department of Forensic Medicine, University of Copenhagen, Denmark

Institute of Forensic Medicine, University Medical Center Cologne, University of Cologne, Germany

National Center for Forensic Science, University of Central Florida (UCF), USA

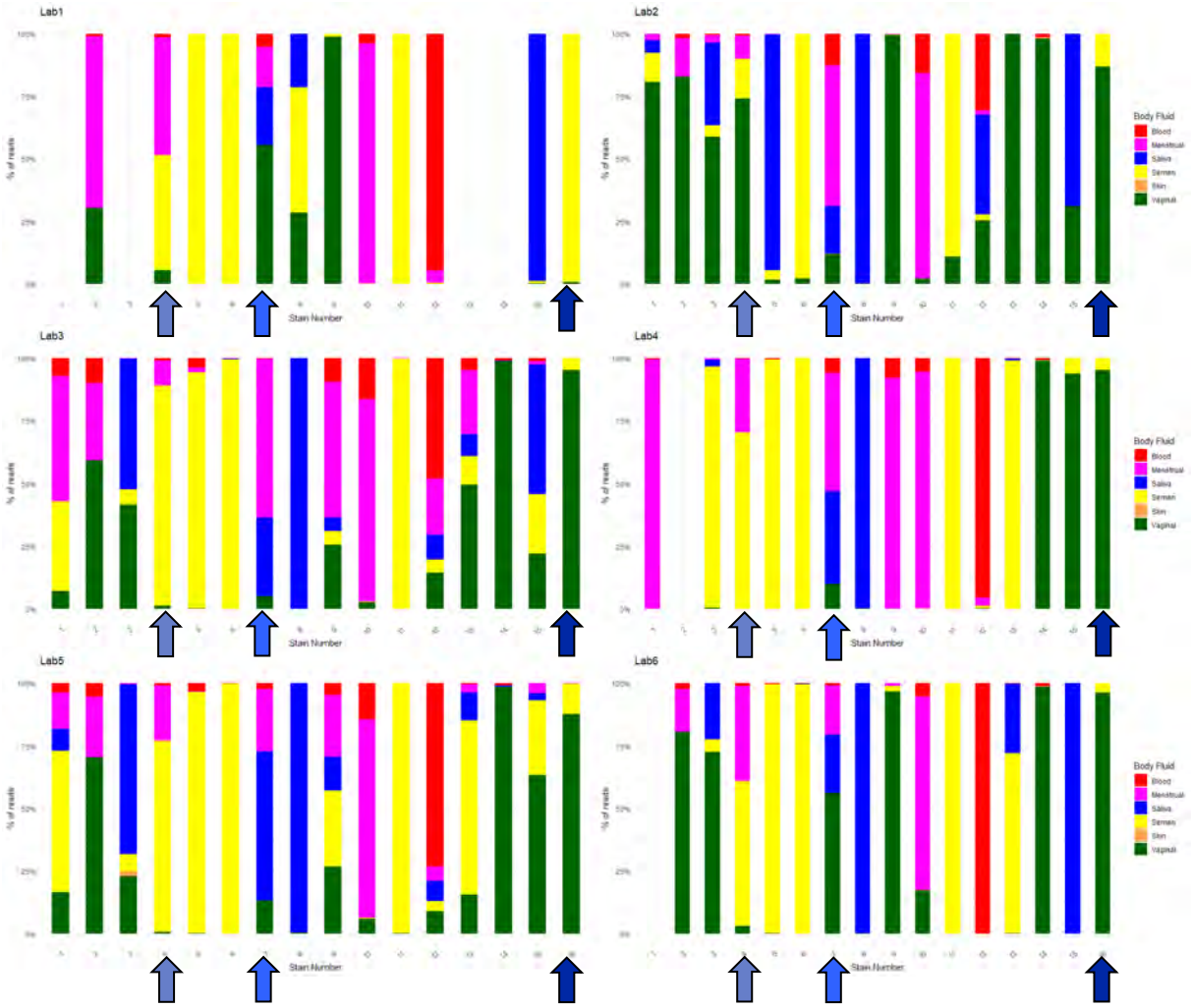
Institute of Forensic Sciences, DNA department, Bavarian State Criminal Police Office, Germany

Departement of Forensic Sciences, Oslo University Hospital, Norway

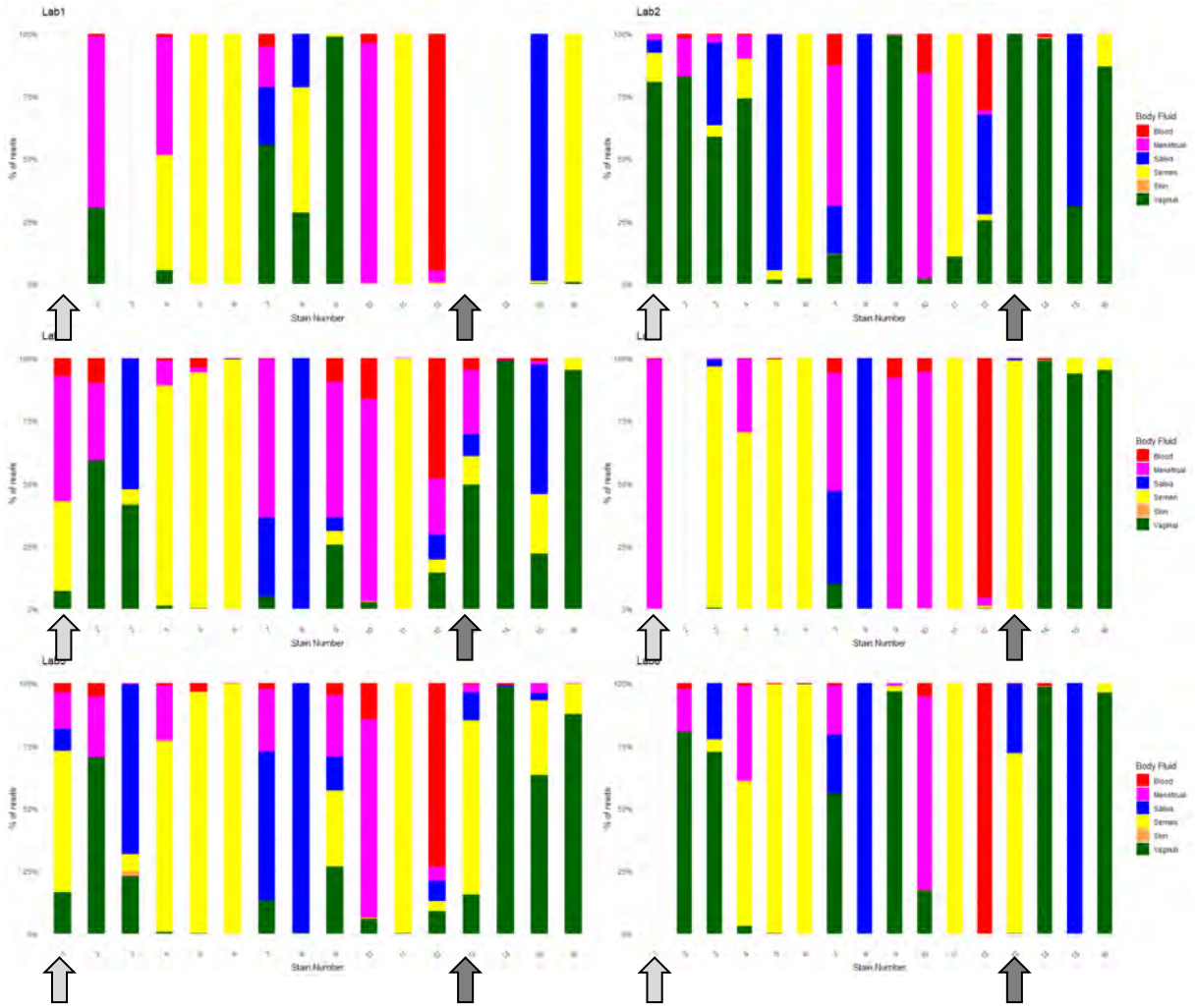
Institute of Legal Medicine, Innsbruck Medical University, Austria

Instituto Nacional de Medicina Legal, I.P., Ministry of Justice, Portugal

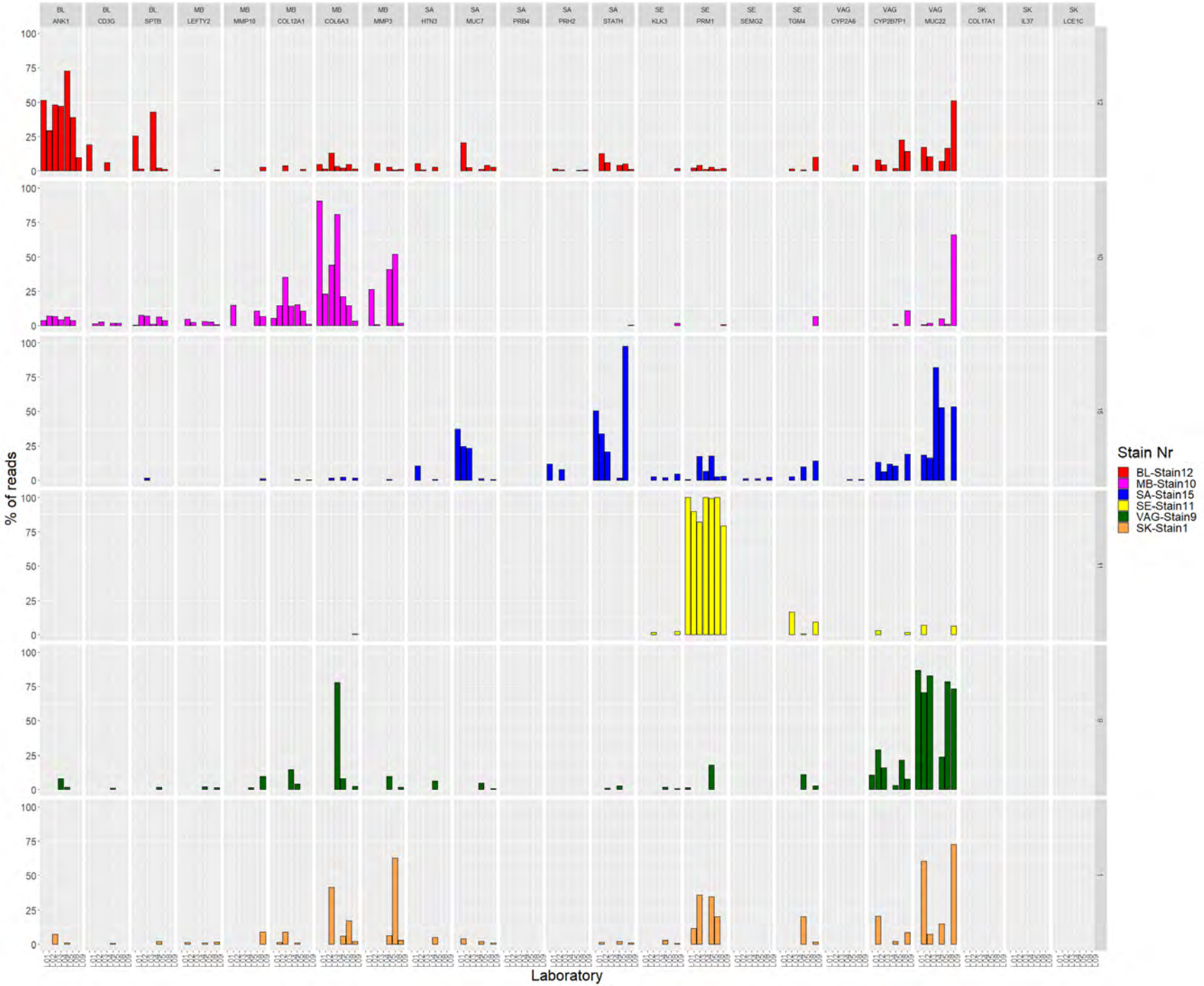
Composition Analysis of Stains by Body Fluid Percentages

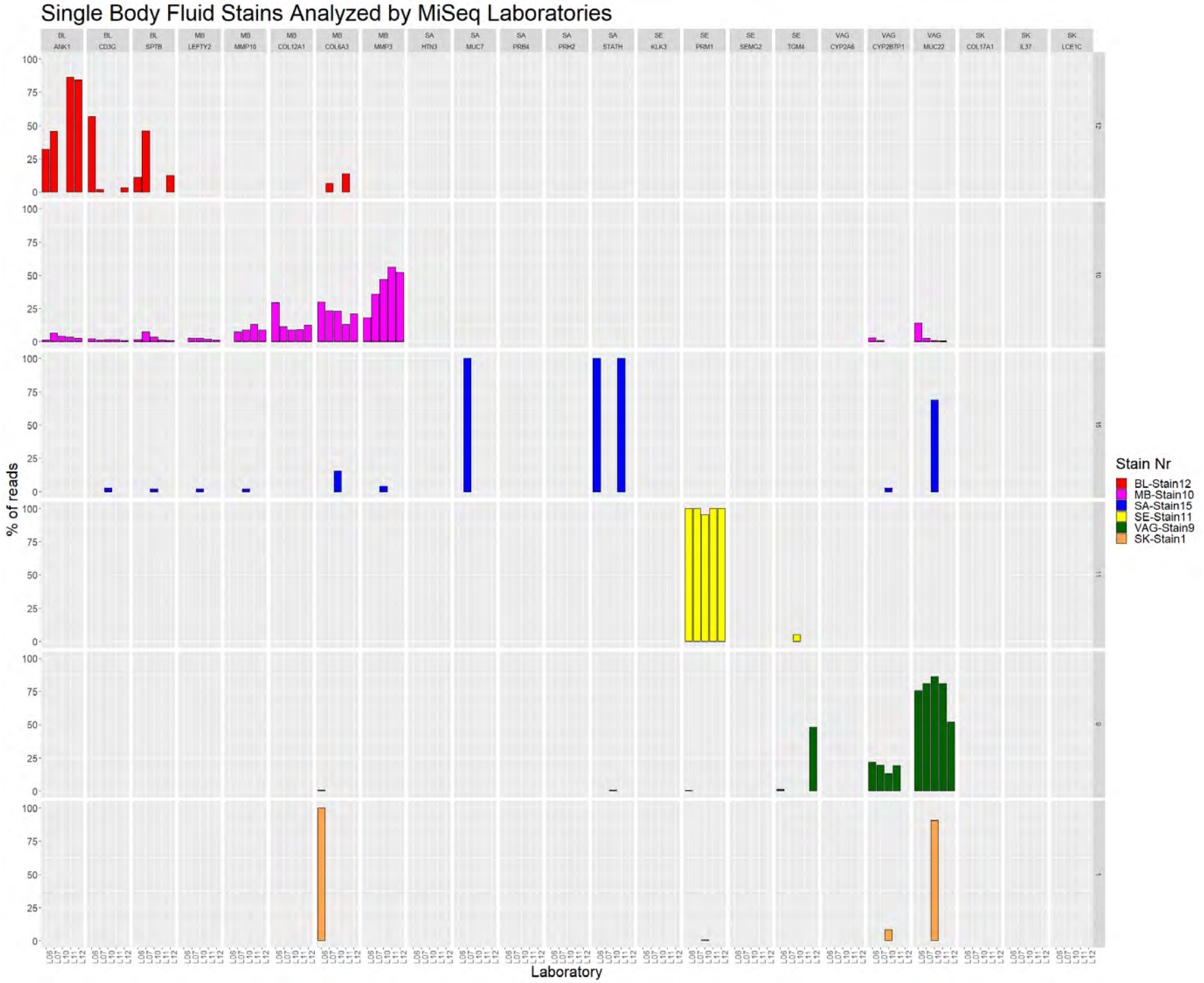


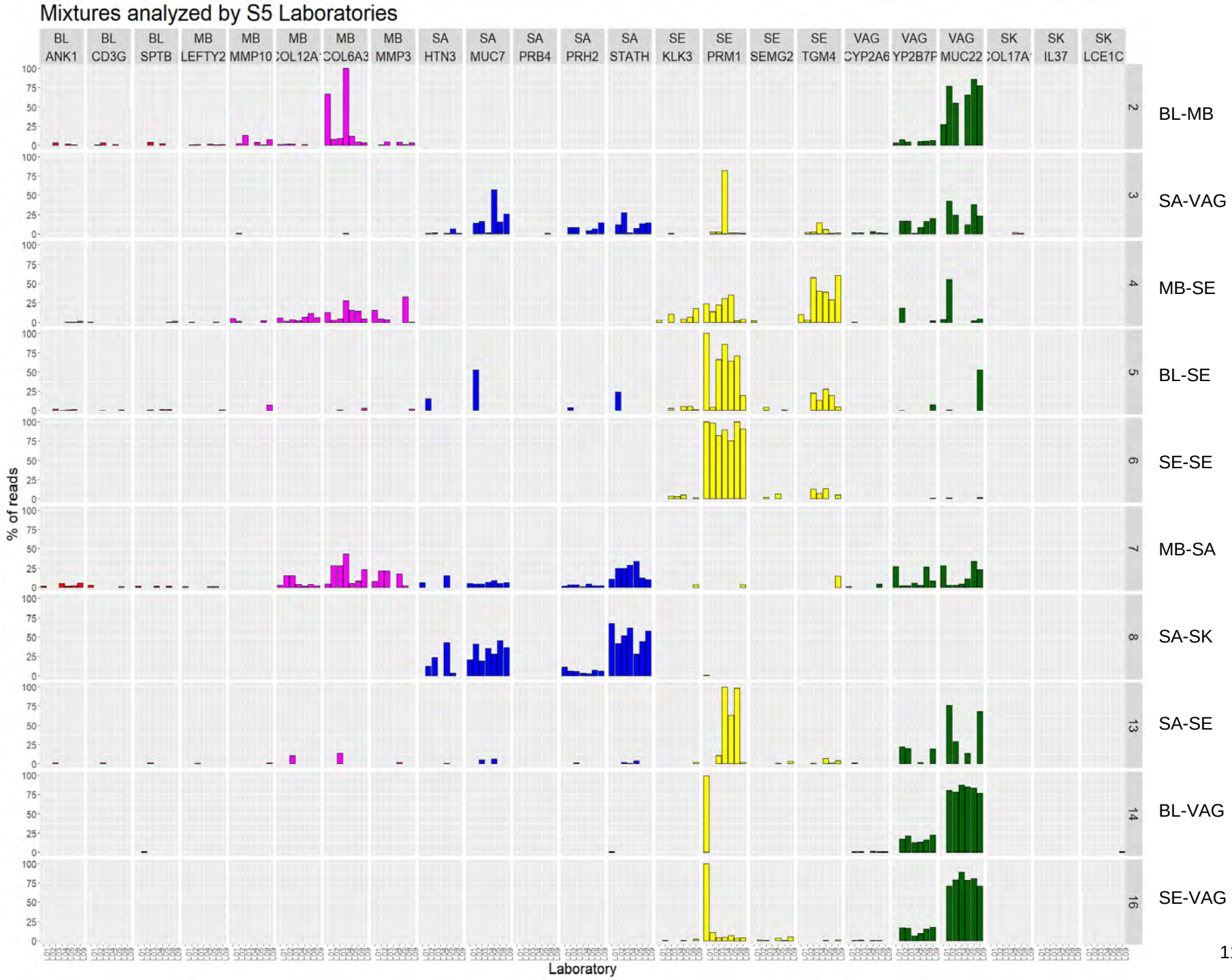
Composition Analysis of Stains by Body Fluid Percentages



Single Body Fluid Stains Analyzed by S5 Laboratories




Single Body Fluid Stains Analyzed by MiSeq Laboratories
BL ANK1
BL CD3G
BL SPTB
MB LEFTY2
MB MMP10
MB COL12A1
MB COL6A3
MB MMP3
SA HTN3
SA MUC7
SA PRB4
SA PRH2
SA STATH
SE KLK3
SE PRM1
SE SEMG2
SE TGM4
VAG CYP2A6
VAG CYP2B7P1
VAG MUC22
SK COL17A1
SK IL37
SK LCE1C
% of reads
Laboratory
Stain Nr
BL-Stain12
MB-Stain10
SA-Stain15
SE-Stain11
VAG-Stain9
SK-Stain1


Mixtures analyzed by S5 Laboratories
BL ANK1
BL CD3G
BL SPTB
MB LEFTY2
MB MMP10
MB COL12A1
MB COL6A3
MB MMP3
SA HTN3
SA MUC7
SA PRB4
SA PRH2
SA STATH
SE KLK3
SE PRM1
SE SEMG2
SE TGM4
VAG CYP2A6
VAG YP2B7P
VAG MUC22
SK COL17A1
SK IL37
SK LCE1C
% of reads
Laboratory
2
3
4
5
6
7
8
13
14
16
BL-MB
SA-VAG
MB-SE
BL-SE
SE-SE
MB-SA
SA-SK
SA-SE
BL-VAG
SE-VAG


Mixtures analyzed by MiSeq Laboratories
% of reads
Laboratory
BL-MB
SA-VAG
MB-SE
BL-SE
SE-SE
MB-SA
SA-SK
SA-SE
BL-VAG
SE-VAG

Association to Donors in Mixed Stains

Stain 3	HTN3	HTN3	HTN3	MUC7	PRB4	PRH2	HTN3	CYP2A6	MUC22	MUC22	MUC22	MUC22	MUC22	MUC22	MUC22	MUC22	MUC22	MUC22
SA-VAG	rs1849937_rs1136515	rs1849937	rs1136515	rs2306948	rs1052808	rs10772391	rs75067954	rs8192721	rs12110470_rs12110785	rs12110470	rs12110785	rs3869098_rs4248153	rs3869098	rs4248153	rs1419664_rs3094672	rs1419664	rs3094672	rs10947121
IonCode_135	TC	T/T	C/C	C/T	G/G	C/C	C/C	C/C	GT	G/G	T/T	AA/GA	A/G	A/A	CT/TA/TC	C/T	T/A	T/T
IonCode_136	CT/CC	C/C	T/C	C/C	G/C	C/C	C/C	C/C	TC/GT	T/G	C/T	AA/GG	A/G	A/G	CA/CC	C/C	A/C	T/C
Lab1 SS - Genotype																		
Lab1 SS - Read Counts																		
Lab2 SS - Genotype	TC/CT	T/C	C/T	C/T	G/G	C/C	C/C		TC/GT	T/G	C/T	CAA/CGG	A/G	A/G	ACA/ACT	C/C	A/T	C/T
Lab2 SS - Read Counts	204\51	204\51	204\51	2185\1911	15\15	2327\2327	204\204		672\600	672\600	672\600	5858\3628	5858\3628	5858\3628	1600\274	1600\1600	1600\274	672\600
Lab3 SS - Genotype	TC/TC	T/T	C/C	C/T		C/C	C/C		TC/GT	T/G	C/T	CGG/CAA	G/A	G/A	ACA	C/C	A/A	C/T
Lab3 SS - Read Counts	215		215\215	1724\1504		1632\1632	215\215		269\200	269\200	269\200	1700\1548	1700\1548	1700\1548	1065	1065\1065	1065\1065	269\200
Lab4 SS - Genotype																		
Lab4 SS - Read Counts																		
Lab5 SS - Genotype	TC/CC	T/C	C/C	T/C		C/C	C/C					CAA/CGG	A/G	A/G	ACT	C/C	T/T	
Lab5 SS - Read Counts	42\17\15	42\17	42\42	4405\3462		560\560	42\42					950\668	950\668	950\668	19	19\19	19\19	
Lab6 MiSeq - Genotype				C/T		C/C			TC/GT	T/G	C/T	CAA/CGG	A/G	A/G	ACA	C/C	A/A	C/T
Lab6 MiSeq - Read Counts				34\19		51\51			76\32	76\32	76\32	103\102	103\102	103\102	250	250\250	250\250	76\32
Lab7 MiSeq - Genotype	TC/TC	T/T	C/C	C/T	G/G	C/C	C/C		TC/GT	T/G	C/T	CAA/CGG	A/G	A/G	ACA	C/C	A/A	C/T
Lab7 MiSeq - Read Counts		14464\14464	14464\14464	21108\17201	2418\2418	8824\8824	14464\14464		8858\8081	8858\8081	8858\8081	27716\18812	27716\18812	27716\18812	30115	30115\30115	30115\30115	8858\8081
Lab8 SS - Genotype	TC/TC	T/T	C/C	C/T	G/G	C/C	C/C		TC/GT	T/G	C/T	CGG/CAA	G/A	G/A	ACA	C/C	A/A	C/T
Lab8 SS - Read Counts	33244	33244\33244	33244\33244	47296\36046	3064\3064	33033\33033	33244\33244		11954\10099	11954\10099	11954\10099	57295\51908	57295\51908	57295\51908	70883	70883\70883	70883\70883	11954\10099
Lab9 SS - Genotype	TC/TC	T/T	C/C	C/T	G/G	C/C	C/C		TC/GT	T/G	C/T	CAA/CGG	A/G	A/G	ACA	C/C	A/A	C/T
Lab9 SS - Read Counts	2536	2536\2536	2536\2536	68781\41249	298\298	60076\60076	2536\2536		6614\3587	6614\3587	6614\3587	30194\23648	30194\23648	30194\23648	36149	36149\36149	36149\36149	6614\3587
Lab10 MiSeq - Genotype	TC/TC	T/T	C/C	C/T	G/G	C/C	C/C		TC/GT	T/G	C/T	CAA/CGG	A/G	A/G	ACA	C/C	A/A	C/T
Lab10 MiSeq - Read Counts	311	311\311	311\311	1017\559	5\5	329\329	311\311		197\71	197\71	197\71	1740\1328	1740\1328	1740\1328	761	761\761	761\761	197\71
Lab11 MiSeq - Genotype	TC/TC	T/T	C/C	C/T	G/G	C/C	C/C		GT/TC	G/T	T/C	CAA/CGG	A/G	A/G	ACA	C/C	A/A	T/C
Lab11 MiSeq - Read Counts	2971	2971\2971	2971\2971	3854\3124	268\268	3107\3107	2971\2971		1978\1777	1978\1777	1978\1777	9956\9835	9956\9835	9956\9835	7131	7131\7131	7131\7131	1978\1777
Lab12 MiSeq - Genotype	TC/TC	T/T	C/C	C/T		C/C	C/C					CAA/CGG	A/G	A/G	ACA	C/C	A/A	
Lab12 MiSeq - Read Counts	36	36\36	36\36	120\110		15\15	36\36					106\69	106\69	106\69	14	14\14	14\14	

Stain 3 (SA-VAG):
- high number of reads
- RNA cSNP genotype mostly reflects donor genotypes

Stain 14	CYP2A6	MUC22.0	MUC22.1	MUC22.2	MUC22.3	MUC22.4	MUC22.5	MUC22.6	MUC22.7	MUC22.8	MUC22.9	ANK1.0	ANK1.1	CD3G	SPTB.0	SPTB.1	SPTB.2	SPTB.3	SPTB.4
VAG-BL	rs8192721	rs12110470_rs12110785	rs12110470	rs12110785	rs3869098_rs4248153	rs3869098	rs4248153	rs1419664_rs3094672	rs1419664	rs3094672	rs10947121	rs504574	rs7816734	rs3753059	rs1741488_rs1741487	rs1741488	rs1741487	rs229592	rs229586
IonCode_139	C/C	GT	T/T	T/T	AA/GA	A/G	A/A	CT/CA/CC	C/C	T/A	T/T	C/G	G/G	T/T	CA	C/C	A/A	A/A	C/T
IonCode_147	C/C	TC/TT	T/T	C/T	GG/AG	G/A	G/G	CA/TA/CC/TC	C/T	A/A	C/T	G/C	G/G	T/T	CA/TG	C/T	A/G	A/G	C/C
Lab1 SS - Genotype																			
Lab1 SS - Read Counts																			
Lab2 SS - Genotype		GT/GT	G/G	T/T	CAA/CGA	A/G	A/A	ACT/ACA	C/C	T/A	T/T	C/G	G/G	T/T	ATG/ACA	T/C	G/A	G/A	C/C
Lab2 SS - Read Counts		16600	16600\16600	16600\16600	76026\60510	76026\60510	76026\76026	38149\36457	38149\38149	38149\36457	16600\16600	530\375	868\868	413\413	244\228	244\228	244\228	304\60	1511\1511
Lab3 SS - Genotype		GT/GT	G/G	T/T	CAA/CGA	A/G	A/A	ACA/ACT	C/C	A/T	T/T	C/G	G/G	T/T	ATG/ACA	T/C	G/A	G/A	C/C
Lab3 SS - Read Counts		6370	6370\6370	6370\6370	25976\23531	25976\23531	25976\25976	18803\18118	18803\18803	18803\18118	6370\6370	84\18	421\421	326\326	82\20	82\20	82\20	31\11	312\312
Lab4 SS - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACA/ACT	C/C	A/T	T/T	C/G	G/G	T/T	ATG/ACA/AAG	T/C	G/A		C/C
Lab4 SS - Read Counts		39616	39616\39616	39616\39616	179565\170161	179565\170161	179565\179565	39416\35836	39416\39416	39416\35836	39616\39616	798\682	56\56	1055\1055	99\64\8	99\64	99\64		1604\1604
Lab5 SS - Genotype		GT/GT	G/G	T/T	CAA/CGA	A/G	A/A	ACT/ACA	C/C	T/A	T/T	C/C	G/G	T/T	ACA	C/C	A/A	G/G	C/C
Lab5 SS - Read Counts		39196	39196\39196	39196\39196	200156\191512	200156\191512	200156\200156	79575\73926	79575\79575	79575\73926	39196\39196	932\932	1226\1226	1665\1665	735	735\735	735\735	479\479	2034\2034
Lab6 MiSeq - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACA/ACT	C/C	A/T	T/T	G/C	G/G	T/T	ATG/ACA	T/C	G/A	G/A	C/T
Lab6 MiSeq - Read Counts		18646	18646\18646	18646\18646	30610\12528	30610\12528	30610\30610	17009\8256	17009\8256	17009\8256	18646\18646	7770\7770	1125\1125	14020\14020	1100	1100\1100	1100\1100	24\24	9180\9180
Lab7 MiSeq - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACT/ACA	C/C	T/A	T/T	G/C	G/G	T/T	ACA/ATG	C/T	A/G	G/A	C/C
Lab7 MiSeq - Read Counts		21248	21248\21248	21248\21248	77165\76895	77165\76895	77165\77165	46196\46069	46196\46196	46196\46069	21248\21248	299\175	274\274	485\485	177\22	177\22	177\22	21\19	588\588
Lab8 SS - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACT/ACA	C/C	T/A	T/T	C/G	G/G	T/T	ACA/ATG	C/T	A/G	A/A	C/C
Lab8 SS - Read Counts		36495	36495\36495	36495\36495	187101\175195	187101\175195	187101\187101	158341\153043	158341\158341	158341\153043	36495\36495	1263\261	189\189	2392\2392	248\231	248\231	248\231	253\253	661\661
Lab9 SS - Genotype		GT/GT	G/G	T/T	CAA/CGA	A/G	A/A	ACT/ACA	C/C	T/A	T/T	G/C	G/G	T/T	ACA	C/C	A/A	G/G	C/C
Lab9 SS - Read Counts		57011	57011\57011	57011\57011	154560\153950	154560\153950	154560\154560	157605\149433	157605\157605	157605\149433	57011\57011	378\364	550\550	1372\1372	301	301\301	301\301	181\181	2274\2274
Lab10 MiSeq - Genotype		GT/GT	G/G	T/T	CAA/CGA	A/G	A/A	ACA/ACT	C/C	A/T	T/T	G/C	G/G	T/T	ACA	C/C	A/A		C/C
Lab10 MiSeq - Read Counts		1106	1106\1106	1106\1106	8971\8630	8971\8630	8971\8971	3612\3557	3612\3612	3612\3557	1106\1106	16\16	26\26	49\49	6	6\6	6\6		24\24
Lab11 MiSeq - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACA/ACT	C/C	A/T	T/T	C/G	G/G	T/T	ATG/ACA	T/C	G/A	A/A	C/C
Lab11 MiSeq - Read Counts		8562	8562\8562	8562\8562	45238\45119	45238\45119	45238\45238	26448\24013	26448\26448	26448\24013	8562\8562	102\42	133\133	77\77	35\31	35\31	35\31	13\13	53\53
Lab12 MiSeq - Genotype		GT/GT	G/G	T/T	CGA/CAA	G/A	A/A	ACA/ACT	C/C	A/T	T/T	G/C	G/G	T/T	ACA	C/C	A/A	G/G	C/C
Lab12 MiSeq - Read Counts		1697	1697\1697	1697\1697	10233\9976	10233\9976	10233\10233	6486\4946	6486\6486	6486\4946	1697\1697	59\9	120\120	49\49	13	13\13	13\13	14\14	25\25

Stain 14 (VAG-BL):
- high number of reads in most markers
- RNA cSNP genotype reflects donor genotypes

Results BFID + Donor Association

Stain 1-16:

BFID

•11/16 stains were predicted correctly
0/2 low input stains correctly predicted

•5/16 stains could not be predicted
1/5 one body fluid was missing
1/5 skin generally difficult

•Difficulties arise because of various (misleading) reads in stains with low number of total reads

cSNPs

•Performance dependent on how many markers are detected per body fluid

Own Stains of the Laboratories:

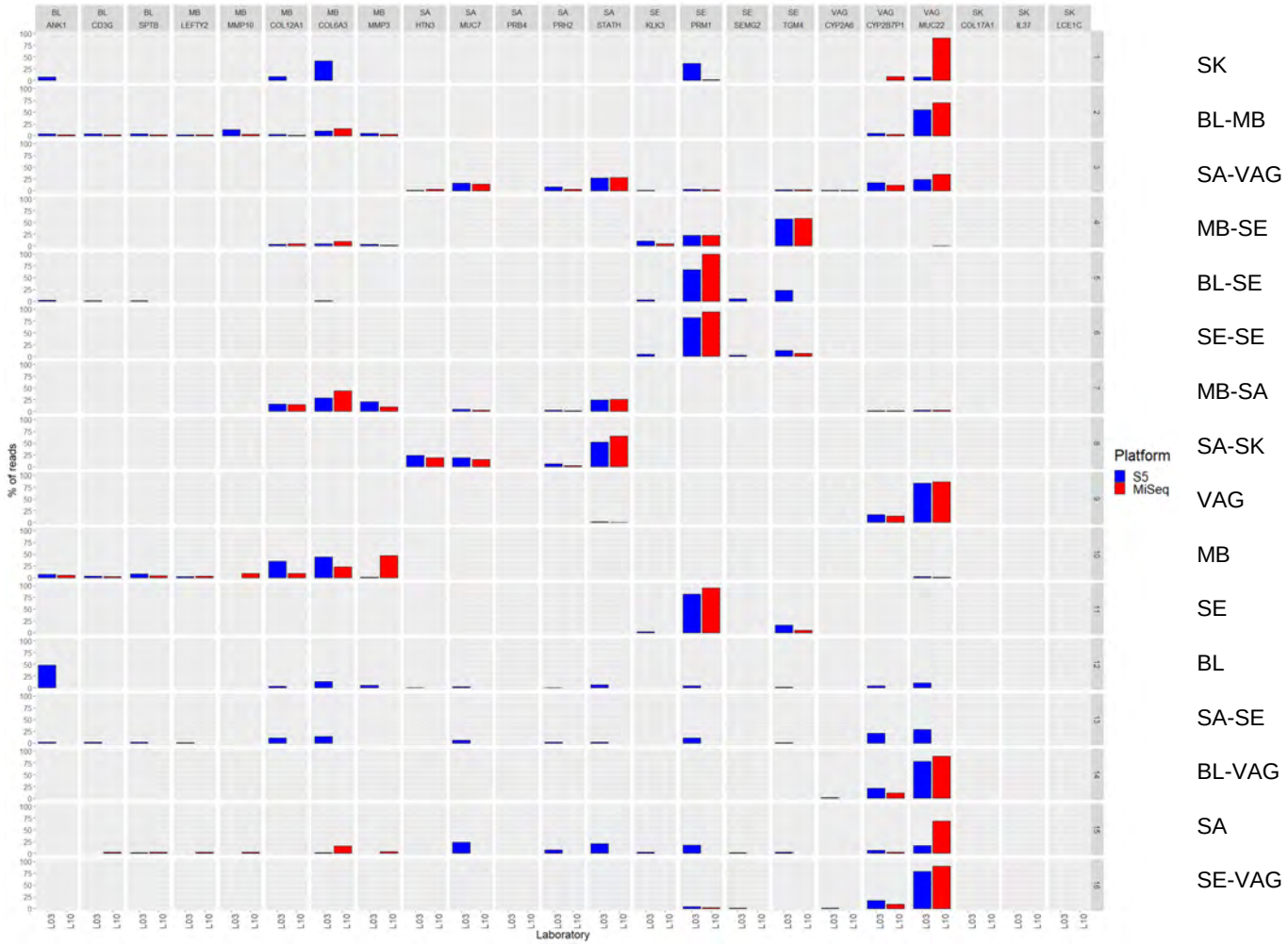
BFID

•Overall we could predict 41/62 stains (74%)

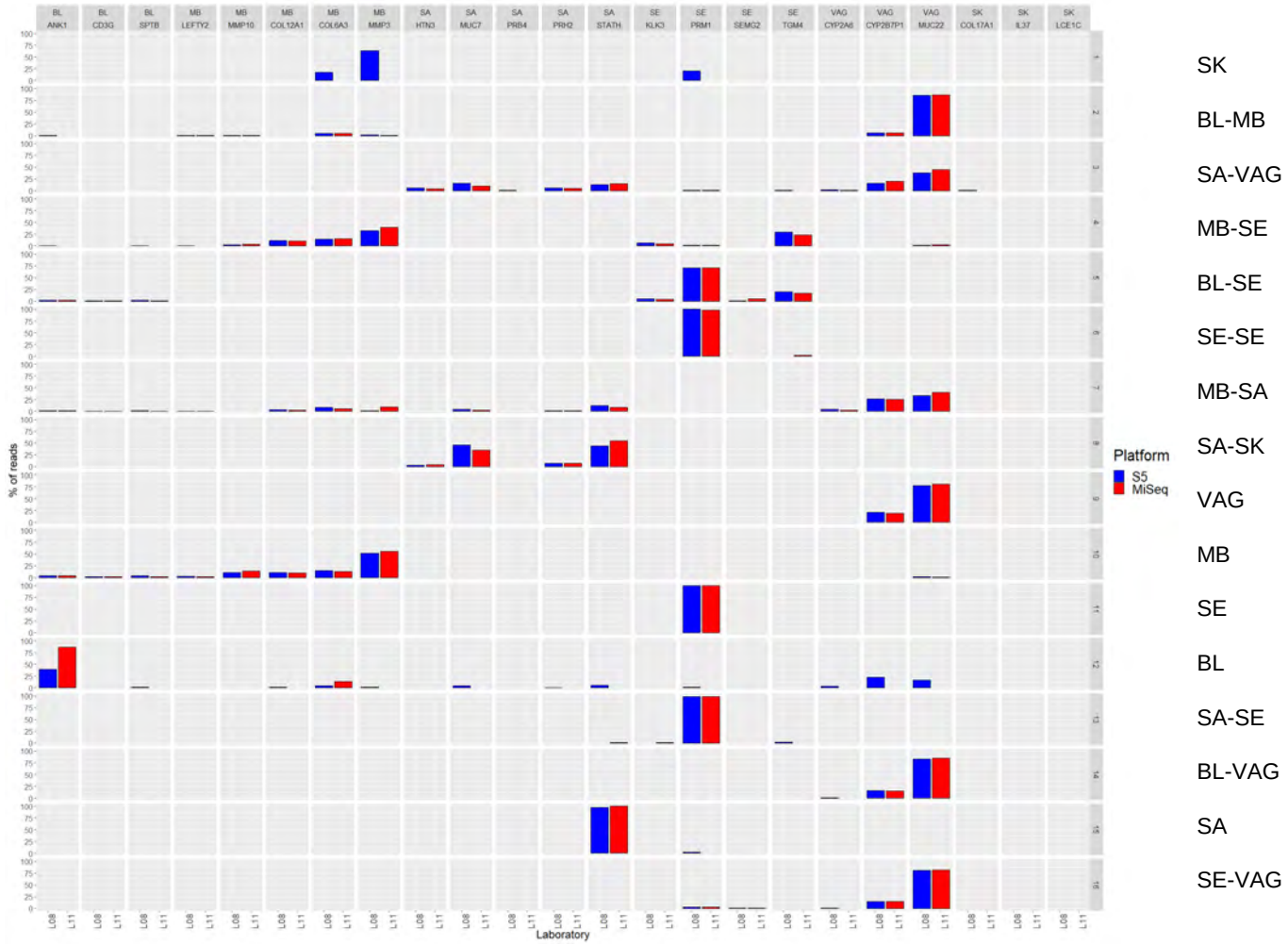
cSNPs

•Performance dependent on how many reads per RNA cSNP were detected
→ the more, the more accurate/complete the reflection of DNA genotypes

Comparison of Sequencing Platforms (Lab1)



Comparison of Sequencing Platforms (Lab2)

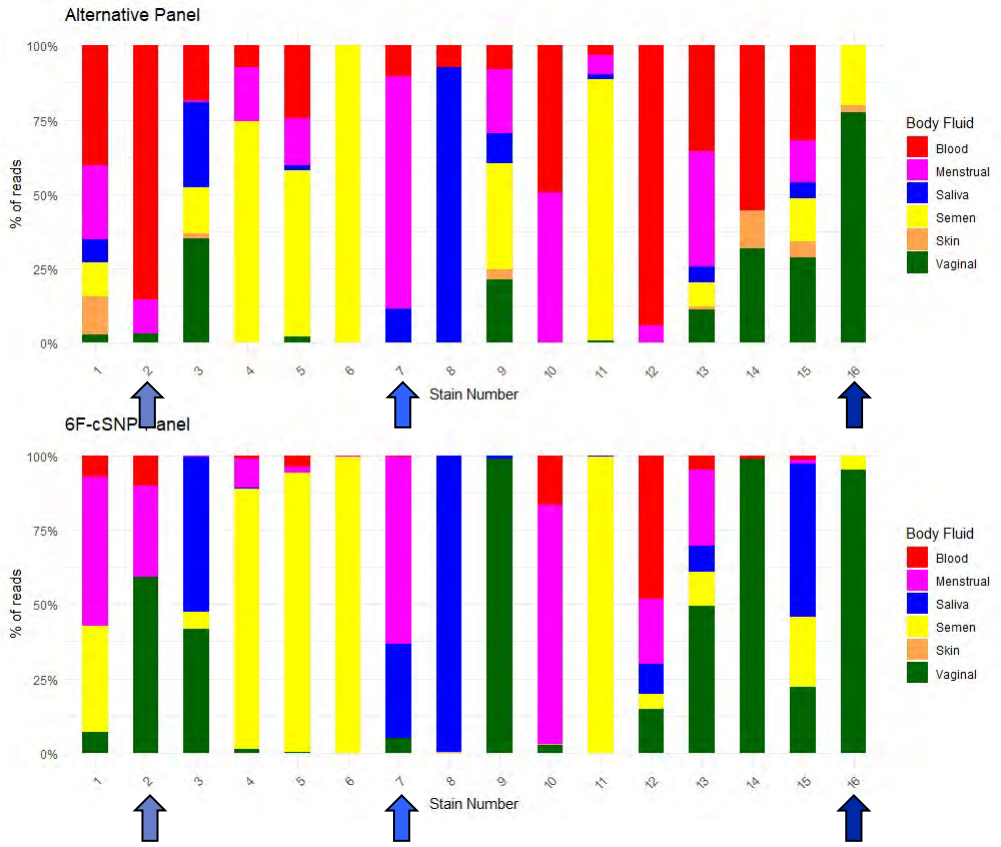


Evaluating an Alternative cSNP Panel (Cologne)

- Target amplification from the same cDNA

Stain Nr.	BF/T	6F-cSNP-Panel	alternative cSNP Panel
1	SK	?	?
2	BL-MB	BL-MB	BL-MB
3	SA-VAG	VAG-SA (SE in VAG?)	?
4	MB-SE	MB-SE	MB-SE
5	BL-SE	SE-BL	?
6	SE-SE	SE-SE	SE-SE
7	MB-SA	MB-SA	MB-SA
8	SA-SK	SA	SA-BL
9	VAG	VAG	?
10	MB	MB	MB
11	SE	SE	SE
12	BL	?	MB
13	SA-SE	?	?
14	VAG-BL	VAG	VAG-BL
15	SA	SA(?)	?
16	VAG-SE	VAG-SE	VAG-SE

Composition Analysis of Stains by Body Fluid Percentages

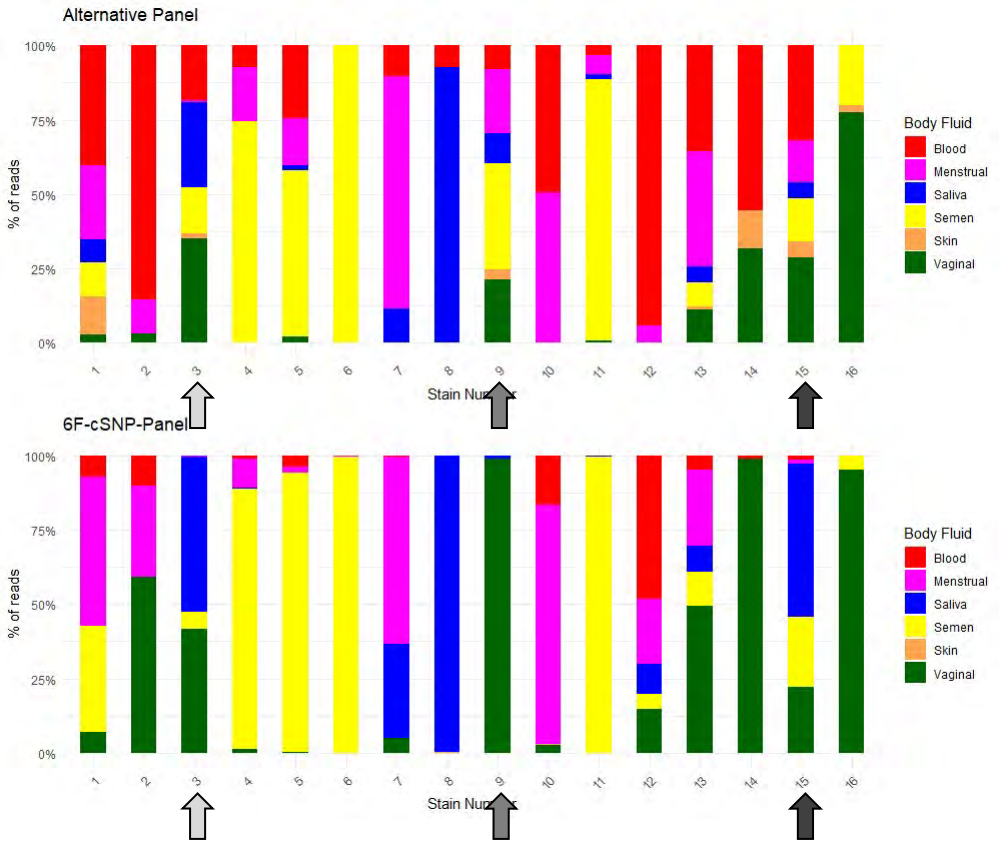


Evaluating an Alternative cSNP Panel (Cologne)

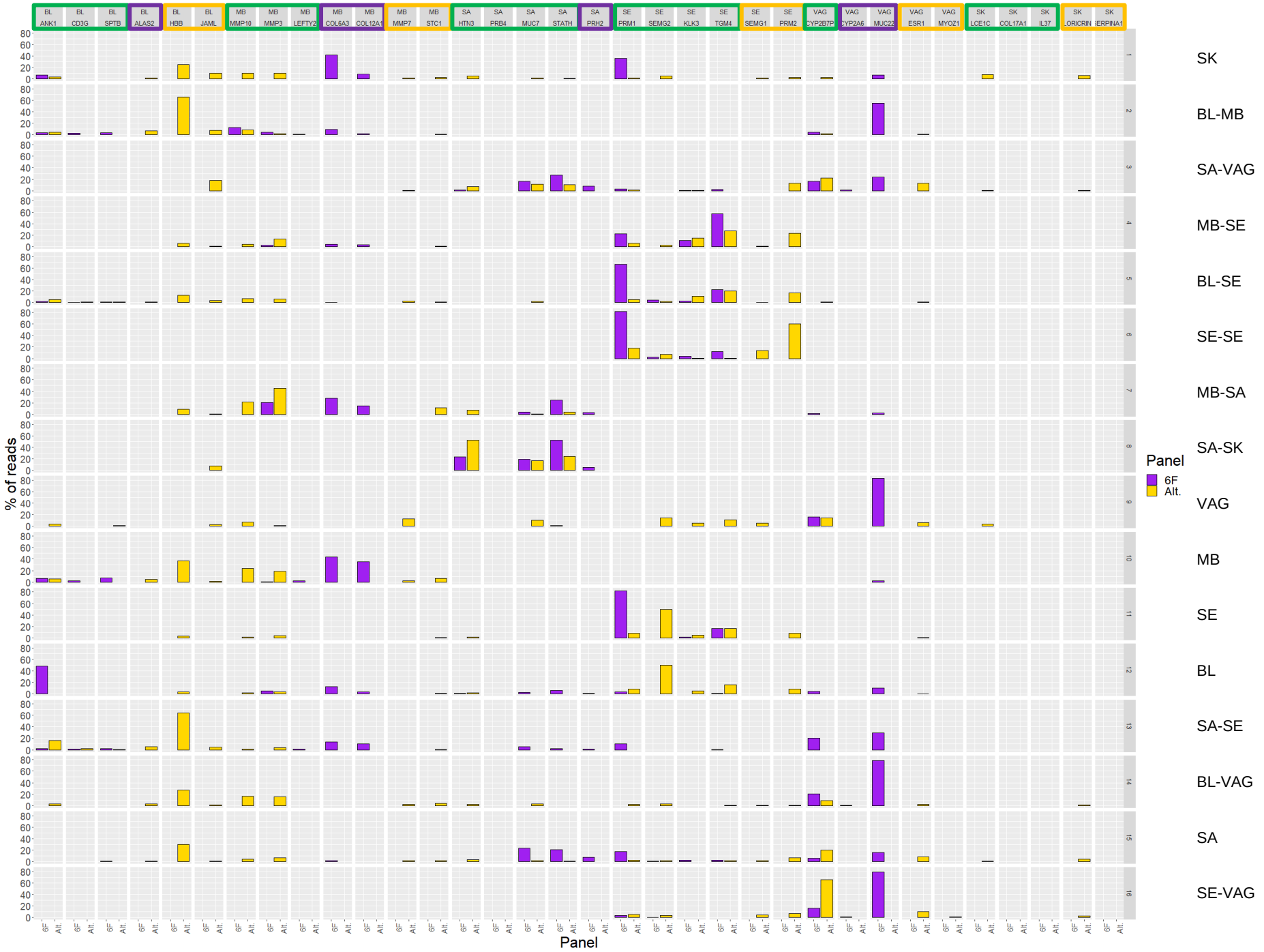
- Target amplification from the same cDNA

Stain Nr.	BF/T	6F-cSNP-Panel	alternative cSNP Panel
1	SK	?	?
2	BL-MB	BL-MB	BL-MB
3	SA-VAG	VAG-SA (SE in VAG?)	?
4	MB-SE	MB-SE	MB-SE
5	BL-SE	SE-BL	?
6	SE-SE	SE-SE	SE-SE
7	MB-SA	MB-SA	MB-SA
8	SA-SK	SA	SA-BL
9	VAG	VAG	?
10	MB	MB	MB
11	SE	SE	SE
12	BL	?	MB
13	SA-SE	?	?
14	VAG-BL	VAG	VAG-BL
15	SA	SA(?)	?
16	VAG-SE	VAG-SE	VAG-SE

Composition Analysis of Stains by Body Fluid Percentages



Comparison of the 2 panels



Summary

- Overall promising results
- Some participants did not (fully) follow the recommendations
- Laboratories with limited RNA experience also achieved good results
- Results were quite consistent across different laboratories
- The cSNP panels performed well on both sequencing platforms
- Comparison with Cologne cSNP panel (31 body fluid markers, 80 cSNPs)

➢ Manuscript draft will be circulated among participants soon

➢ Poster presentation at 30. ISFG congress

Acknowledgements

University of Zurich:
Research team

Nadescha Hänggi

University of Central Florida:
Jack Ballantyne, Erin Hanson

Thermofisher:
Robert Lagace, Chantal Roth

Email from Roland van Oorshot, 27 May 2024

Hi All

A brief update:

Many of you submitted the requested data some time ago, others only submitted them very recently. We held off commencing collation and analyses of the submitted data multiple times at the request of some potential committed laboratories in anticipation of additional data. We gratefully received additional data from some of these laboratories, while some were ultimately, due to various understandable circumstance, unable to deliver. We appreciate their attempts to provide the intended data and understand the difficulties they encountered that prevented submission. So, it is only very recently that we have the final set of submitted data to work with.

Received:

- Submissions from 18 labs.
- 16 labs submitted data for tool handles and gloves; 1 lab submitted data for tool handles only; 1 lab submitted data for gloves only.
- Data relating to a total of 1427 tool handle samples: Average 84 per lab; 10 labs 100-130; 5 labs 50-99; 2 labs 10-49
- Data relating to a total of 1357 glove samples: Average 80 per lab; 9 labs 100-160; 4 labs 50-99; 4 labs 5-49.

It was our intent to now commence collation and analyses of the submitted data. However, due to serious health issues being experiences by one of our team (Bianca) there will be a few months delay in progressing this.

Niels / Bo Simonsen, you are welcome to relay this information during the upcoming EDNAP meeting and/or add to the meeting's notes/minutes.

Kind regards
Bianca, Bas & Roland

THE SERIES OF EXERCISES RELATING TO DNA TRANSFER

DATA SUBMITTED FROM 18 LABORATORIES

	LABORATORIES
TOOL HANDLES AND GLOVES	16
TOOL HANDLES ONLY	1
GLOVES ONLY	1

	SAMPLES
TOOL HANDLE SAMPLES	1,427
AVERAGE PER LAB	84
GLOVE SAMPLES	1,357
AVERAGE PER LAB	80

A comparison of CE and MPS using typical ‘trace’ DNA

From a paper by: Maria Martin Agudo[1,2*], Chiara Fantinato[1,2], Arne Roseth[1], Håvard Aanes[1], Peter Gill[1,2], Ane Elida Fonneløp[1] , Øyvind Bleka[1].

[1]Department of Forensic Sciences, Oslo University Hospital, Oslo, Norway

[2]Department of Forensic Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

A Comparison of Likelihood Ratios Calculated from Surface DNA Mixtures Using Mps and CE Technologies

Abstract

This study evaluates the performance of analyzing surface DNA samples using massively parallel sequencing (MPS) compared to traditional capillary electrophoresis (CE). A total of 30 samples were collected from various surfaces in an office environment and were analyzed with CE and MPS. These were compared against 60 reference samples (office inhabitants). To identify contributors, likelihood ratios (LRs) were calculated for MPS and CE data using the probabilistic genotyping software MPSproto and EuroForMix respectively. Although a higher number of sequences/peaks were observed per DNA profile in MPS compared to CE, LR values were found to be lower for MPS data formats. This might be the result of the increased complexity of MPS data, along with a possible elevation of unknown alleles and/or artefacts. The study highlights avenues for improving MPS data quality and analysis to facilitate more robust interpretation of challenging casework-like samples.

Preprint available:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4715225

Sample collection

- We used samples collected from our institute as part of an eDNA project
  - Dust samples from uncleaned surfaces such as door ledges
  - Swabs from surfaces such as table tops, light switches etc
- Reference samples collected from staff members
- EuroForMix was used to calculate CE based likelihood ratios (LRs) of mixtures, conditioned on staff members
- An extension of EuroForMix called MPSProto (also open source) was used to analyse MPS based LRs

METHODS

Study design

- Indoor premises at Oslo University Hospital:

  14 offices, **2** meeting rooms, **5** laboratories.

- Reference database: **55** of **64** employees.

- DNA profiles compared to reference of participants, considering their occupancy status.

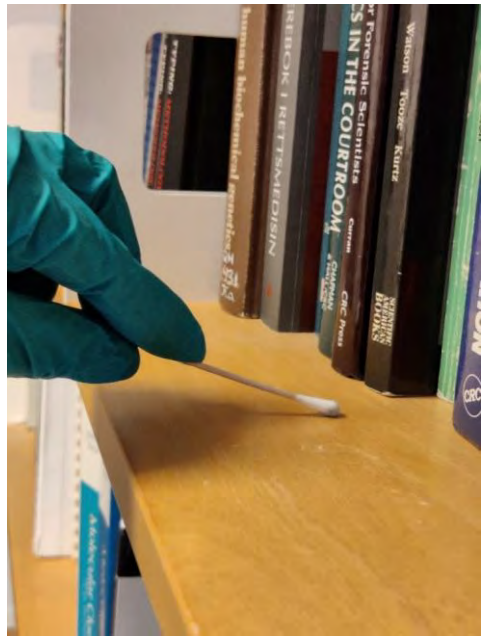
METHODS

Samples collection

40 air samples and **144** dust samples.

AirPrep ACD220 electret filter air sampler (Innovaprep®).

Time after occupancy: **0 h** and **16 h**. **2 h** collection time.



AirPrep ACD220 (Innovaprep®)

Dust samples: moistened cotton swabs from **undisturbed surfaces**.

eDNA method described by Chiara Fantinato

scientific reports

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific reports › articles › article

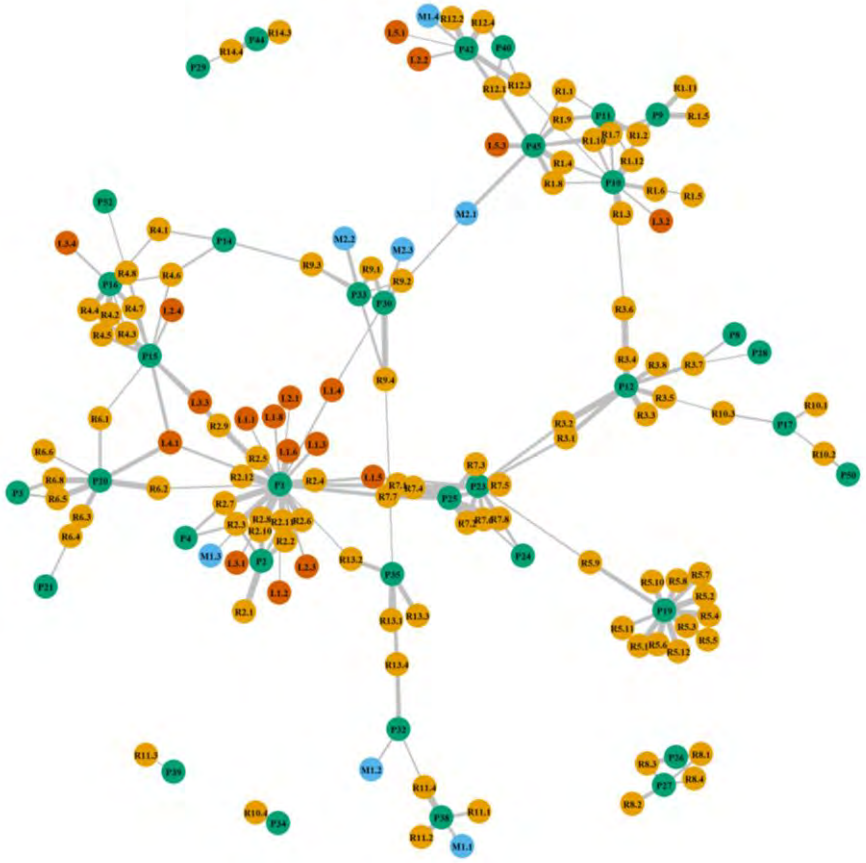
Article | Open access | Published: 04 November 2023

The invisible witness: air and dust as DNA evidence of human occupancy in indoor premises

Chiara Fantinato ✉, Ane Elida Fonneløp, Øyvind Bleka, Magnus Dehli Vigeland & Peter Gill

Scientific Reports **13**, Article number: 19059 (2023) | Cite this article

Match network dust samples



Motivation for CE/MPS comparison

- Traditional comparative studies are often based upon contrived samples where DNA extracts (or body fluids) are simply mixed together.
  - Advantage: The contributors are known; proportions are known
  - Disadvantage: Does not simulate typical case samples
- The method employed in this study utilises samples that are realistically encountered in casework e.g. a burglary
  - Advantage: Simulates casework
  - Disadvantage: Ground truth is not strictly known
- However, there is a high probability that DNA from office inhabitants will be recovered
- High LRs achieved with dnamatch2 gives us the rationale to infer contributor identity $Pr(H|E)$ since we show prior probability c. 0.5

Methods compared

- traditional **CE- based with CE nomenclature**
- **MPS-based with flanks**
- **MPS-based without flanks**
- **MPS-based converted to CE nomenclature** (to allow direct comparison - any differences are down to method)

Allele formats

Marker TH01

Traditional CE	Allele 6
MPS with flanks	5´ TGCAGGTCACAGGGAACACAGACTCCATGTTG [AATG]6 AGGGAAATAAGG 3´
MPS without flanks	[AATG]6
MPS-CE	Allele 6

MPS analysis

- Sequencing libraries were prepared from 30 surface DNA samples and screened against staff database
- POIs were identified in 27 of these samples, using software dnamatch2
  - dnamatch2 carries out rapid comparison of large databases of reference samples and casework samples to identify potential contributors

What are the challenges of interpreting MPS mixtures

- The majority of samples analysed are mixtures. Therefore, to maximise the benefits of MPS, it is necessary to utilise software that is able to interpret mixtures
- MPSproto is based on EuroForMix, but there are important differences.
  - Stutter detection is sequence-based i.e. multiple stutters are possible (not necessarily based on LUS)
  - Noise model replaces drop-in model and this enables us to reduce the analytical threshold – hence we retrieve more information

Noise sequences - details

- Sequences are similar to true alleles or stutters, but they are the consequence of molecular errors, such as insertions, deletions or substitutions that are introduced during PCR

Marker TH01	Length-Based	Reads	Sequence
Allele	**6**	1796	TGCAGGTCAC**A**GGGAACACAGACTCCATGGTG**AATGA** **ATGAATGAATGAATGAATG**AGGGAAATAAGG
Stutter	**5**	117	TGCAGGTCACAGGGAACACAGACTCCATGGTG**AATGA** **ATGAATGAATGAATG**AGGGAAATAAGG
PCR/Sequencing error	**6**	11	TGCAGGTCAC**G**GGGAACACAGACTCCATGGTG**AATGA** **ATGAATGAATGAATGAATG**AGGGAAATAAGG

Noise sequences - details

- Sequences are similar to true alleles or stutters, but they are the consequence of molecular errors, such as insertions, deletions or substitutions that are introduced during PCR

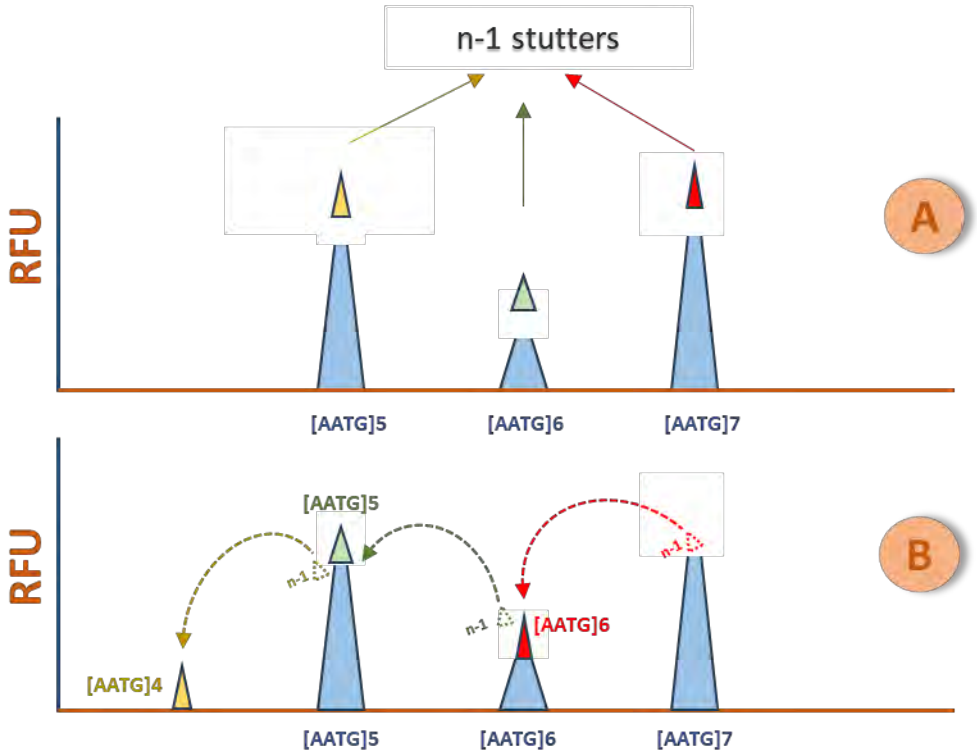
Marker TH01	**Length-Based**	**Reads**	**Sequence**
Allele	**6**	1796	TGCAGGTCAC**A**GGGAACACAGACTCCATGGTG**AATGAATG** **AATGAATGAATGAATG**AGGGAAATAAGG
Stutter	**5**	117	TGCAGGTCACAGGGAACACAGACTCCATGGTG**AATGAATG** **AATGAATGAATG**AGGGAAATAAGG
PCR/Sequencing error	**6**	11	TGCAGGTCAC**G**GGGAACACAGACTCCATGGTG**AATGAATG** **AATGAATGAATGAATG**AGGGAAATAAGG

Noise sequences

- Noise sequences are typically accommodated by increasing the threshold read number
- But this has the effect of eliminating low level sequences from true alleles, and is therefore inefficient.
- MPSproto models noise sequence so that the threshold can be minimised (we use 11 reads)
  - There is one parameter for noise frequency and one for noise size

Stutters

- Euroformix model was adapted for MPSproto
- MPSproto models n+1, n+2, n-1 n-2, n0 stutters
- Stutters do not always originate from LUS, and this is an additional layer of complexity to model



Analytical settings for LR calculations

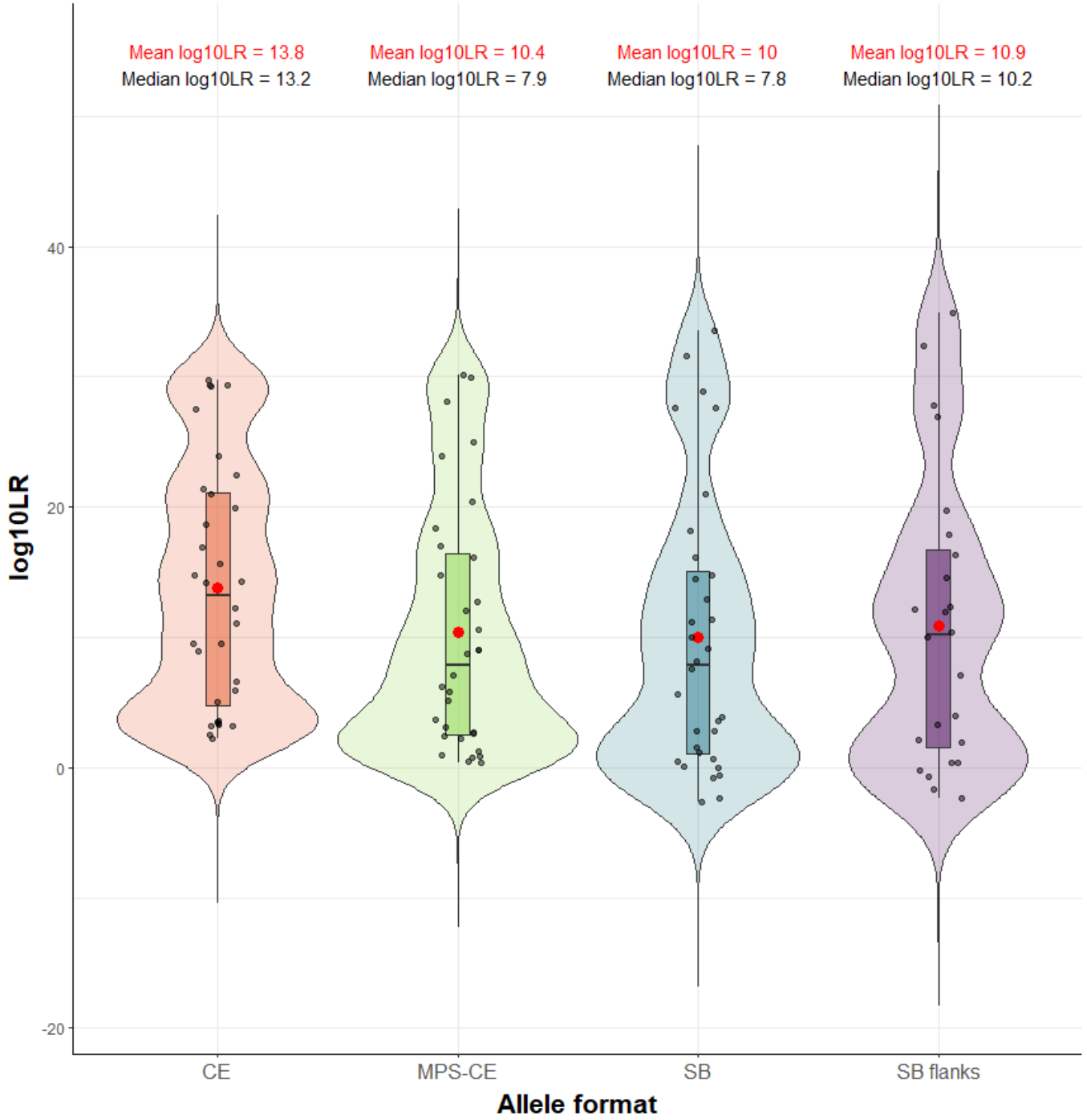
Settings	ShinyRFU (EFM)		MPSproto
	CE data (RFU)	**MPS data (reads)**	**MPS data (reads)**
Allele format	CE format	MPS-CE	SB and SB flanks
Analytical threshold (AT)	50	11	11
Fst	0.01	0.01	0.01
Degradation	Yes	Yes	Yes
Drop-in	λ = 0.05	λ= 0.05	Noise model
Stutter model	EFM (BW and FW)	EFM (BW and FW)	Pre-calibrated
Number of optimisations (MLE)	3	3	1

Results

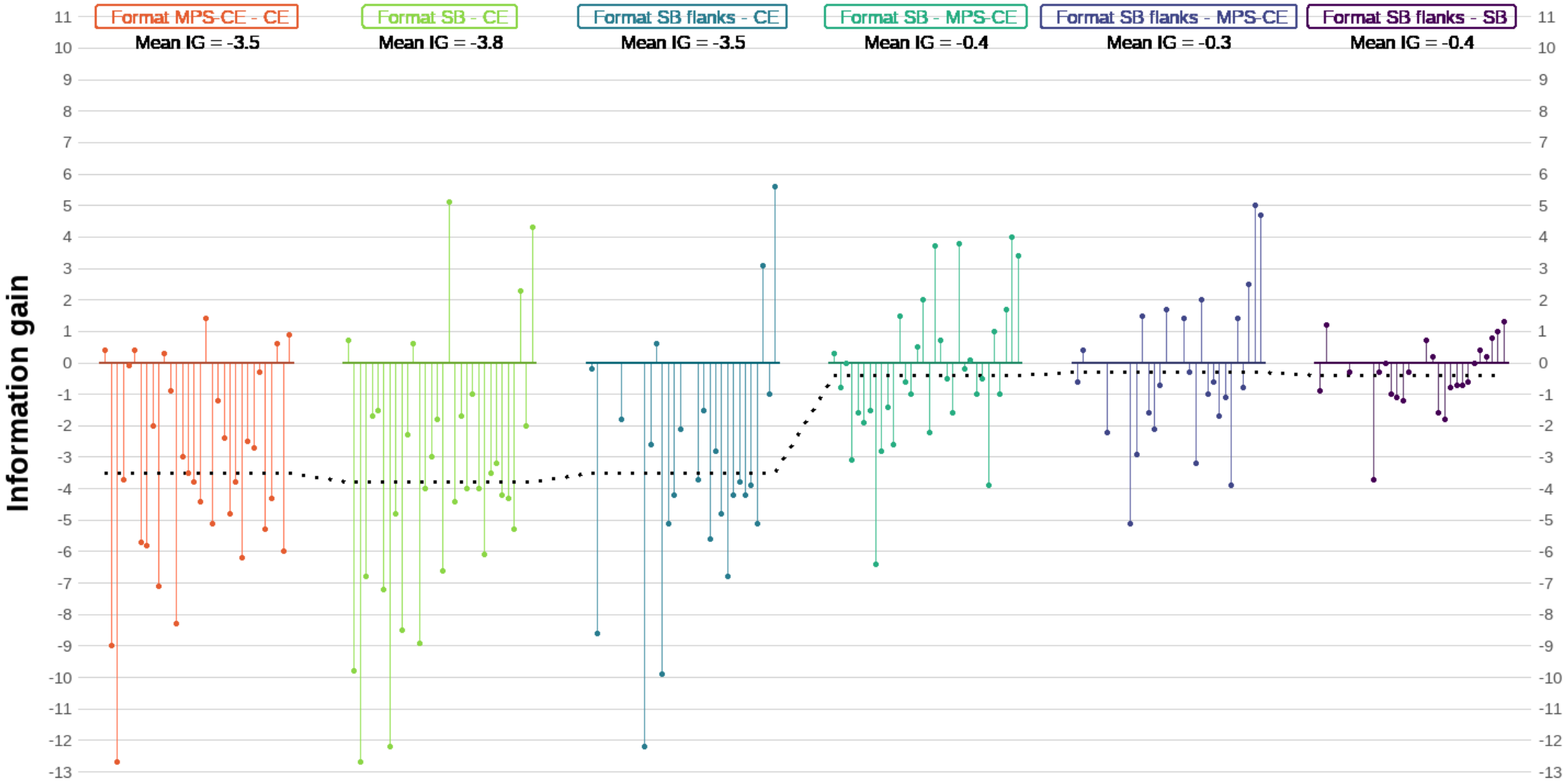
LR results for 4 different methods

Red dot=mean value

Note CE performs best



Information gain (log10)



Real casework: Birgitte Tengs case MPS vs CE YSTR tests

YDNA site examined	576	389 I	448	389 II	19	391	481	549	533	438	437	570	635	390	439	392	643	393	458	385	456	YGA	627	460	518	449	387s1	505	612	461
Y sites #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
REFERENCE BC20001445																														
BC20001445																														
PowerPlex Y23	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	not tested	not tested	not tested	not tested	not tested	not tested	not tested	not tested
YFilerPlus	R	R	R	R	R	R	R	not tested	R	R	R	R	R	R	R	R	not tested	R	R	R	R	R	R	R	R	R	R	not tested	not tested	not tested
Combined profile: YDNA	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	not tested	not tested	not tested
MPS	R					R			R			R	R		R	R	R			R		R		R				R	R	R
TIGHTS A-12-F																														
GMI																														
YFilerPlus	R	R	R	R	R	R	R	not tested	F	R	R	R	R	R	R	R	not tested	R	R	R	R	R	R	R	R	R	R			
PowerPlex Y23 (test 1)	R	R	R	R	R	R	R	R	R	R	F	R	R	F	R	R	F	R	R	R	R	R	not tested	not tested	not tested	not tested	not tested	not tested	not tested	not tested
PowerPlex Y23 (test 2)	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	F	R	R	R	R	R	not tested	not tested	not tested	not tested	not tested	not tested	not tested	not tested
Combined profile: YDNA	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	F	R	R	R	R	R	R	R	R	R	F	not tested	not tested	not tested
MPS above threshold	R					R						R				R	R					R		R				R	R	R
MPS below threshold									R				R		R					R										

nr = no result obtained

F= failure to produce a result

MPS = massively parallel sequencing

[one observation]

Generalised reasons for poor MPS performance

- The main advantage of MPS is allele diversity which should increase discriminating power
- Unfortunately, this is negated by several factors
  - More complex stutters.
  - Larger stutter ratios with lower DNA inputs exaggerated with MPS
  - Note that a priori we cannot distinguish between true alleles and stutter, hence their presence will always reduce the LR
  - Sequence errors observed more frequently than with CE
  - Higher imbalance of loci – both inter and intra loci
  - Differences between batches
  - Attempting to negate effects by using a high analytical threshold negates the likelihood ratio because information is ignored

Conclusion

- On one hand MPS increases the amount of genetic diversity to be detected which drives up the information content (or higher LRs if Hp is true).
- On the other hand, for poor quality samples (mixtures), MPS leads to greater numbers of artefacts which drives the information content downward in the opposite direction.
- Coupled with the lower sensitivity of MPS, this means that CE ultimately provides higher information content for compromised, low quality samples.
- The poor results obtained with MPS, indicate a need to improve the biochemistry to a) increase sensitivity b) to decrease the artefact sequences.
- Whereas this research has indicated that model improvement is also needed, there are limits to the extent that statistics can accommodate the negative outcomes of poor biochemistry.
- However, modelling sequences with software like MPSproto will need to be mainstream before MPS can reach its full potential – and compete with CE as method of choice in routine casework
- Expectations related to MPS need to be managed

Acknowledgements

- Maria Martin Agudo (PhD thesis)
- Chiara Fantinato
- Arne Roseth
- Håvard Aanes
- Ane Elida Fonneløp
- Øyvind Bleka

CERTAIN-FORS Project
"Competency, Education, Research, Testing, Accreditation, and Innovation in Forensic Scienc
ISFP-2020-AG-IBA-ENFSI

FOR-FUTURE Project
"Forensic Fundamentals, Technology, Multidisciplinarity, Research, Evaluation"
ISF-2023-TF2-AG-ENFSI-IBA-2

The ReAct Project: An update

Peter Gill

This project was funded by the European Union's Internal Security Fund — Police. The content of this presentation represents the views of the author only and is his/her sole responsibility. The European Commission does not accept any responsibility for use that may be made of the information it contains.

ReAct

Aims

- To collect data from 23 laboratories for a number of experimental designs which simulate common casework requests for help
- To analyse the data, a comparative study between laboratories and between experiments
- To prepare open-source software to analyse the data using Bayesian Networks
- To make non-sensitive data available for further study
- To gain knowledge how to coordinate large collaborative exercises

Outline of case circumstances (experiment 2)

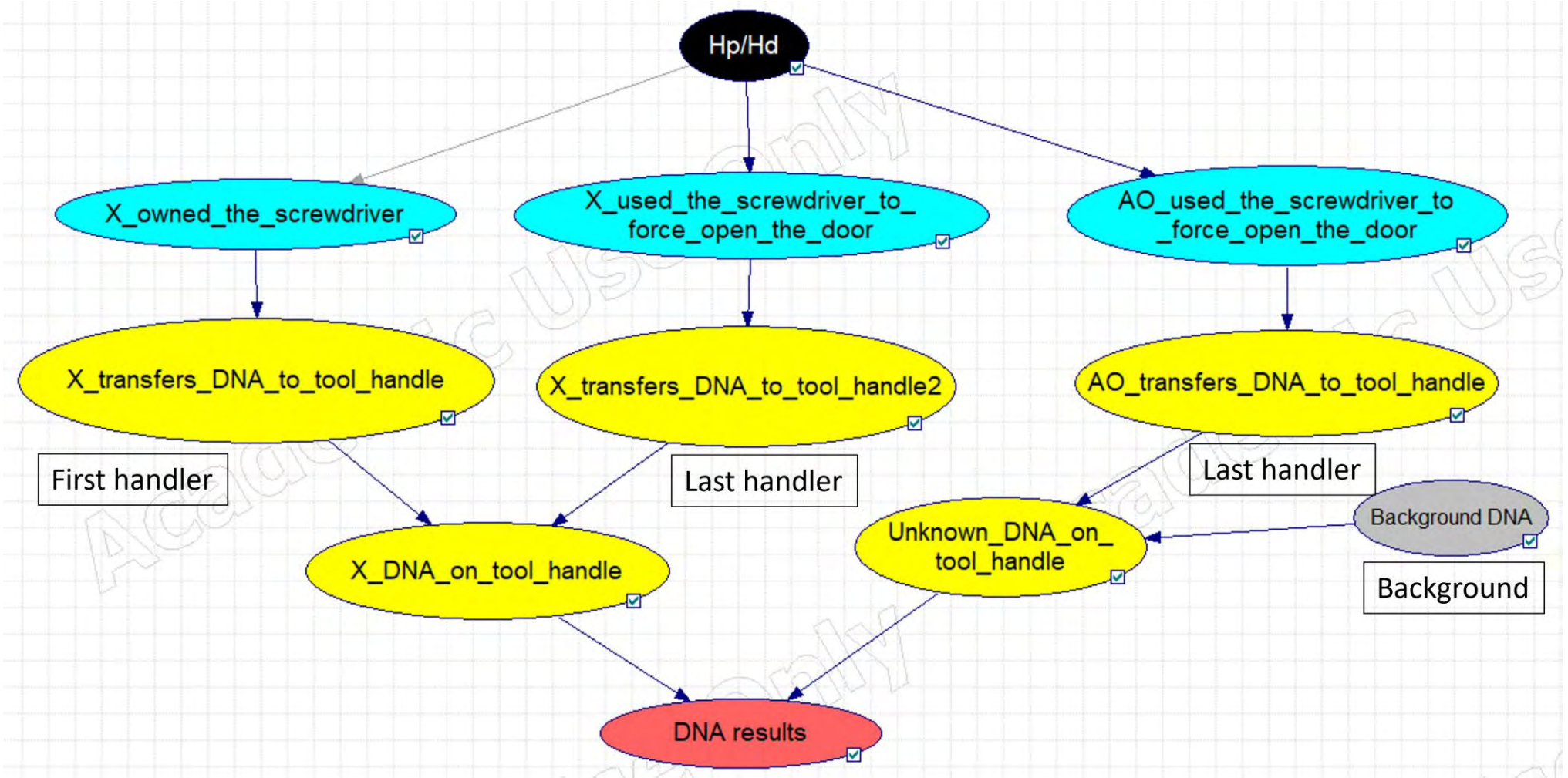
(note that only simplest examples are used here)

- A tool (screwdriver) is used to force open a door in a burglary. The tool has been left at the crime scene, and there is no doubt that it was used by the perpetrator.
- Suspect X (known individual) is arrested and accused of the crime. He states that it is his tool, but that it had been recently stolen, and that he did not force the door in the burglary.
- Findings: The DNA aligns with X at sub-source level. There may or may not be DNA from an unknown contributor(s) present

Propositions

- Either: the suspect (known individual) handled the tool at the crime-scene
- Or: an unknown person handled the tool at the crime scene
- The analysis evaluates the probability of the evidence if the suspect was either the last handler (H_p) of the tool or the first handler (H_d)

Bayesian network for experiment 2



Notice how the probabilities of FirstH, LastH and Background are used to inform the nodes

Experiment 3 case circumstances

- Mr X is accused of forcing a door. He denies that the screwdriver belonged to him and at the time of the offence he was at home watching TV. Since he lives alone there was no way to verify the alibi.
- DNA is recovered that aligns with Mr. X

A note on propositions when none is forthcoming from the defence

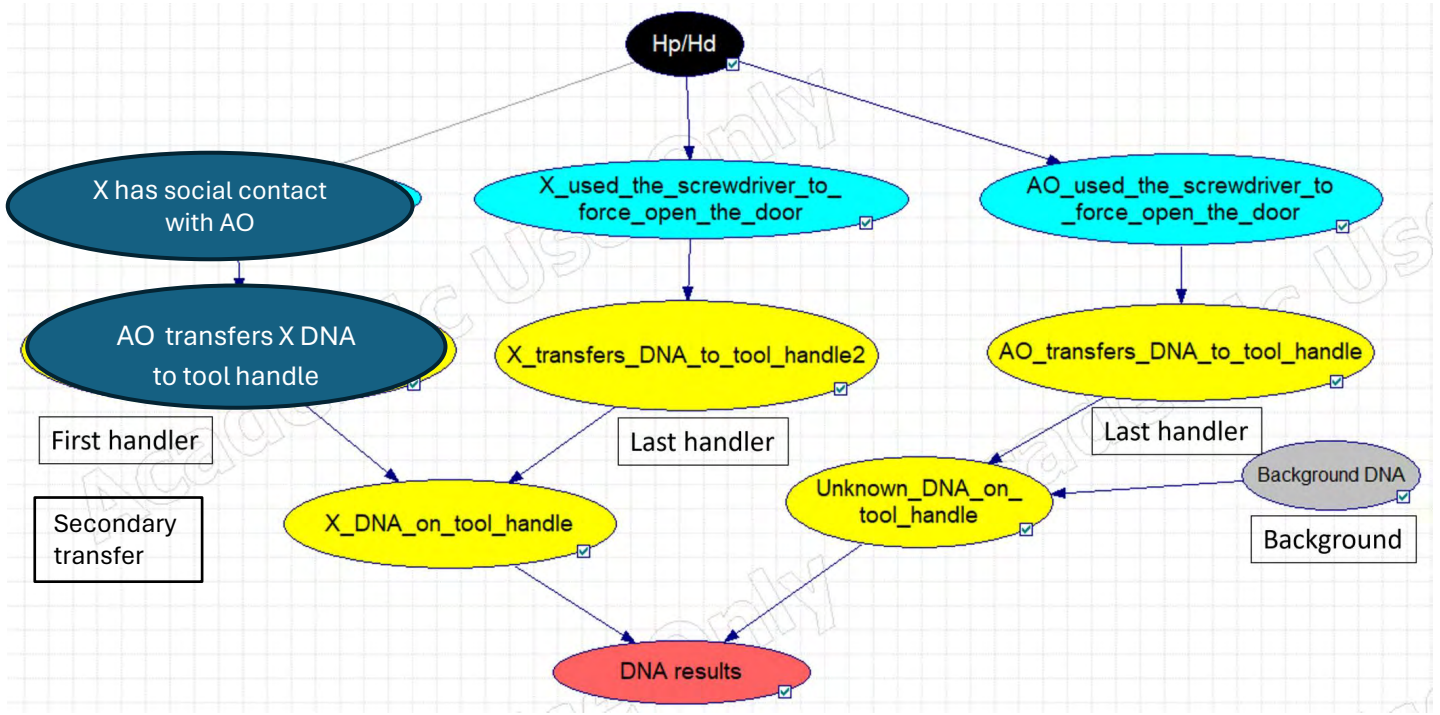
- In case work, the prosecution proposition is usually straightforward
- The defence proposition is often problematic to formulate.
- Note: The defence are under no obligation to provide an alternative proposition.
- However, If the defendant is innocent, then his DNA has transferred by the alternative route of secondary transfer.
- This can only occur via some unspecified social interaction - either a direct contact with an unknown individual (the perpetrator) under H_d e.g. shaking hands; or by contact with an item that has been handled by the perpetrator.
- Here we simulate shaking hands proposition

Experiment 3: propositions

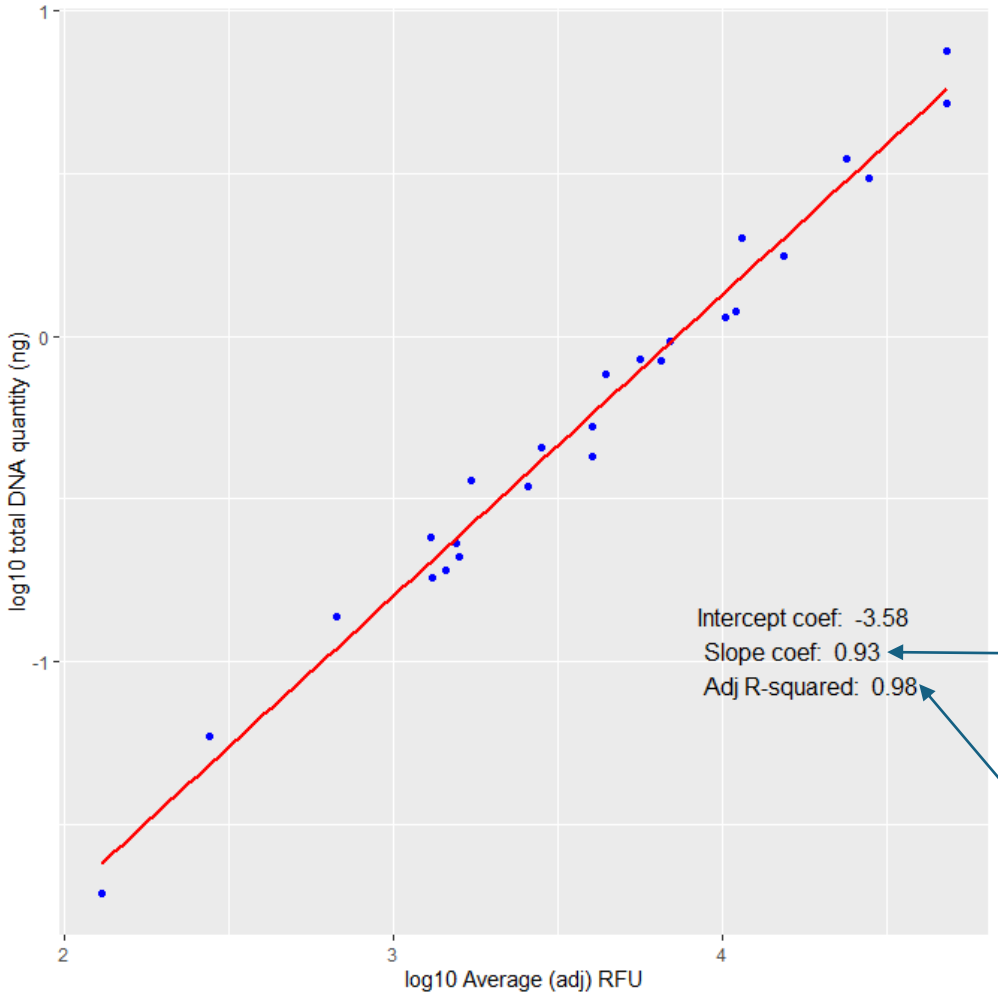
- Note that the handshake experiment will maximise the Pr(E|Hd) since alternatives that involve tertiary transfer will always manifest lower recoveries of DNA.
- There were three different times analysed – time 0, 1h, 2h.
- There is a prior expectation of loss of DNA over time
- If timing is disputed then an analysis that uses results from 0h maximises Pr(E|Hd)
- Therefore, in the absence of a clear alternative Hd proposition, we can introduce a proposition that is designed to maximise Pr(E|Hd) and acts as a proxy for multiple alternatives that involve secondary and tertiary transfer.

Experiment 3: BN

- Same structure as for expt 2 with the difference that X has had social contact with AO under Hd
- Hp asserts that X is present because of direct contact with screwdriver whereas Hd asserts no direct contact



There is a linear log log relationship of average RFU with quantity (from expt 1)



Note that a lab can easily convert quant into RFU and vice versa

This should be close to one

This should be close to one (good fit)

Note we record quantities of DNA recovered for each contributor

Standardisation

- Many methods used in the literature
- Different multiplexes
- Difficult to compare results
- All labs measure quantities of DNA
- Mixtures are very common
- So we need to assign a quantity of DNA to a contributor
- Only way to do this is with probabilistic genotyping software
  - We use EuroForMix

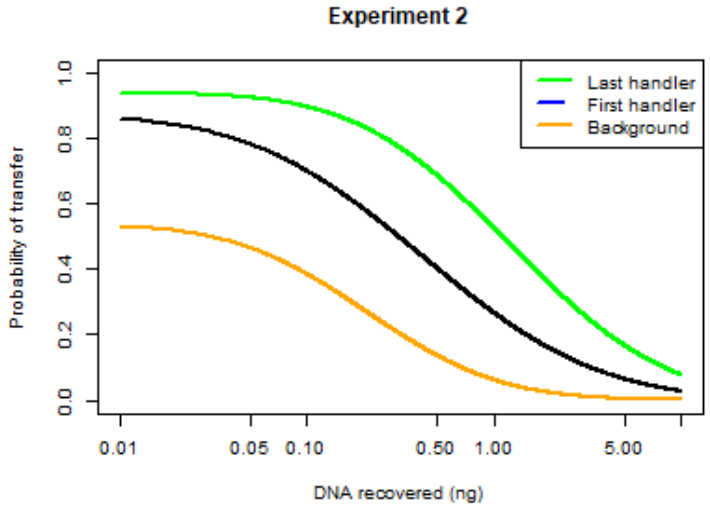
Collection of data

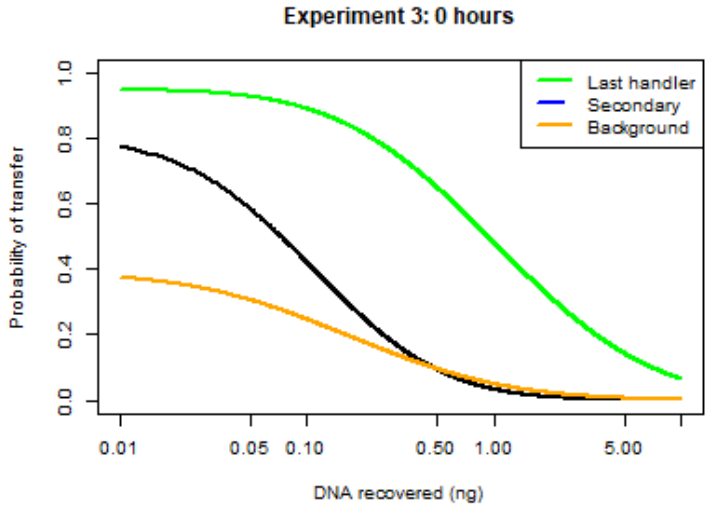
- More than 2700 sets of data from 23+ labs
- How to coordinate?
- Data analysis must be automated
  - Data collected and analysed using a genotyper macro
  - Once it is in the right format then analysed with ShinyRFU() which is an R program to automate analysis and extract information
  - These details combined with a spreadsheet supplied by the lab which includes quant value
- This produces an output file which is analysed with ShinyReAct() which outputs probabilities and likelihood ratios

Analysis

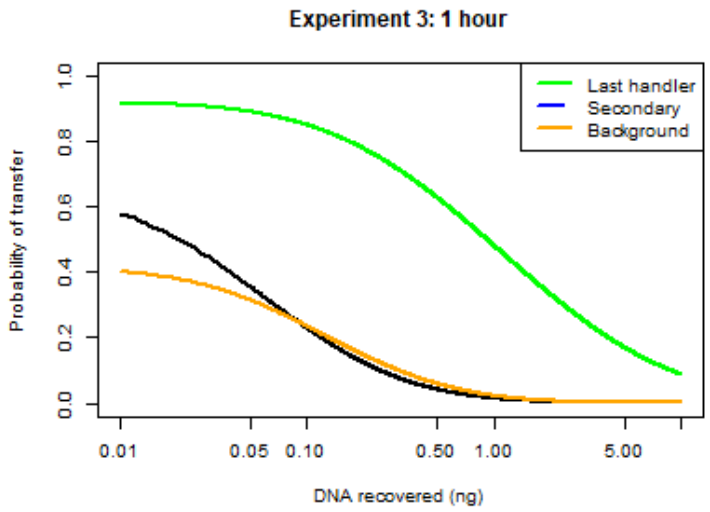
- Software has been prepared in house – nothing external has been used
- Bayesian network programs like Hugin and Genie are available but commercial licenses are needed
- All the BNs used in this work were raw programmed
  - Calculations are fast
  - Open source
  - No commercial restrictions
- All software used in this project is open source
- Data are open access (not genotypes)
- We are interested in three probabilities: First Handler, Second Handler and Background

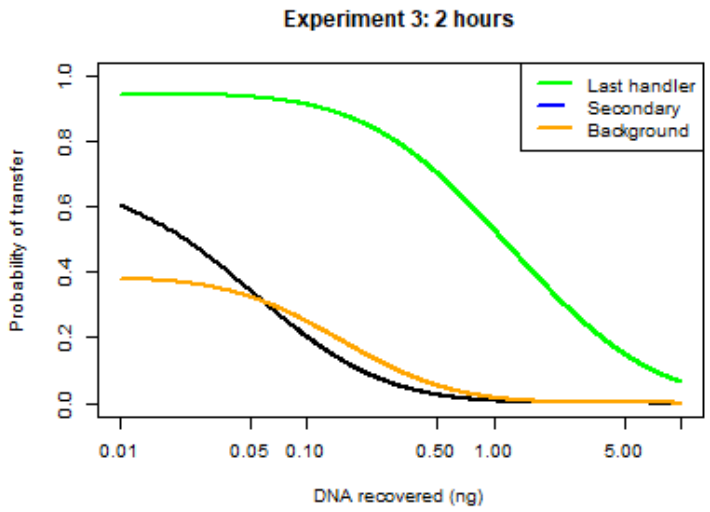
An example of probability distributions (1-cdf) (lab 201 – all data)

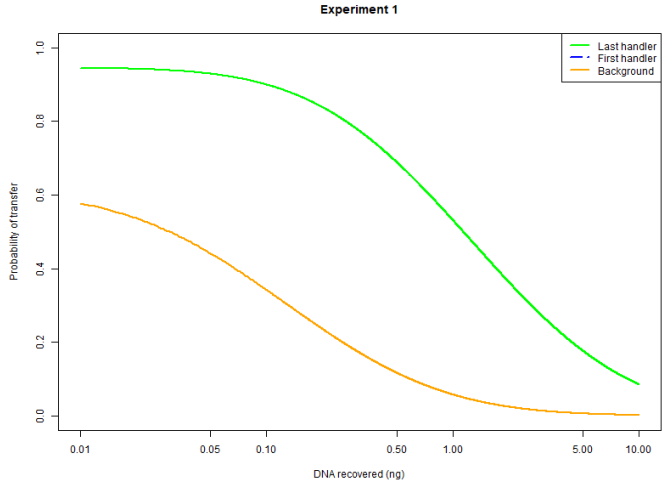




Once we have found quantities of DNA recovered, then probabilities are calculated as shown, using log normal distributions, dependent on the experiment, and plugged into the relevant BN.



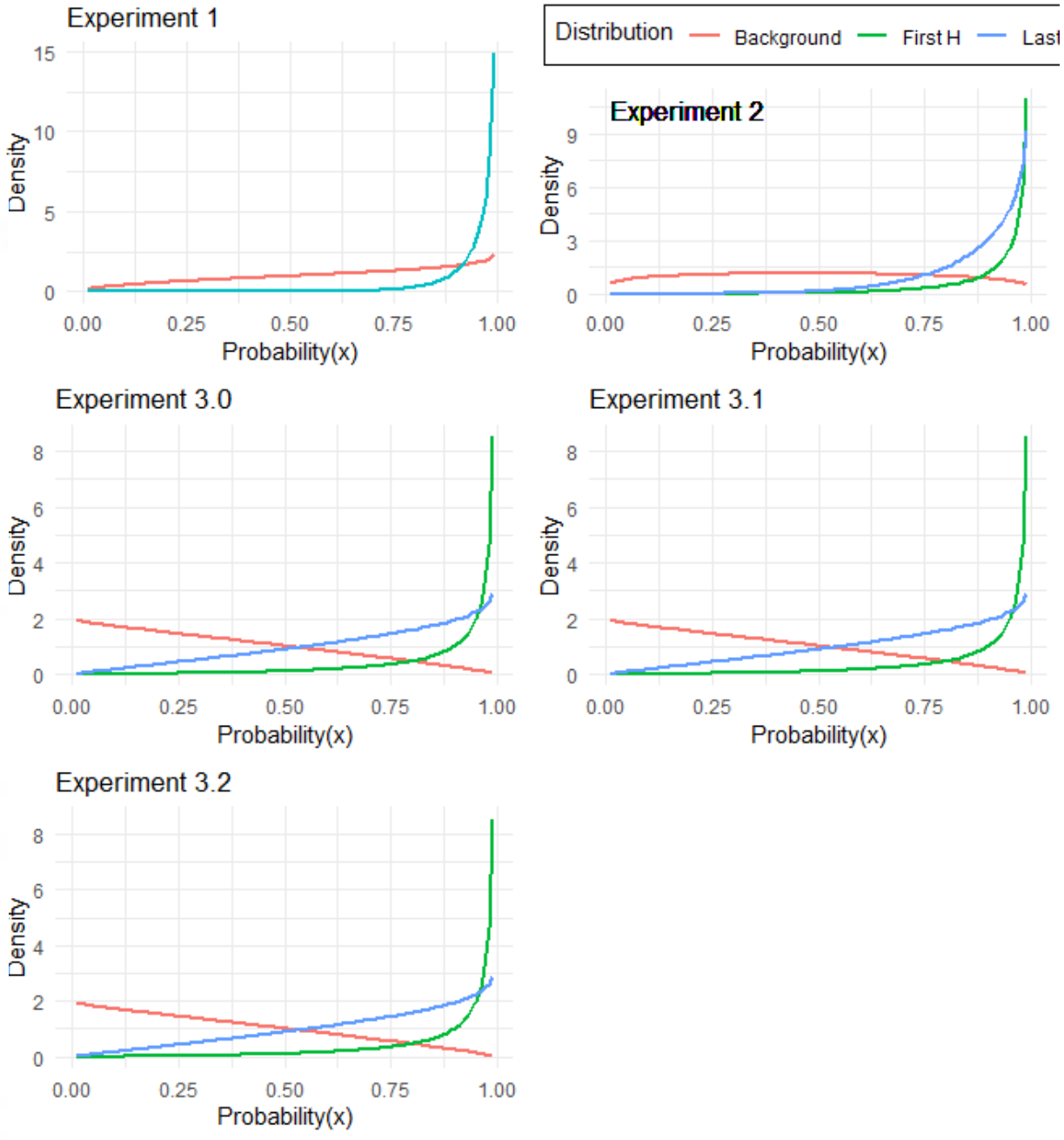




Beta distributions

LastH	Ex_1	Ex_2	Ex_3.0	Ex_3.1	Ex_3.2
alpha	9.37	4.58	2.45	1.3	1.51
beta	0.29	0.19	0.14	0.1	0.11
mean	0.97	0.96	0.95	0.93	0.93
variance	0.003	0.01	0.01	0.02	0.02
FirstH					
alpha	NA	5.40	1.96	1.77	1.51
beta	NA	0.71	0.87	1.94	1.17
mean	NA	0.88	0.69	0.48	0.46
variance	NA	0.01	0.06	0.05	0.06
Background					
alpha	1.49	1.22	0.98	0.83	0.74
beta	0.86	1.26	1.86	1.16	1.17
mean	0.63	0.49	0.35	0.42	0.39
variance	0.09	0.07	0.06	0.09	0.08

Table 3: Parameters for the beta binomial, along with expected means and variance, calculated from 23 laboratories for all experiments tab:BetaBinomial



DNA recovery across laboratories ranked by LastH, Exp 1, using raw medians

Laboratory	LastH Ex1	FirstH Ex2	FirstH Ex3.0	FirstH Ex3.1	Background Ex1
lab_10	202	1	2	1	1
lab_12	222	20	22	16	35
lab_6	301	117	23	5	56
lab_2_ESI	352	558	145	98	165
lab_15	383	81	1	1	1
lab_1_ESS	483	83	42	7	35
lab_1_NGM	488	90	41	9	30
lab_4_ESI	750	314	90	79	40
lab_4_NGM	750	275	111	68	45
lab_7	773	178	104	85	44
lab_21	954	34	64	1	88
lab_9	1073	252	76	1	7
lab_14	1088	1440	87	13	82
lab_8	1921	279	148	25	35
lab_3_ESX	2077	218	114	28	59
lab_3_NGM	2087	310	183	34	6
lab_16	2169	1073	1	1	1
lab_18_F6C	2301	344	73	1	77
lab_18_GOF	2326	289	56	19	58
lab_5	2713	143	1	1	1
lab_13	3143	787	193	91	564
lab_2_NGM	4012	233	41	1	1
lab_17	4959	942	169	113	10
medians	1073	252	73	13	35

Median recoveries (pg) are illustrated per laboratory

LRs from expt 2: Where Only the POI is recovered

Quantity of DNA from POI in ng

Laboratory	0.01	1	3	5	7	9
lab_21	5	7	6	5	5	5
lab_6	9	16	13	11	10	9
lab_13	8	14	18	19	19	19
lab_17	12	18	23	25	25	25
lab_11	22	47	58	60	59	57
lab_14	23	29	45	54	59	62
lab_12	23	79	79	75	71	67
lab_18_F6C	29	61	73	74	73	71
lab_16	24	35	57	69	75	79
lab_3_ESX	22	64	80	85	86	86
lab_9	27	90	107	106	102	98
lab_19	43	78	97	101	101	99
lab_23	26	50	72	86	96	105
lab_3_NGM	22	61	85	97	105	111
lab_22	54	73	117	154	186	215
lab_2_NGM	23	83	171	234	285	328
lab_4_ESI	12	46	113	181	249	320
lab_2_ESI	23	63	157	251	346	443
lab_1_NGM	25	132	269	388	499	607
lab_1_ESS	25	143	314	471	626	780
lab_4_NGM	42	149	391	667	978	1321
lab_7	22	127	371	665	1007	1391
lab_8	26	134	420	755	1126	1525
lab_5	12	116	590	1273	2095	3015
lab_10	7	5				
lab_15	5	90	2086			
Medians	**23**	**63**	**91**	**99**	**101**	**102**

- The LR data can be ranked and divided into tertiles. Mid tertile is in yellow.
- Note wide range of LRs across laboratories
- Note median value in final row
- Note that high/low LRs do not necessarily correspond to high/low recovery in median polish
- Labs that only recover small amounts of DNA will tend to record higher LRs because the majority of tests are negative – i.e. rare events are less likely to occur and therefore when they happen, they result in higher LRs e.g. lab_5

Data set combinations investigated (available in the online folders)

- Set_100_median_LR_Exp_2
- Lab_201 (all labs except v. low recovery labs)
- Lab_500_High_LR_Exp 2
- Lab_600_low_LR_Exp 2
- Lab_300: An extended dataset based on experiment 3

There are differences between laboratories, so how can we standardise across laboratories

- One way will be to combine data and calculate distributions
- Increases dataset size from approx. 20 to 615 (exp 2) and 356 (exp 3)

A comparison of two methods

- Continuous method
- Binary method (thanks to Tacha for rule-set) based on mixture proportions (Mx values) where POIs are categorised into
  - Single >99%
  - Major >70% and <99%
  - Balanced <70% and >1%
  - Absent <1%

LR results

- Experiment 2 (combined data across labs N=615)

Continuous method

	POI quant (ng)	0.01	0.5	1	2	4	6	10
Single	**0**	27	50	61	73	84	89	94
	0.05	0.6	1.0	1.2	1.5	1.7	1.8	1.9
Unknown	**0.1**	0.5	0.8	1.0	1.2	1.4	1.5	1.6
quants (ng)	**0.2**	0.3	0.6	0.8	0.9	1.0	1.1	1.2
	0.3	0.3	0.5	0.6	0.7	0.8	0.9	0.9
	0.5	0.2	0.4	0.4	0.5	0.6	0.6	0.7
	1	0.1	0.2	0.3	0.3	0.4	0.4	0.4
	2	0.1	0.1	0.1	0.2	0.2	0.2	0.2

Single contributor only

Red type = major
Black type = balanced

Mx method (median value across all labs)		combined data
Single	55.0	191.0
Balanced	0.8	0.8
Major	2.6	4

Note LRs are sensitive to sample size

LR results experiment 3.0 (combined data N=356)

Continuous method

	POI quant (ng)	0.01	0.5	1	2	3	4	5
Single	0	20	106	207	450	742	1077	1453
	0.05	0	2	4	9	15	22	30
Unknown	0.1	0	2	3	7	12	18	24
quants (ng)	0.15	0	1	3	6	10	14	19
	2	0.017	0.1	0.2	0.4	0.6	0.9	1.2
	3	0.009	0.05	0.1	0.2	0.3	0.5	0.6
	5	0.004	0.02	0.04	0.1	0.1	0.2	0.3

Single contributor only

Red type = major
Black type = balanced

Mx method (median value across all labs)		combined data
Single	60.0	552
Balanced	0.24	0.3
Major	4	8

LR results experiment 3.1 (combined data N=356)

Continuous method

	POI quant (ng)	0.01	0.5	1	2	3	4	5
Single	0	16	143	299	691	1173	1739	2382
	0.05	1	5	11	25	42	63	86
Unknown	0.1	0.5	4	8	19	33	49	67
quants (ng)	0.15	0.4	3	7	16	27	40	54
	2	0.04	0.3	1	2	3	4	5
	3	0.02	0.2	0.4	1	2	2	3
	5	0.01	0.1	0.2	0.4	1	1	1

Single contributor only

Red type = major
Black type = balanced

Mx method (median value across all labs)		combined data
Single	84.0	533
Balanced	0.23	0.2
Major	8	49

LR results experiment 3.2 (combined data N=356)

Continuous method

	POI quant (ng)	0.01	0.5	1	2	3	4	5
Single	0	24	407	1087	3229	6361	10461	15522
	0.05	1	10	26	77	152	250	370
Unknown	0.1	0.5	8	20	61	119	196	291
quants (ng)	0.15	0.4	6	16	49	96	159	235
	2	0.02	0.4	1	3	6	10	15
	3	0.01	0.2	1	2	3	5	8
	5	0.01	0.1	0.2	1	1	2	3

Single contributor only

Red type = major
Black type = balanced

Mx method (median value across all labs)		combined data
Single	117.0	433
Balanced	0.2	0.23
Major	8.0	73

Conclusions

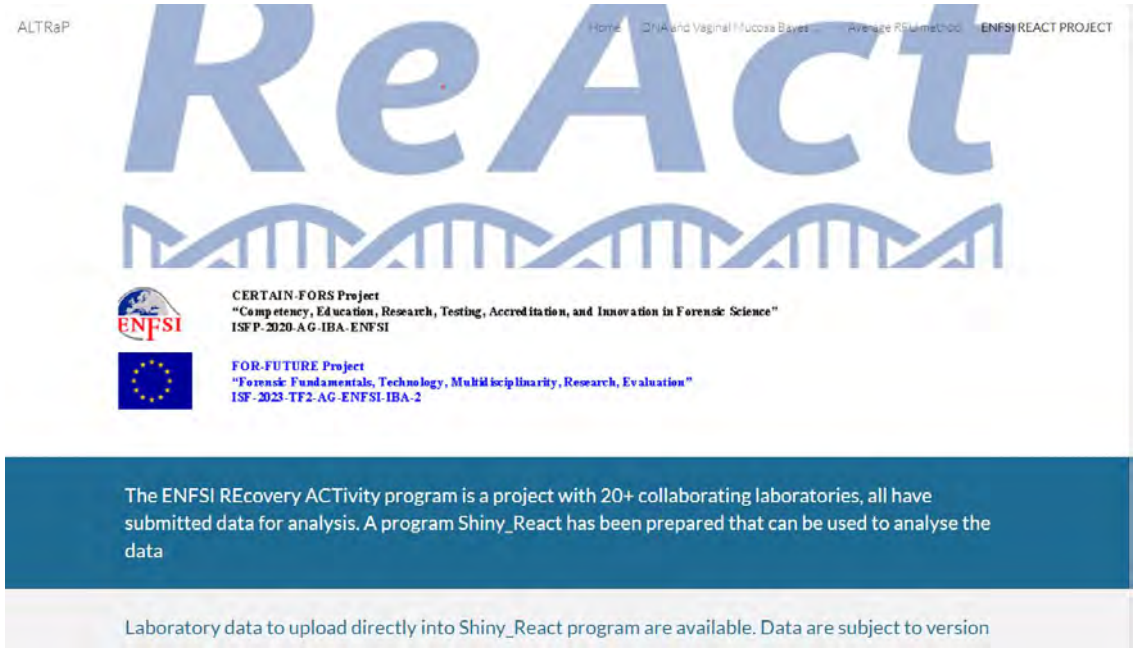
- For all experiments, the presence of the POI as a single profile is probative evidence (LR>1); LR can be substantial
  - However, note that the great majority of results are mixtures:
- For all experiments, the presence of the POI as a balanced contributor between 1% and 70% is always neutral or supports Hd proposition (depending on quants)
- For experiment 2, the presence of the POI as a major contributor >70% tends to neutrality (the evidence does not help to distinguish between first and last handler of the screwdriver)
- For experiments 3, the presence of the POI as a major contributor >70% supports Hp proposition
- LRs increase as time since handshake increases, since probability of recovering DNA from handshaker reduces quickly over time

Next steps

- It was shown that there was great variability between labs in their effectiveness of recovering DNA
- This has implications that go beyond this exercise
  - If a method is sub-optimal then the lab will recovery of DNA may be compromised
  - Therefore it is useful to try to find out reasons for differences
  - This is a focus of ReAct II
- Aim is to devise simple experiments so that labs can determine their recovery efficiency. This information may be useful to help identify lab(s) that perform in a similar way to each other, so that their data may be used to help inform LR values

Resources available

- See my website for latest details and links (still under construction)
  - https://sites.google.com/view/altrap/enfsi-react-project



- All data-sets (not genotypes) are available
- Shiny React app (link provided) to review the data and to carry out calculations
- User manual in preparation
- Publications in preparation

Compilation of data

- File Grand_compilation.xlsx contains all data from all labs
- A total of 2735 rows of samples
- This acts as a searchable database

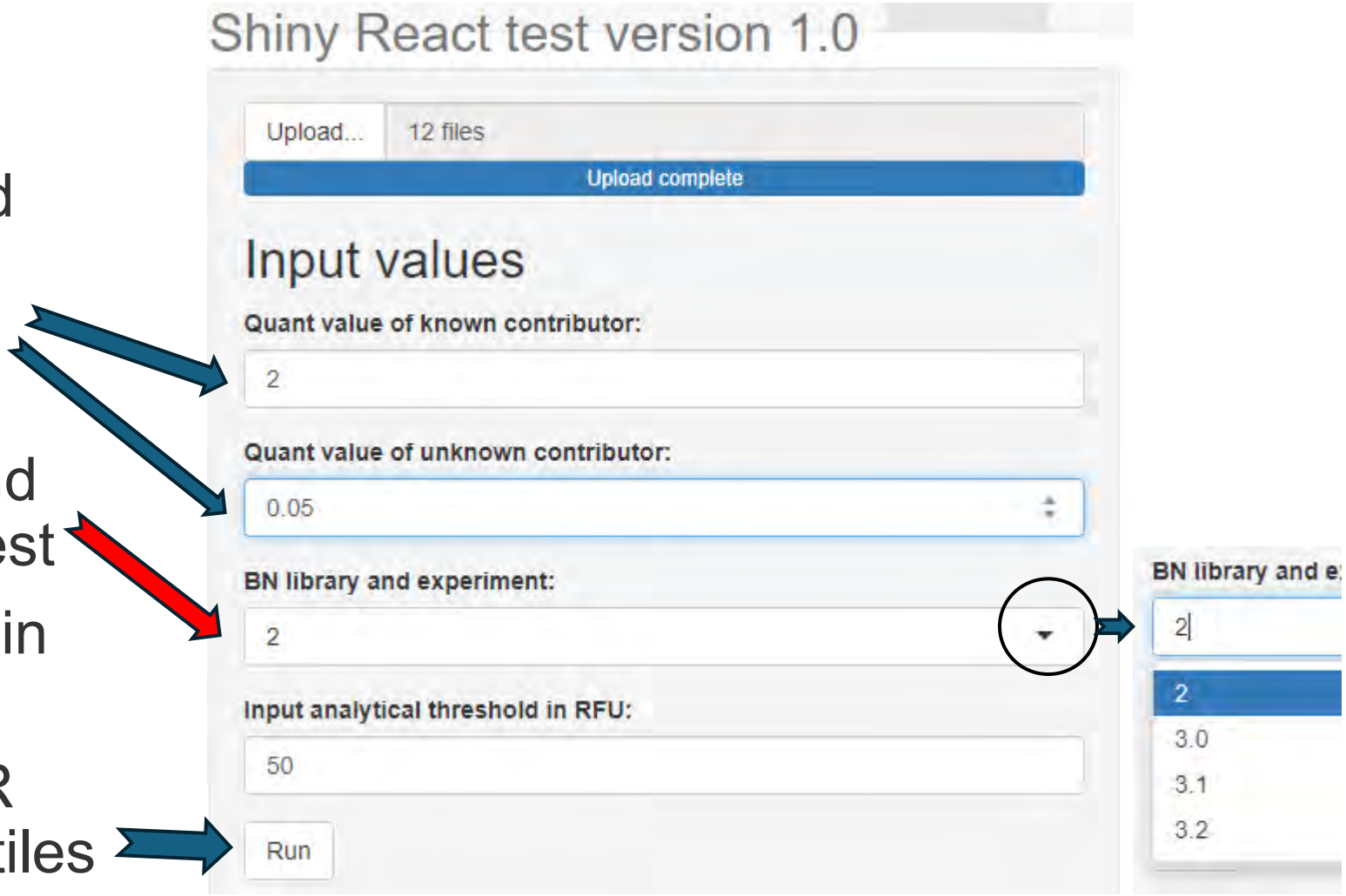
- Data location:
- Sites.google.com/view/altrap/
- Navigate to ENFSI Shiny React site

Shiny React

- The program is accessed from this link: http://cchampod.shinyapps.io/Shiny_React_App
- Thanks to Christophe Champod for hosting

Input

- Input quant values in nanograms for known and unknown contributors. If there is no unknown contributor, input 0
- Select BN (Bayes Net) and experiment you want to test
- Input analytical threshold in RFU
- Click 'Run' to calculate LR and bootstrapped percentiles



Output

Results Output

	BN	Known.Quant	Unknown.Quant	LRTog	LRPOI	Qual.LR	Sample.Size
Results	2	2	0.05	4.3e+00	NA	6.3e-01	24

Bootstrapped LR Results (Percentiles)

	2.5%	5%	25%	50%	75%	95%	97.5%
Percentiles	7.9e-01	9.9e-01	2.0e+00	4.9e+00	2.2e+01	5.3e+02	1.3e+03

```
Failed bootstraps = 15
```

```
Check Pr sums Hp = 1 ;;;  Check Pr sums Hd = 1
```

- “Results Output” gives a summary of input values
- “LRTog” means POI+U model is used hence LRPOI only model is NA
- Qual.LR is the qualitative LR from the binary model

Summary of resources

- A folder containing analysed data for every laboratory
- Results are compiled into a searchable database
- A program Shiny_React can be used to analyse data and report LRs using bootstrap confidence intervals
- Program has a modular design to allow new data and BNs to be added.
- Data files will have version numbers to allow for updates (new data and corrections)

Thoughts

- Can expand to include new datasets and act as a repository of Bayesian networks that can be selected by the user.
- We need to think about standardisation of methods
- There are many publications, but data are generally not available to use – also different methods are in use
- Big question is whether it is valid for laboratory A to use data of laboratory B, unless recovery rates / LRs can be demonstrated to be similar
- Should the same effort be put into activity level as with frequency databases?

Acknowledgements

- Core group
  - Ingo Bastisch
  - Stavroulla Xenophontos
  - Marios Cariolou
  - Tacha Hicks Champod
  - Elida Fonneløp
  - Iris Buckel
  - Roland van Oorschot
  - The ReAct Consortium