

Forensic DNA phenotyping

Angel Carracedo

**Fundación Pública Galega de Medicina Xenómica- SERGAS-
Xunta de Galicia**



**Instituto de Ciencias Forenses- Universidade de Santiago de
Compostela**



Genome variation



99.8% identity

85% variation within individuals/local,

8% within continents,

7% within populations of the same continent



FORENSIC GENETICS

CASEWORK:

Paternity testing

Criminal casework

Identification of human remains

Criminal DNA databases

Non human DNA typing



OTHERS:

Toxicogenetics

Forensic Molecular Pathology

DNA Typing in Forensic Analysis

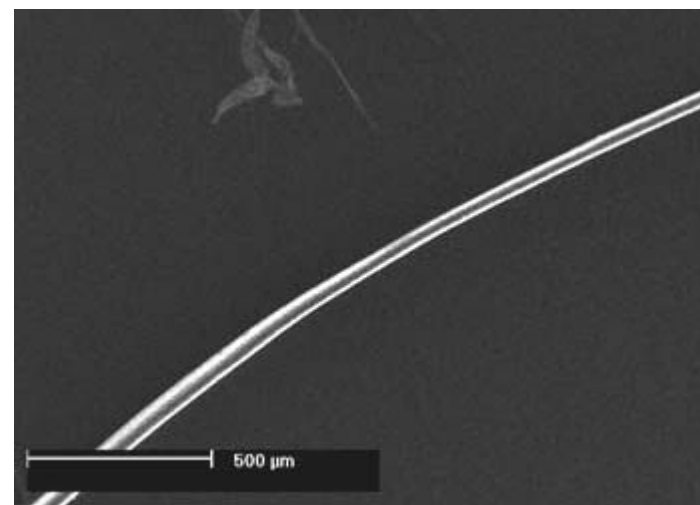
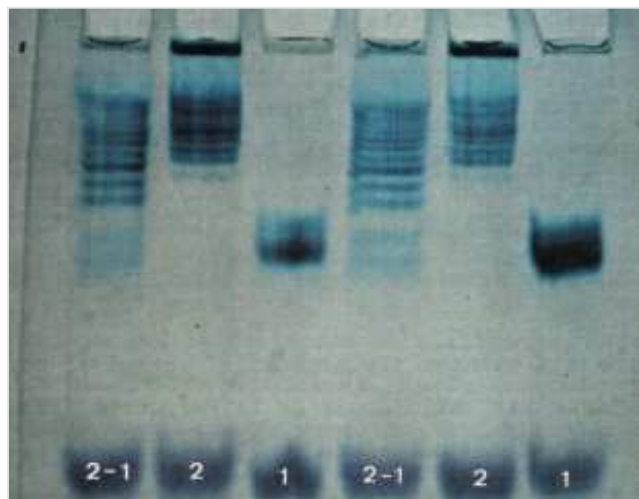


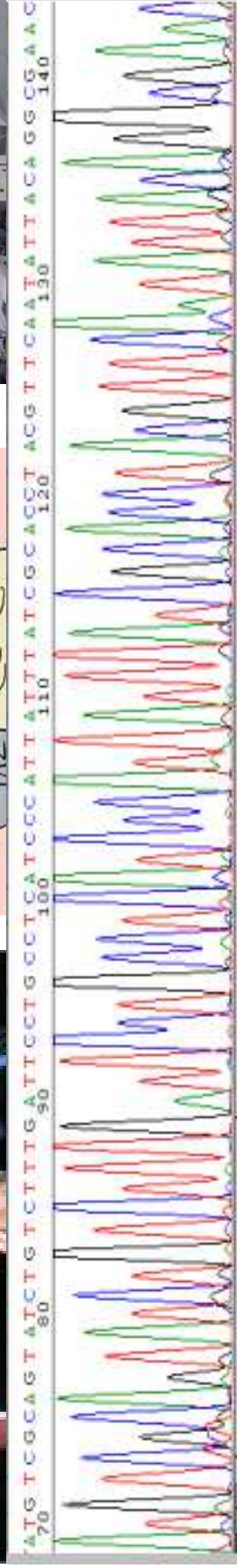
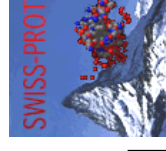
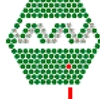
Blood groups

HLA

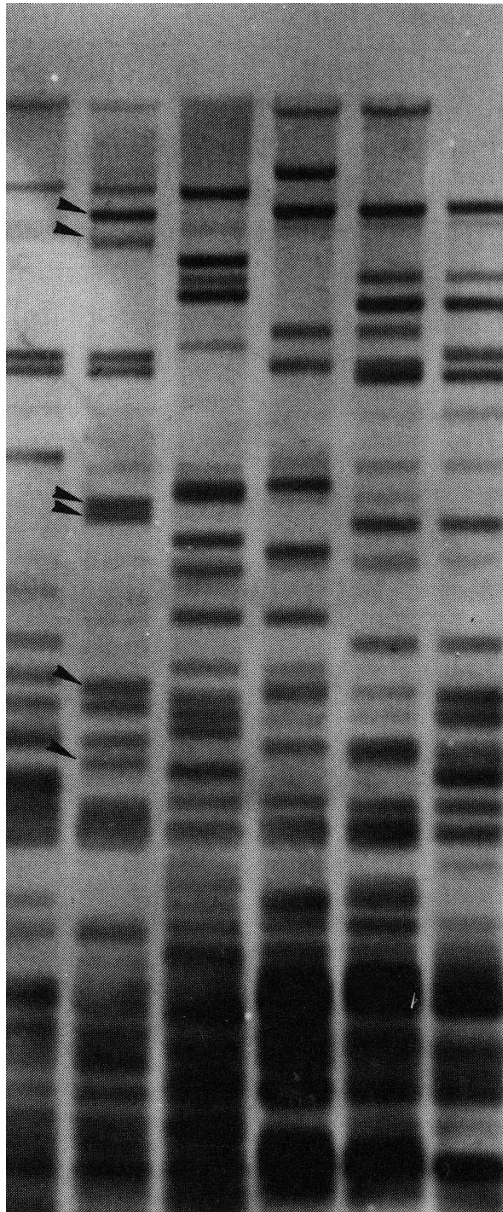
Serum proteins

Enzymes



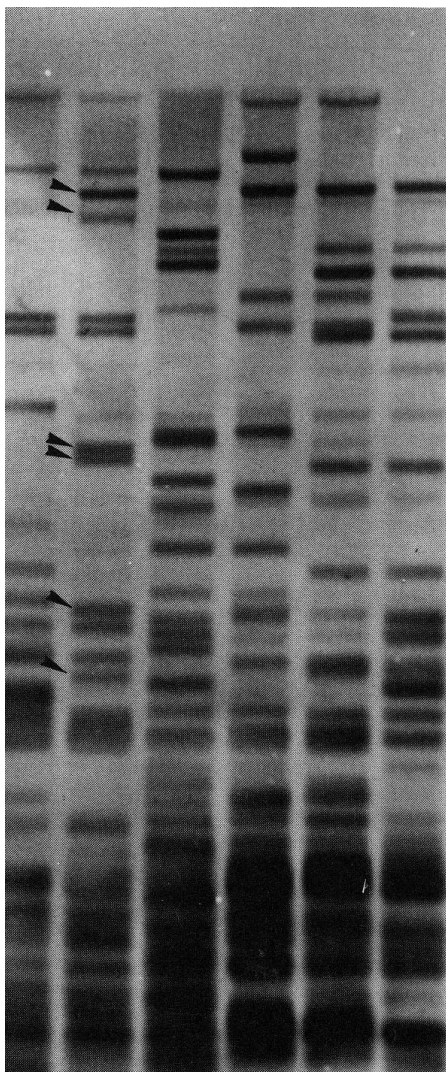


Alec Jeffreys DNA fingerprint 1985

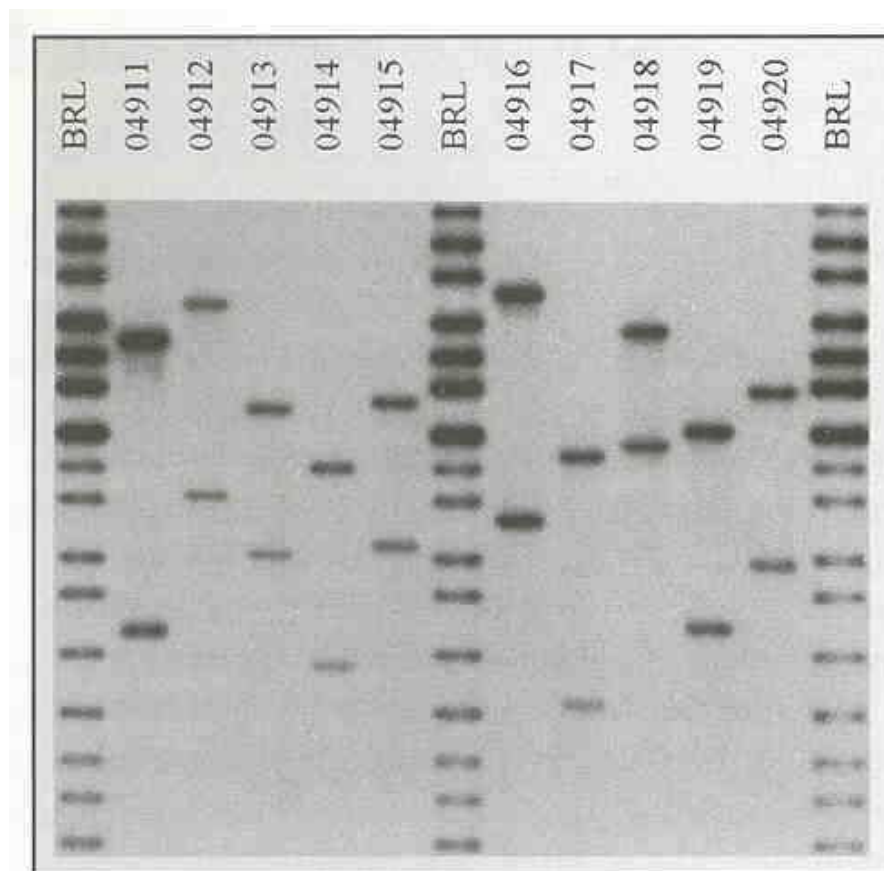


Jeffreys, A.J., Wilson, V., and Thein, S.L.
Hypervariable minisatellite regions in human
DNA. Nature 314 (1985) 67-73.

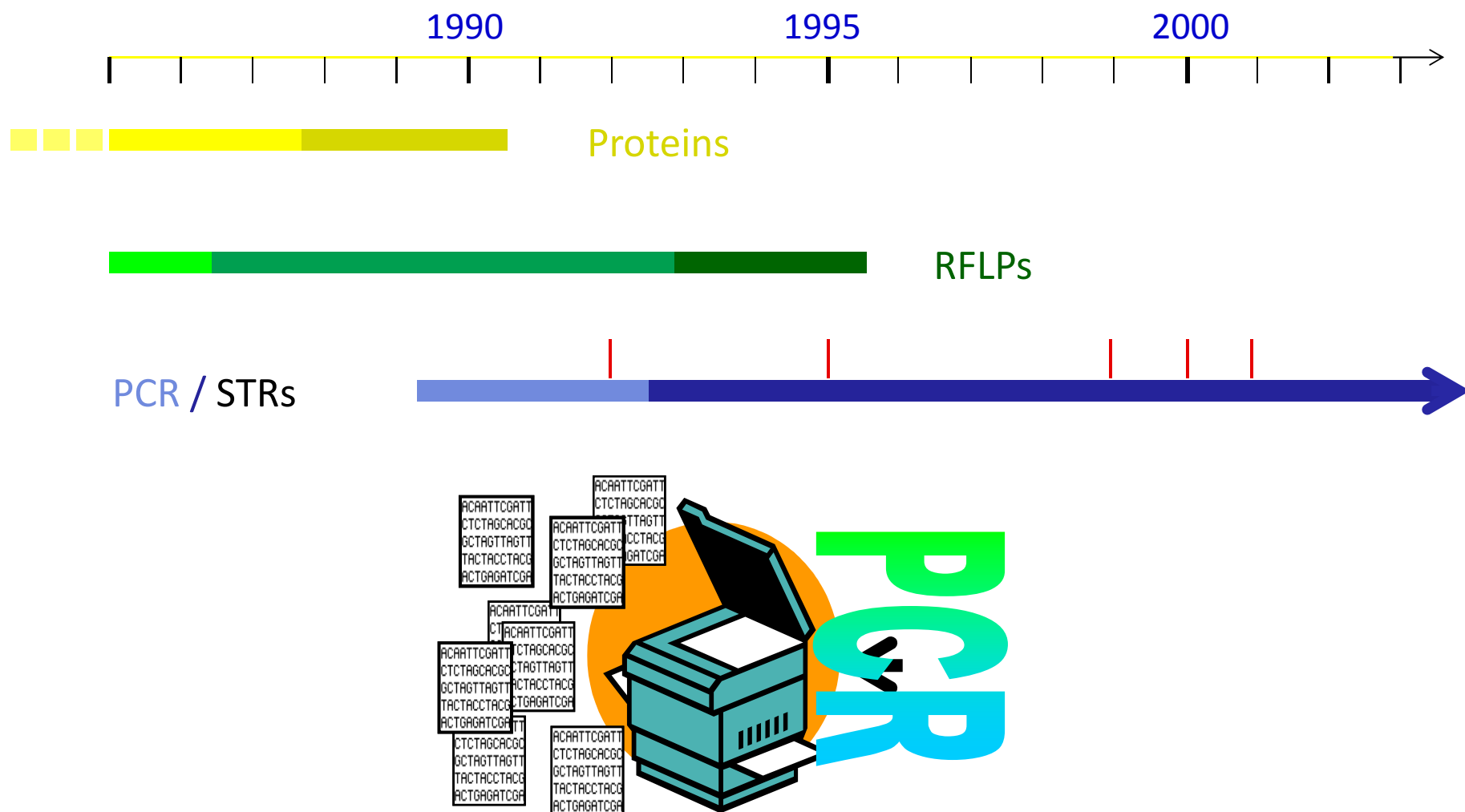
DETECTION OF MINISATELLITES USING MLPs



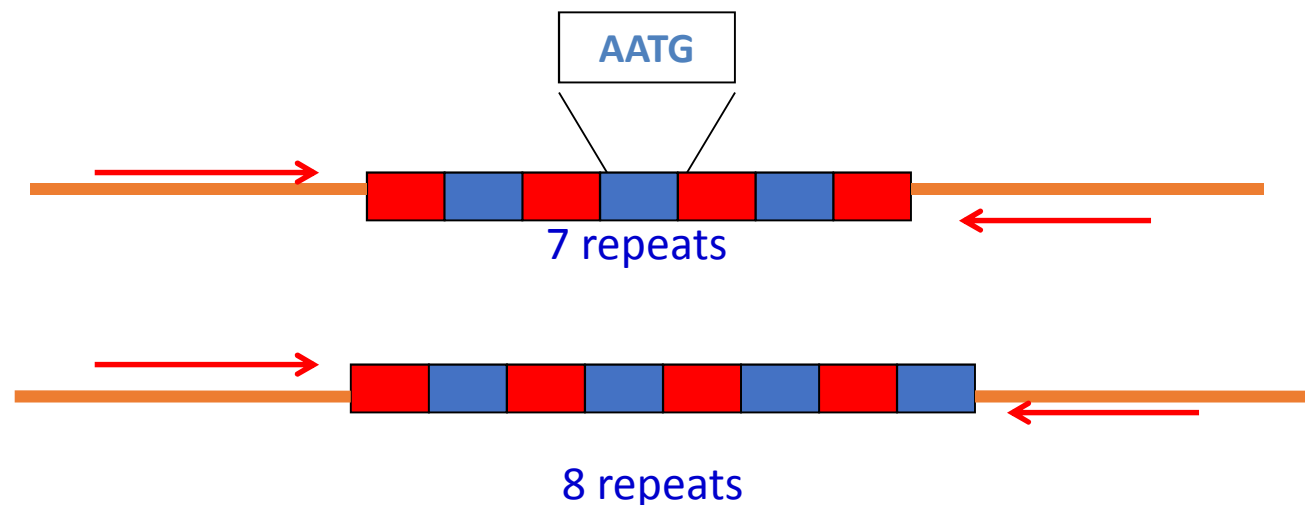
DETECTION OF MINISATELLITES USING SLPs



DNA Typing in Forensic Analysis



Short Tandem Repeats (STRs)



ADVANTAGES OF STRs OVER SLPs

Amount of DNA required

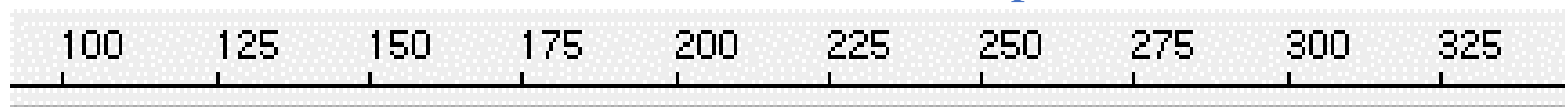
Analysis of degraded samples

Time of analysis

Standardization and value of the evidence

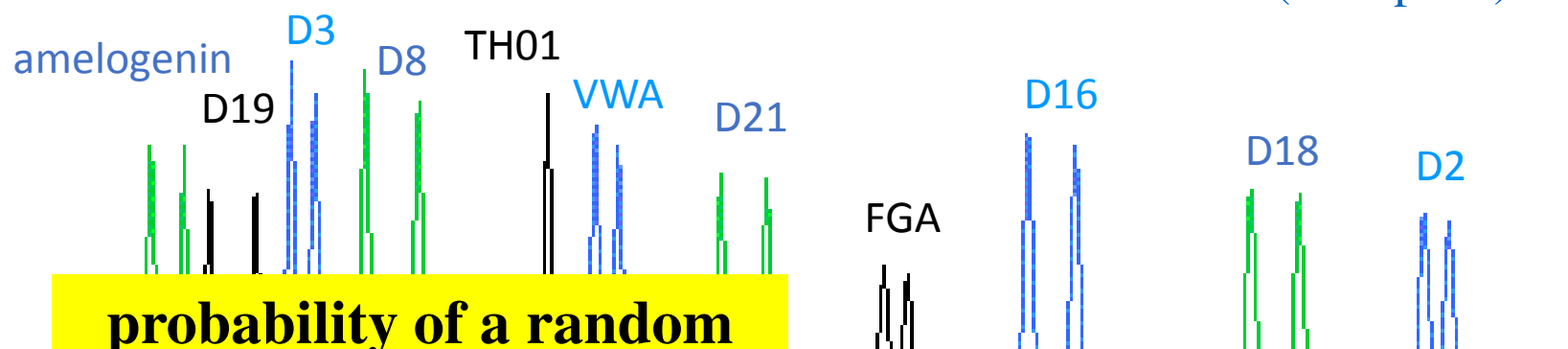
Human Identity Testing with Multiplex STRs

AmpFlSTR® SGM Plus™ kit



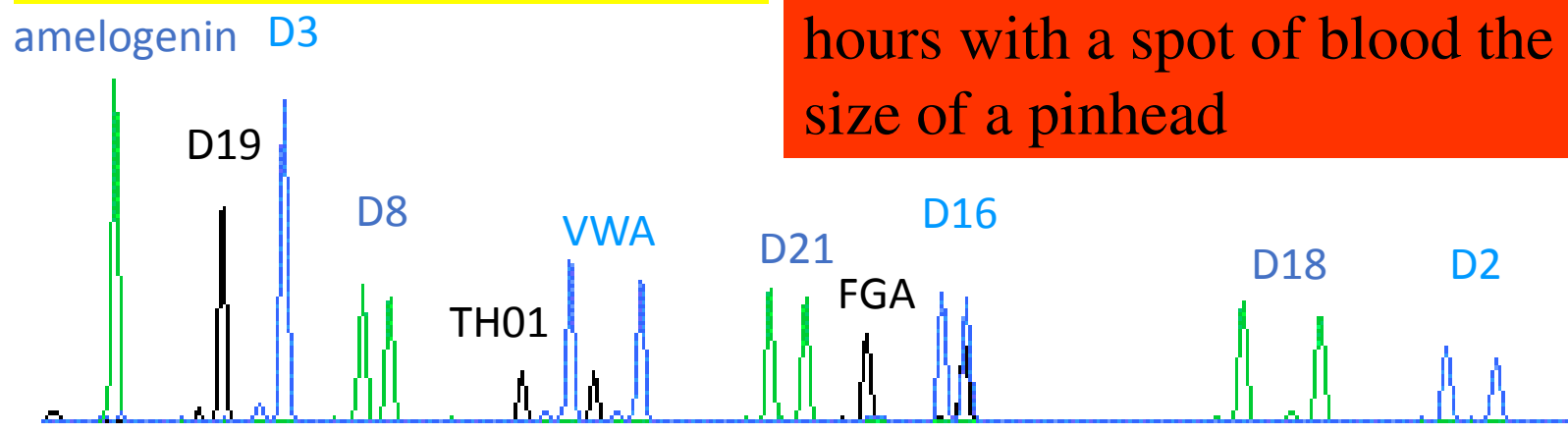
DNA Size (base pairs)

Two different individuals

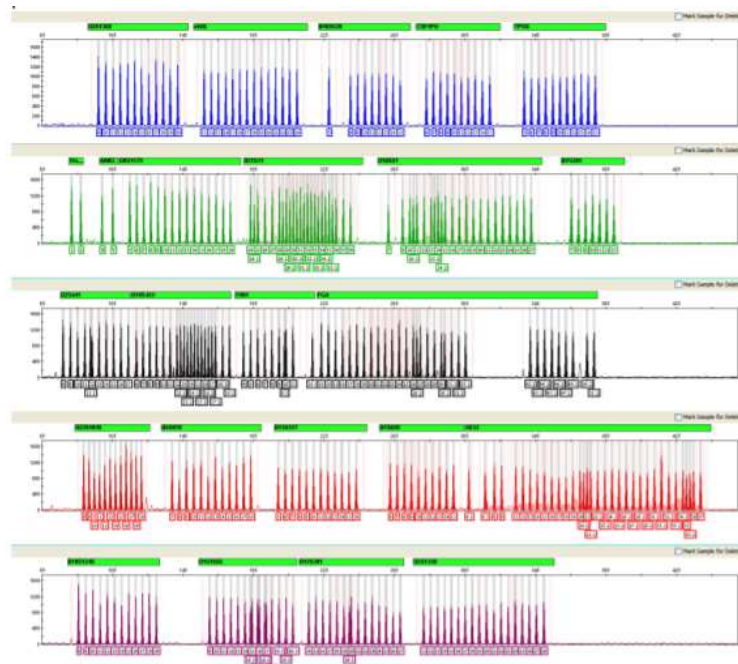


**probability of a random
match: ~1 in 3 trillion**

**Results obtained in less than 5
hours with a spot of blood the
size of a pinhead**



Simultaneous Analysis of 9 STRs and Gender ID



GlobalFiler (AB)

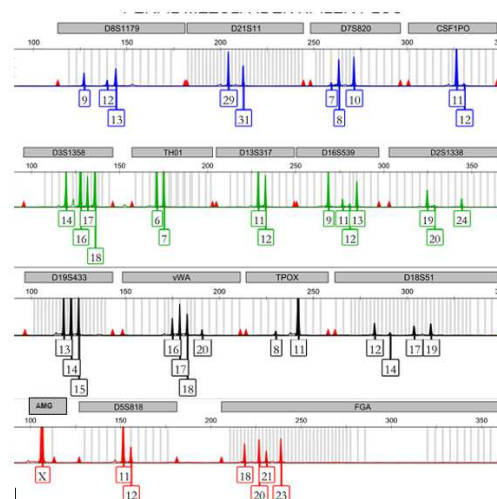


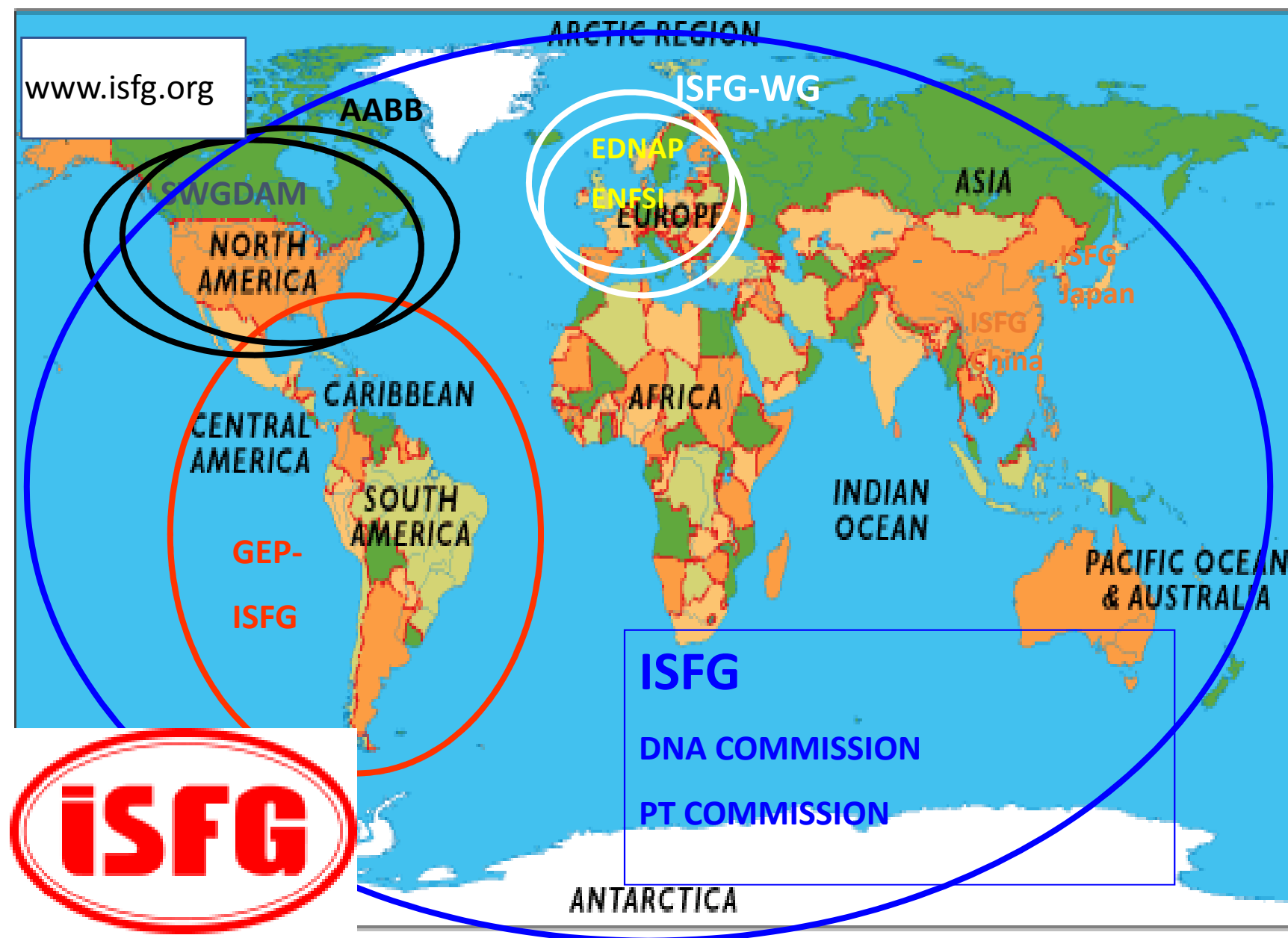
Imagen perfil mezcla de ADN de dos mujeres



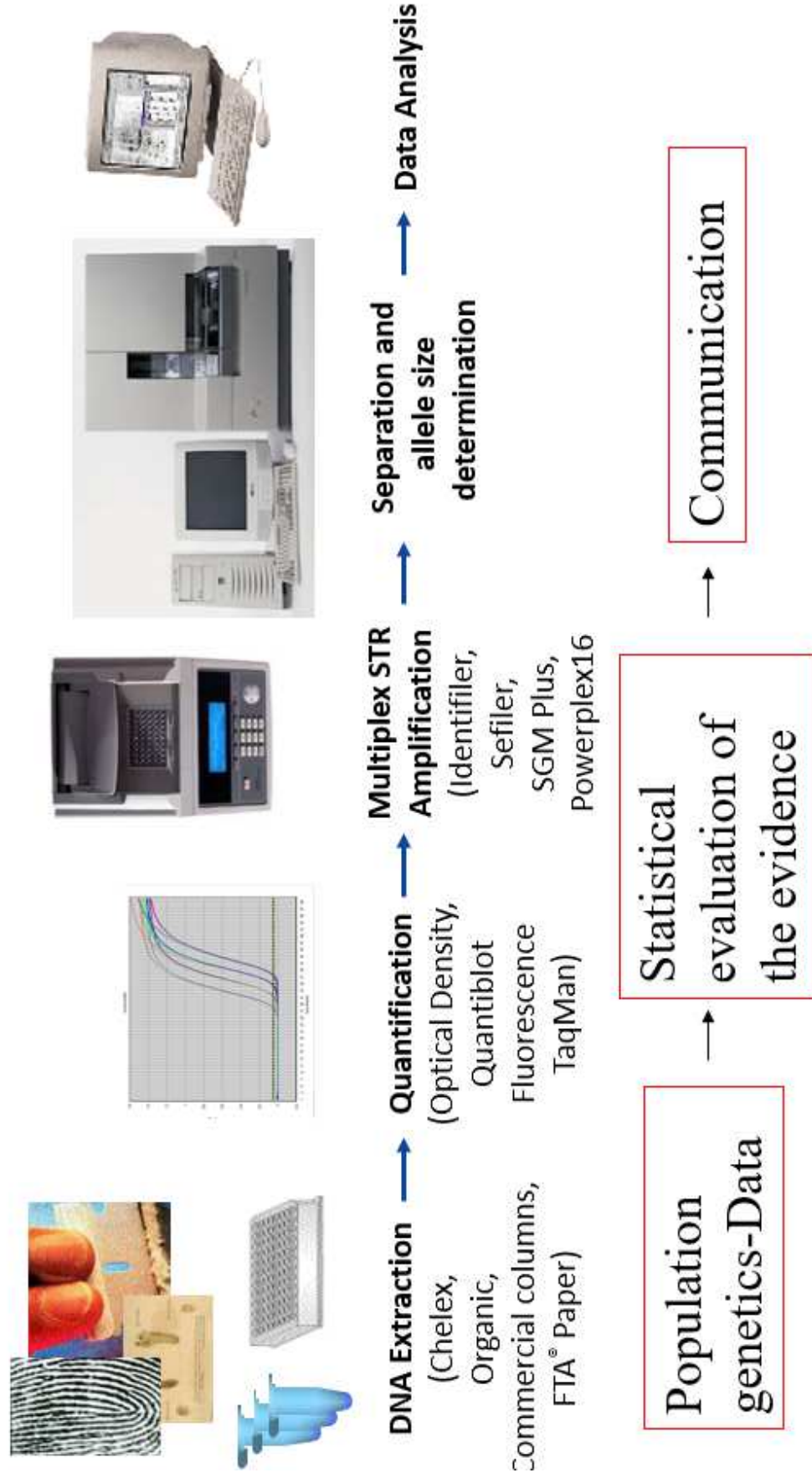
En Europa aunque hay diferencias entre países en todos es obligatorio usar el denominado

| STRs | ESS | CODIS |
|-------------|-----|-------|
| D3S1358 | X | X |
| VWA | X | X |
| FGA | X | X |
| Amelogenina | (X) | |
| THO 1 | X | X |
| TPOX | | X |
| CSF1PO | | X |
| D5S818 | | X |
| D13S317 | | X |
| D7S820 | | X |
| D2S1338 | | |
| D8S1179 | X | X |
| D18S51 | X | X |
| D21S11 | X | X |
| D16S539 | | X |
| D19S433 | | |
| Penta D | | |
| Penta E | | |
| SE 33 | | |
| D22S1045 | X | |
| D2S441 | X | |
| D10S1248 | X | |
| D12S391 | X | |
| D1S1656 | X | |





The DNA analysis procedure in forensics



Revolution: technology and interpretation

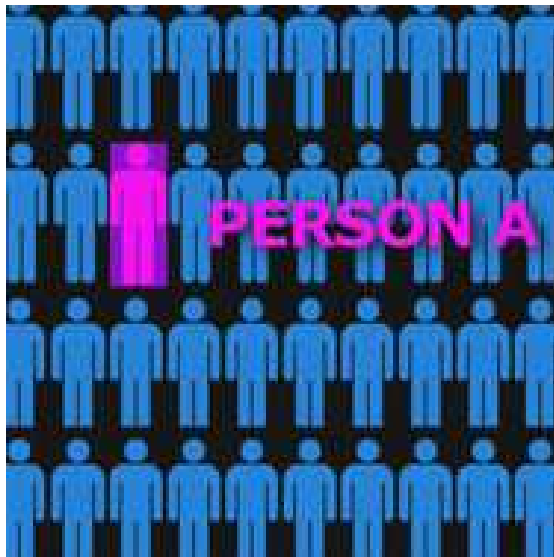


Technology revolution:

Evolution of genetic markers

Evolution of equipment

Shorter response time



Revolution in interpretation:

Scientists are convinced that any opinion they give has a degree of uncertainty and that it is their duty and obligation to inform the judge of that **uncertainty**. In order to do that, a standard called **probability** exists.

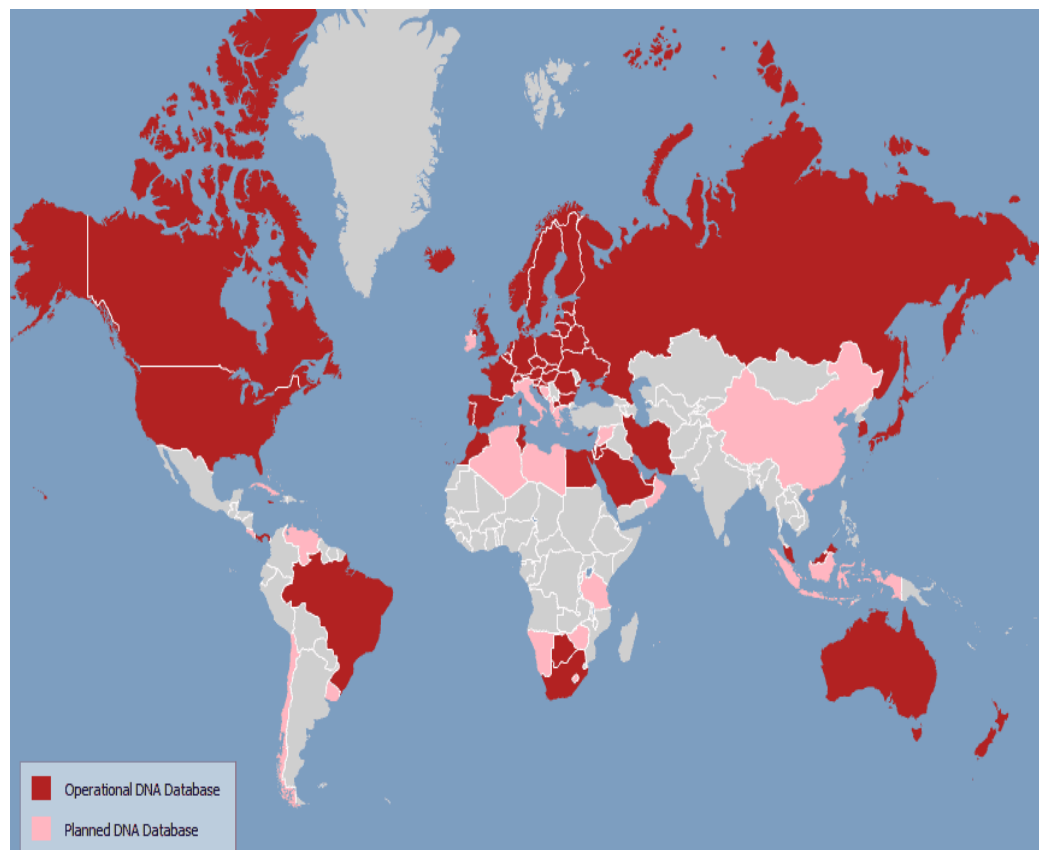
Statistics, interpretation and communication

The single most important advance in forensic science thinking is the realisation that the scientist should address the probability of the evidence

Still the area with most important challenges and need of standards

complex samples (partial genetic profiles, stochastic effects, DNA mixtures, low level DNA samples)

complex scenarios (massive identifications)



Kenya Government mandates DNA-linked national ID, without data protection law

Alice Muryua | February 8, 2019 | [No responses yet](#)

Last month, the Kenya Parliament passed a seriously concerning amendment to the country's national ID law, making Kenya home to the most privacy-invasive national ID system in the world. The rebranded, National Integrated Identity Management System (NIIMS) now requires all Kenyans, immigrants, and refugees to turn over their DNA, GPS coordinates of their residential address, retina scans, iris pattern, voice waves,



Información sobre este sitio web

THEEPOCHTIMES.COM

22/ Rwanda Plans to Collect DNA From 12 Million Citizens in World's First Nationwide Database

The New York Times

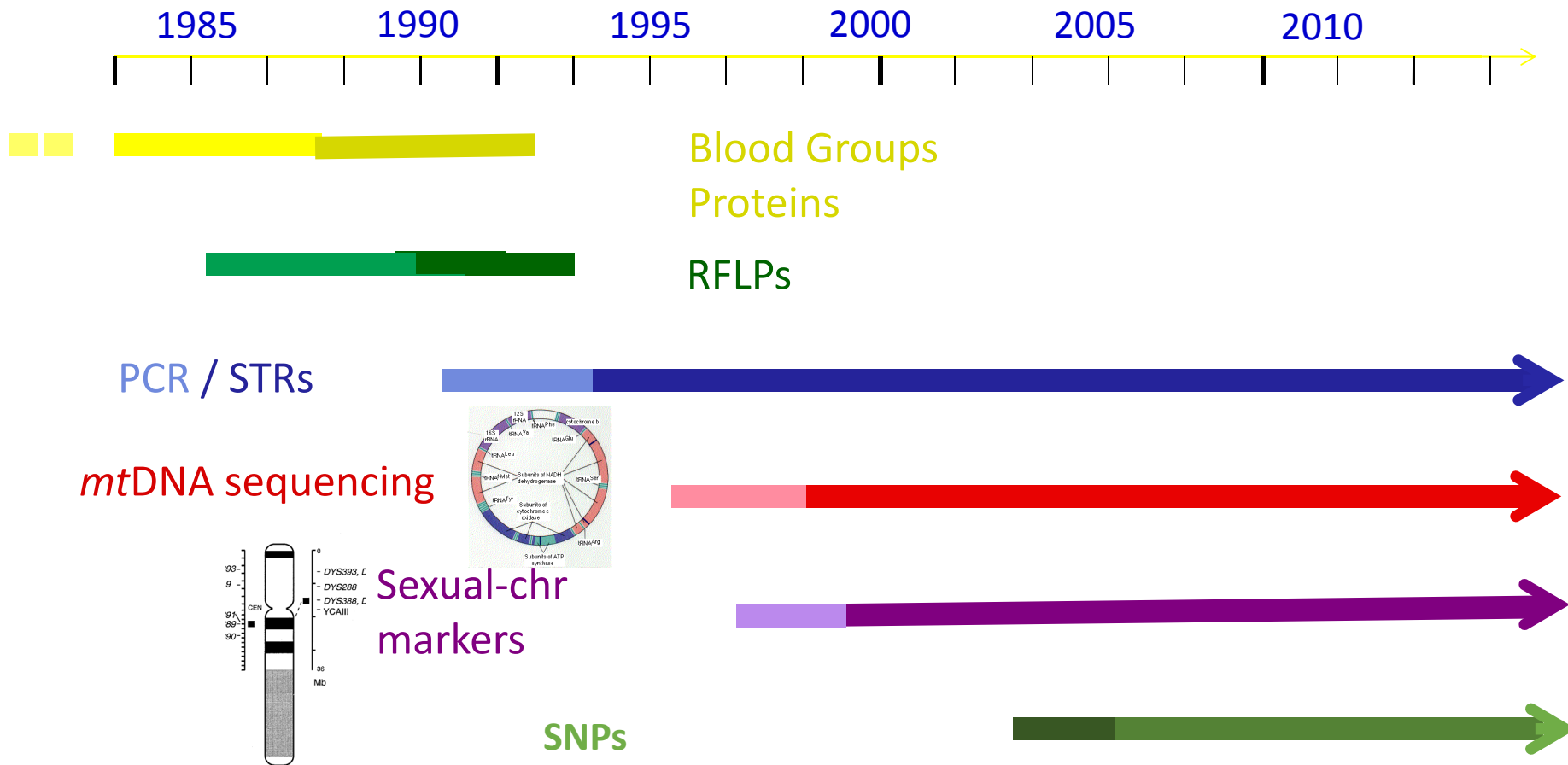
China Uses DNA to Track Its People, With the Help of American Expertise

The Chinese authorities turned to a Massachusetts company and a prominent Yale researcher as they built an enormous system of surveillance and control.

first country in the world to collect DNA from all citizens and visitors:...

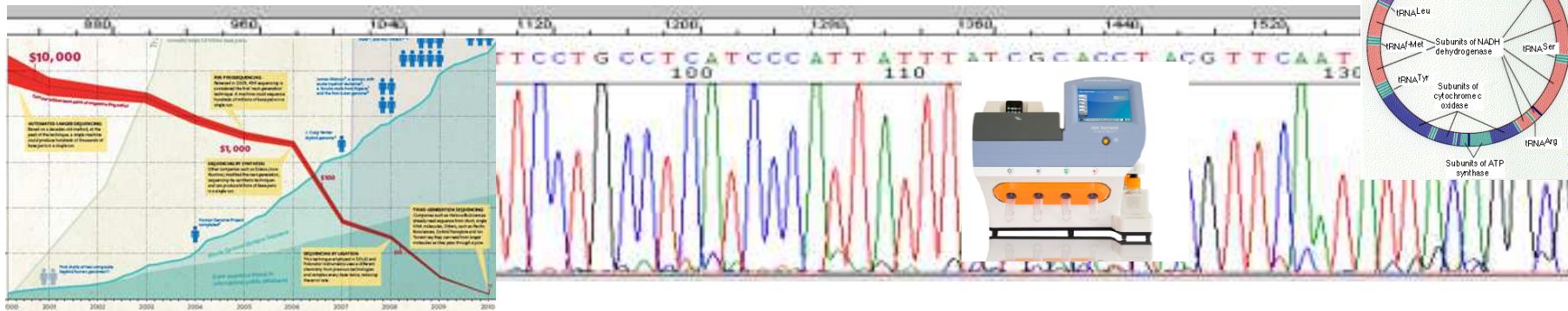
ment power to collect DNA samples is for national security, waiti official.

DNA Typing in Forensic Analysis



mtDNA

There are thousands of copies of mtDNA in each cell compared to two copies of nucDNA, making mtDNA analysis a more sensitive assay, and thus, more successful on highly degraded specimens (e.g., old skeletal material and hair shafts)



Y STRs

Banaheia case
Kristiansand,
Norway

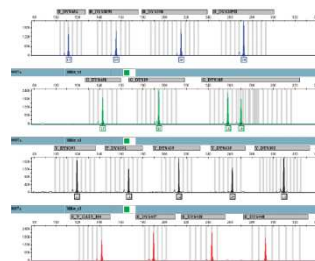
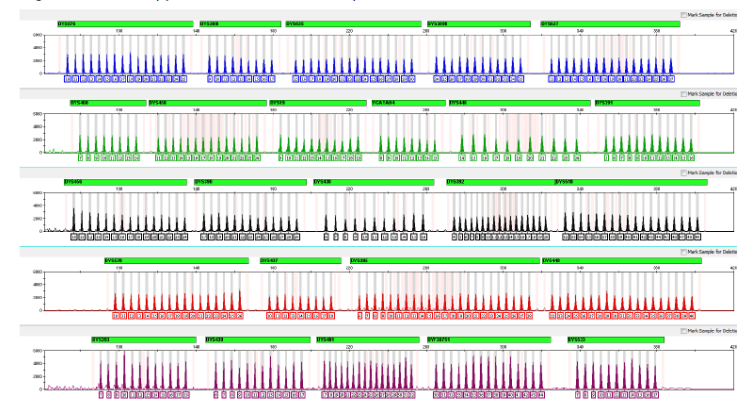


Figure 1 GeneMapper® ID-X Software v1.4 plot of the Yfiler® Plus Allelic Ladder





Mitochondrial DNA Control Region Database

Introduction

The high copy number per cell, the stability against degradation and the maternal mode of inheritance make the mitochondrial (mt) genome particularly suitable for palaeo-, medical and forensic-genetic investigations. Its increased evolutionary rate led to sequence variation that has been generated by sequential accumulation of new mutations along radiating maternal lineages during human dispersal into different parts of the world. Forensic molecular biology takes advantage of this variation for human identity testing by sequence analysis of hypervariable segments within the mtDNA control region. MtDNA analysis is a powerful tool for identifying individuals from the same individual/matriline. If two samples cannot be distinguished by mtDNA analysis, the significance of the mtDNA match needs to be assessed by making reference to the frequency with which that particular mtDNA sequence (= mtDNA haplotype) has been observed in a relevant population.

Concept

The **EMPOP Database** aims at the collection, quality control and the searchable presence of mtDNA control region haplotypes from all over the world. The EMPOP project is a scientific collaboration between the Institute of Legal Medicine (GMI), Innsbruck Medical University and laboratories performing mtDNA research.

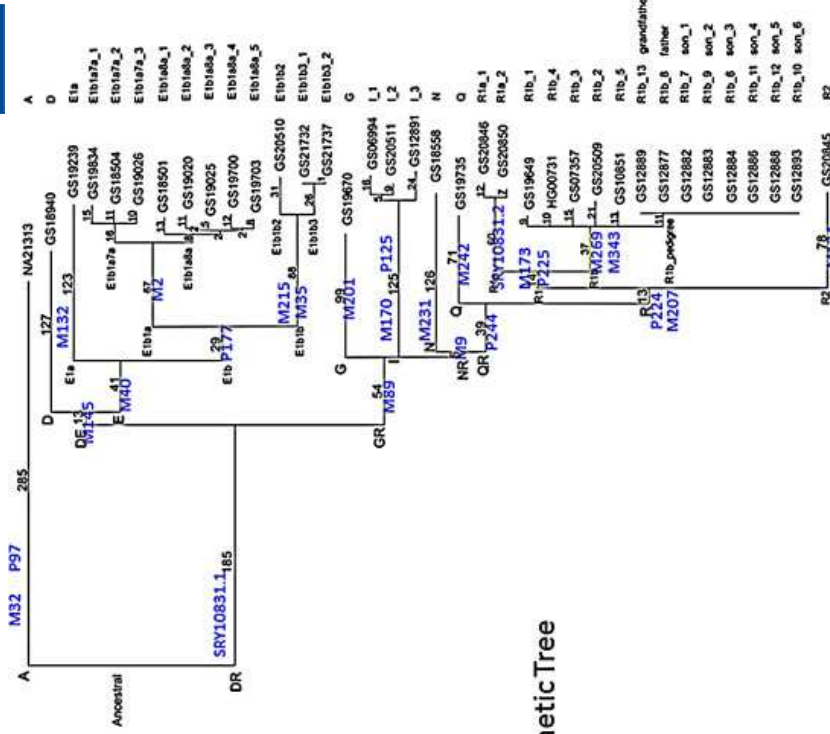


Release 22:
52,655 Haplotypes
in 464 populations



About the "YHRD - Y Chromosome Haplotype Reference Database"

Please cite the database as follows: **Willuweit S, Roewer L, et al. (2007) Y chromosome haplotype reference database (YHRD): Update, Forensic Science International: Genetics 1(2) 83-87** ([external link](#))

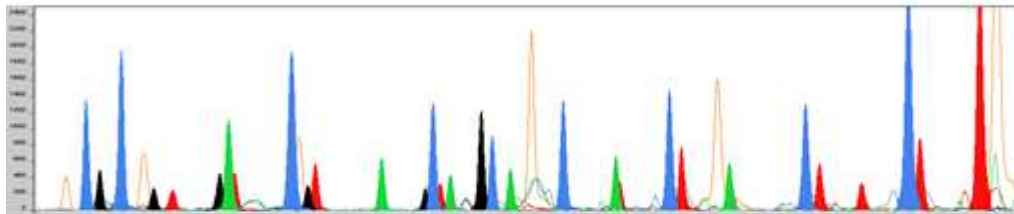


SNP: SINGLE NUCLEOTIDE POLYMORPHISM

ATCGGCGTACCTGATTCCGAATCCGTATCG
ATCGGCGTACCTGAATCCGAATCCGTATCG

The SNP for ID Consortium

The First Autosomal Multiplex



Electrophoresis 2008, 27, 1713-1724

Juan J. Sanchez¹
Chris Phillips²
Clare Bævre³
Kings Baboo⁴
Magdalena Bogus⁵
Manuel Fondeville⁶
Cheryl D. Harrison⁷
Esther McGuire-Brown⁸
Antonio Salas⁹
Denise Syndercombe-Court¹⁰
Peter M. Schneider¹¹
Angel Carracedo¹²
Nils Morling¹³

¹Department of Forensic Genetics,
Institute of Forensic Medicine,
University of Copenhagen,
Copenhagen, Denmark
²Institute of Legal Medicine,
University of Santiago de
Compostela,
Santiago de Compostela, Spain
³Institute of Legal Medicine,
Johannes Gutenberg University,
Mainz, Germany
⁴Centre for Haemoglobinopathy
Research, Barts and The London,
Queen Mary's School of Medicine
and Dentistry,
London, UK

Research Article

A multiplex assay with 52 single nucleotide polymorphisms for human identification

A total of 52 SNPs reported to be polymorphic in European, Asian and African populations were selected. Of these, 42 were from the distal regions of each autosome (except chromosome 15), heavily all selected SNPs were located at least 100 kb distant from known genes and commonly used STRs. We established a highly sensitive and reproducible SNP-typing method with amplification of all 52 DNA fragments in one PCR reaction followed by detection of the SNPs with two single base extension reactions analysed using CE. The amplicons ranged from 88 to 115 bp in length. Complete SNP profiles were obtained from 500 pg DNA. The 52 loci were efficiently amplified from degraded samples where previously only partial STR profiles had been obtained. A total of 700 individuals from Denmark, Greenland, Somalia, Turkey, China, Germany, Taiwan, Thailand and Japan were typed, and the allele frequencies estimated. All 52 SNPs were polymorphic in the three major population groups. The mean match probability was at least 5.0×10^{-14} in the populations studied. Typical paternity indices ranged from 336 000 in Asians to 549 000 in Europeans. Details of the 52 SNP loci and population data generated in this work are freely available at <http://www.snpsforid.org>.

Keywords: Autosomes / Human identification / Multiplex PCR / Single base extension / Single nucleotide polymorphism
DOI 10.1002/elps.200500671

JOHANNES
GUTENBERG
UNIVERSITÄT
MAINZ



Queen Mary
University of London

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

small is 
beautiful



SNPforID

Degraded samples

Parental testing:
Distant relationships

DNA Phenotyping

SNP performance in critical samples

ISFG'05 Azores

SNPforID consortium

Saliva stains degraded for 147 days

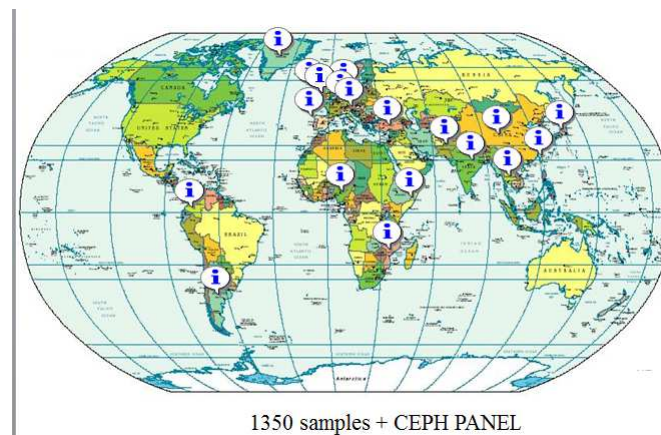
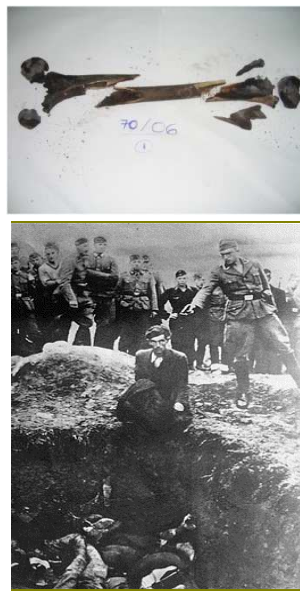
81% complete SNP profile /STRs were only 18% complete

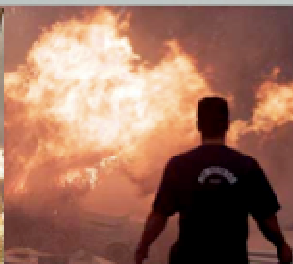
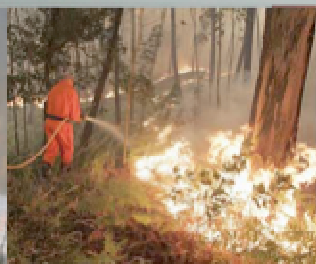
Blood degraded for 243 days

100% full SNP profiles /only 9% with STRs

Low copy number from personal belongings

Full STR profile 24% / complete SNP profiles 92%





Casework application challenging DNA analysis



- Charred human bones found by fire investigators following a severe forest fire 8-06
- An adult male had been reported missing in the area 10 years previously
- A surviving daughter was sampled to allow comparison of any genotypes obtained
- Modified phenol-chloroform extraction of N_2 pulverized dust from white femur
- Identifiler - PP16 - mini NC01 & NC02 - Auto1 & Auto2 - PS-34plex - mtDNA

A new SNP assay for identification of highly degraded human DNA

A. Freire-Aradas^a, M. Fondevila^a, A.-K. Kriegel^b, C. Phillips^{a,*}, P. Gill^{c,d}, L. Prieto^e,
P.M. Schneider^b, Á. Carracedo^a, M.V. Lareu^a

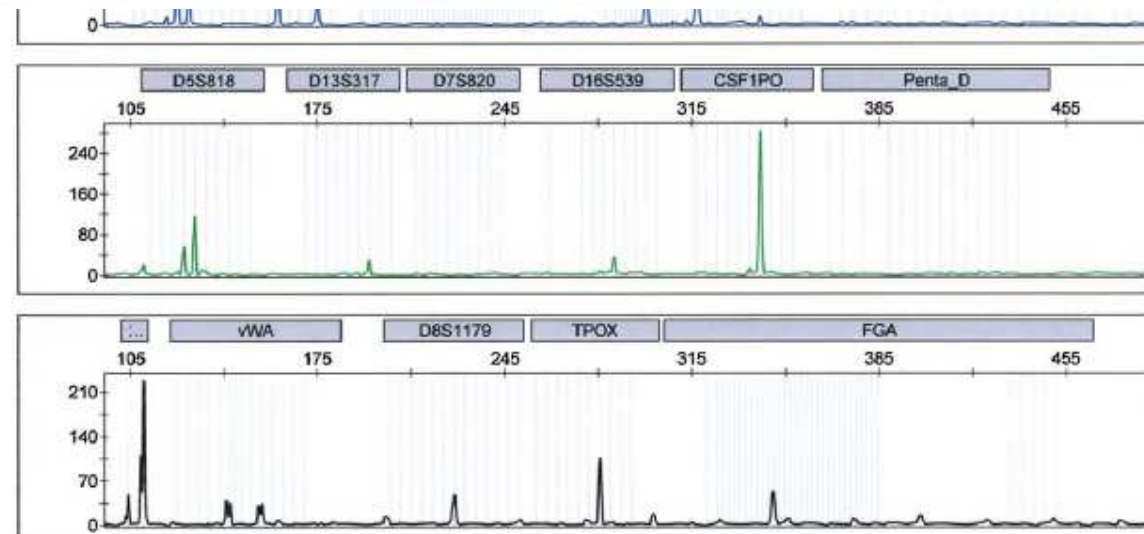
^a Forensic Genetics Unit, Institute of Legal Medicine, University of Santiago de Compostela, Spain

^b Institute of Legal Medicine, University Hospital, University of Cologne, Germany

^c Institute of Legal Medicine, University of Oslo, Rikshospitalet, Norway

^d Forensic Science Centre, University of Strathclyde, Glasgow, UK

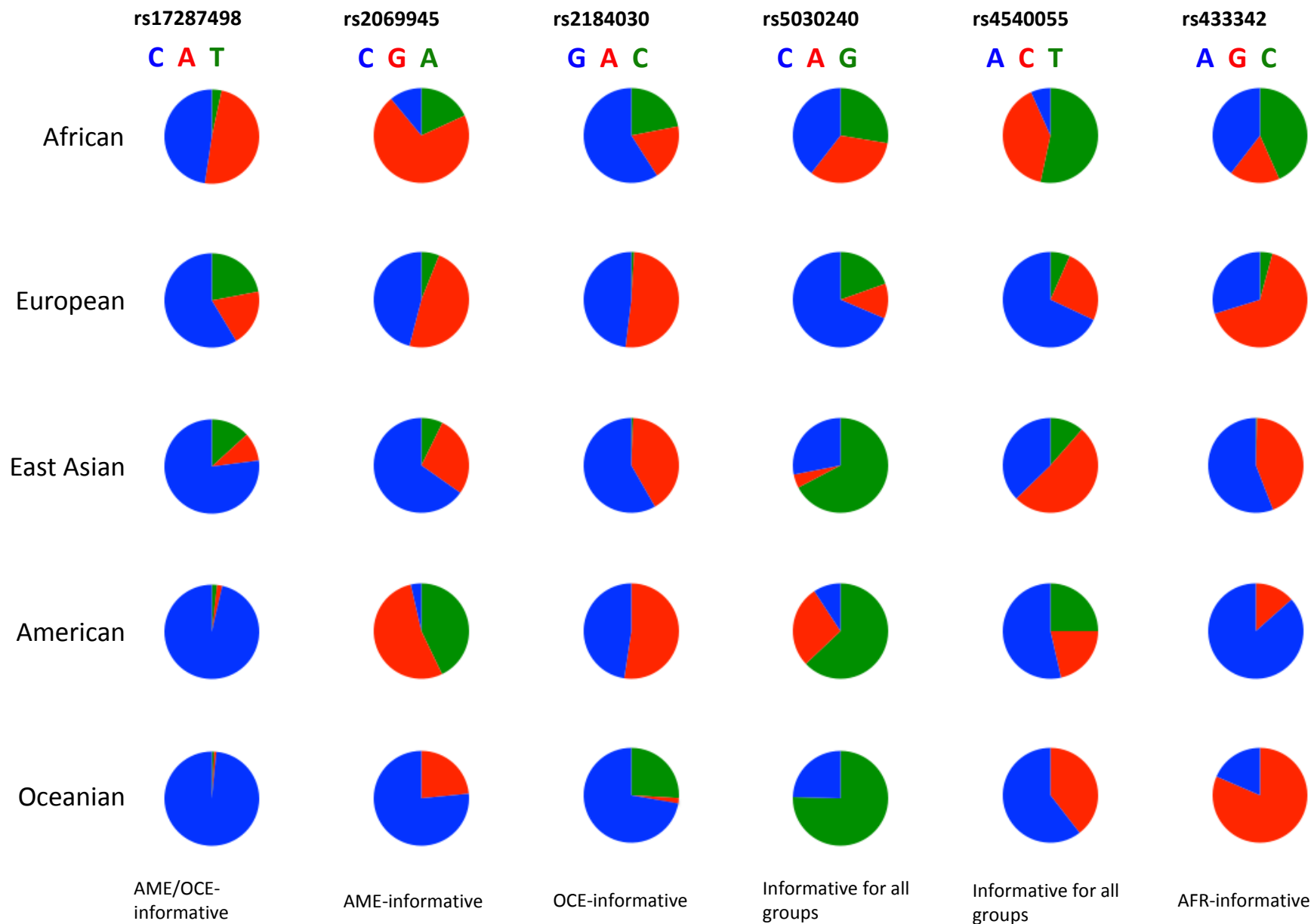
^e University Institute of Research Police Sciences (IUICP), DNA Laboratory, Comisaría General de Policía Científica, Madrid, Spain



Electrophoresis. 2009 Nov;30(21):3682-90.

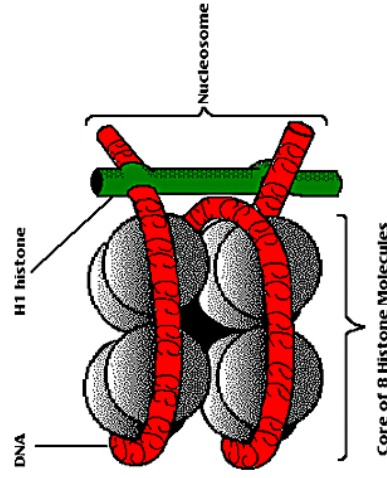
A new multiplex for human identification using insertion/deletion polymorphisms.

[Pereira R](#), [Phillips C](#), [Alves C](#), [Amorim A](#), [Carracedo A](#), [Gusmão L](#).

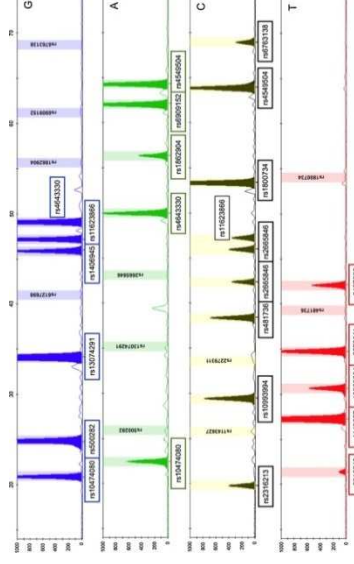
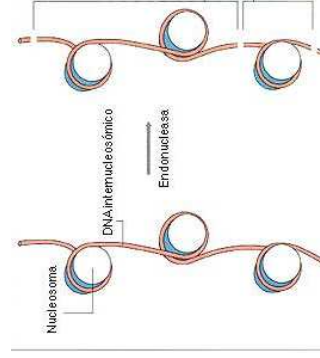


A new SNP assay for identification of highly degraded human DNA

A. Freire-Aradas^a, M. Fondevilla^a, A.-K. Kriegel^b, C. Phillips^{a,*}, P. Gill^{c,d}, L. Prieto^e,
P.M. Schneider^b, Á. Carracedo^a, M.V. Lareu^a




Nucleosome

[illegible]

The screenshot shows a window titled "Example" with a white background. It contains the following elements:

- A text area with the text: "Enter sequence in plain format
 ○ from Screen (*exit de jargon*)..."
- A "Scan" button at the bottom left.
- A "Clear" button at the bottom center.
- A "Confidence level" label above a text input field containing "0.95".
- A "Standard deviation by dispersion" checkbox, which is checked.
- A "From File..." button at the bottom right.
- An "Examine..." button at the bottom right, below the "From File..." button.



Single Nucleotide Polymorphism

Pooled Nucleotide Profiles: Genome, Structure, PolySnp, Transome, QMM, Books, SNP

Search for SNP on NCBI Reference Assembly

Search EnzSnp: SNP

Reference SNP(rsNP) Cluster Report: rs1143627

rsNP ID: rs1143627
 Organism: human (Homo sapiens)
 Allele: SNP
 Variation: C>G
 Single nucleotide polymorphism
 RefSNP Alleles: C/T
 Ancestral Allele: C
 Clinical Association: unknown

Molecule Type: Genomic
 Created/Updated in build: 86129
 Map to Genome Build: 36.3
 Citation: [PubMed](#)

Should not be frightened of difficult DNA

DNA from Pre-Clovis Human Coprolites in Oregon, North America

M. Thomas P. Gilbert,^{1,*} Dennis L. Jenkins,^{2,*} Anders Götherström,³ Nuria Naveran,⁴ Juan J. Sanchez,⁵ Michael Hofreiter,⁶ Philip Francis Thomsen,¹ Jonas Biniadze,¹ Thomas F. G. Higham,⁷ Robert M. Yohe II,⁸ Robert Parr,⁹ Linda Scott Cummings,⁹ Eske Willerslev^{1,†}

9 MAY 2008 VOL 320 SCIENCE



Available online at www.sciencedirect.com

ScienceDirect

Forensic Science International: Genetics 3 (2008) 212–218



www.elsevier.com/locate/FSIG

Case report: Identification of skeletal remains using short-amplicon marker analysis of severely degraded DNA extracted from a decomposed and charred femur

M. Fondevila^a, C. Phillips^{a,b,*}, N. Naveran^a, I. Fernandez^a, M. Cerezo^a, A. Salas^{a,b}, A. Carracedo^{a,b}, M.V. Lareu^{a,b}

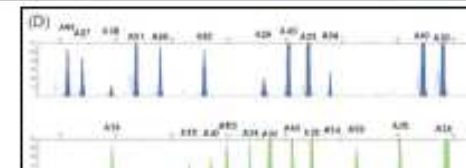
14KYA 5-plex SNaPshot
2-SNP pyrosequencing mt codin

WTC (New York)



Double-degraded DNA 29-plex, 23-plex & 4-plex SNaPshot

used ancient DNA extraction system



Int J Legal Med (2009) 123:315–325
DOI 10.1007/s00414-009-0248-5

ORIGINAL ARTICLE

Pigment phenotype and biogeographical and from ancient skeletal remains: inferences from multiplexed autosomal SNP

Caroline Brunkage · Christine Keyser · Eric Croubey · Daniel Montagnon · Bertrand Ludes

up to 5KYA 1+8+1-plex SNaPshot
AIMs and pigmentation SNPs

SNPs improves mini-STR sizes
to combat degradative process



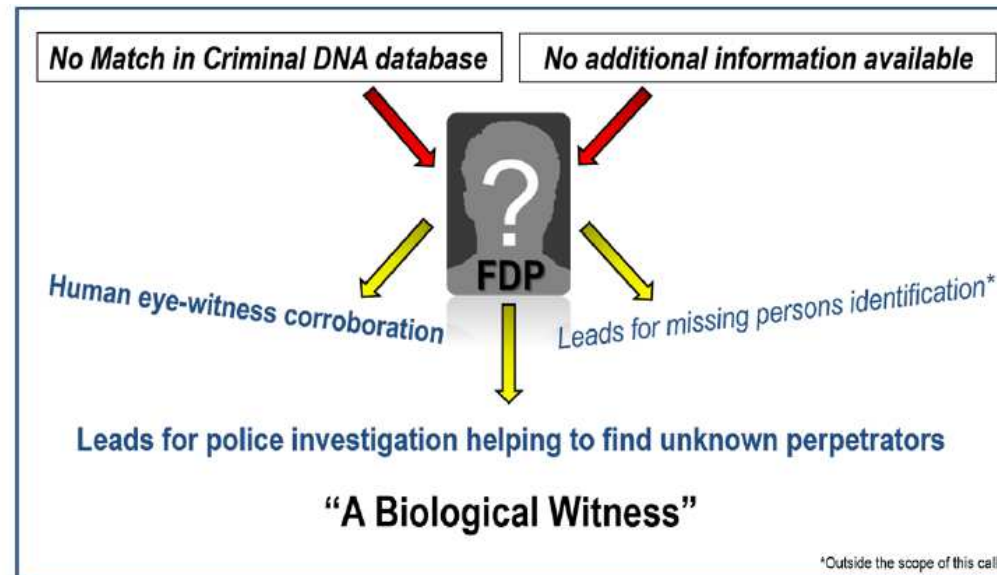
Tsunami



What is Forensic DNA Phenotyping

- Witness of a crime is usually asked about sex, appearance and age of the perpetrator
- DNA variants can be informative in terms of appearance phenotype and age prediction
- **FDP** refers to DNA prediction of externally visible traits of unknown sample donor from biological traces collected at the scene of a crime

FDP is an investigative tool



- Matching DNA profiles only works for known suspects
 - reference DNA profile match
 - DNA database profile match
- DNA mass screenings – when no match is found
- FDP allows for more precise selection of target group when searching for unknown suspects
- FDP may be useful in missing persons identification

The Innocence Project, founded in 1992 by Peter Neufeld and Barry Scheck at Cardozo School of Law, exonerates the wrongly convicted through DNA testing and reforms the criminal justice system to prevent future injustice.



"Injustice anywhere is a threat to justice everywhere." Martin Luther King, Jr.

342

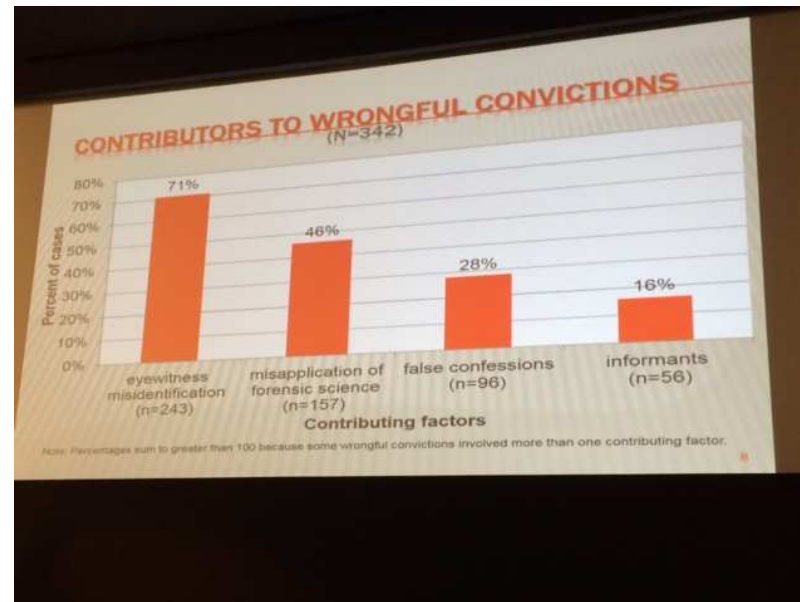
Number of DNA
Exonerations

147

Number of Real
Perpetrators found



71%

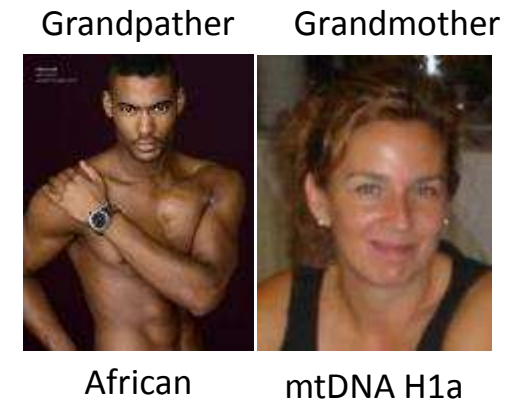
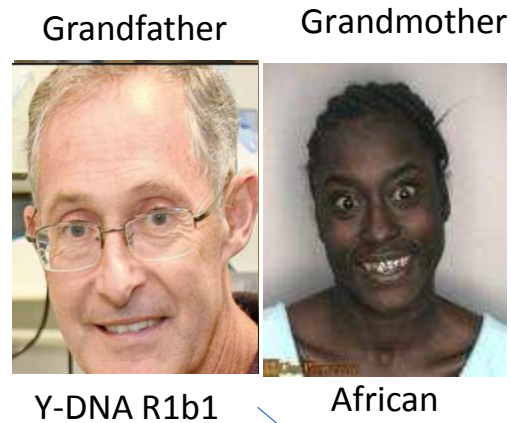


Forensic DNA Phenotyping

1. External visible characteristics
EVC: Pigmentation traits, body height, hair morphology, ear morphology, facial shape, and for some age-related appearance traits such as hair loss, pigmented age spots.
2. Bio- geographic ancestry i.e., the geographic region of origin of a person's biological ancestors
3. Estimation of age



Why mtDNA and Y chr are not ideal markers for forensic DNA phenotyping



I'm European
and white! Look
at my DNA!!

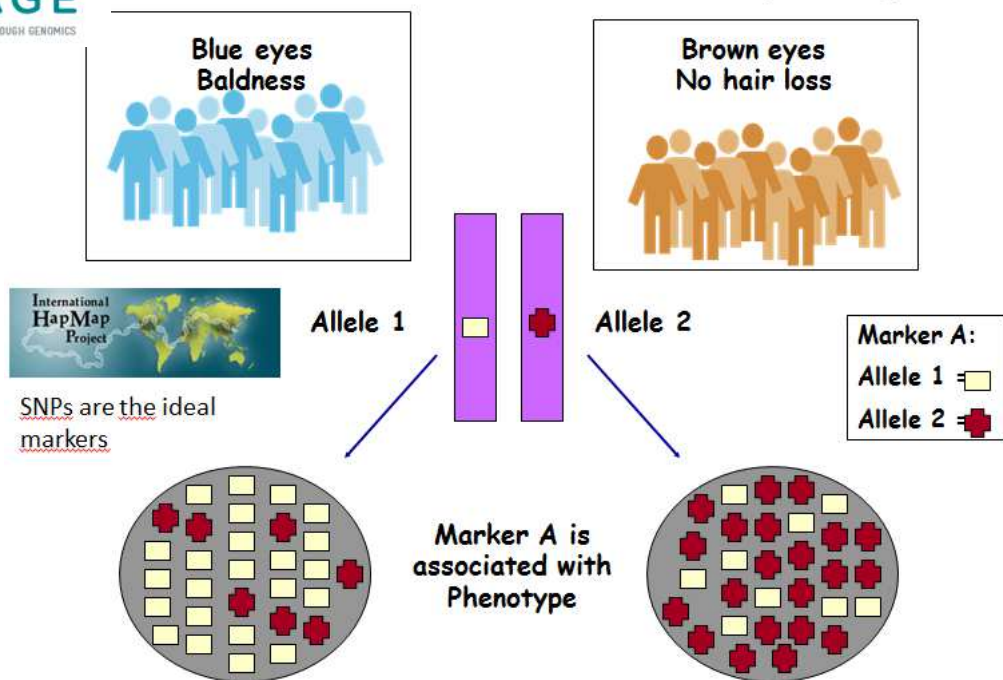
Genetics of Phenotypic Traits

- **Monogenic traits: One gene explains all heritability**
 - Mendelian inheritance (Mendelian traits)
 - E.g. Lactose intolerance (Chr2, *MCM6-LCT*)
- **Polygenic traits: Many genes, all with small effects**
 - Complex pattern of inheritance
 - Most appearance traits, e.g. body height (700 SNP: 16% variance Nat Genet 2014)
- **The middle class: One or few major gene(s) + many additional genes with minor effects**
 - Semi-Mendelian inheritance
 - E.g. blue / brown eye color (major effect: Chr15, *HERC2*)
 - E.g. male pattern baldness (major effect: ChrX, *AR*)

How To Find Genes / Predictive Markers

- **Linkage Analysis:** use of related individuals / family samples
 - suitable to identify rare alleles with relatively large phenotypic effects (Mendelian traits)
 - Candidate region or genome-wide approach
- **Association Testing:** use of unrelated individuals / population samples
 - Suitable to identify common alleles with medium to small phenotypic effects (polygenic / complex traits)
 - Candidate gene or **genome-wide association study (GWAS)**
 - Large sample size to *safely* identify small effect genes / SNPs
 - Association threshold for SNP array GWAS: $P < 5 \times 10^{-8}$
 - Replication in independent samples needed to trust results

Human Genetic Association Study Design



SNPs are the ideal markers

Instituto de Salud Carlos III

DE ECONOMIA Y COMPETITIVIDAD

Scientific advisory board

Board

Coordination **USC**

NODE 1
Santiago de Compostela (USC)

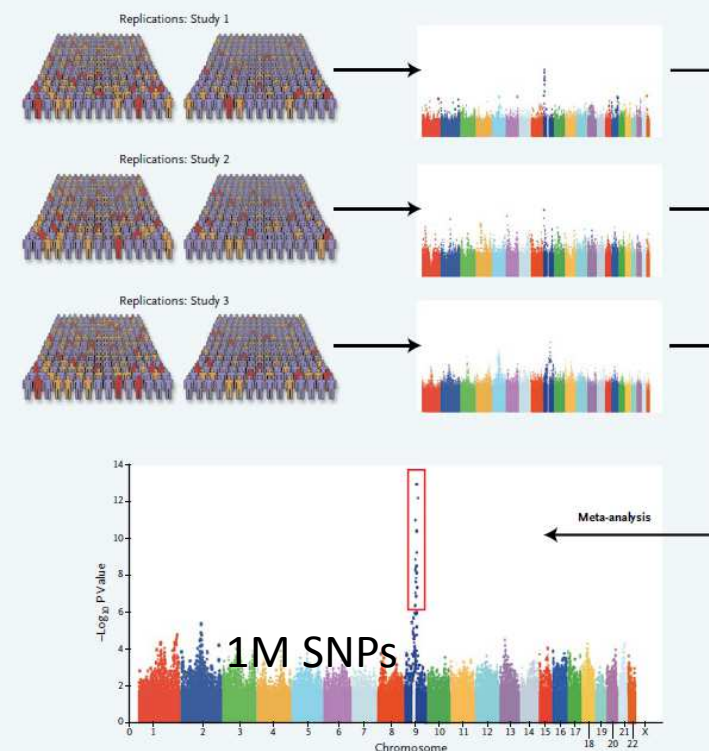
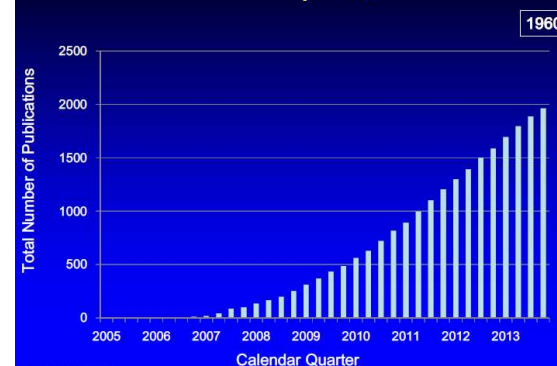
USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Management unit

NODE 2
Madrid (CNIO)

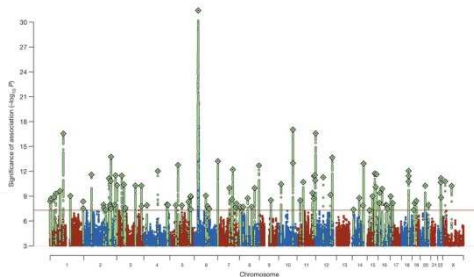
CNIO
Centro Nacional de Investigaciones Oncológicas

Published GWA Reports, 2005 – 2013



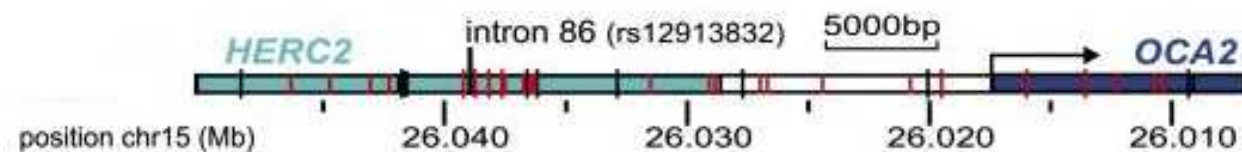
Discovery of pigmentation markers

- Animal models and albinism affected people indicate **candidate genes in humans**
- ***MC1R*** – the red hair color and fair skin gene
- ***TYR, OCA2, TYRP1, SLC45A2*** – albinism genes, candidates for natural eye, hair and skin color variation
- >10 Successful **GWA studies**:
 - Confirmation of already known loci associated with pigmentation (*TYR, MC1R, OCA2, TYRP1, ASIP, SLC45A2, SLC24A5*)
 - Novel pigmentation and non-pigmentation genes
 - Eye colour: *HERC2, SLC24A4, IRF4, LYST, DSCR9*
 - Hair colour: *KITLG, IRF4, EXOC2, TPCN2, HERC2*
 - Skin colour and type: *IRF4, LOC401937*



HERC2 – powerful eye colour regulator

- **rs12913832** in *HERC2* – key predictor in all eye colour prediction models
- This polymorphism regulates expression of *OCA2* – the important pigmentation gene



[HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter.](#)

Visser M, Kayser M, Palstra RJ.

Genome Res. 2012 Mar;22(3):446-55.

From the SNP to the causal gene – Functional studies

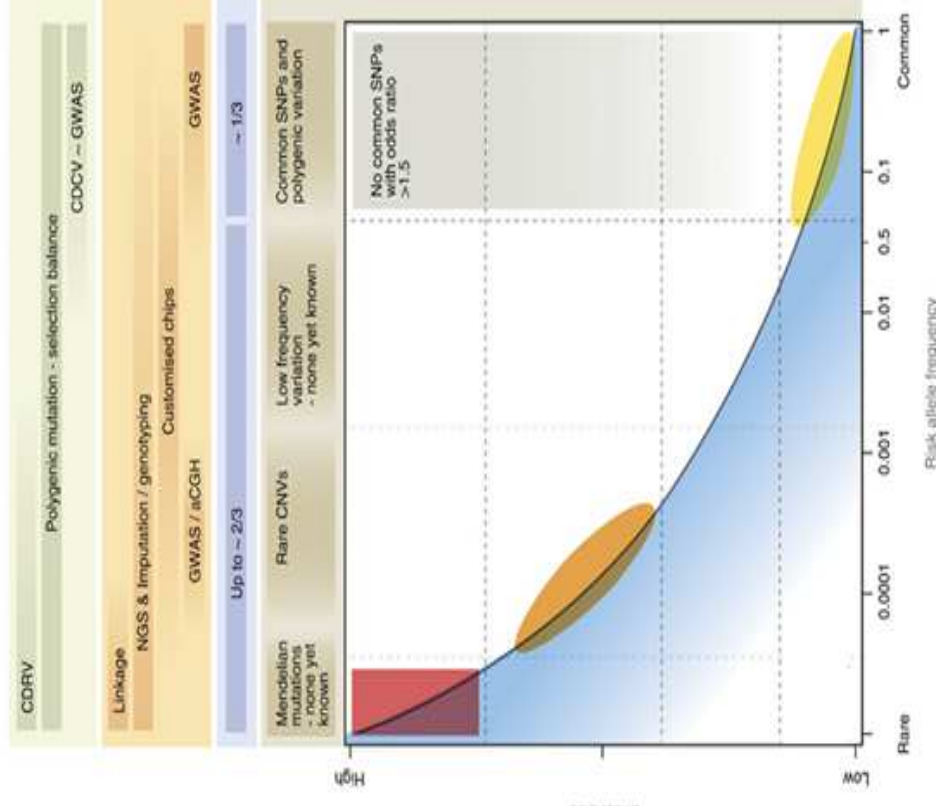
- **Associated SNP is the causal DNA variant**
 - mutation in a codon leads to aminoacid change that changes the protein to influences the trait
- **Most SNPs found associated via SNP arrays not causal**
 - because of e.g. SNP chip design biased towards no-gene regions
 - May have regulatory function i.e. a mutation changes a binding site of a transcription factor which regulates the expression of a gene that directly influences the trait
 - Maybe indirectly associated because of linked inheritance with causal or regulatory SNPs that are not represented on the SNP chip used

Power of GWAS depends on:

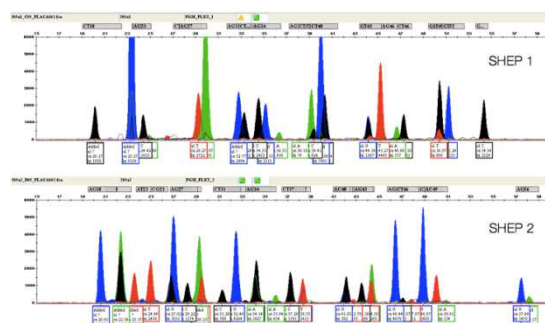
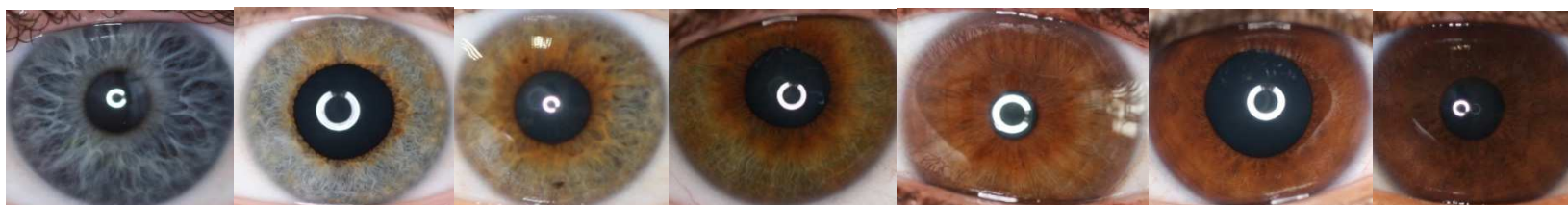
- Trait heritability (environmental impact)
- Genotyping tools / marker types used (SNP arrays versus sequencing)
- Sample size
- Allele effect size
- Effect allele frequency
- Genetic model used
- Reliable association results: $p < 5 \times 10^{-8}$ and replicated in independent samples
- At best with causal / regulatory effect experimentally proven

What's not found by SNP array GWAS?

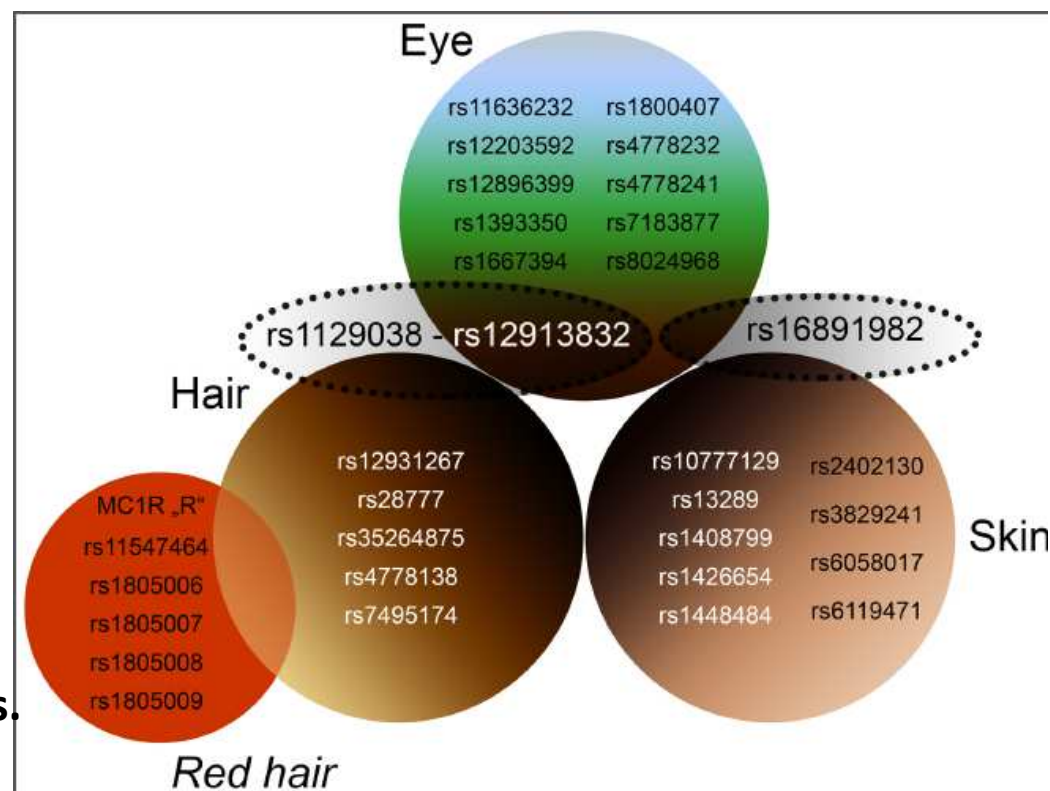
- DNA variants with very small effects
- Rare DNA variants
- Recessive DNA variants
- Allelic heterogeneity
- Genetic interaction effects
- Compound heterozygote effects
- = **Missing heritability of complex traits**



Kettunen et al. 2012



[Forensic Sci Int Genet.](#) 2015
Sep;18:33-48. **Forensic DNA
Phenotyping: Predicting human
appearance from crime scene
material for investigative purposes.**
[Kayser M](#)



The Genetics of Skin, Hair, and Eye Color Variation and Its Relevance to Forensic Pigmentation Predictive Tests

O. Maroñas¹, J. Söchtig¹, Y. Ruiz², C. Phillips^{1*}, Á. Carracedo^{1,3,4}, M. V. Lareu¹

Welcome to the Department of Genetic Identification of Erasmus MC

With the advancement of DNA phenotyping as a tool in Forensic and Anthropological usage, we now provide an easy to use interactive website to predict eye, hair and skin colour from DNA using the IrisPlex, HlrisPlex and HlrisPlex-S systems.

This work is in collaboration with the [Walsh laboratory](#) of Indiana-University-Purdue-University-Indianapolis (IUPUI), USA.

The HlrisPlex-S System



| Gene | SNP | Allele | No. of Alleles |
|------------------------|-------------|--------|----------------|
| 1 <i>MC1R</i> | rs312262906 | A | 0 1 2 NA |
| 2 <i>MC1R</i> | rs11547464 | A | 0 1 2 NA |
| 3 <i>MC1R</i> | rs885479 | T | 0 1 2 NA |
| 4 <i>MC1R</i> | rs1805008 | T | 0 1 2 NA |
| 5 <i>MC1R</i> | rs1805005 | T | 0 1 2 NA |
| 6 <i>MC1R</i> | rs1805006 | A | 0 1 2 NA |
| 7 <i>MC1R</i> | rs1805007 | T | 0 1 2 NA |
| 8 <i>TUBB3</i> | rs1805009 | C | 0 1 2 NA |
| 9 <i>MC1R</i> | rs201326893 | A | 0 1 2 NA |
| 10 <i>MC1R</i> | rs2228479 | A | 0 1 2 NA |
| 11 <i>MC1R</i> | rs1110400 | C | 0 1 2 NA |
| 12 <i>SLC45A2</i> | rs28777 | C | 0 1 2 NA |
| 13 <i>SLC45A2</i> | rs16891982 | C | 0 1 2 NA |
| 14 <i>KITLG</i> | rs12821256 | G | 0 1 2 NA |
| 15 <i>LOC105374875</i> | rs4959270 | A | 0 1 2 NA |
| 16 <i>IRF4</i> | rs12203592 | T | 0 1 2 NA |
| 17 <i>TYR</i> | rs1042602 | T | 0 1 2 NA |
| 18 <i>OCA2</i> | rs1800407 | A | 0 1 2 NA |
| 19 <i>SLC24A4</i> | rs2402130 | G | 0 1 2 NA |
| 20 <i>HERC2</i> | rs12913832 | T | 0 1 2 NA |
| 21 <i>PIGU</i> | rs2378249 | C | 0 1 2 NA |
| 22 <i>LOC105370627</i> | rs12896399 | T | 0 1 2 NA |
| 23 <i>TYR</i> | rs1393350 | T | 0 1 2 NA |
| 24 <i>TYRP1</i> | rs683 | G | 0 1 2 NA |
| 25 <i>ANKRD11</i> | rs3114908 | T | 0 1 2 NA |

| | | | |
|-------------------|------------|---|----------|
| 26 <i>OCA2</i> | rs1800414 | C | 0 1 2 NA |
| 27 <i>BNC2</i> | rs10756819 | G | 0 1 2 NA |
| 28 <i>HERC2</i> | rs2238289 | C | 0 1 2 NA |
| 29 <i>SLC24A4</i> | rs17128291 | C | 0 1 2 NA |
| 30 <i>HERC2</i> | rs6497292 | C | 0 1 2 NA |
| 31 <i>HERC2</i> | rs1129038 | G | 0 1 2 NA |
| 32 <i>HERC2</i> | rs1667394 | C | 0 1 2 NA |
| 33 <i>TYR</i> | rs1126809 | A | 0 1 2 NA |
| 34 <i>OCA2</i> | rs1470608 | A | 0 1 2 NA |
| 35 <i>SLC24A5</i> | rs1426654 | G | 0 1 2 NA |
| 36 <i>ASIP</i> | rs6119471 | C | 0 1 2 NA |
| 37 <i>OCA2</i> | rs1545397 | T | 0 1 2 NA |
| 38 <i>RALY</i> | rs6059655 | T | 0 1 2 NA |
| 39 <i>OCA2</i> | rs12441727 | A | 0 1 2 NA |
| 40 <i>MC1R</i> | rs3212355 | A | 0 1 2 NA |
| 41 <i>DEF8</i> | rs8051733 | C | 0 1 2 NA |

Display Predicted Phenotype

Download Predicted Phenotype

<https://hlrisplex.erasmusmc.nl/>

Human height – small additive effects

- 70-90% of variance of human height may be due to genetic component
- 35 GWA studies and meta-analyses on height until 2016
- Additive effects of multiple loci with small effect (OR ~ 1.1) explain variance of human height
- **Lango-Allen** et al. (2010): 10% of height variance explained by 180 loci in the human genome (prediction of extremely tall stature with **AUC=0.75**)
- **Wood** et al. (2014) – 9500 SNPs explains 29% of height variance

Update on the predictability of tall stature from DNA markers in Europeans

Fan Liu^{a,b,*}, Kaiyin Zhong^b, Xiaoxi Jing^a, André G. Uitterlinden^{c,d}, A. Emile J. Hendriks^e, Stenvert L.S. Drop^f, Manfred Kayser^{b,*}

DOI: <https://doi.org/10.1016/j.fsigen.2019.05.006>

Article Info



Abstract

Full Text

Images

References

Supplemental Materials

Highlights

- Current forensic DNA phenotyping focuses on pigmentation traits.
- Update on DNA predictability of tall stature in Europeans is presented.
- 689 SNPs provided AUC of 0.79, while a subset of 412 SNPs achieved 0.76.
- New models improved prediction accuracy compared to previous ones.



Facial morphology

- Genes involved in cleft lip and other craniofacial pathologies – candidate loci for natural variation (e.g. **Boehringer** et al. 2011)
- **Liu** et al. (2012) – GWAS identified *PRDM16*, *PAX3*, *TP63*, *C5orf50* and *COL17A1*
- Small effect size of the identified DNA variants
- **Claes** et al. (2014a, b) – complex mathematical approach for predictive modeling involving ancestry, sex and genes
 - 24 SNPs improve by 1% accuracy of face morphology prediction

Male pattern baldness

- Very high estimated heritability 85-95%
- **Brockschmidt** et al. (2011) – *AR-EDA2R*, 20p11.22
- **Li** et al. (2012) – *AR-EDA2R*, *PAX1*, *HDAC4*, *HDAC9*, *TARDBP*, 17q21.31, *AUTS2*, *SETBP1*
- *AR-EDA2R* – confirmed in Europeans in several studies
- **MPB risk is predictable at an accuracy level of 0.74 when 14 SNPs were included in the model (Liu et al, 2016)**
- EUROFORGEN-NoE project on baldness up to 0.80

[Liu et al. Eur J Hum Genet.](#) 2016 Jun;24(6):895-902..

Prediction of male-pattern baldness from genotypes.

Hair morphology

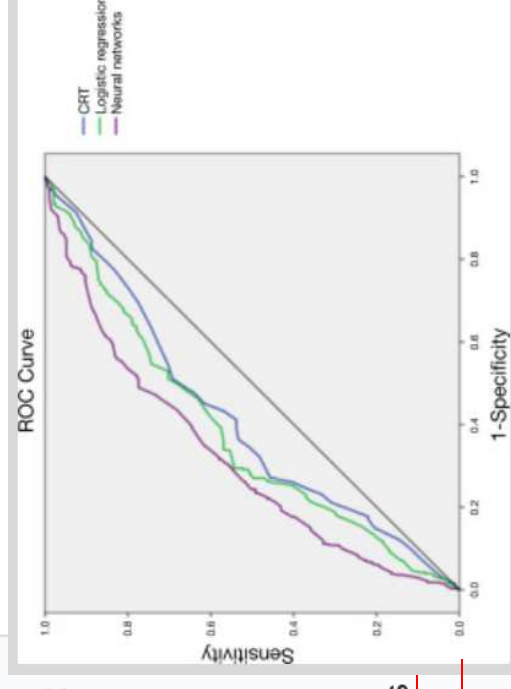
- Heritability 85-95% as estimated from twins study
- Very variable among Europeans (straight-curly-wavy)
- One GWA study and several candidate studies
- **Medland** et al. (2009) – strongest association with *TCHH* on 1q21.3
- Merely 6% of variance of hair morphology explained
- **Eriksson** et al. (2010) - other candidate genes: *WNT10A* (chr 2), *FRAS1* (chr 4), *LCE3E* (chr 1), *OFCC1* (chr 6)

Research paper

Evaluation of the predictive capacity of DNA variants associated with straight hair in Europeans

Ewelina Pośpiech^{a,*}, Joanna Karłowska-Pik^b, Magdalena Marcińska^c, Sarah Abidi^d, Jeppe Dyrberg Andersen^e, Margreet van den Berge^f, Ángel Carracedo^{g,h}, Mayra Eduardoffⁱ, Ana Freire-Aradas^g, Niels Morling^e, Titia Sijen^f, Małgorzata Skowron^j, Jens Söchtig^g, Denise Syndercombe-Court^d, Natalie Weiler^f, Peter M. Schneider^k, David Ballard^d, Claus Børsting^e, Walther Parson^{l,1}, Chris Phillips^g, Wojciech Branicki^{a,2}

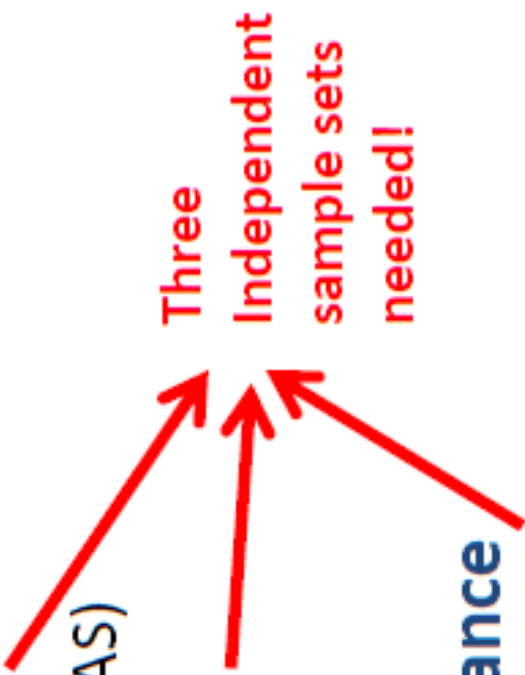
- A replication study was made of SNPs most closely associated with hair morphology variation in Europeans (assigning straight, wavy and curly hair phenotypes).
- Analysis of 670 samples from seven European populations revealed the strongest association for rs11803731 in *TCHH*, rs7349332 in *WNT10A* and rs1268789 in *FRAS1*.
- Applying three different mathematical models to assess the predictive capacity of the SNPs indicated neural networks gives the best performing model to predict straight hair with high sensitivity and better specificity than logistic regression and CRT methods.
- The combined TTGGGG SNP genotype (rs11803731-rs7349332-rs1268789) was identified as the best predictor, giving greater than 80% probability of straight hair.
- The reported study is the first assessment of the forensic suitability of hair morphology as an externally visible characteristic. The results provide the basis for extended analyses of SNPs associated with this trait.



Association is **NOT** Prediction

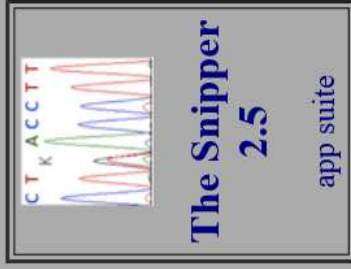
- Prediction accuracy is determined by:
 - (Sample-size)
 - Allele effect size
 - Effect allele frequency
 - Genetic model
- Prediction considers the joint effect of multiple SNPs
- SNPs with high association but strong linkage disequilibrium provide no independent / additional prediction value

Three Steps in Prediction Studies

- **1. Identify DNA predictors**
 - E.g. Association studies (GWAS)
 - **2. Build prediction models**
 - Linear / logistic regressions
 - Other models
 - **3. Evaluate model performance**
 - Categorical traits: Sensitivity, specificity, area under the receiver operating characteristic curves
 - Continuous traits: Explained variance (R^2), mean squared error (MSE)
- Three Independent sample sets needed!
- 

Prediction modeling in forensics

- Regression methods, Bayesian statistics, classification trees
- Prediction model building
 - samples with known phenotype – training set (usually 70-80%)
 - selecting the most accurate model
- Prediction model testing
 - Independent samples with known phenotype – testing set (usually 20-30%)
- Parameters defining performance of the model
 - sensitivity, specificity, PPV, NPV, AUC



[Need help?](#)




Input search string

Search Snipper

[Main Home](#)

Binary AIM classification of individuals

You are entering our portal on SNP-Indel classification, hosting the *Snipper App suite* version 2.5. It is the companion site to [several papers](#).

- **Faster algorithms.** 
- Improved **Multinomial Logistic Regression** (MLR) implementation. 
- Revised **density plots** and **3D views** no longer dependent on Java. 
- A new **tool to predict age** is available. 

Green colour in links means fixed training sets. On the other hand, **red** means custom training sets. Available tasks are:

- 🚩 [Classification as Europe-East Asia-Africa-America-Oceania \(34 SNPs, 46 Indels, or both sets\)](#)
- 🚩 [Classification as individual having black-intermediate-white skin](#)
- 🚩 [Classification as individual having fair-dark or red-blond-brown-black hair](#)
- 🚩 [Classification as individual having blue-greenhazel-brown eyes](#)
- 🚩 [Classification with a custom Excel file of populations](#)
- 🚩 [Classification of multiple profiles with a custom Excel file of populations](#)
- 🚩 [Tune up your training/testing set](#)
- 🚩 [Classification with the 32 STR training set or a custom Excel file of frequencies](#)
- 🚩 [Thorough analysis of population data of a custom Excel file](#)
- 🚩 [Profile generator for fixed or custom training sets](#)
- 🚩 [Plot some or all the populations of a custom Excel file](#)

Age prediction in forensics

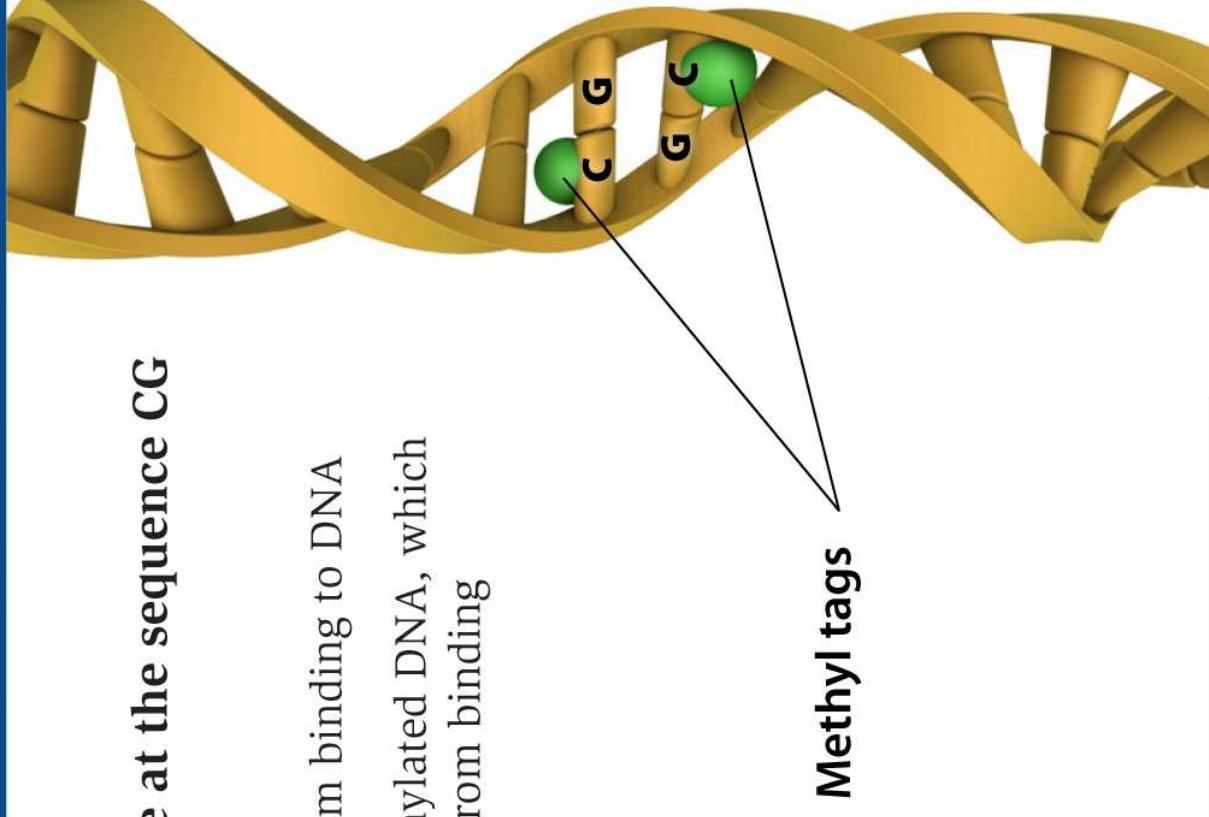
- Age estimation may significantly improve precision of a portray of an unknown individual
- Various methods considered in forensic science
 - shortening of telomeres
 - accumulation of 4977 bp deletion in mitochondrial DNA
- Decrease in the number of sjTREC molecules caused by thymus involution occurring in the course of human life
- **Zubakov** et al. (2010) – accurate prediction of age groups based on sjTREC quantification
- DNA methylation of markers linearly correlated with age

Methyl tags usually turn genes off

Methyl tags are added to a cytosine at the sequence CG

They silence genes by:

- Blocking transcription machinery from binding to DNA
- Recruiting proteins that bind to methylated DNA, which then block transcription machinery from binding



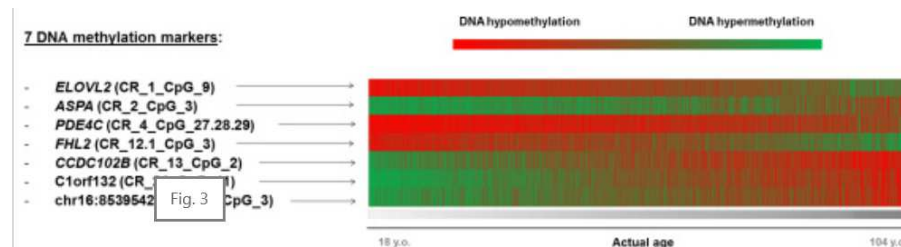
Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system

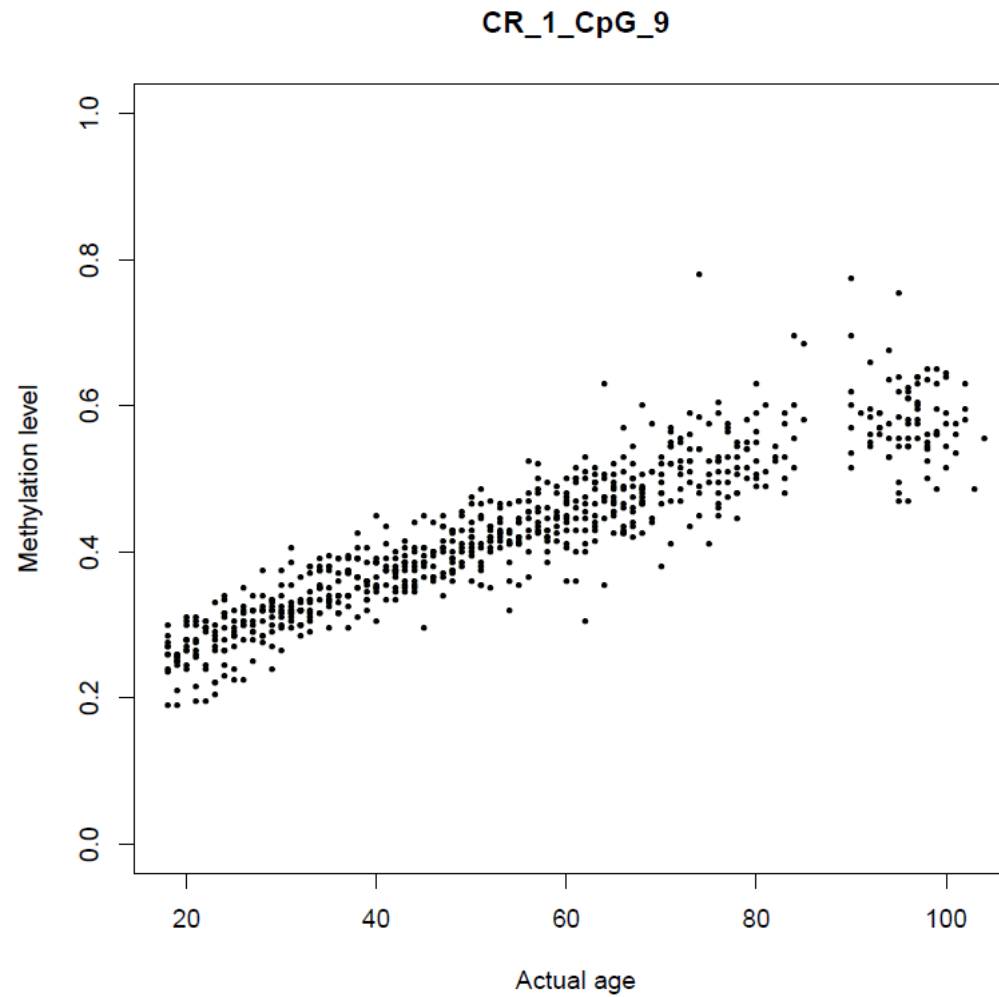


[A. Freire-Aradas](#), [C. Phillips](#), [A. Mosquera-Miguel](#), [L. Girón-Santamaría](#), [A. Gómez-Tato](#), [M. Casares de Cal](#), [J. Álvarez-Dios](#), [J. Ansedo-Bermejo](#), [M. Torres-Español](#), [P.M. Schneider](#), [E. Pośpiech](#), [W. Branicki](#), [Á. Carracedo](#), [M.V. Lareu](#)

FSI Genetics 24: 65-74 (2016)

- Illumina HumanMethylation 450 BeadChip data from populations all over the world.
- 177 CpG sites with informative methylation patterns analyzed with EpiTYPER technology
- *Most informative ELOVL2, ASPA, PDE4C, FHL2, CCDC102B, C1orf132 and chr16:85395429*
- Training set of 725 European individuals (18-104 y) and a test set of 52 monozygotic twin pairs
- Multivariate quantile regression age predictor placed in *Snipper*





- Median absolute age prediction error of ± 3.07 years
- Percentage of prediction error relative to the age of 6.3%

Freire-Arada et al.,
FSIGEN 2016

Development of methylation marker sets for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. Freire-Aradas et al. FSIGEN (2016)

Predicted age with the
panel: 64.48
Real age: 63.50

Age prediction with DNA methylation

Input the methylation values for the seven CpG's shown on the table below. Blanks will be ignored:

| Gene | Internal code | CpGID | GRCh38 chromosome position |
|---------------|-------------------|------------|----------------------------|
| ELOVL2 | CR_1_CpG_9 | cg21572722 | chr6:11044661 |
| ASPA | CR_2_CpG_3 | cg02228185 | chr17:3476273 |
| PDE4C | CR_4_CpG_27.28.29 | none | chr19:18233127/31/33 |
| FHL2 | CR_12_1_CpG_3 | cg06639320 | chr2:105399282 |
| CCDC102B | CR_13_CpG_2 | cg19283806 | chr18:68722183 |
| C1orf132 | CR_21_CpG_11 | none | chr1:207823715 |
| no associated | CR_23_CpG_3 | cg07082267 | chr16:85395429 |

| Name | CR_1_CpG_9 | CR_2_CpG_3 | CR_4_CpG_27.28.29 | CR_12_1_CpG_3 | CR_13_CpG_2 | CR_21_CpG_11 | CR_23_CpG_3 |
|--------|------------|------------|-------------------|---------------|-------------|--------------|-------------|
| Values | 0.37 | 0.535 | 0.21 | 0.485 | 0.18 | 0.455 | 0.415 |

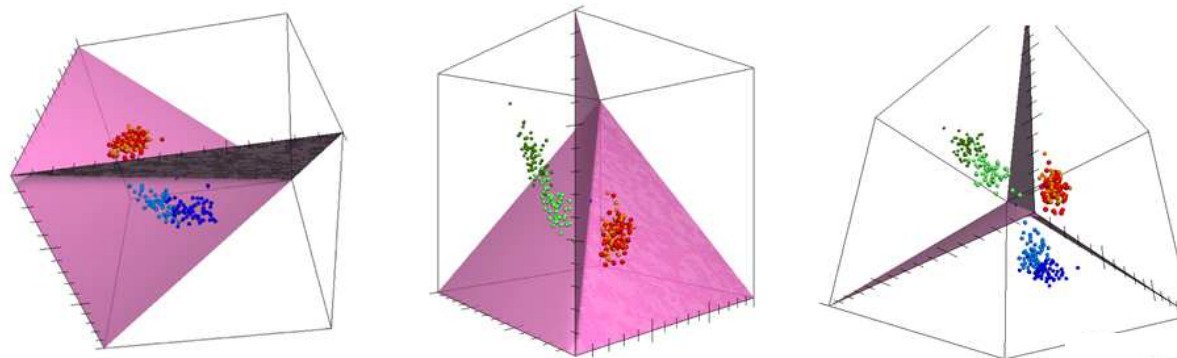
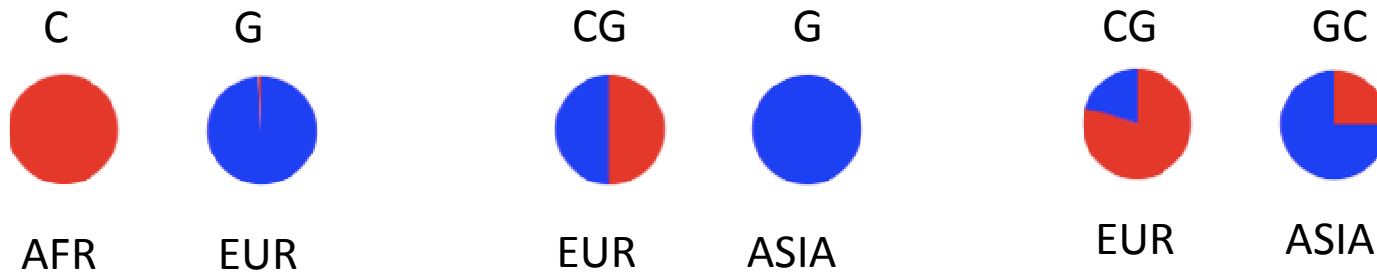
Clear values

Calculate

Age prediction interval is [51.64, 69.91] (0.1 and 0.9 quantiles, respectively) and the predicted age is 60.08 (in years).

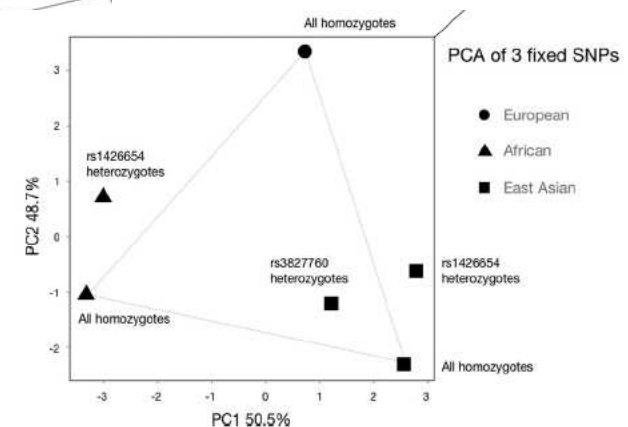


Ancestry informative makers: AIMS



- Chinese - Taiwanese
- Mozambican - Somali
- Danish - Galician

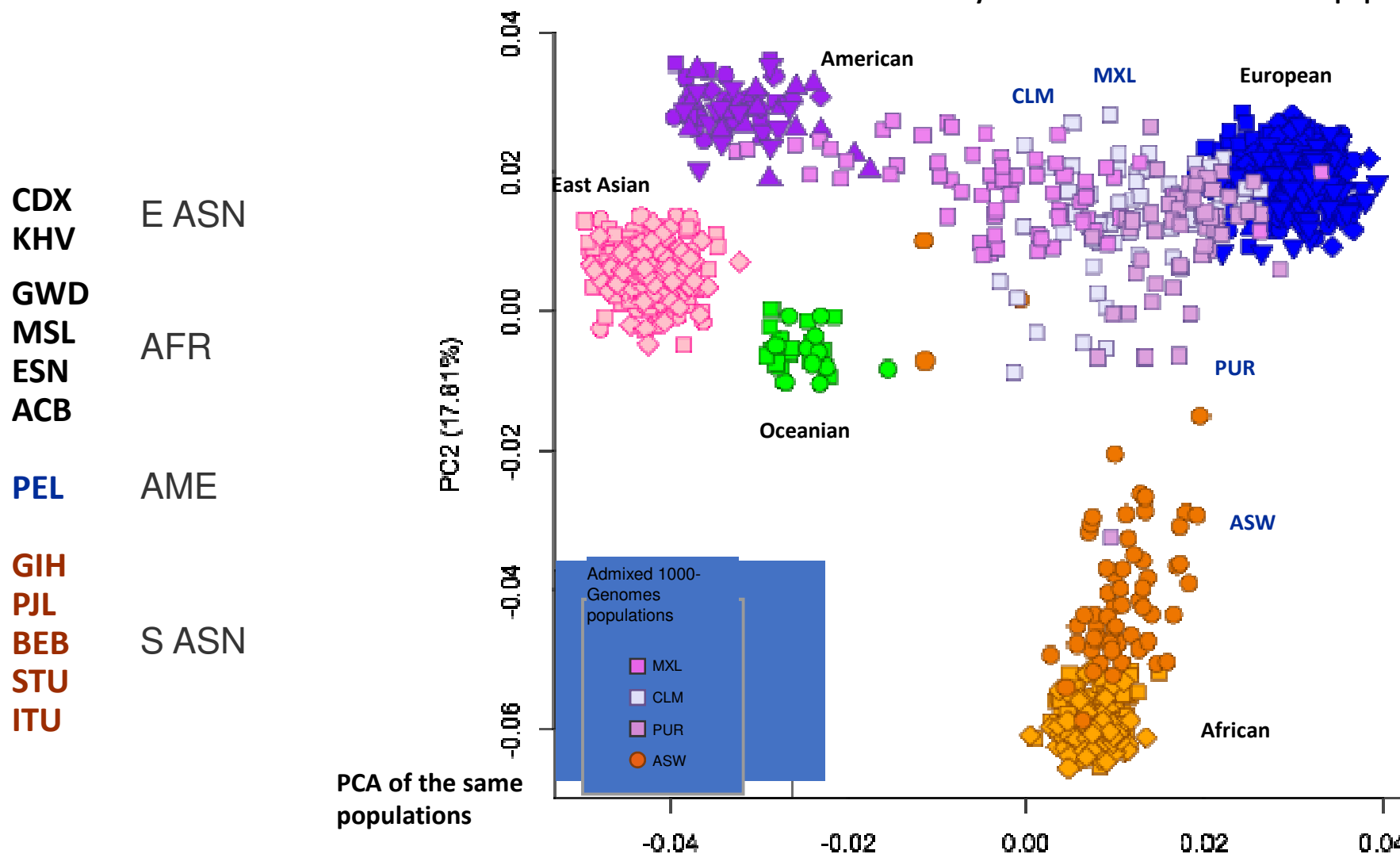
USC 34plex panel
Phillips et al., FSIGEN
2010



C. Phillips. Forensic genetic analysis of bio-geographical ancestry . FSI Genetics 18 (2015) 49-65



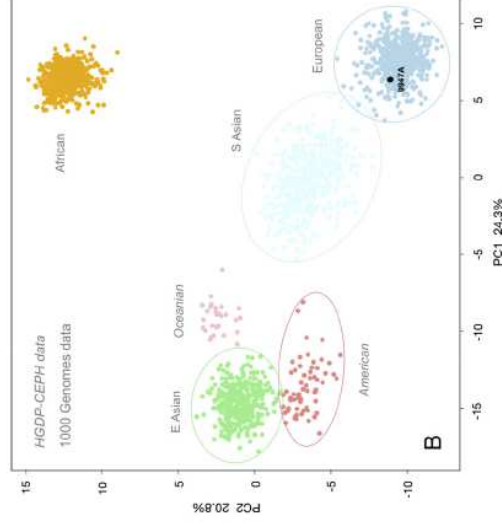
STRUCTURE analysis of admixed 1000-Genomes populations



Research paper

Inter-laboratory evaluation of the EUROFORGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM™

M. Eduardoff^{a,1}, T.E. Gross^{b,1}, C. Santos^c, M. de la Puente^c, D. Ballard^d, C. Strobl^a, C. Børsting^e, N. Morling^e, L. Fusco^e, C. Hussing^e, B. Egyed^f, L. Souto^g, J. Uacyisrael^h, D. Syndercombe Court^d, Á. Carracedo^{c,1}, M.V. Lareu^c, P.M. Schneider^b, W. Parson^{a,i}, C. Phillips^c,  The EUROFORGEN-NoE



Evaluation of the Global AIM-SNP panel with Ion PGM™

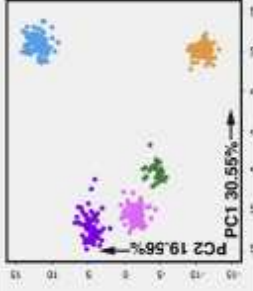
Global AIMS markers

↓
97.6% assay
conversion rate

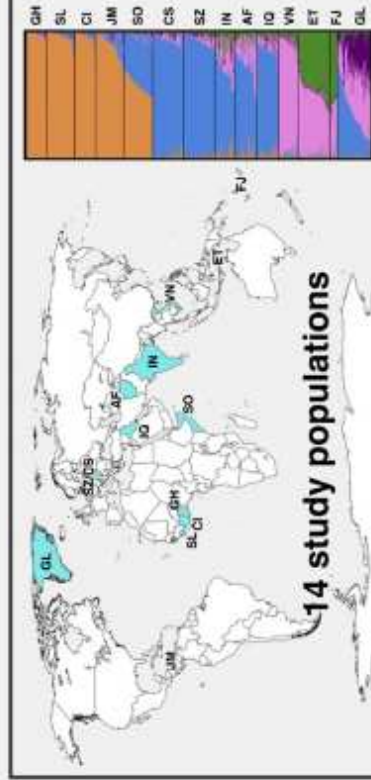
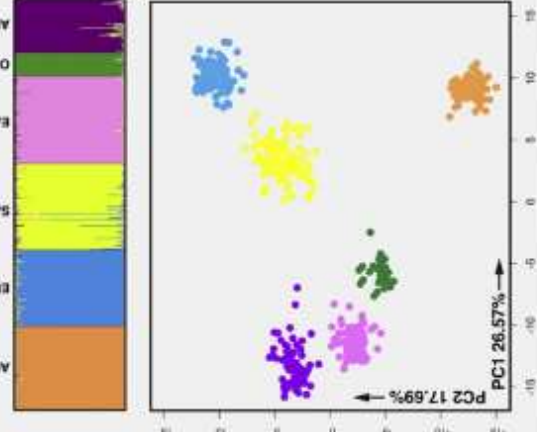
↓
Inter-laboratory
validation
(5 collaborators)

↓
inter-laboratory
database
concordance > **99.8%**
full concordance down to **100pg of DNA**
mixtures detected up to **1:9** ratio

Designed for
5 population
groups



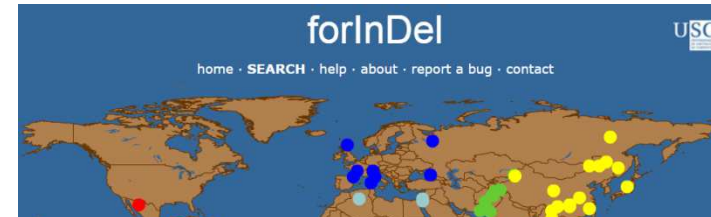
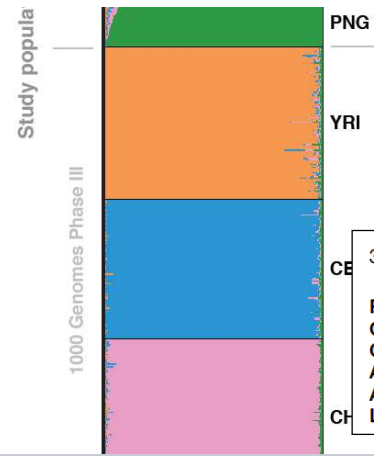
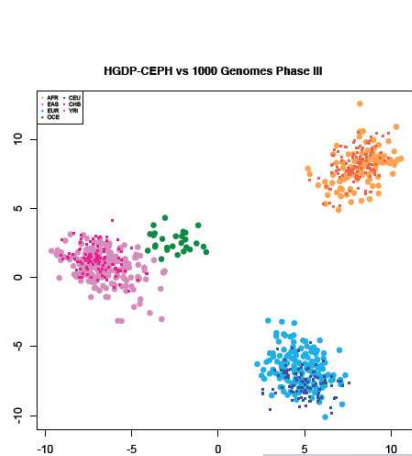
Possibility to differentiate
up to 6 population groups



14 study populations

PacifiPlex: an ancestry-informative SNP panel centred on Australia and the Pacific region.

Santos C¹, Phillips C², Fondevila M², Daniel R³, van Oorschot RA³, Burchard EG⁴, Schanfield MS⁵, Souto L⁶, Uacyisrael J⁷, Via M⁸, Carracedo Á⁹, Lareu MV².



3682
Rui Pereira^{1,2}
Christopher Phillips²
Cintia Alves¹
António Amorim^{1,3}
Ángel Carracedo^{2,4}
Leonor Gusmão¹

Research Article

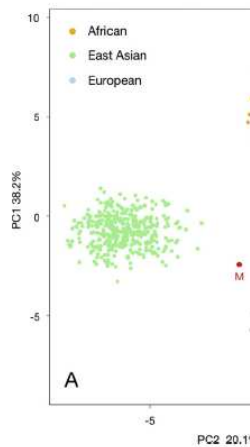
A new multiplex for human identification using insertion/deletion polymorphisms

Electrophoresis 2009, 30, 3682–3690

SPSsmart: v5.1.1 · dbSNP version: build 132

selected database: IPATIMUP-USC 46-plex · data source: GEO-IPATIMUP · data version: Feb 2014

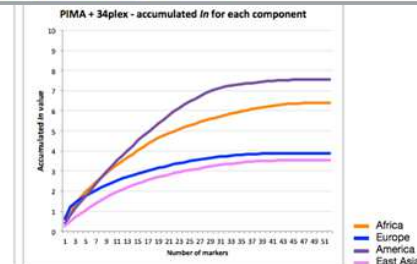
available databases on SPSmart: 1000_Genomes · HapMap · Perlegen · HGDP-CEPH · copyright and disclaimer



MULTIPLEX PIMA (plex de identificación de mezcla americana)

| dbSNP reference | AFRICA (n=102) | EUROPE (n=158) | Population Set 3 (n=42) | EAST ASIA (n=229) | HGDP-CEPH data | HapMap reference |
|-----------------|----------------|----------------|-------------------------|-------------------|----------------|------------------|
| rs12402499 | G: 1.000 | G: 0.946 | G: 0.369 | G: 0.987 | G: 0.987 | G: 0.987 |
| rs10122740 | A: 0.000 | A: 0.054 | A: 0.631 | A: 0.013 | A: 0.013 | A: 0.013 |
| rs1834619 | T: 0.015 | T: 0.084 | T: 1.000 | T: 0.054 | T: 0.054 | T: 0.054 |
| rs12498138 | C: 0.985 | C: 0.916 | C: 0.000 | C: 0.146 | C: 0.146 | C: 0.146 |
| rs6880206 | G: 1.000 | G: 0.927 | G: 0.060 | G: 0.277 | G: 0.277 | G: 0.277 |
| rs252155 | A: 0.000 | A: 0.073 | A: 0.940 | A: 0.723 | A: 0.723 | A: 0.723 |
| rs2471532 | A: 0.010 | A: 0.590 | A: 0.143 | A: 0.910 | A: 0.910 | A: 0.910 |
| rs12242080 | G: 0.980 | G: 0.408 | G: 0.857 | G: 0.080 | G: 0.080 | G: 0.080 |
| rs6993205 | C: 0.020 | C: 0.443 | C: 0.952 | C: 0.266 | C: 0.266 | C: 0.266 |
| rs2471532 | C: 0.980 | C: 0.557 | C: 0.048 | C: 0.734 | C: 0.734 | C: 0.734 |
| rs12242080 | C: 0.018 | C: 0.180 | C: 0.952 | C: 0.219 | C: 0.219 | C: 0.219 |
| rs6993205 | C: 1.000 | C: 0.921 | C: 0.167 | C: 0.493 | C: 0.493 | C: 0.493 |
| rs4979224 | C: 0.999 | C: 0.987 | C: 0.833 | C: 0.945 | C: 0.945 | C: 0.945 |
| rs3014614 | C: 0.000 | C: 0.018 | C: 0.000 | C: 0.052 | C: 0.052 | C: 0.052 |
| rs12130385 | C: 0.127 | C: 0.087 | C: 0.131 | C: 0.341 | C: 0.341 | C: 0.341 |
| rs3212862 | C: 0.873 | C: 0.513 | C: 0.912 | C: 0.659 | C: 0.659 | C: 0.659 |
| rs2051822 | C: 0.973 | C: 0.949 | C: 0.131 | C: 0.736 | C: 0.736 | C: 0.736 |
| rs12434073 | C: 0.025 | C: 0.051 | C: 0.869 | C: 0.432 | C: 0.432 | C: 0.432 |
| rs1280476 | C: 1.000 | C: 0.924 | C: 0.869 | C: 0.264 | C: 0.264 | C: 0.264 |
| rs132649 | C: 0.000 | C: 0.076 | C: 0.845 | C: 0.255 | C: 0.255 | C: 0.255 |
| rs279839 | C: 0.966 | C: 0.949 | C: 0.155 | C: 0.897 | C: 0.897 | C: 0.897 |
| rs135041 | C: 0.034 | C: 0.051 | C: 0.845 | C: 0.103 | C: 0.103 | C: 0.103 |
| rs1377273 | C: 0.931 | C: 0.639 | C: 0.960 | C: 0.900 | C: 0.900 | C: 0.900 |
| | A: 0.069 | A: 0.361 | A: 0.040 | A: 0.100 | A: 0.100 | A: 0.100 |
| | A: 0.123 | A: 0.250 | A: 0.940 | A: 0.282 | A: 0.282 | A: 0.282 |
| | A: 0.877 | A: 0.345 | A: 0.060 | A: 0.718 | A: 0.718 | A: 0.718 |
| | A: 0.995 | A: 0.655 | A: 0.036 | A: 0.136 | A: 0.136 | A: 0.136 |
| | A: 0.020 | A: 0.345 | A: 0.964 | A: 0.644 | A: 0.644 | A: 0.644 |
| | A: 0.980 | A: 0.655 | A: 0.036 | A: 0.356 | A: 0.356 | A: 0.356 |
| | C: 0.015 | C: 0.506 | C: 0.952 | C: 0.830 | C: 0.830 | C: 0.830 |
| | C: 0.802 | C: 0.494 | C: 0.048 | C: 0.170 | C: 0.170 | C: 0.170 |
| | C: 0.198 | C: 0.506 | C: 0.036 | A: 0.085 | A: 0.085 | A: 0.085 |

PIMA plex
L. Porras et al. 2013



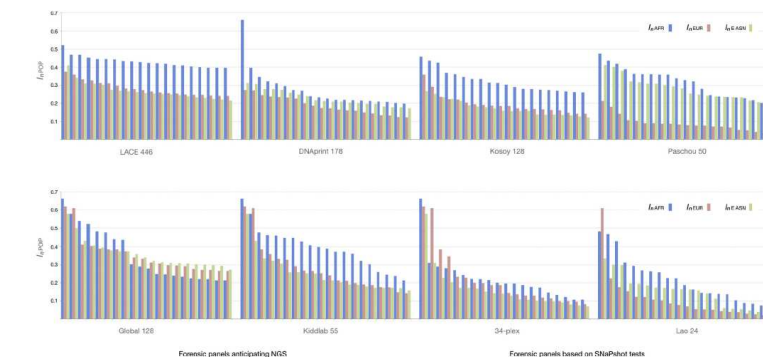
Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

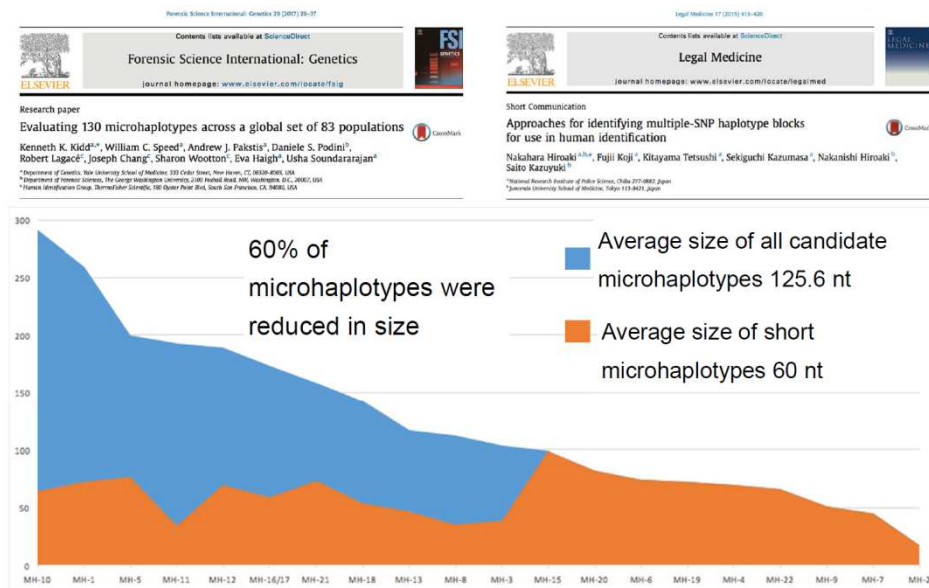
journal homepage: www.elsevier.com/locate/fsig

Eurasiaplex: A forensic SNP assay for differentiating European and South Asian ancestries

C. Phillips ^{a,*}, A. Freire Aradas ^{a,1}, A.K. Kriegel ^{b,1}, M. Fondevila ^a, O. Bulbul ^c, C. Santos ^a, F. Serrulla Rech ^d, M.D. Perez Carceles ^e, Á. Carracedo ^a, P.M. Schneider ^a, M.V. Lareu ^a



Phillips et al. Ancestry markers for Asia-Pacific region

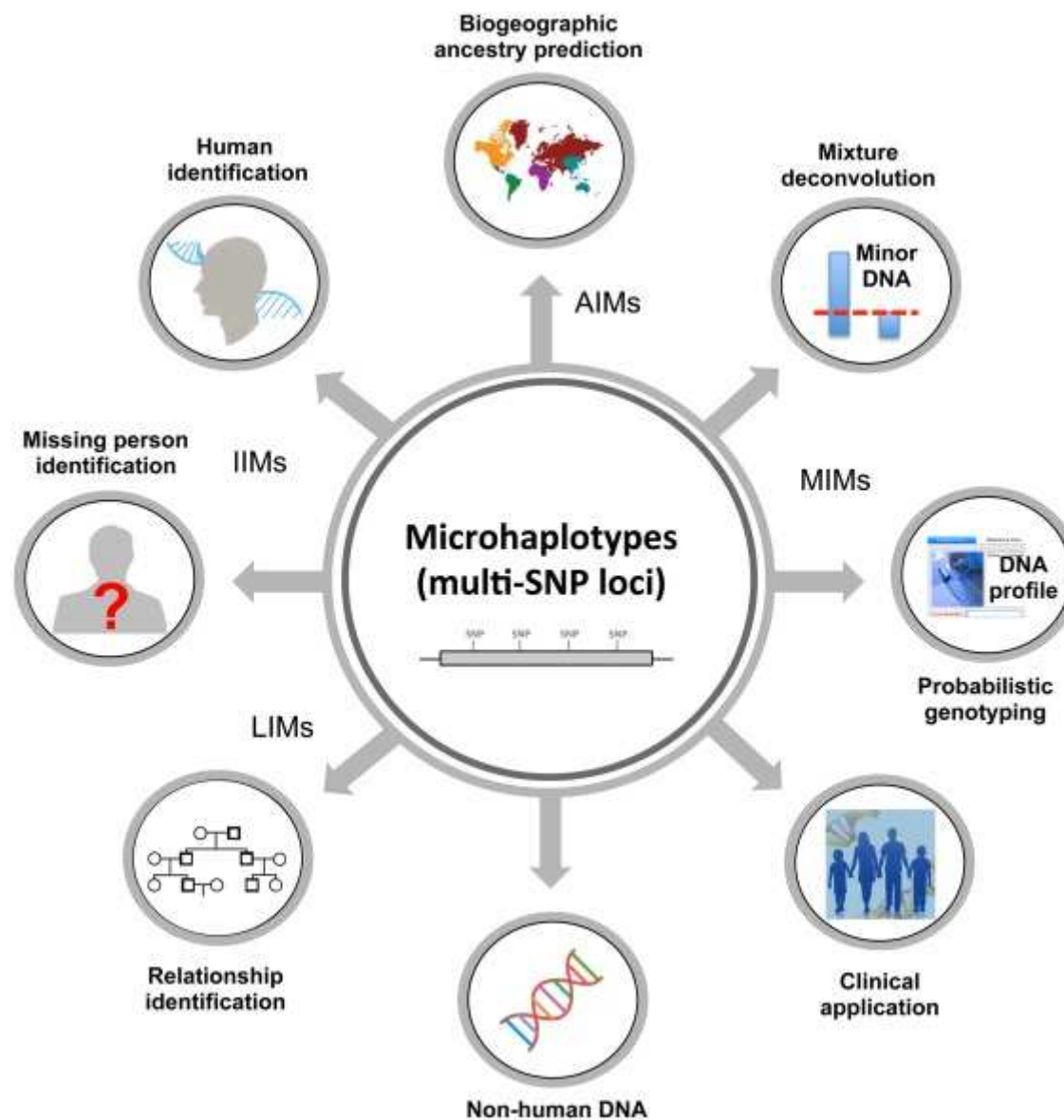


MAPlex Middle East Asia Pacific



| | 34plex (CE) | gAIMs (MPS) | MAPplex (MPS) |
|--|----------------|----------------|------------------|
| Population specific SNPs | 13 | 12 | 122 > 84 |
| Skewed frequency SNPs | 14 | 93 | |
| Fixed difference SNPs | 4 | 17 | |
| Tri allelic SNPs | 3 | 6 | 30 |
| South Asian informative AIMs (EurasiaPlex+) | | | 24 |
| Short amplicon microhaplotype markers (x22) | | | 80 |

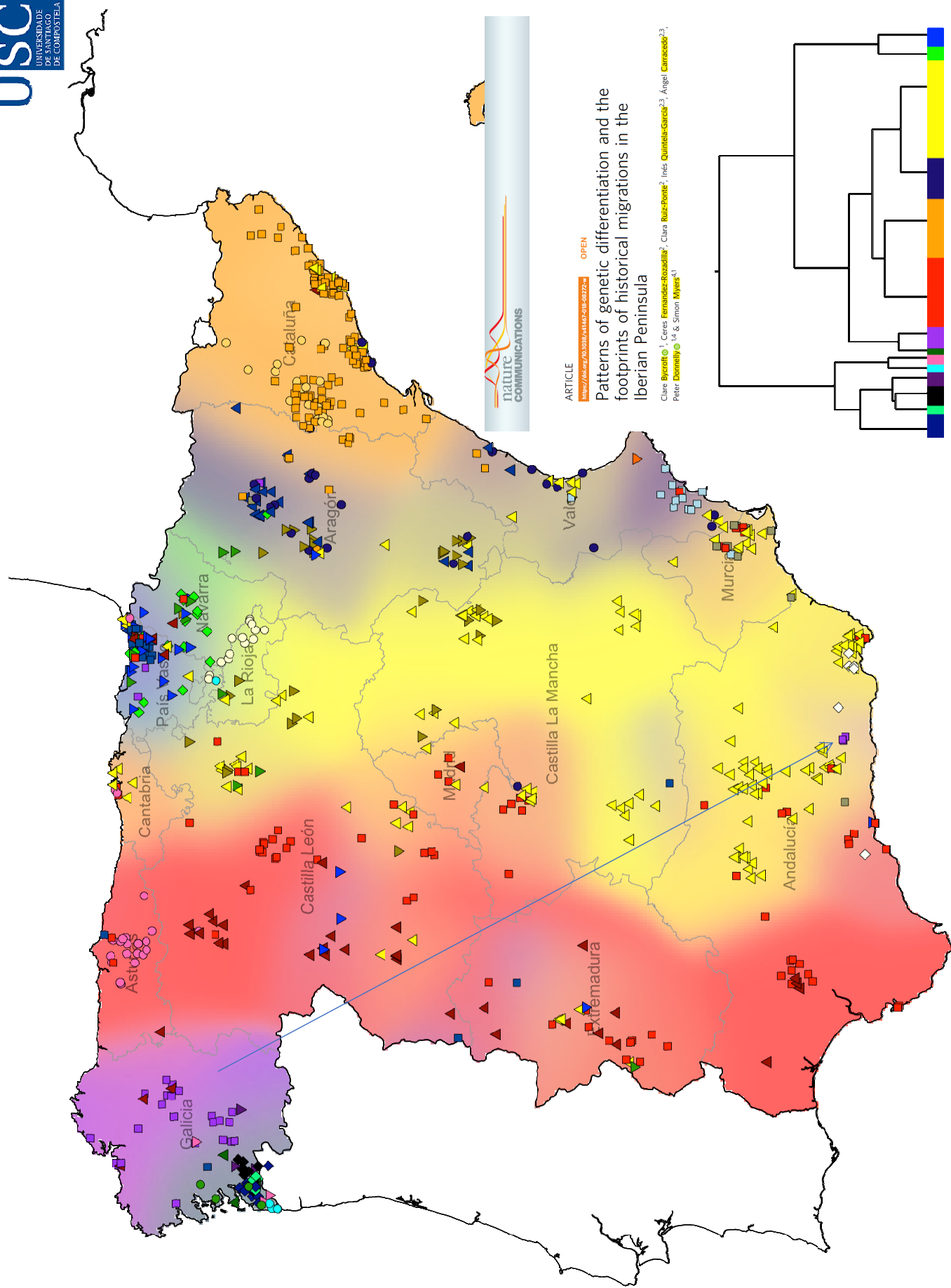




[Microhaplotypes in forensic genetics](#)

Fabio Oldoni, Kenneth K. Kidd, Daniele Podini

FSI: Genetics January 2019

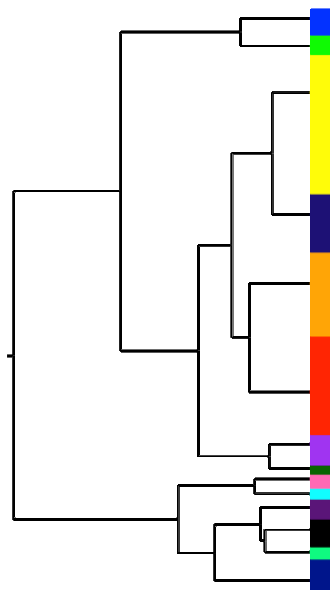


ARTICLE

<https://doi.org/10.1038/s41467-018-0272-z> OPEN

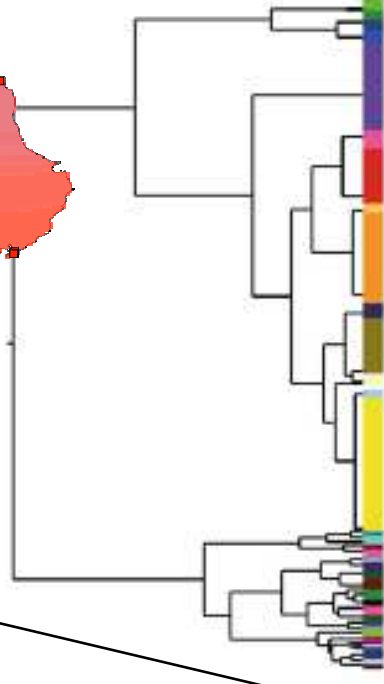
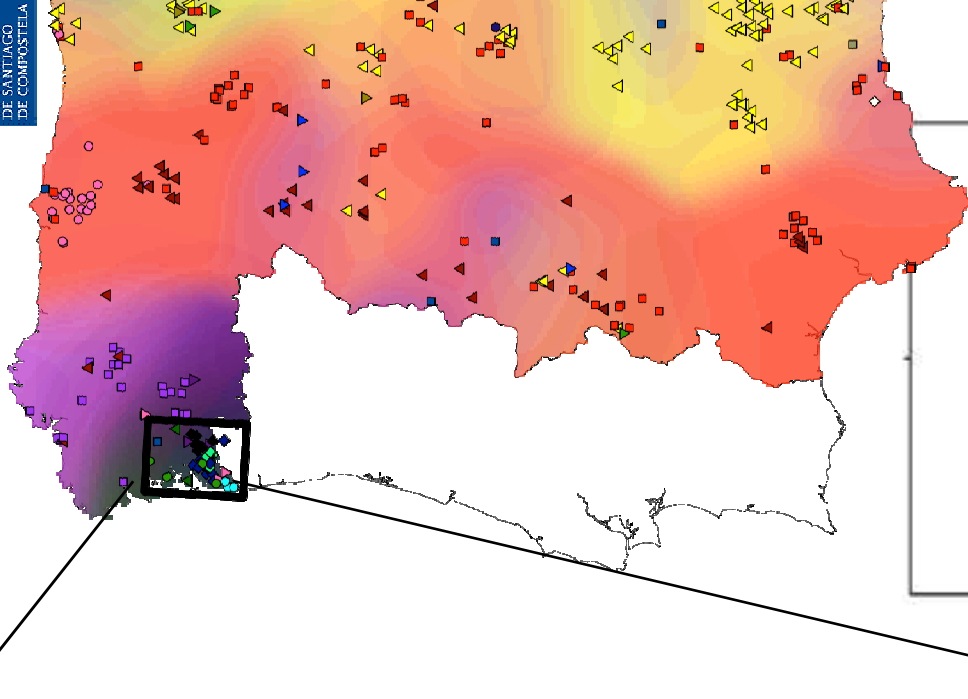
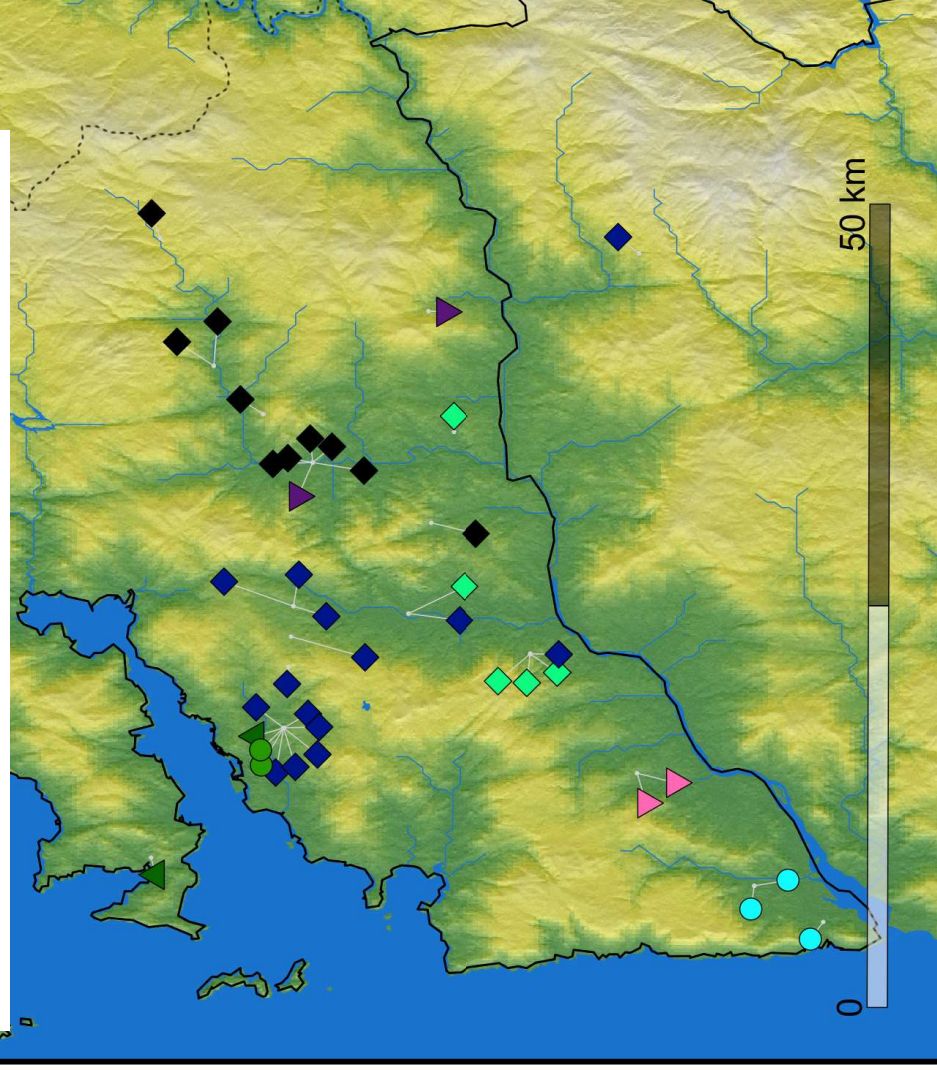
Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula

Clara [Bryce](#)¹, Ceres [Fernández-Rodríguez](#)², Clara [Ruiz-Ponte](#)², Inés [Quintela-García](#)^{2,3}, Ángel [Corrazeiro](#)^{2,3}, Peter [Donnelly](#)⁴ & Simon [Meyer](#)^{4,5}



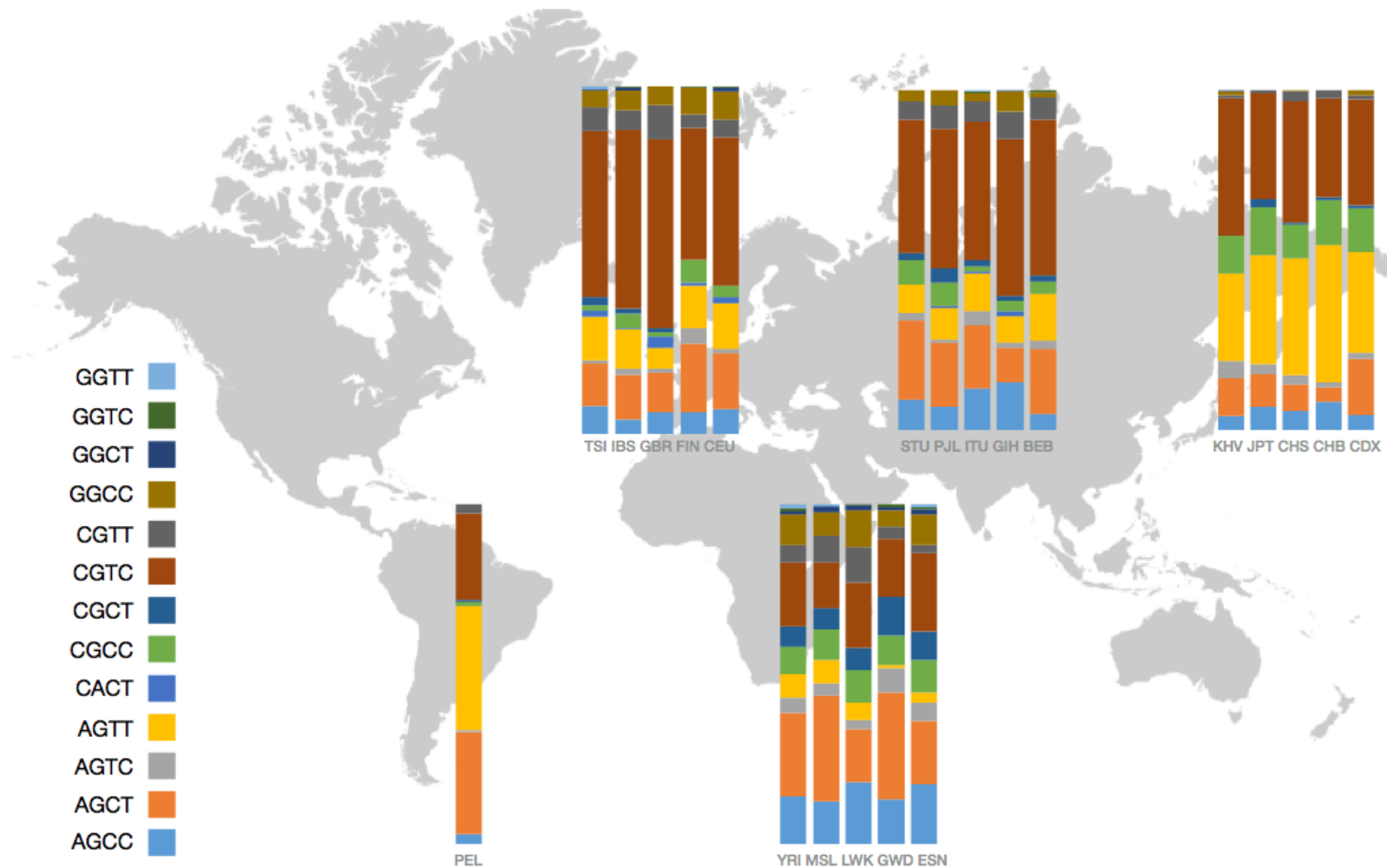
Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula

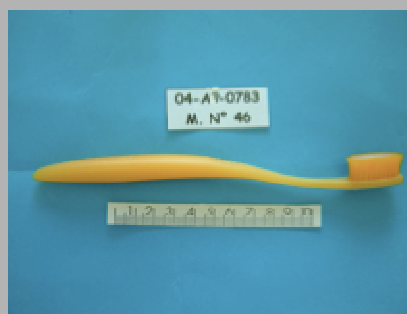
Clare Bycroft¹, Ceres Fernandez-Rozadilla², Clara Ruiz-Ponte², Inés Quintela-García^{2,3}, Ángel Carracedo^{2,3}, Peter Donnelly^{1,4} & Simon Myers^{4,1}



Some microhaplotypes are very informative

MH-21 73nt 4 SNPs rs6517970-rs202132081-rs8131148-rs6517971



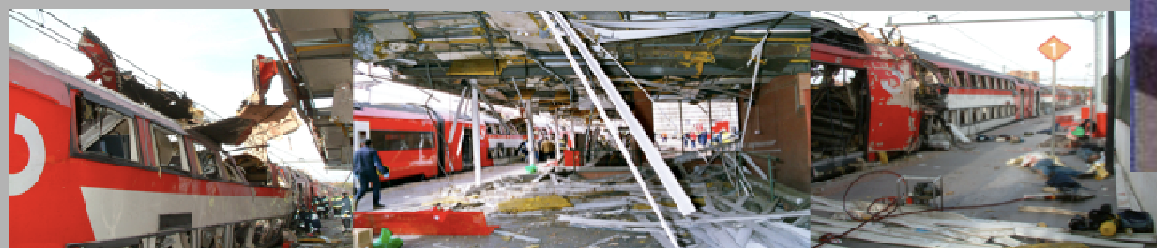


11-M was a 'closed' ancestry analysis



personal items from bomb assembly site in Leganés with DNA profiles that didn't match with the terrorists

un-detonated explosives in a holdall found at El Pozo station



Calculation results

Executing the query with 3 default populations and the 34 SNPs of the individual to classify :

GTCCCCCTAGAACTCCAACCTGGGGTTCCAACCAACCAAGGCCAACCTTACAAGGTTTCGAAGGTTACAC

The $-\log(\text{LIKELIHOOD})$ (lower is best) and PERCENTILE (percent of population samples with lower likelihoods than individual submitted).

| | | |
|----------|-----------|-------|
| Gal-Dani | 41.109466 | 2.50% |
| Moz-Som | 66.935324 | 0.00% |
| Chi-Taiw | 65.629561 | 0.00% |

This person was European



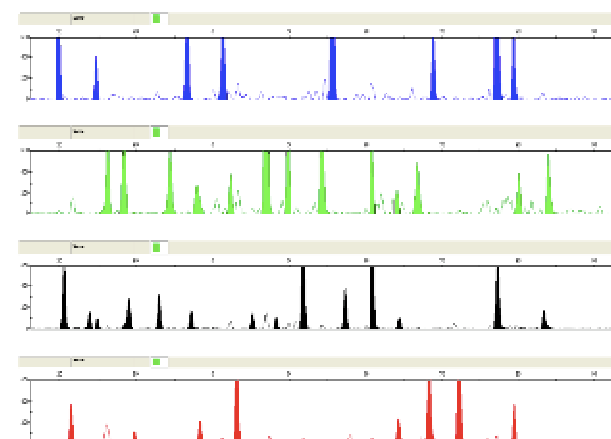
| training set: | -log likelihood | percentile | exp | times more likely to be | verbal predicate |
|---------------|-----------------|------------|-------------|-------------------------|---|
| Gal-Dani | 41.109466 | 2.50% | 1.40083E-18 | EUR not AFR | |
| Moz-Som | 66.935324 | 0.00% | 8.5184E-30 | 1.64448E+11 | 164 billion times more likely to be European than African |
| Gal-Dani | 41.109466 | 2.50% | 1.40083E-18 | EUR not ASN | |
| Chi-Taiw | 65.629561 | 0.00% | 3.14372E-29 | 44559668806 | 44 billion times more likely to be European than Asian |
| Moz-Som | 66.935324 | 0.00% | 8.5184E-30 | AFR not ASN | |
| Chi-Taiw | 65.629561 | 0.00% | 3.14372E-29 | 0.270965709 | |

GT GG GG TT CC AA CC AA CC AA GG CC AA CC TT AC TT GG TT AA CG TT GG AC AC

Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, Alvarez-Dios J, Alonso A, Blanco-Verea A, Brión M, Montesino M, Carracedo A, Lareu MV

Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One. 2009 Aug 11;4(8):e6583.

34plex



Calculation results

Executing the query with 3 default populations and the 34 SNPs of the individual to classify :

GTCCCCCTAGAACTCCAACCTGGGGTTCCAACCAACCAAGGCCAACCTTACAAGGTTTCGAAGGTTACAC

The **-log(LIKELIHOOD)** (lower is best) and **PERCENTILE** (percent of population samples with lower likelihoods than individual submitted).

| | | |
|----------|-----------|-------|
| Gal-Dani | 41.109466 | 2.50% |
| Moz-Som | 66.935324 | 0.00% |
| Chi-Taiw | 65.629561 | 0.00% |

This person was European

SNIPPER

| training set: | -log likelihood | percentile | exp | times more likely to be | verbal predicate |
|---------------|-----------------|------------|-------------|-------------------------|---|
| Gal-Dani | 41.109466 | 2.50% | 1.40083E-18 | EUR not AFR | |
| Moz-Som | 66.935324 | 0.00% | 8.5184E-30 | 1.64448E+11 | 164 billion times more likely to be European than African |
| Gal-Dani | 41.109466 | 2.50% | 1.40083E-18 | EUR not ASN | |
| Chi-Taiw | 65.629561 | 0.00% | 3.14372E-29 | 44559668806 | 44 billion times more likely to be European than Asian |
| Moz-Som | 66.935324 | 0.00% | 8.5184E-30 | AFR not ASN | |
| Chi-Taiw | 65.629561 | 0.00% | 3.14372E-29 | 0.270965709 | |

GT CC CC CT AG AA CT CC AA CT **GG** GG TT CC AA CC AA CC AA GG CC AA CC TT AC TT GG TT AA CG TT GG AC AC

... and had blue eyes: rs12913832 = GG

GG

Blue

BEY1

Green
Hazel

AG

AG

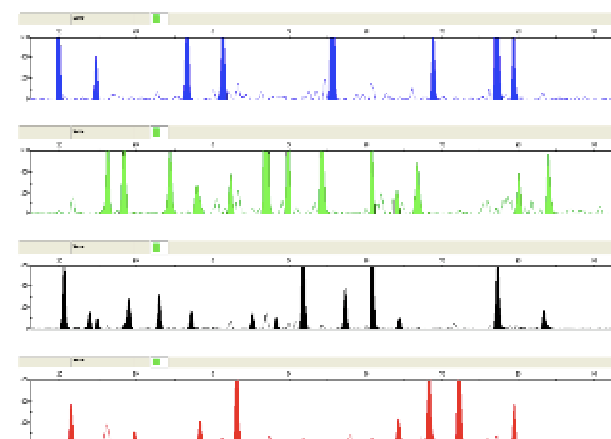
Green
(Unknown)

AG / (AA)

Brown

BEY2

34plex



Operation Minstead

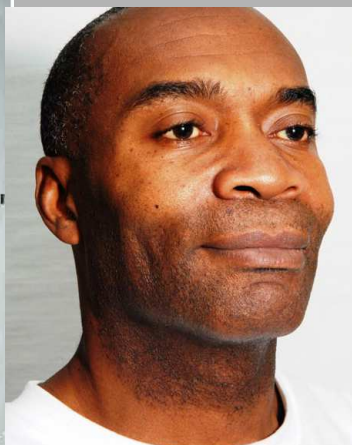


A elderly victim of an attack in May helped police with the e-fit

- 18 year investigation of a series of aggressive rapes of elderly victims in London

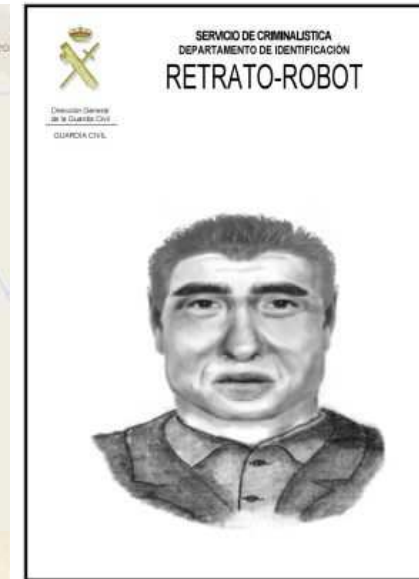
- USC testing ancestry independently plus skin pigmentation SNPs to reduce suspect pool

- 82% AFR, 12% AME 6% EUR and suggest this shows Windward Island origins Y=EUR, mt=AFR



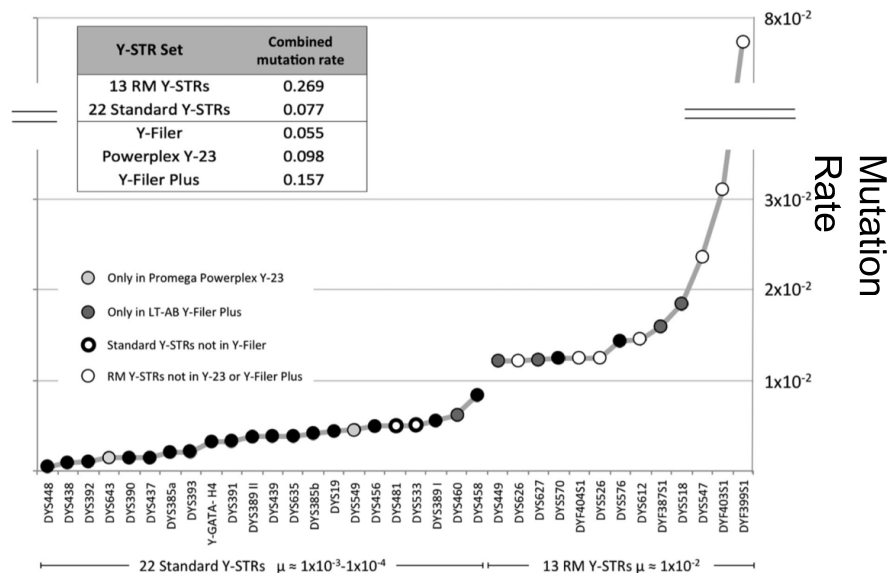
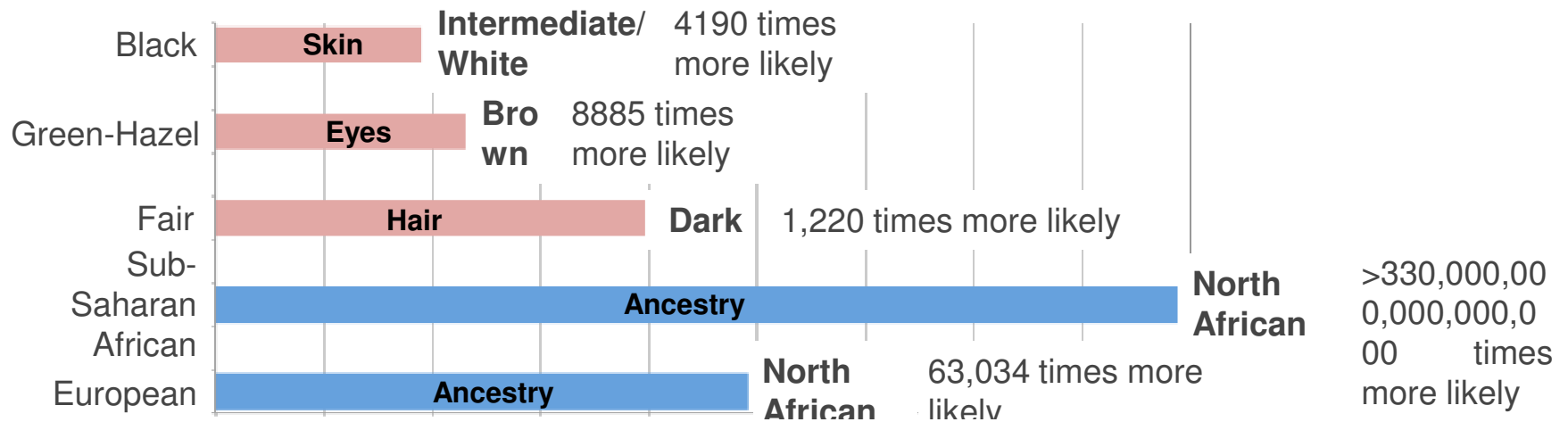
The murder of Eva Blanco Puig - a cold case investigation

- Eva Blanco was raped and murdered in April, 1997
- Heavy rain obscured tyre marks and footprints at scene
- Small semen stain on underwear
- Police requested a mass screen of 2000 males from Algete
- Judge deemed it too expensive and likely to be ineffectual



USC asked to analyse ancestry and pigmentation

- 75 / 80 ancestry-informative SNPs successfully genotyped
- 18 / 19 pigmentation-predictive 'SHEP' SNPs successfully genotyped



- Successful application to screen 300 Moroccans living in a larger area around Algete

- One sample gave a full Y-STR match - led to identification of a brother who had left Spain in 1999 and lived in France

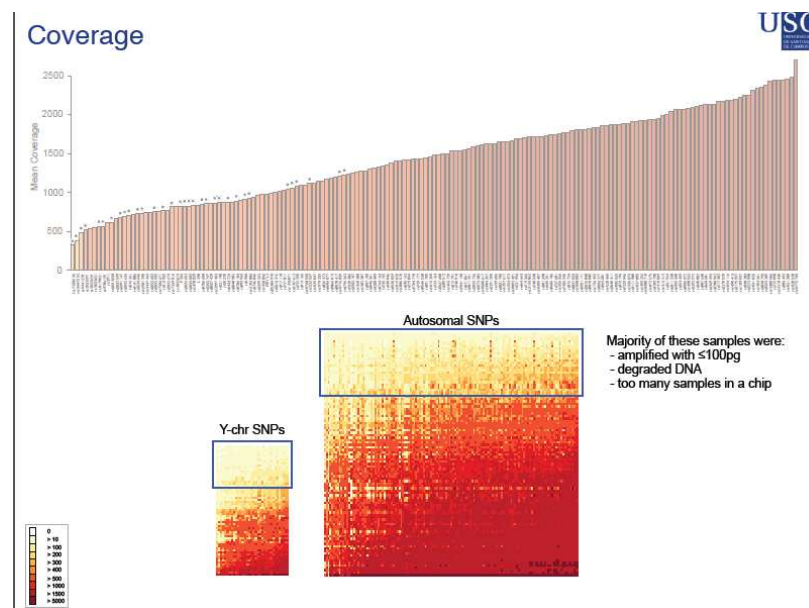
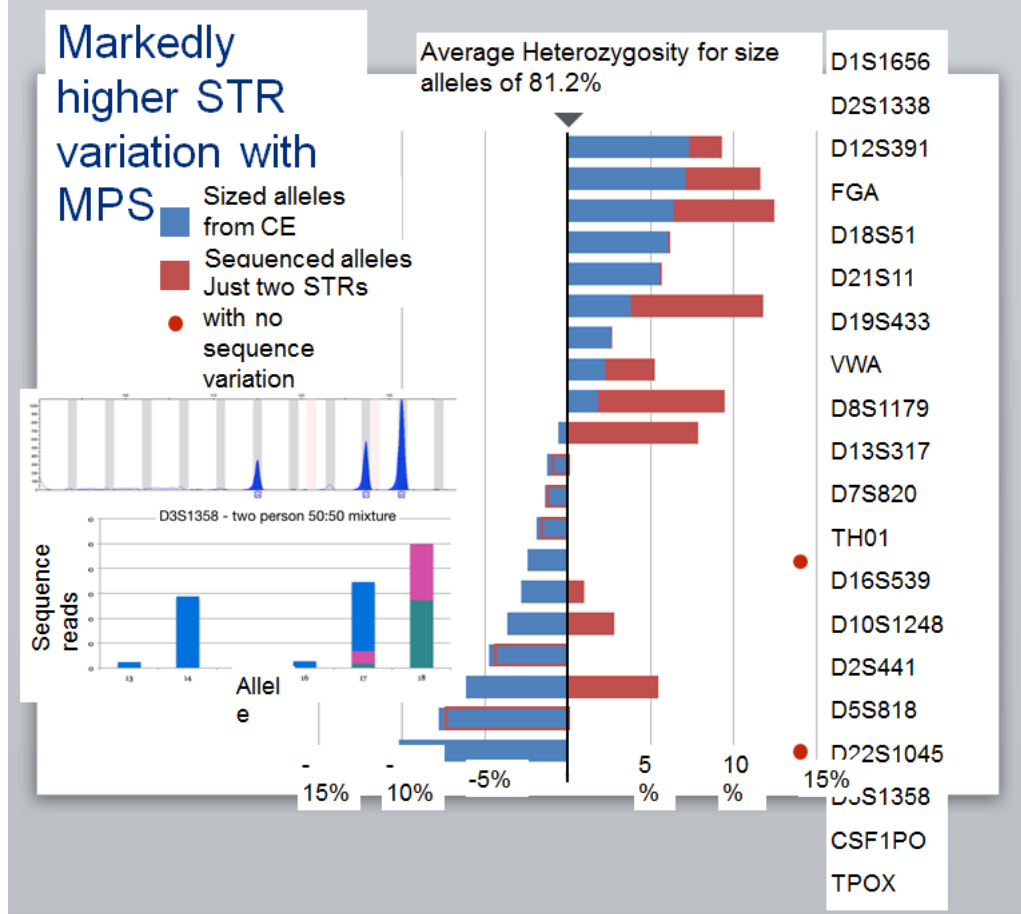


Legal and ethical issues

- Former terms “coding” and “non-coding” referring to DNA sequences lost their meaning
- The issue of amendment of the law allowing for FDP (Netherlands, UK, France)
- Externally visible traits of an individual are publically available
- FDP is useful at the stage of investigation
- Data processed to predict EVC are not stored in forensic databases

| <u>País</u> | Rasgos físicos | Edad | <u>Ancestralidad</u> | Bases de datos |
|-----------------------------|---|---|---|--|
| Austria | No claro- No se hace | No claro- No se hace | No claro. Si se hace con <u>ADNmt</u> y cromosoma Y | No permitido(ADN codificante) |
| Francia | Caso judicial que lo permitió en 2015 - | No se hace. Abierta la posibilidad | Si se hace | No permitido(ADN codificante) |
| Alemania | No permitido explícitamente por el código penal | No permitido explícitamente por el código penal | No permitido explícitamente por el código penal | No permitido(ADN codificante) |
| Polonia | No regulado, permitido y se hace | No regulado, permitido y se hace | No regulado, permitido y se hace | No permitido(ADN codificante) |
| Suecia | No regulado, no prohibido y se subcontrata | No regulado, no prohibido y no se ha hecho | No regulado, no prohibido y se hace para casos de identificación | Poco claro. No se hace |
| Holanda | Permitido explícitamente por la ley | No regulado pero no prohibido- Se ha solicitado | Permitido explícitamente por la ley | No para rasgos físicos y edad, posible para <u>ancestralidad</u> |
| Reino Unido | No regulado: considerado permitido | No regulado: considerado permitido | No regulado: considerado permitido | No regulado: considerado permitido |

NGS



**New ISFG
Nomenclature!!!**

STRs + SNPs validation Santiago Ion Torrent

Next-Generation Sequencing Technologies

Library
preparation



Sequencing



Analysis



NGS platforms



ion torrent
by life technologies™



life technologies™ SOLiD
AB applied biosystems™



Helicos
BioSciences Corporation



PACIFIC BIOSCIENCES™

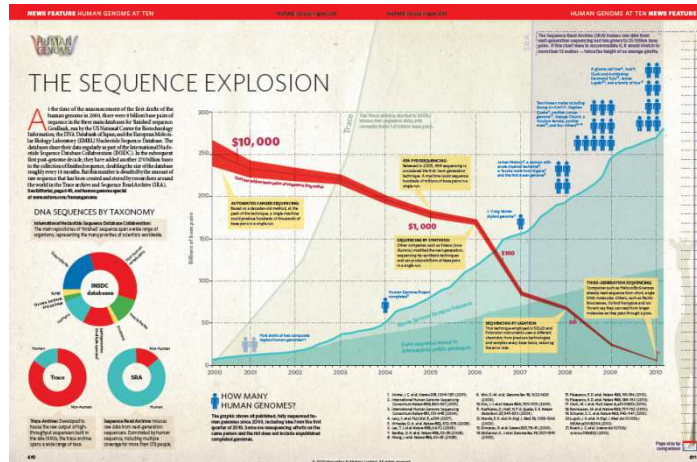


Oxford NANOPORE Technologies®

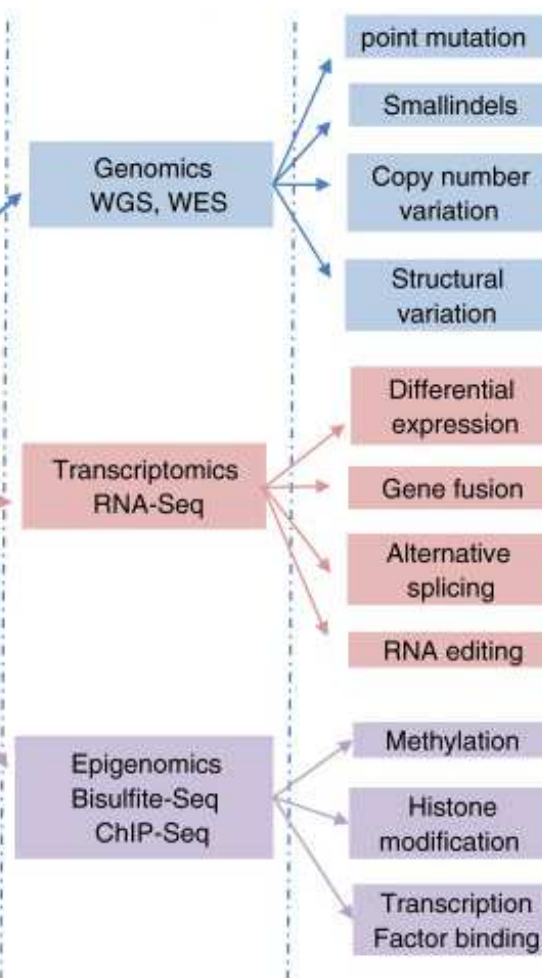
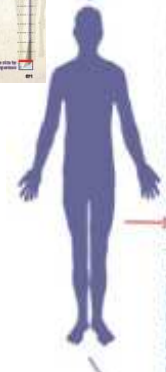
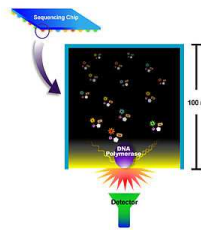


NGS versus Capillary Electrophoresis

| | Advantages | Disadvantages |
|-----|--|---|
| CE | <ul style="list-style-type: none"> Established technology Accepted in court Relatively easy workflow | <ul style="list-style-type: none"> Limited multiplex capability Complex mixture analysis Genotyping based on length only |
| NGS | <ul style="list-style-type: none"> Genotyping based on length and sequence Greater multiplex capability High dynamic range Potential improvement to mixture interpretation Smaller amplicons (degraded DNA) | <ul style="list-style-type: none"> High cost per sample Large amount of data Pooling of samples needed to reduce cost per sample No guidelines available yet More complex workflow Time to result |



Non human DNA typing



The workflow of integrating omics data in forensic research and applications

A.C. 't Hoen et al. Reproducible mRNA and small RNA sequencing across different laboratories. Nature Biotechnology (2013)

Ten times more bacterial DNA than human DNA in the human body

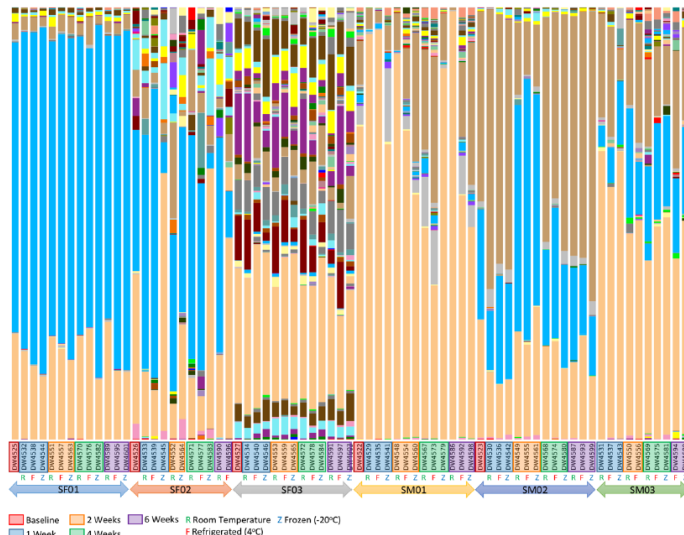
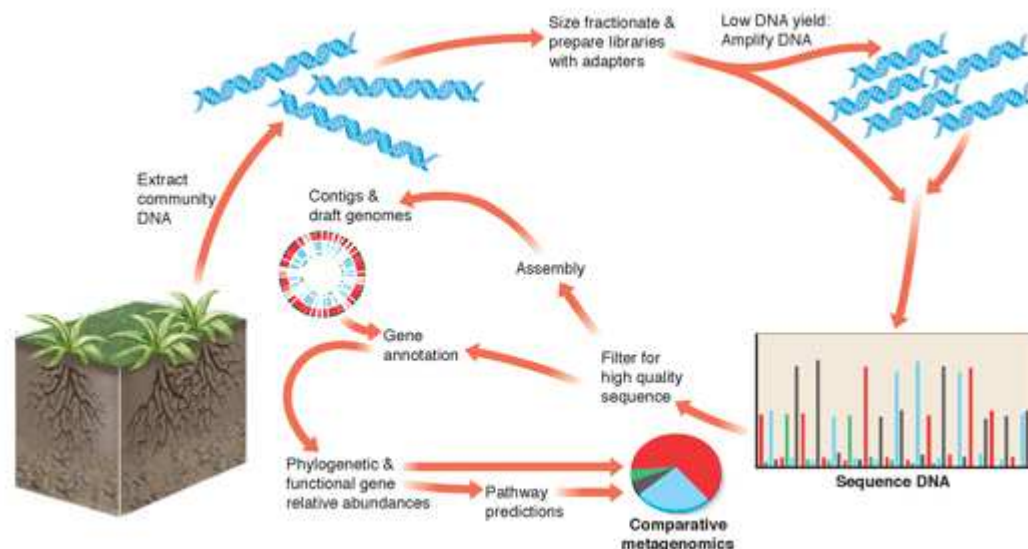


Figure 1: Taxonomic distribution at the genus level across storage time and temperature conditions. Sorted by individuals. Samples were normalized to 900 reads per sample.

Effects of storage time and temperature on pubic hair bacterial communities. Williams & Gibson FSIGEN (in press)

- Short term storage of pubic hair at various temperatures does not significantly affect the recover bacterial microbiome.
- The greatest sources of variation in the pubic hair microbiome are attributed to the individual and gender.



Soil metagenomics

Intraindividual variations
Validation
Statistical challenges



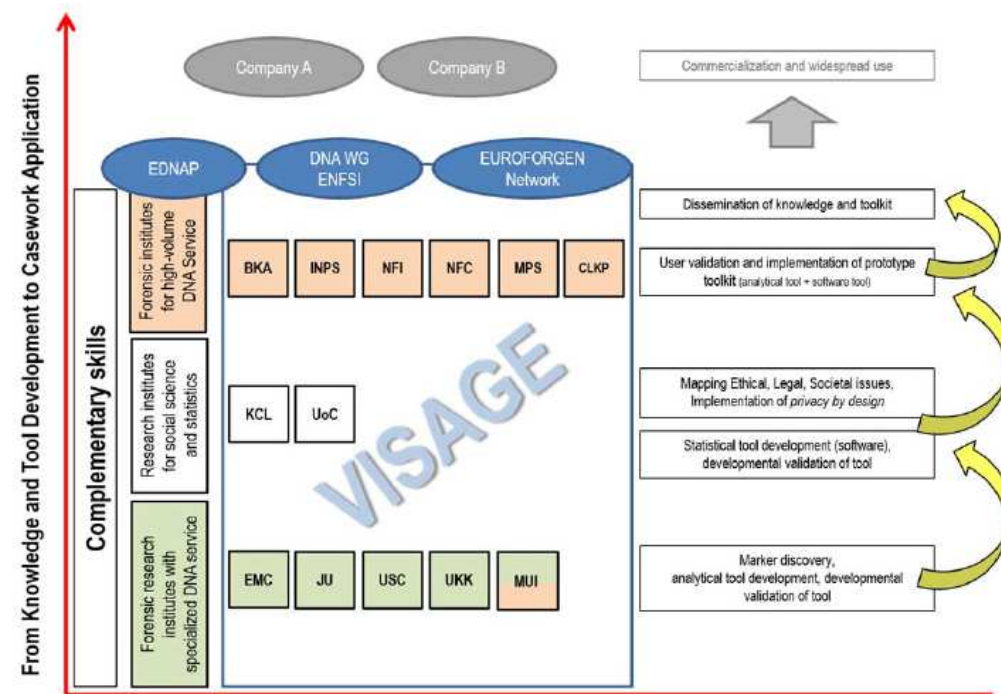
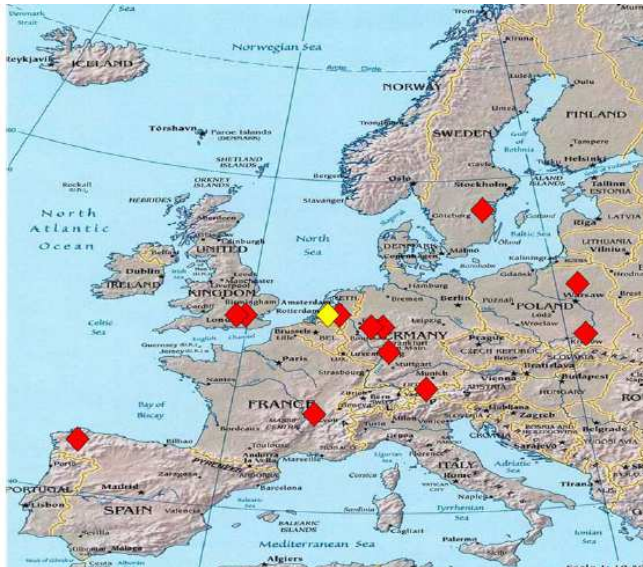
We are good with some physical traits, geographic origin and age but still far from photofit portrait

Low frequency- Intermediate penetrance variants

Gene-gene interactions, GxE interactions

Methylation-MicroRNA-... other approaches

Ethical/Legal problems for EVCs in some countries



Constructing Composite Sketches from DNA Traces for Finding Unknown Perpetrators

Figure 6. VISAGE expertise along the value chain (drawn in a simplified way).



H2020 EU research and innovation
programme-grant agreement No 740580.

Obtaining good results is not always possible

Forensic tests are infallible
Results are obtained immediately

THE CSI EFFECT

Creates false expectations

Creates a distorted perception of the forensic test in judges, prosecutors and jurors



The hunt for the "phantom" – a critical reappraisal of DNA database reality

- In 2007, a female police officer was shot in her car, her male colleague was severely wounded
- A female DNA profile was identified in the car
- The profile generated multiple hits in the DNA databases of Germany and Austria
- The profile resurfaced 39 times in subsequent analyses of other cases (from burglaries to homicides)

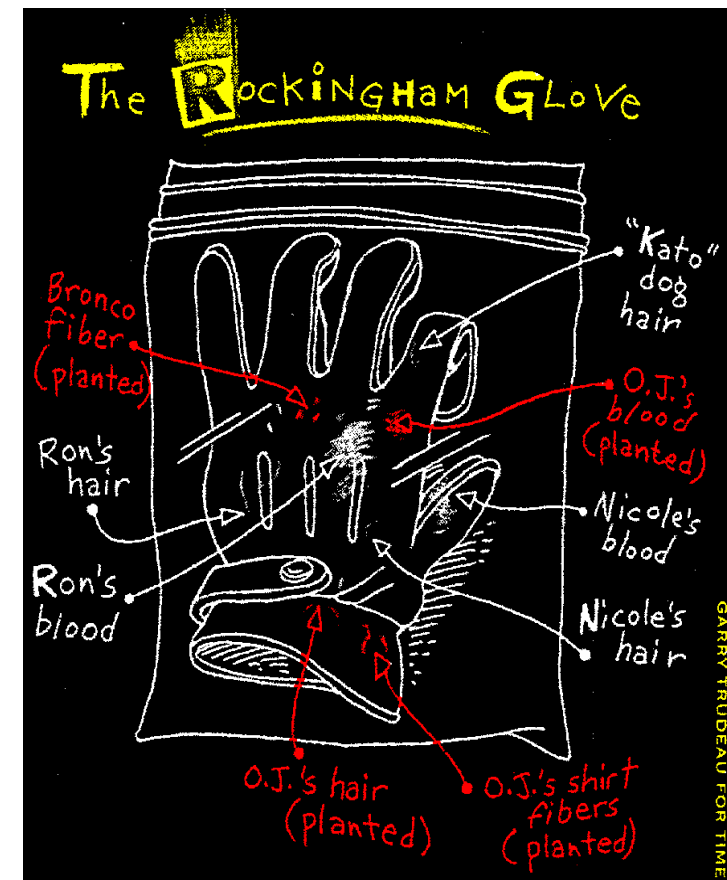
Hier wurden die Spuren des Phantoms entdeckt

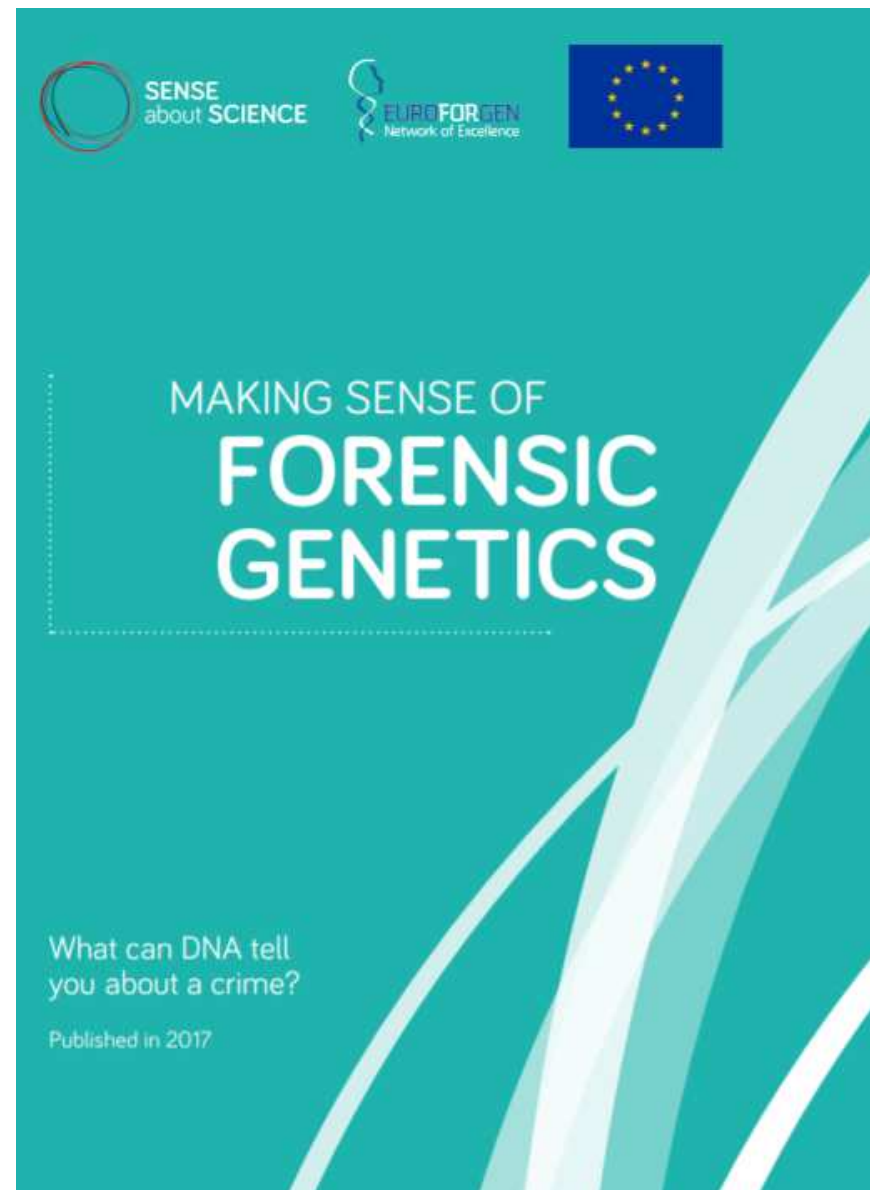
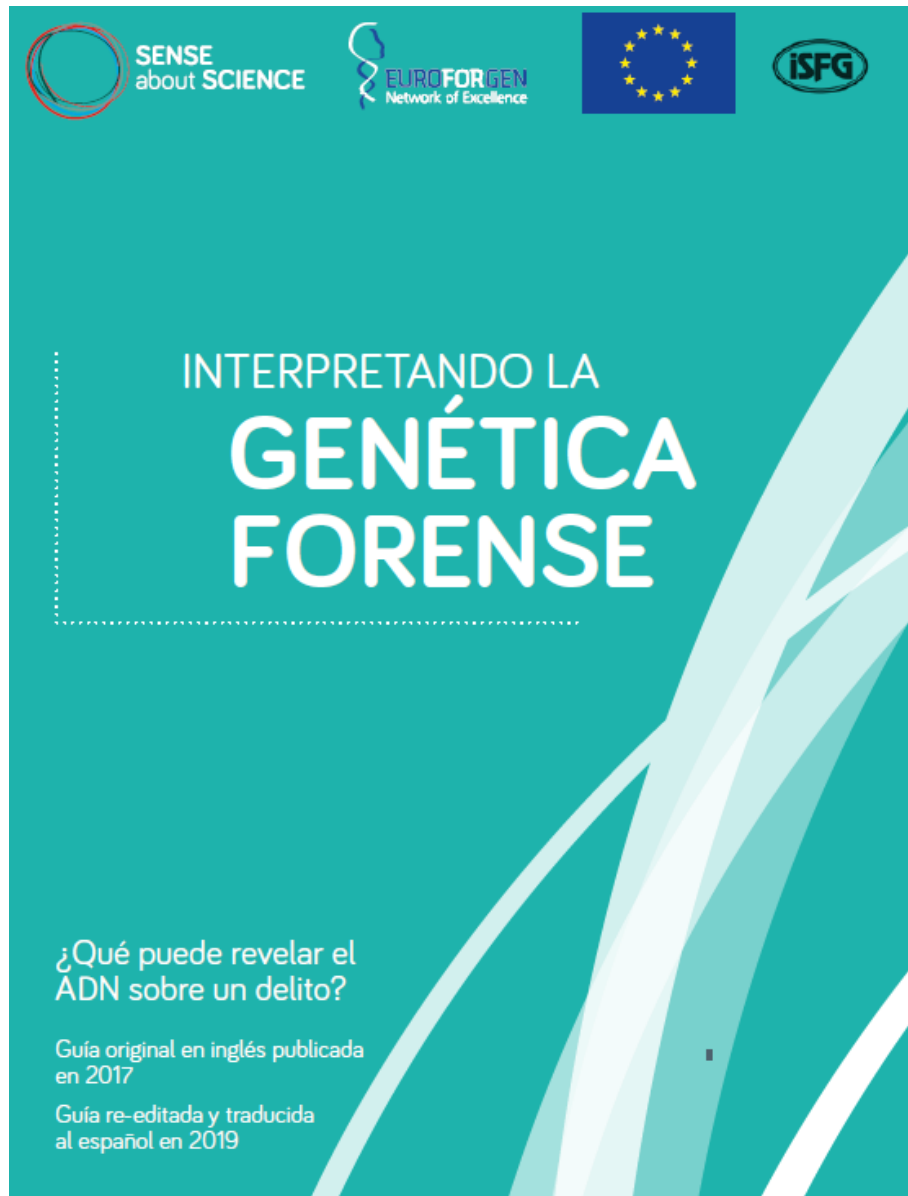


What happened then?

- The profile reappeared later in two more cases where the DNA source was a known male person
- Contamination was now seriously considered
- Cotton swabs were identified as potential source
- The supplier provided reference samples from the factory workers – and a retired 71-year old woman was found to match the phantom's DNA profile

It is not the science itself which is flawed but the interpretation and context in which these results may be used.





<https://senseaboutscience.org>

LA FIESTA CON QUE SE CELEBRA
LA LLEGADA DE NUESTROS AMI-
GOS, ES MAGNIFICA... Y SI ASURAN-
CETURIX NO HUBIERA SUFRIDO UN
ACCIDENTE, HABRIA AMENIZADO
LA VELADA CON SUS CANCIONES...

