

# Genetics in forensics



**Professor Ángel Carracedo**

**Institute of Legal Medicine**

**University of Santiago de Compostela, Spain**



# 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

1990

1995

2000

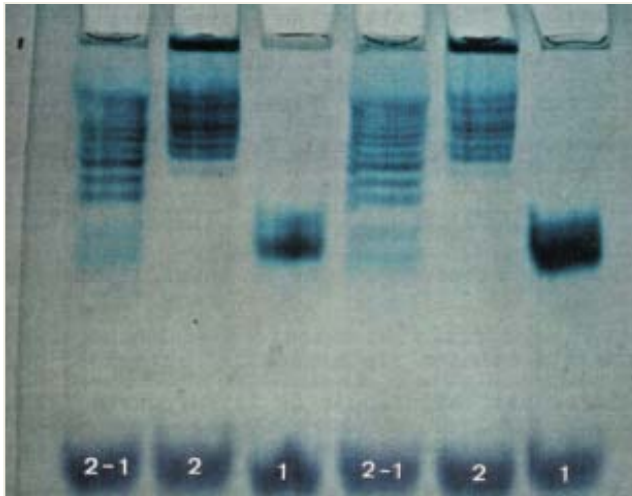


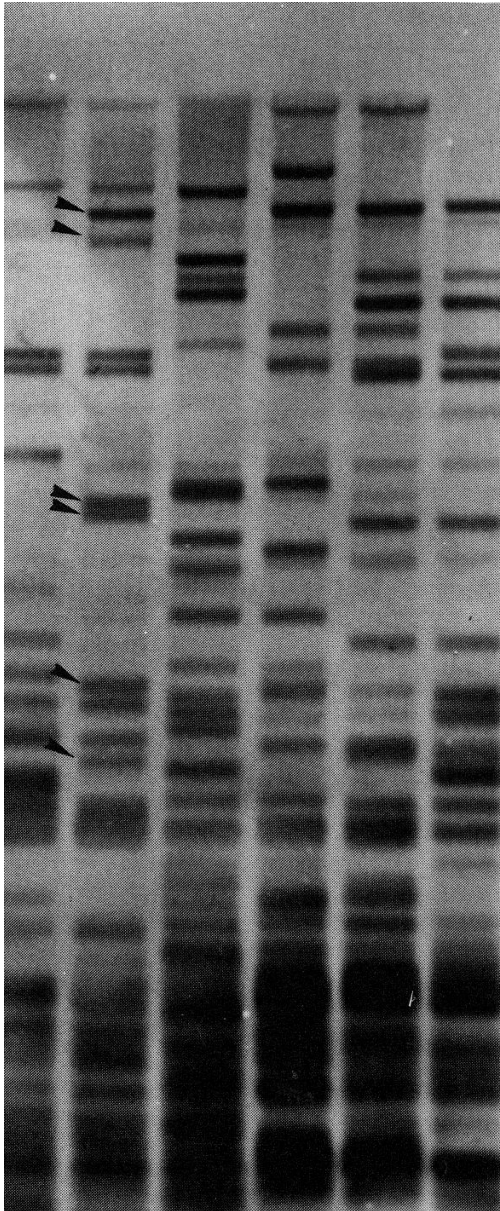
Blood groups

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.

ATCGGCTAGCTGATCGACGATGACCGTAG  
CGTTGATCGGTAGGCTAGCTGAAACTTAAC  
GGA

SNPs

INDELS

ATCTACGGATGGCTGACTGATG

ATCTACGGATGGCTGACTGATG

ATCTACGGATGGCTGACTGATG

Minisatellites

ATCTACGGATGGCTGAGATG Indels

ATTACTGATCGGTAGCTGAGCCAATGGCA  
GTGATGGATGGTAGCTGAGTGCTGGACAT

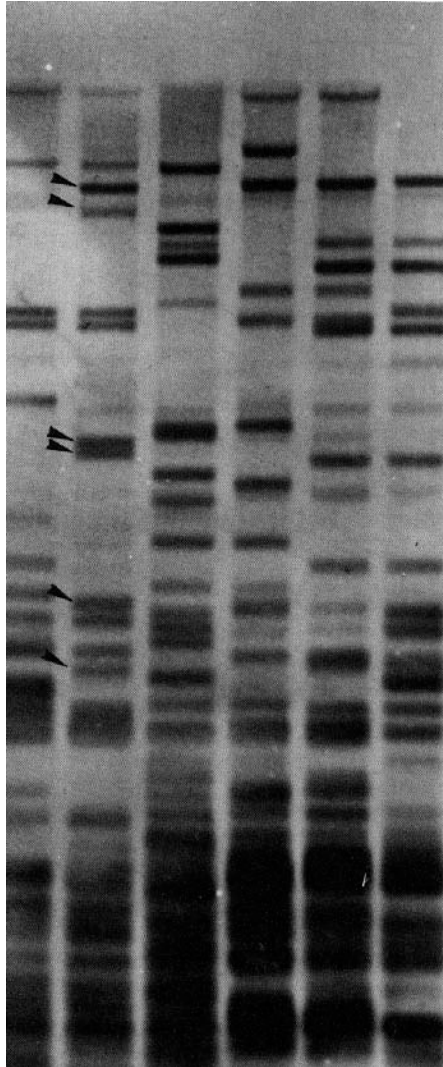
GATA GATA GATA GATA GATA GATA GATA  
GATA

Microsatellites or STRs

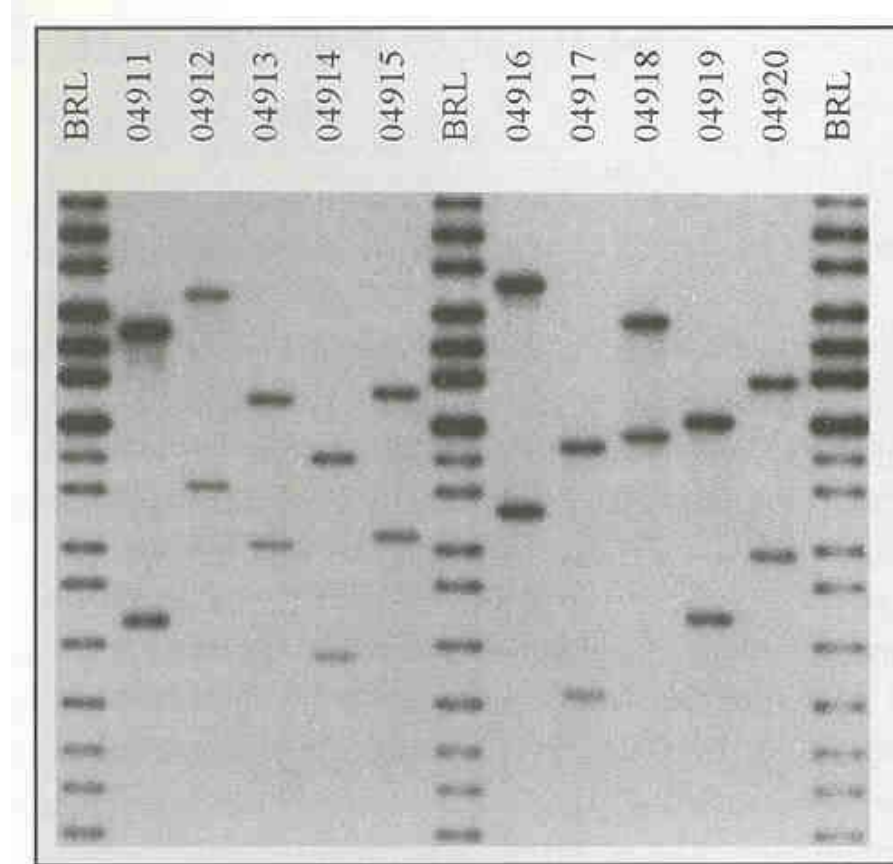
**Copy Number Variants: CNVs represents a copy number change involving a DNA fragment that is ~1 Kb or larger- Feuk et al 2006 Nature Genetics**



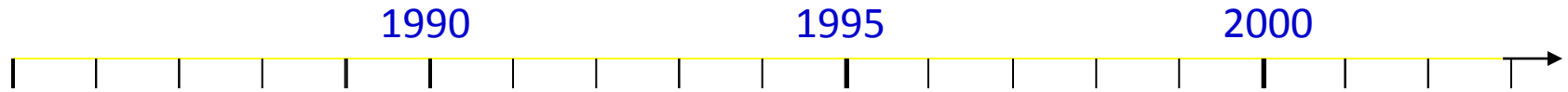
# DETECTION OF MINISATELLITES USING MLPs



# DETECTION OF MINISATELLITES USING SLPs



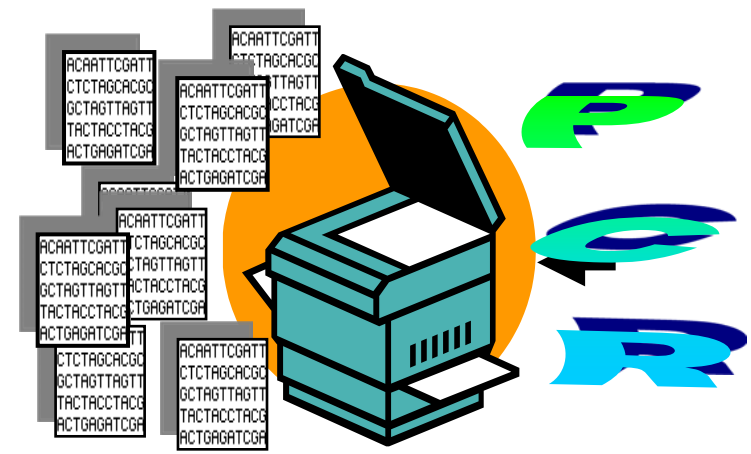
# DNA Typing in Forensic Analysis



Proteins

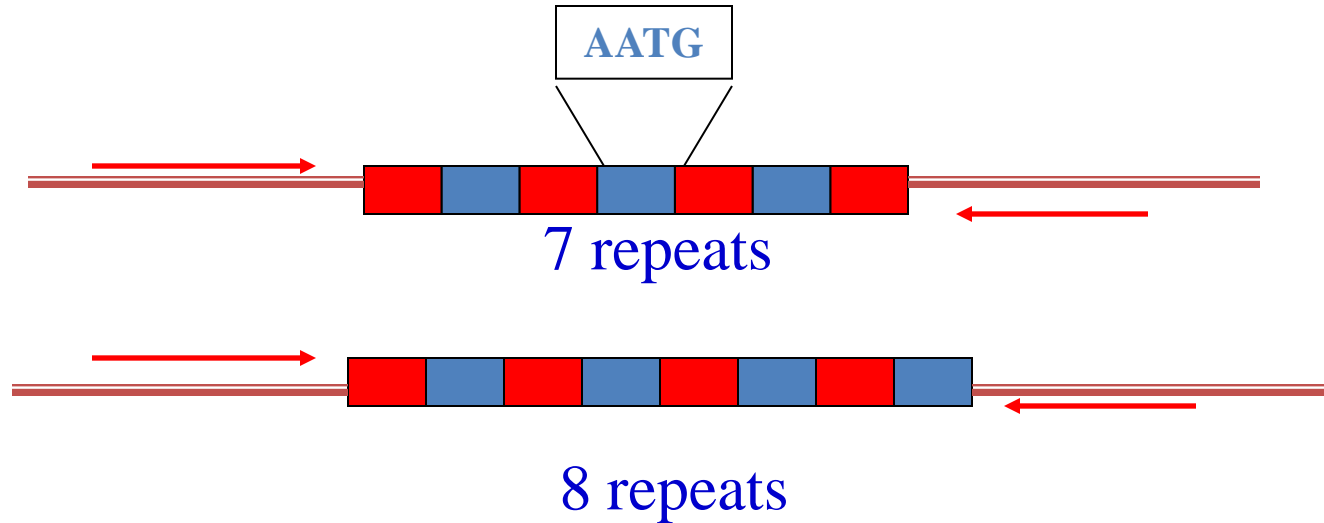
RFLPs

PCR / STRs





# Short Tandem Repeats (STRs)



## ADVANTAGES OF STRs OVER SLPs

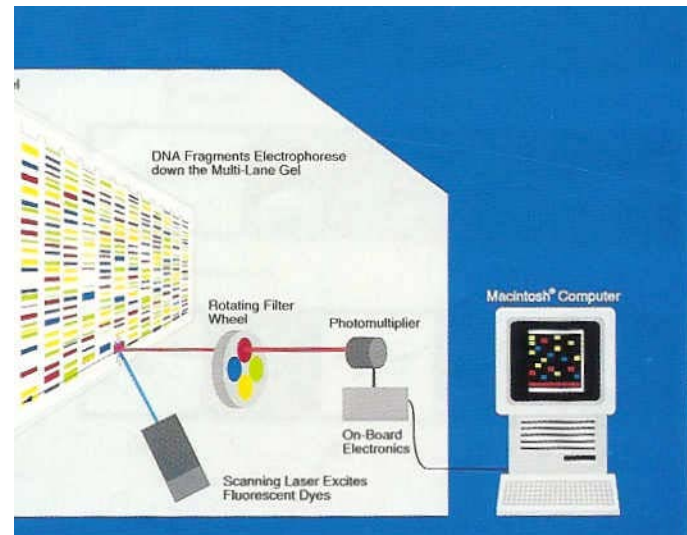
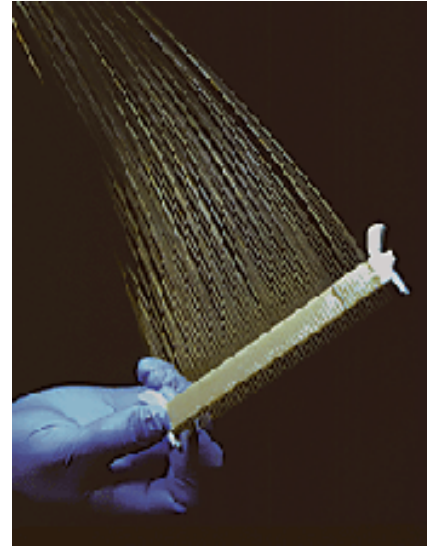
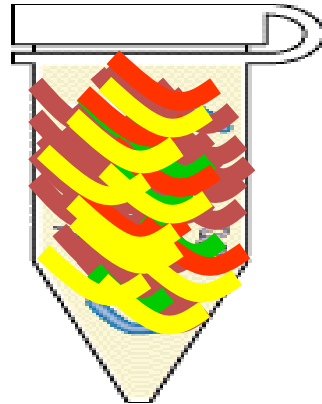
**Amount of DNA required**

**Analysis of degraded samples**

**Time of analysis**

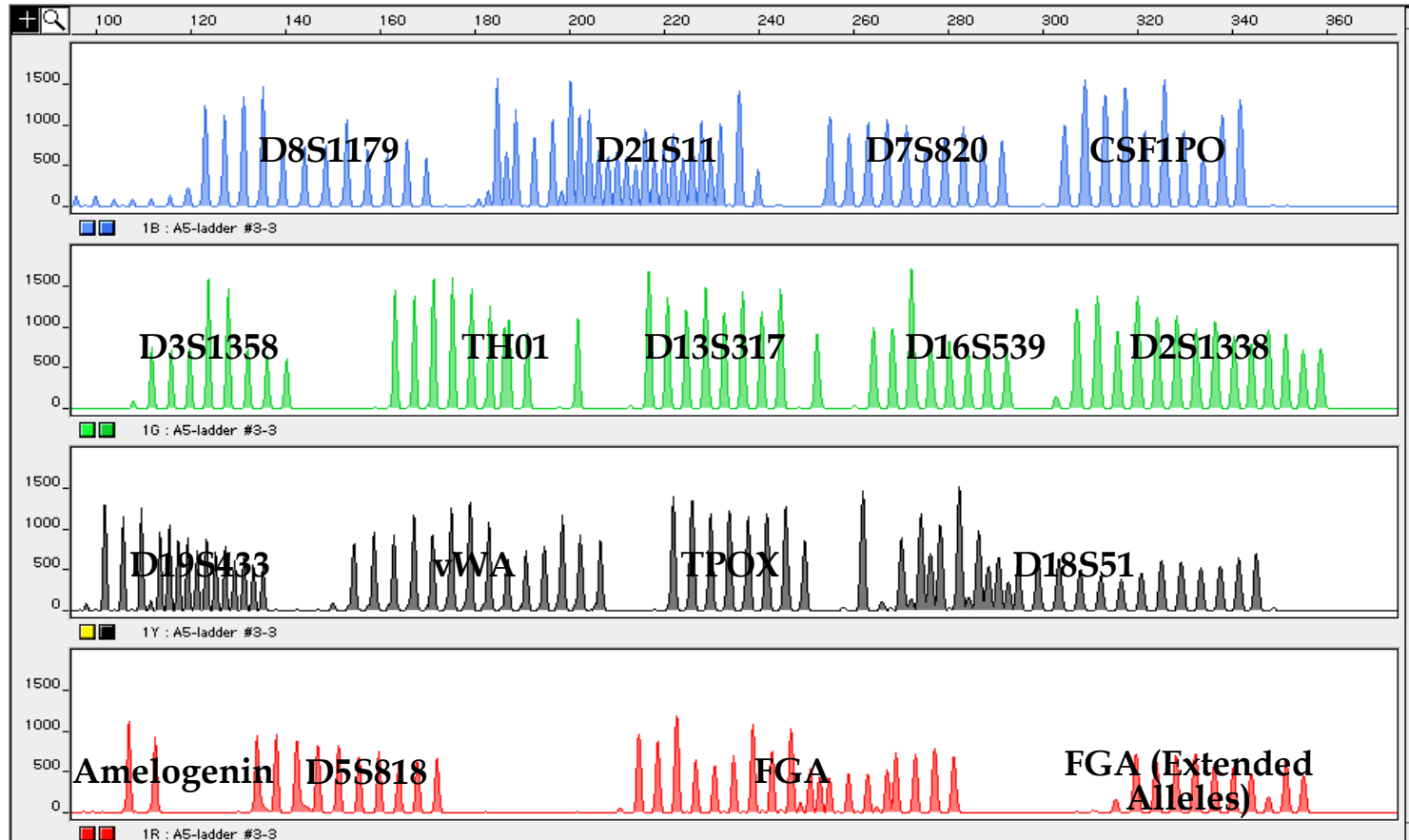
**Standardization and value of the evidence**

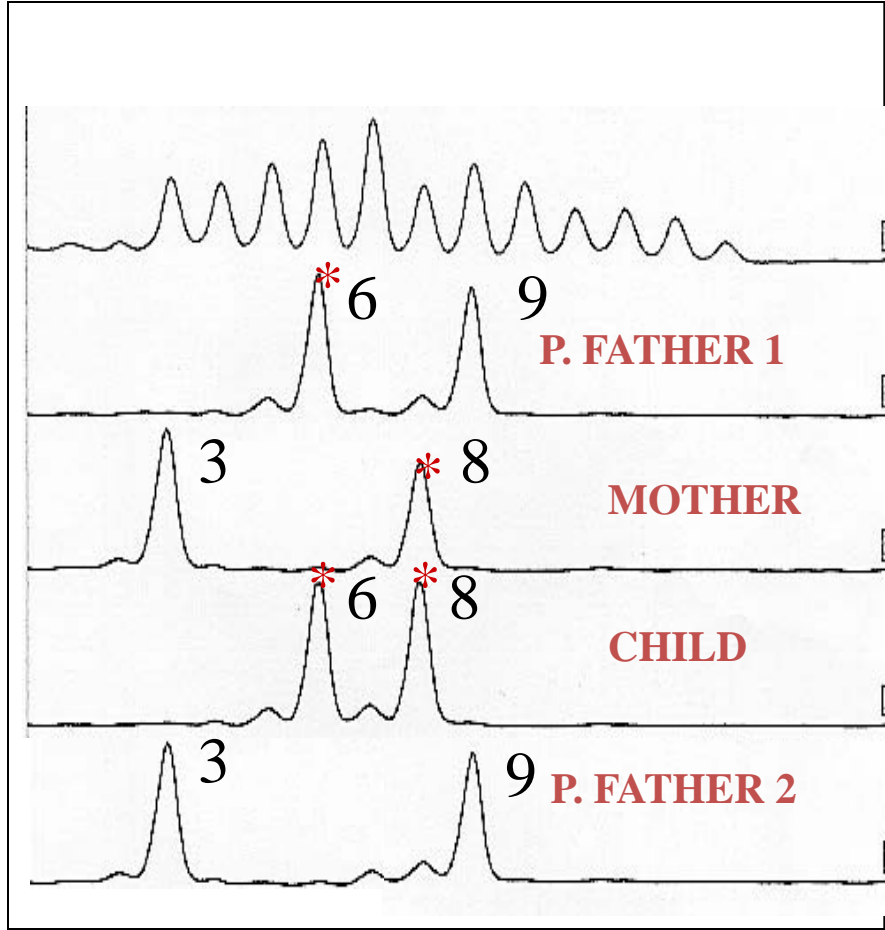
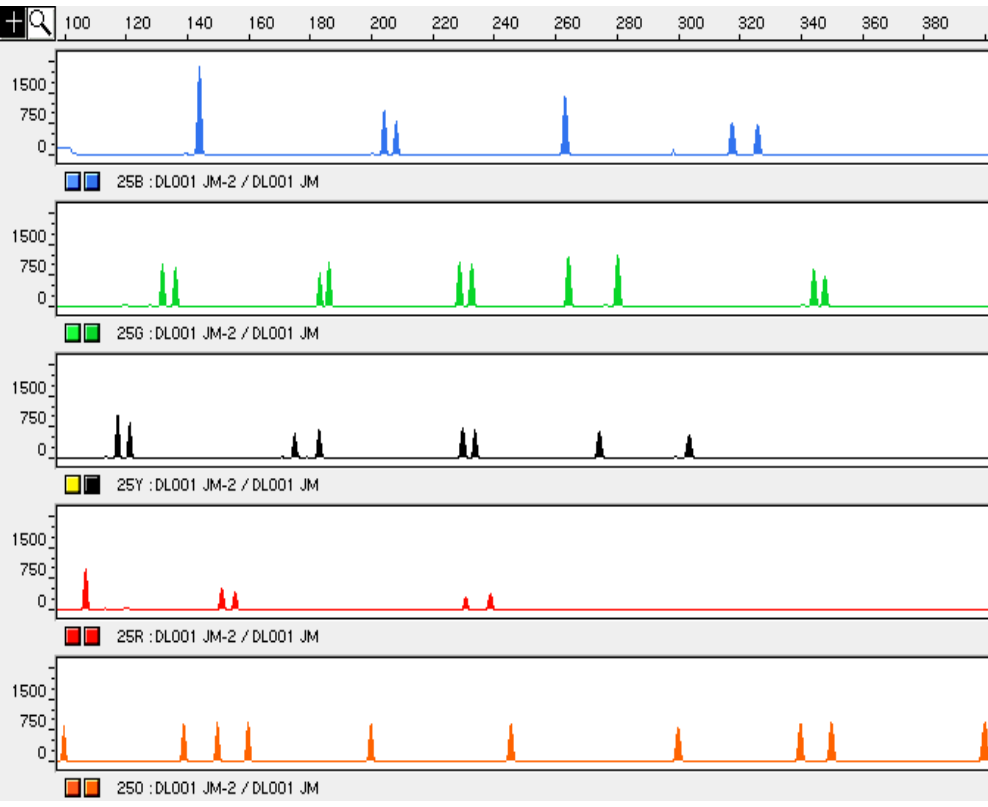
# Fluorochrome technology-PCR multiplex Capillary electrophoresis-Automation



# Multiplexing STRs

16 PLEX





# Commonly Used STR Loci

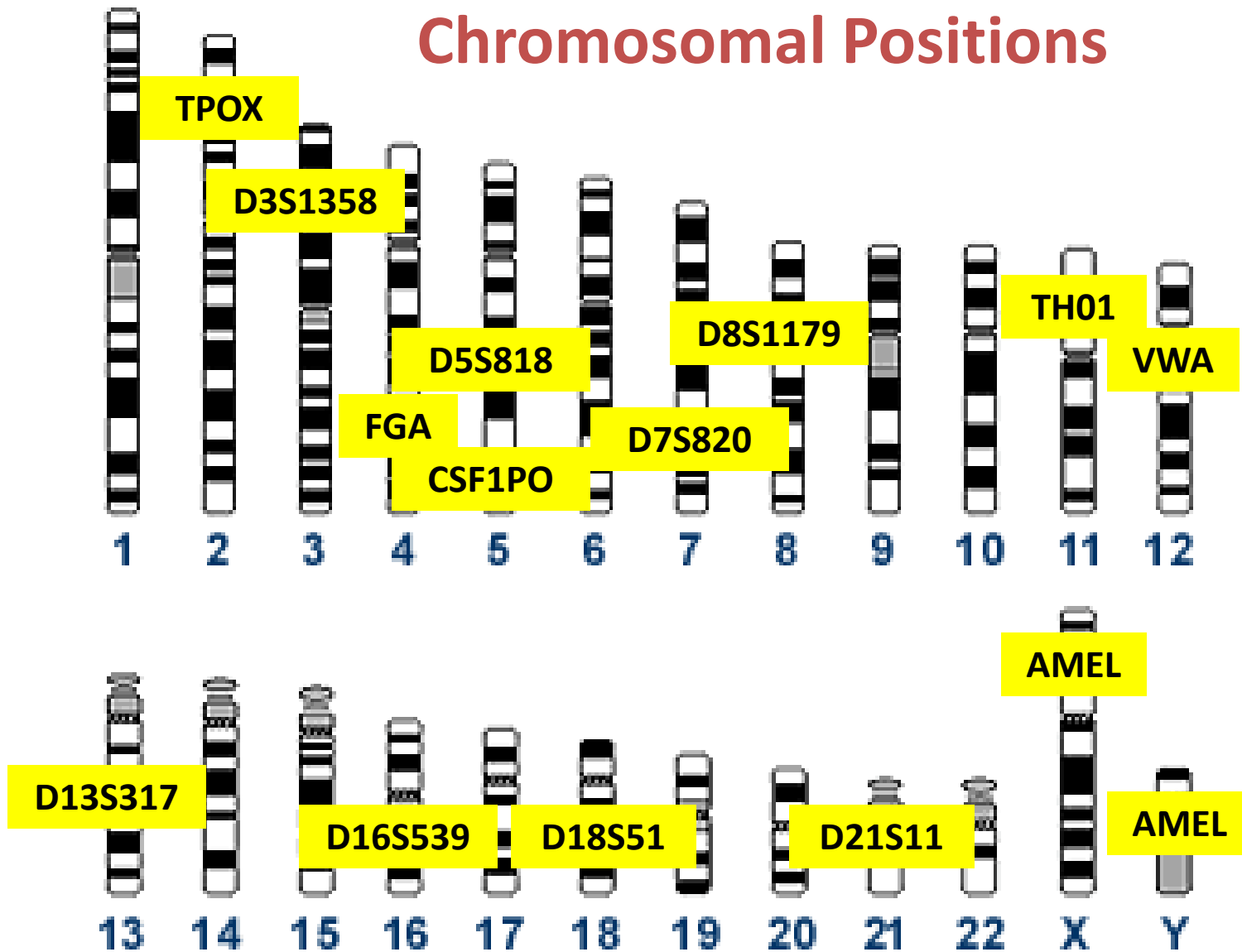
STRs	ESS/ISSOL	CODIS	DAD	Profiler	Prof. Plus	Cofiler	SGM Plus	PowPlex16	Identifiler
D3S1358	x	x	x	x	x	x	x	x	x
VWA	x	x	x	x	x		x	x	x
FGA	x	x	x	x	x		x	x	x
Amelogenin	(x)		x	x	x	x	x	x	x
THO 1	x	x	x	x		x	x	x	x
TPOX		x		x		x		x	x
CSF1PO		x		x		x		x	x
D5S818		x		x	x			x	x
D13S317		x		x	x			x	x
D7S820		x		x	x	x		x	x
D2S1338							x		x
D8S1179	x	x	x		x		x	x	x
D18S51	x	x	x		x		x	x	x
D21S11	x	x	x		x		x	x	x
D16S539		x				x	x	x	x
D19S433							x		x
Penta D								x	
Penta E								x	
SE 33			x				(x)	(x)	

**4 bp-5bp in repeat unit**

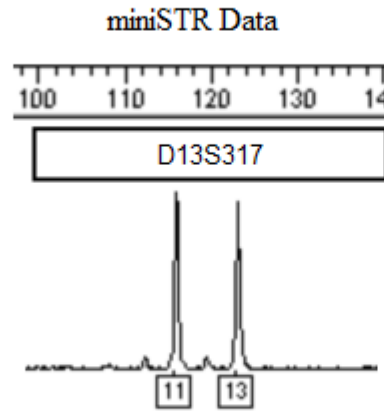
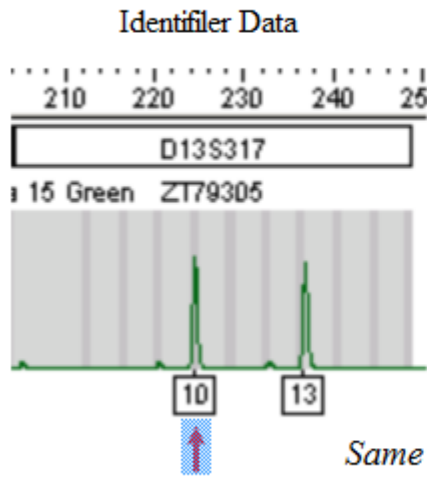
(SE-Filer / PowerPlex ES)

ESS-ISSOL: THO1, VWA, FGA, D21S11, D8S1179, D3S1358, D18S51

# 13 CODIS Core STR Loci with Chromosomal Positions

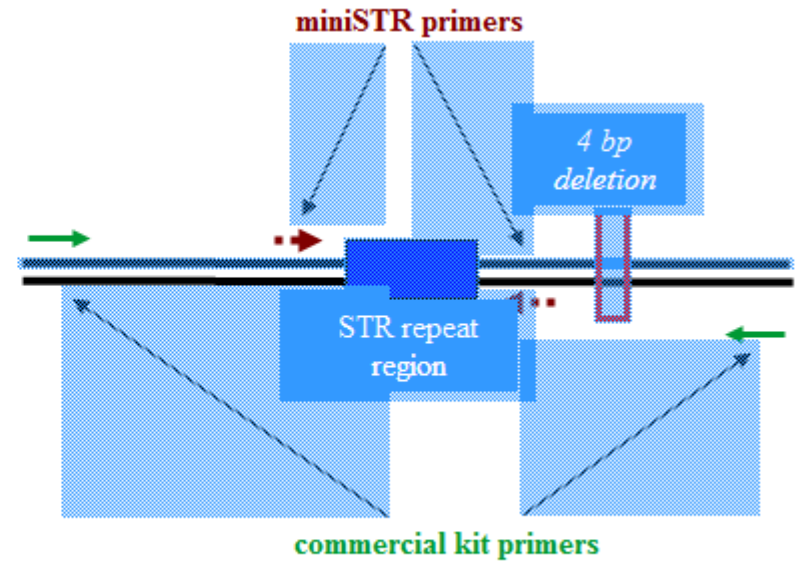


# MINISTRs



Same DNA Sample

really "11 repeats -4 bp"



**Group I Group II**  
 D10S1248 D12S391  
 D22S1045 D1S1656  
 D2S441 TPOX

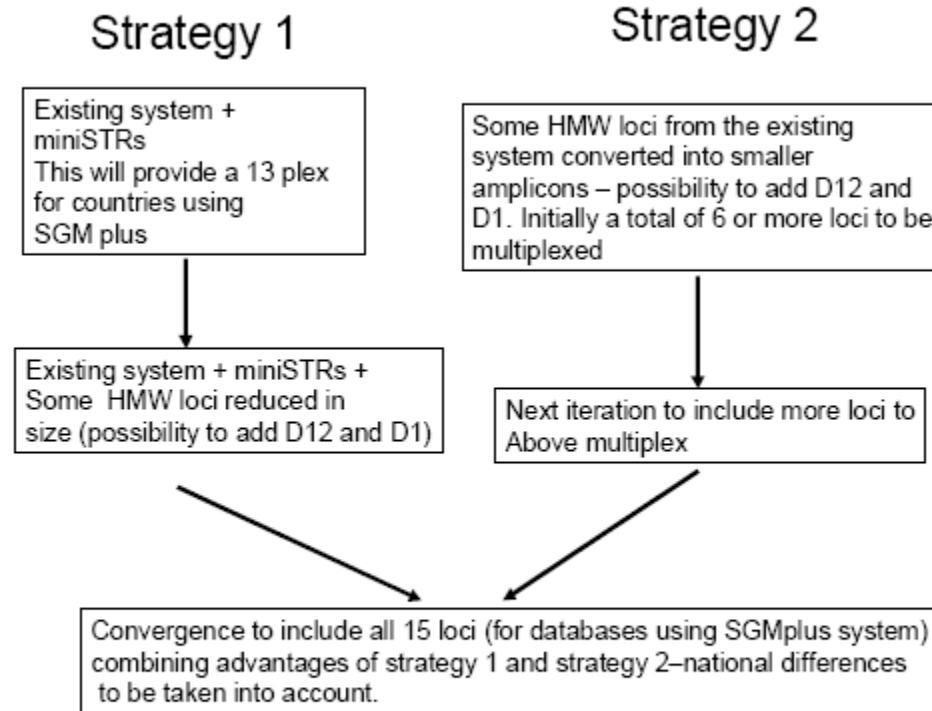


Fig 1: Two multiplex strategies, showing path to convergence.

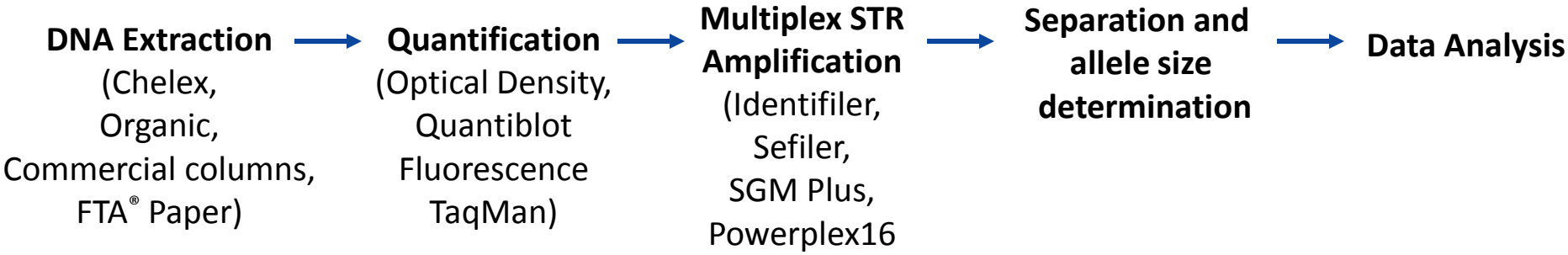
[Gill P, Fereday L, Morling N, Schneider PM.](#)

New multiplexes for Europe-amendments and clarification of strategic development.

Forensic Sci Int. 2006 Nov 10;163(1-2):155-7.



# The DNA analysis procedure in forensics



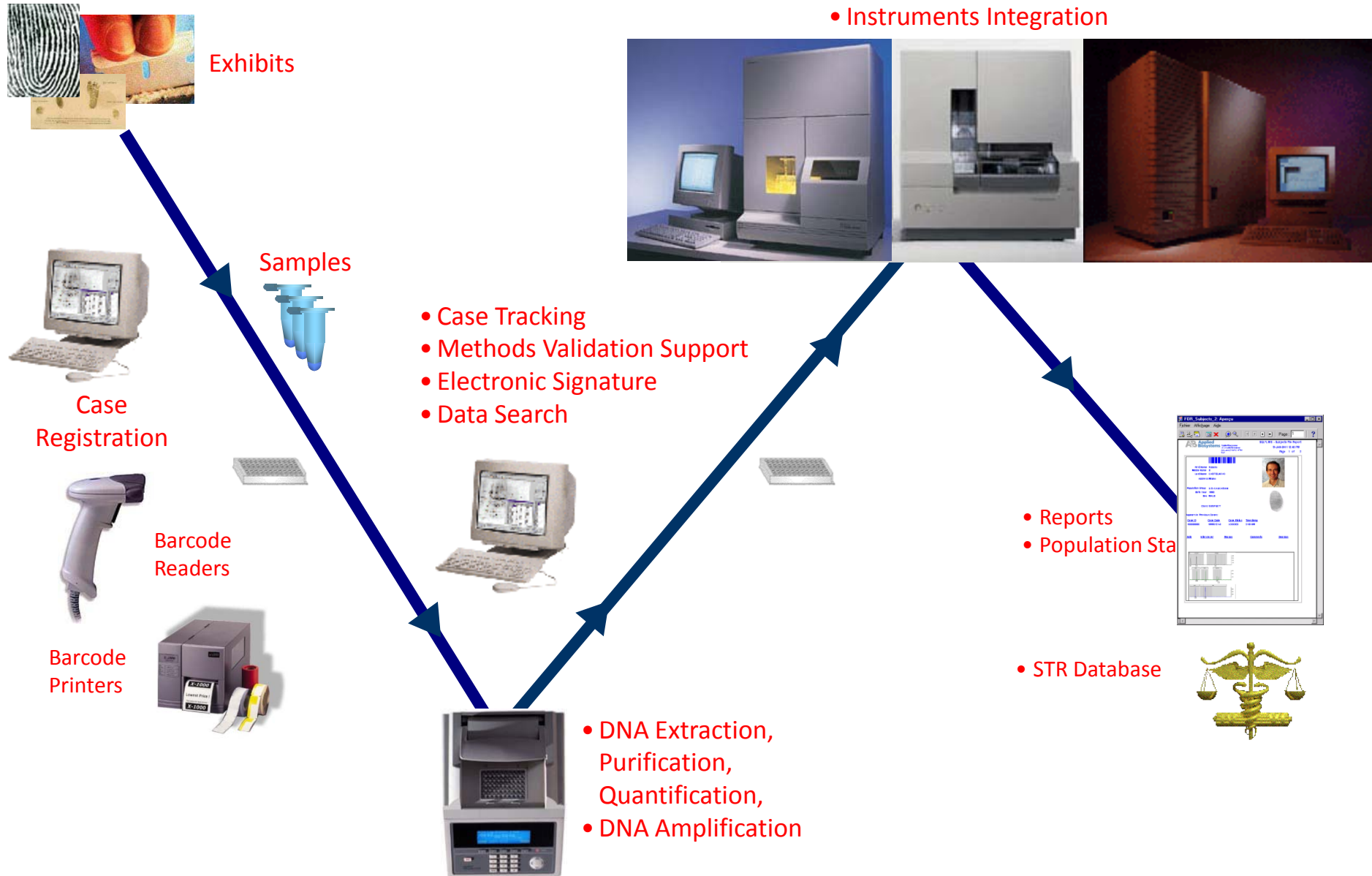
Population genetics-Data

Statistical evaluation of the evidence

Communication

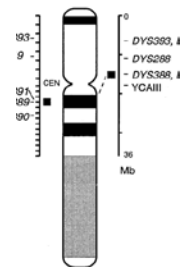
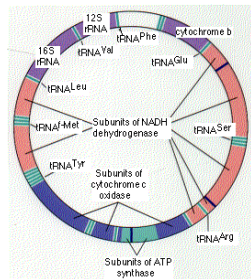
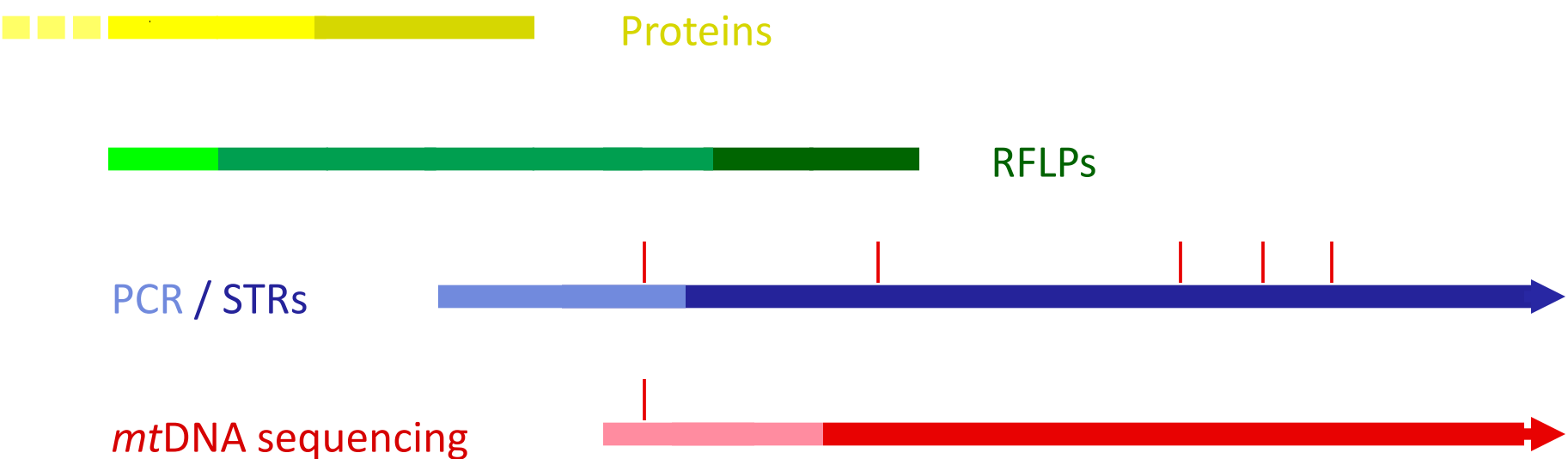


# Forensic LIMS





# DNA Typing in Forensic Analysis

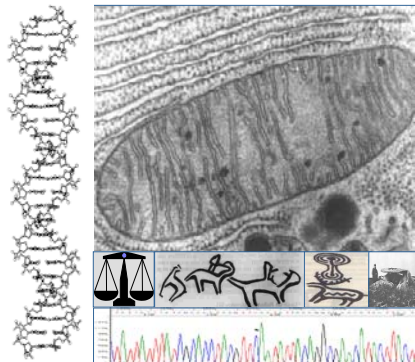
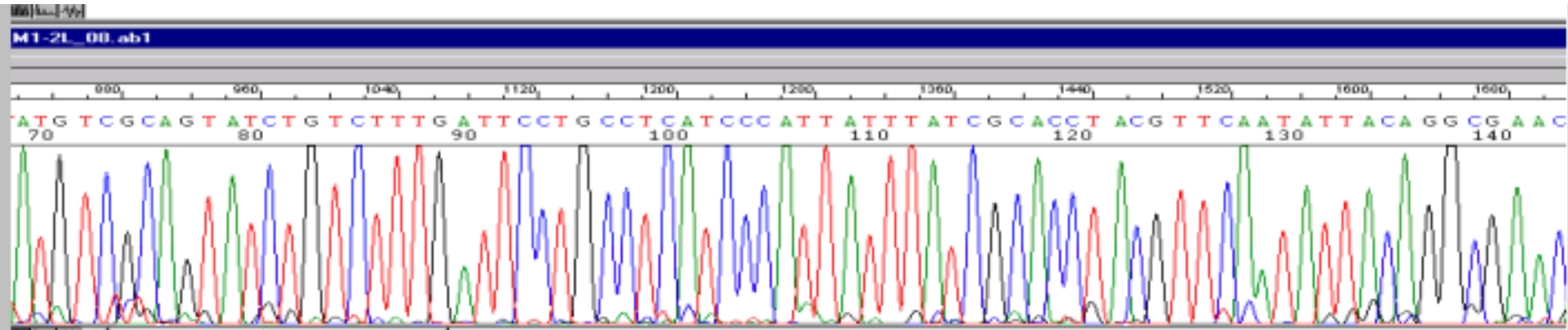


Y-chr markers

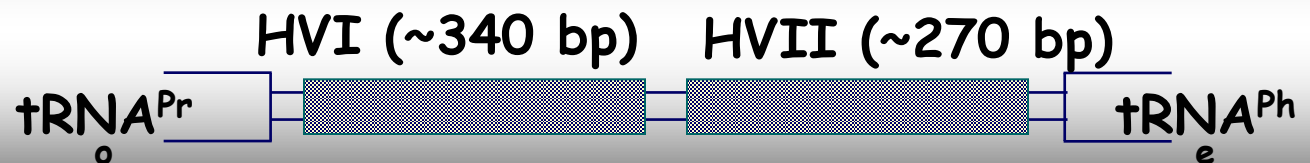


# Advantages of mtDNA analysis over nucDNA

*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)*



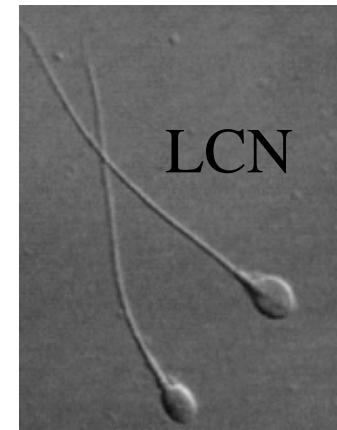
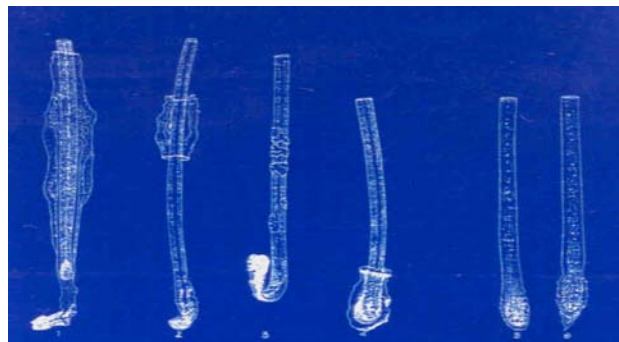
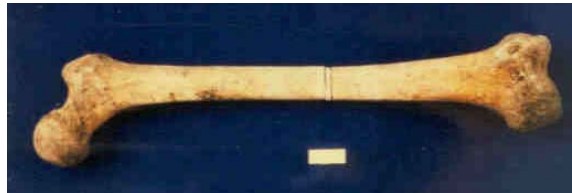
D-loop: Control region



# Applications of mtDNA analysis in forensic genetics

**Criminal casework: Low copy number specimens or lack of nuclear DNA- HAIR SHAFTS OR TELOGENIC HAIRS**

**Identification: Old skeletal remains-Highly degraded material- Civil cases and historical cases-Reconstruction of maternal lineages**

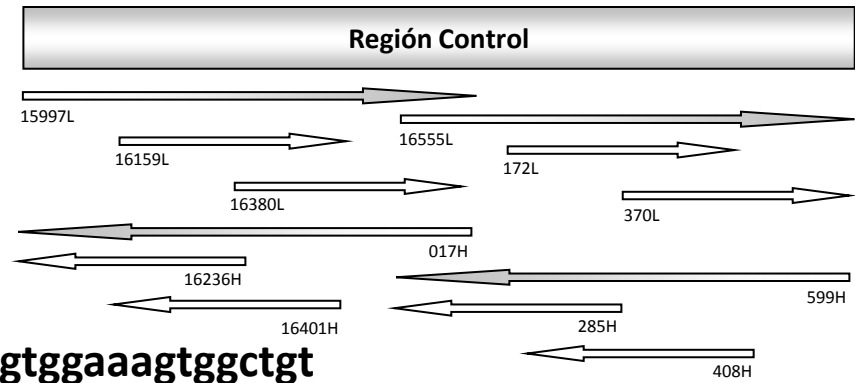


**Primer L (5'-3')**

**Primer H Secuencia (5'-3')**

**HVSI**

L15997 caccattagcacccaaagct H16142 ttgtacggtaccataaatac  
 L16055 gaagcagatttgggtaccac H16157 actacaggtggtcaagtatttatggt  
 L16121 tactgccagccacatgaat H16218 tacaagcaagtacagcaatc  
 L16131 caccatgaatattgtacggt H16236 ggctttggagttgcagttgatg  
 L16159 tacttgaccacctgtagtac H16260 ttggtatcctagtggggtgagg  
 L16185 accaatccacatcaaaacc H16261 ctgcaactccaaagccacc  
 L16209 cccatgcttacaagcaagta H16281 ttggtatcctagtggggtgagg  
 L16247 ctatcacacatcaactgcaa H16306 tgtacggtaaattggctttatgtactatg  
 L16254 cacatcaactgcaactccaaa H16313 ctatgtacggtaaattggctttatg  
 L16275 caccctcaccactagga H16380 gtcaagggaccctatctgag  
 L16296 cccacccttaacagtacatagtacataa H16382 tggtaagggaccctatct  
 H16401 tgatttcacggaggatggtg  
 H16410 gtcccttgaccaccatcctc



**HVSII**

L034 gagctctccatgcatttgggt H255 tctgtgtggaaagtggctgt  
 L127 gagcaccctatgtcgcagta H405 ttttggcggatgcactttt

**Primers used for PCR amplification of mtDNA control region**







## Mitochondrial DNA Control Region Database

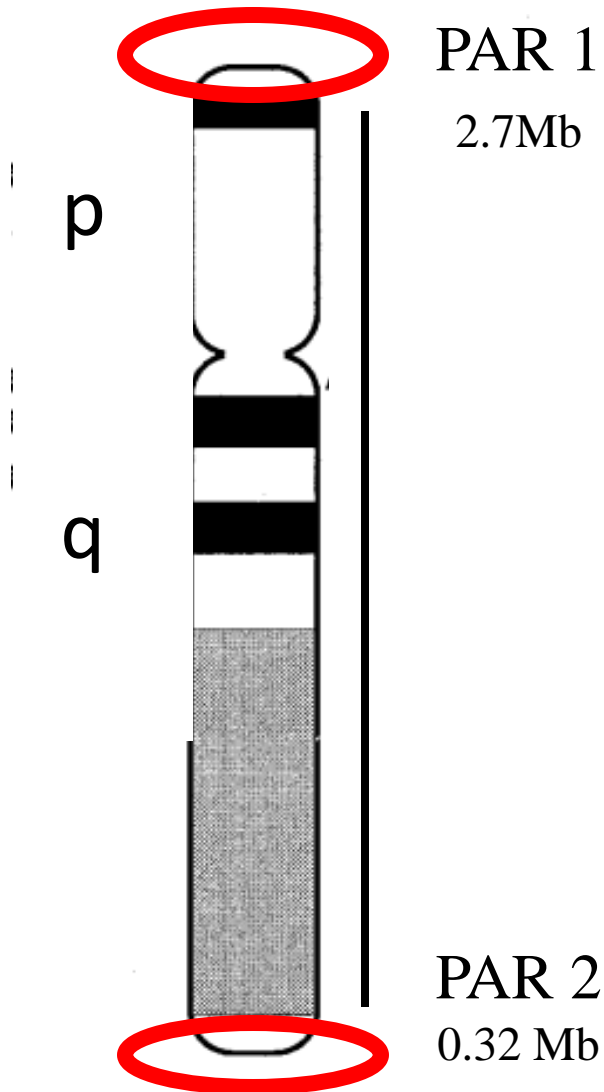
### Introduction

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 to exclude samples as originating from the same individual/matriline. If two samples cannot be excluded 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 presentation 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.

# Inheritance properties of the Y Chromosome



Except for PAR 1 and 2

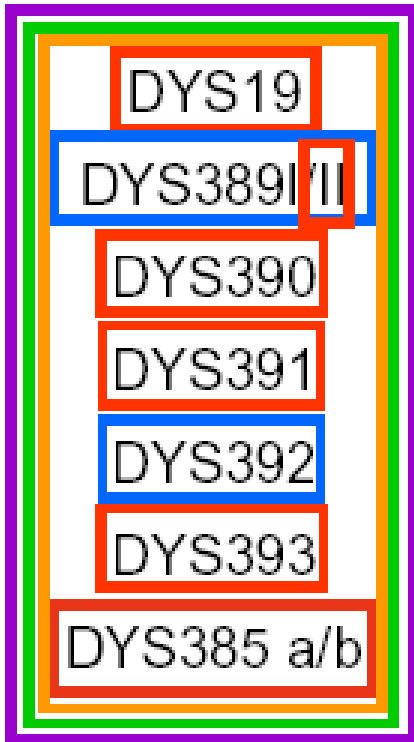
Y chromosome:

- is male specific
- haploid
- transmitted from father to son unchanged unless a mutational event takes place



(Minimal/extended haplotype)

# Commercial Y-STR Kits



(White et al.)

A7.1 (DYS460)

A7.2 (DYS461)

A10

C4

H4

(Ayub et al.)

DYS434

DYS435

DYS436

DYS437

DYS438

DYS439

(Iida et al.)

DYS441

DYS442

DYS443

DYS444

DYS445

(Redd et al.)

DYS446

DYS447

DYS448

DYS449

DYS450

DYS452

DYS453

DYS454

DYS455

DYS456

DYS458

DYS459 a/b

DYS463

DYS464 a/b/c/d

(Bosch et al.)

G09411 (DYS462)

43 (51) Y-STRs  
(217 with Manfred's)

YCAII a/b

DYS388

DYS425

DYS426

YCAIII a/b

Y-PLEX 6 (ReliaGene)

Y-PLEX 5 (ReliaGene)

Y-PLEX 12 (ReliaGene)

PowerPlex Y (Promega)

Yfiler (Applied Biosystems)

DYS468-DYS645

166 new Y STRs

(Manfred Kayser GDB entries)

## About the "YHRD - Y Chromosome Haplotype Reference Database"

as pdf email link

### Release 22:

52,655 Haplotypes  
in 464 populations



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



To view the metapopulation structure of the YHRD (Popsearch) click in the map

The ability to identify male-specific DNA renders polymorphic Y-chromosomal sequences an invaluable addition to the standard panel of autosomal loci used in forensic genetics. Y-STR haplotyping is particularly important for sensitive typing of male DNA in mixed stains as well as for rapid assortment of biological crime scene evidence.

Moreover, Y chromosomal profiling can trace back paternal lineages into the past and has thus been proven a useful tool in genealogical and kinship testing.

The individuality of the male-specific part of the Y chromosome can be optimally explored by the Y haplotype analysis using a set of highly variable short tandem repeat markers approved by the forensic scientific community.

### Latest news

#### October 10 6th Forensic Y Chromosome User Workshop (Lutz Roewer)

The 6th workshop as part of the congress "DNA in Forensics 2008" will take place for the first time in Italy, in Ancona, May 27-30, 2008. ([read more](#))

#### September 14 Online SNPY Database and Encyclopedia (Sascha Willuweit)

We proudly announce the launch of [SNPY.org](#), a database built on the Wiki principles to support the continuous curation and validation of the Y chromosome phylogenetic tree. ([read more](#))

#### August 10 YHRD update (Lutz Roewer, Sascha Willuweit)

Release 22 is out with 52,655 haplotypes in 464 populations. 50,867 haplotypes of these are completely typed for 9 (MinHt) and 23,981 for 11 loci (ExtHt). ([read more](#))

Search database

Frequency calculation

About this project

YHRD contributors

How to submit?

Haplotype characteristics

Mutations

Population analyses

Statistics

News

6th Y User workshop 2008

Downloads



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)



Forensic Science International: Genetics 1 (2007) 93–99



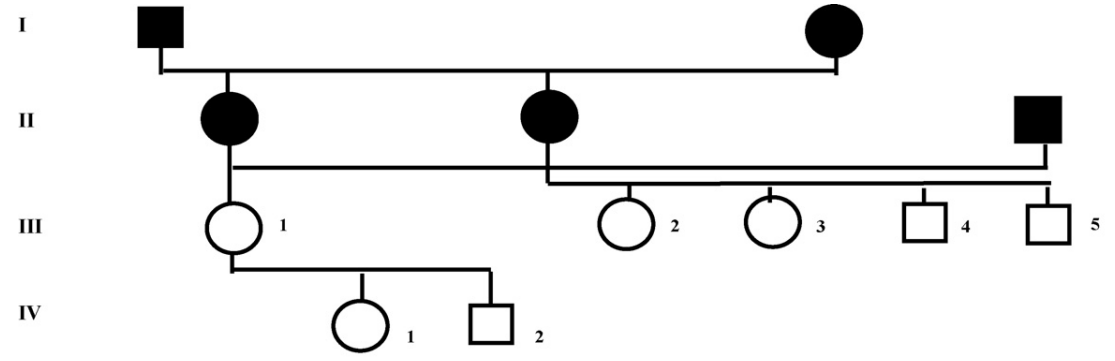
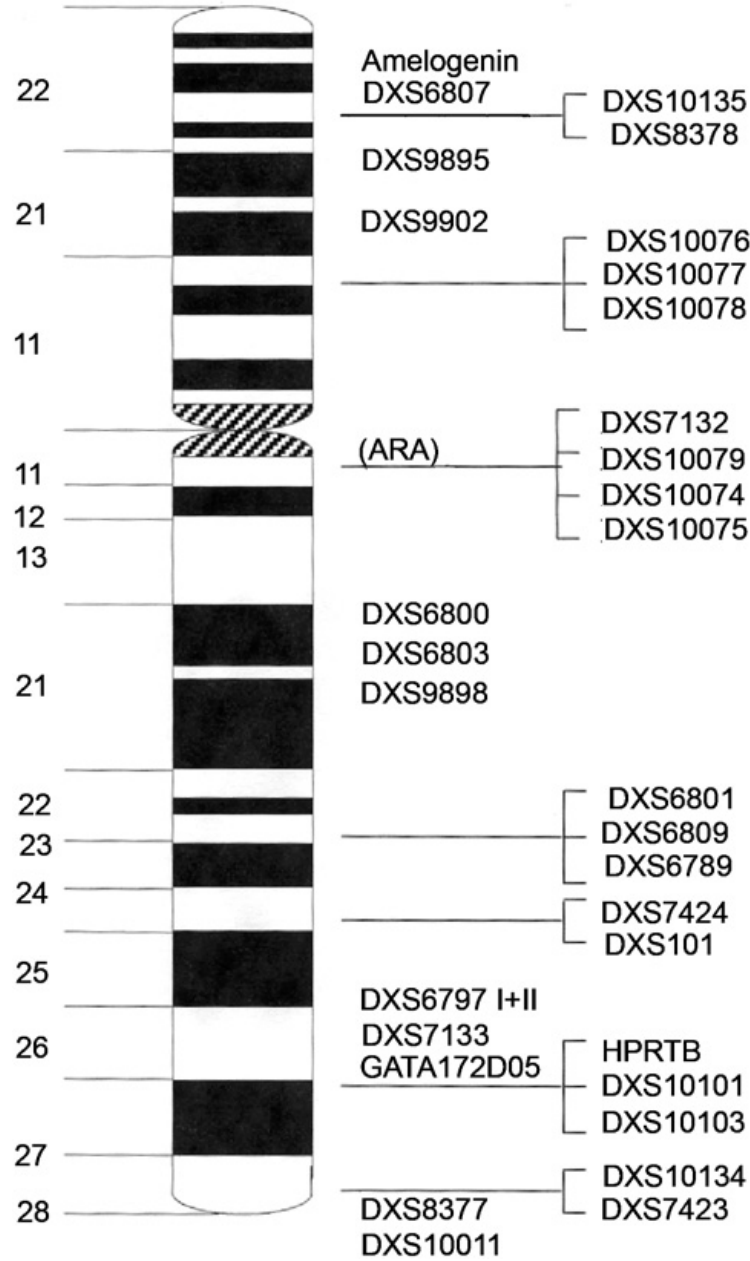
[www.elsevier.com/locate/fig](http://www.elsevier.com/locate/fig)

Review

# X-chromosomal markers: Past, present and future

Reinhard Szibor\*

*Institut für Rechtsmedizin, Otto-von-Guericke-Universität Magdeburg, Leipziger Straße 44, D-39120 Magdeburg, Germany*



	III-1		IV-1		IV-2		III-2		III-3		III-4		III-5	
DXS10135	19 ▼	27	19	29	27	19 ▼	26	●	19 ▼	26	●	26	●	26
DXS8378	13 ▼	12	13	11	12	13 ▼	11	●	13 ▼	11	●	11	●	11
DXS7132	14 ▼	15	●	14	15	15	●	14 ▼	12	●	12	●	12	●
DXS10074	17 ▼	19	●	17	13	19	●	17 ▼	19	●	17 ▼	16	●	16
HPRTB	13 ▼	12	●	12	●	13	12	●	13 ▼	12	●	12	●	12
DXS10101	31 ▼	28.2	●	28.2	●	31	28.2	●	31 ▼	29.2	●	31 ▼	28.2	●
DXS10102	12 ▼	10	●	10	●	10	10	●	12 ▼	10	●	10	●	11
DXS10134	37 ▼	35	●	35	●	38	35	●	37 ▼	34	■	37 ▼	35	●
DXS7423	16 ▼	15	●	15	●	15	15	●	16 ▼	14	■	16 ▼	15	●
DXS10146	24 ▼	18	●	18	●	20	18	●	24 ▼	18	■	24 ▼	18	●

# DNA Typing in Forensic Analysis

1985

2000

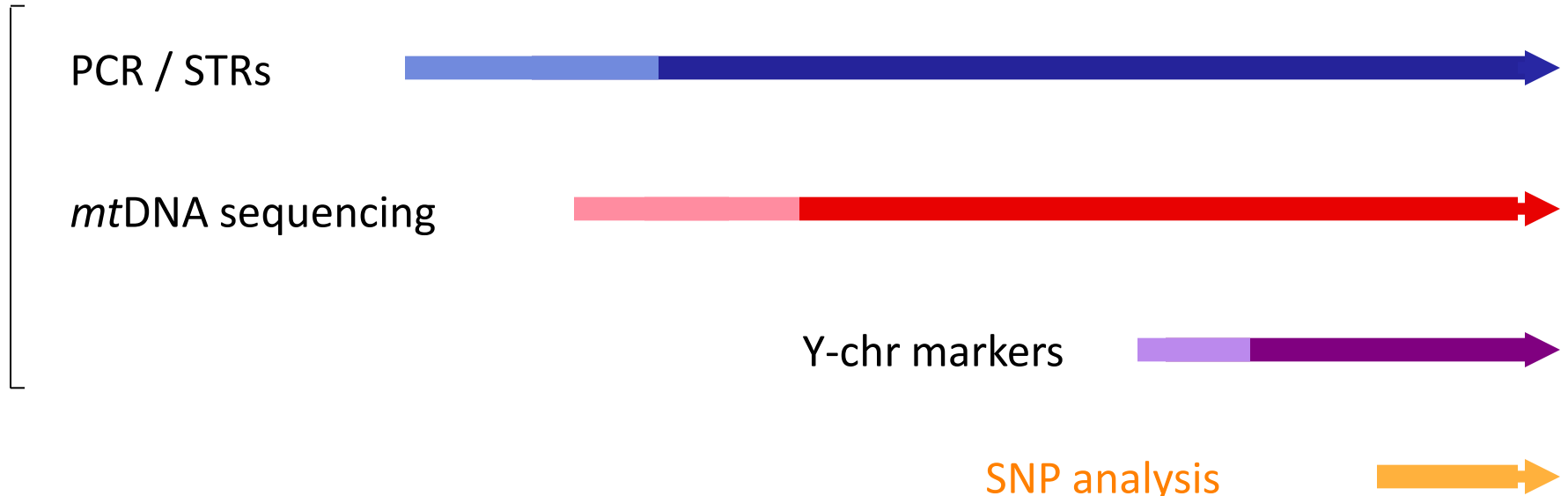
- Complicated paternity testing cases
- Degraded DNA
- Automated high throughput typing
- Trait analysis (character prediction)
- Geographical origin

PCR / STRs

*mtDNA* sequencing

Y-chr markers

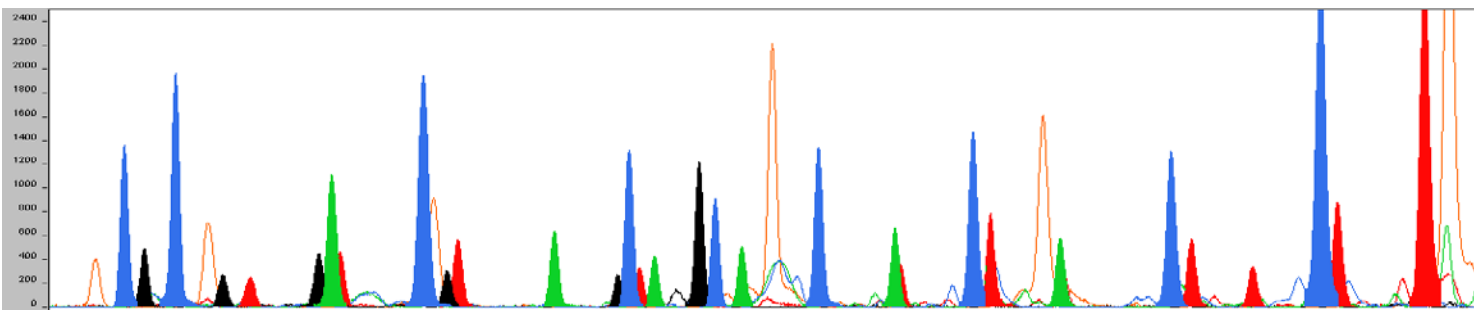
SNP analysis



# The SNP for ID Consortium



## The First Autosomal Multiplex



*Electrophoresis* 2006, 27, 1713–1724

1713

Juan J. Sanchez<sup>1</sup>  
Chris Phillips<sup>2</sup>  
Claus Børsting<sup>1</sup>  
Kinga Balogh<sup>3</sup>  
Magdalena Bogus<sup>3</sup>  
Manuel Fondevila<sup>2</sup>  
Cheryl D. Harrison<sup>4</sup>  
Esther Musgrave-Brown<sup>4</sup>  
Antonio Salas<sup>2</sup>  
Denise Syndercombe-Court<sup>4</sup>  
Peter M. Schneider<sup>3</sup>  
Angel Carracedo<sup>2</sup>  
Niels Morling<sup>1</sup>

<sup>1</sup>Department of Forensic Genetics,  
Institute of Forensic Medicine,  
University of Copenhagen,  
Copenhagen, Denmark

<sup>2</sup>Institute of Legal Medicine,  
University of Santiago de  
Compostela,

Santiago de Compostela, Spain

<sup>3</sup>Institute of Legal Medicine,  
Johannes Gutenberg University,  
Mainz, Germany

<sup>4</sup>Centre for Haematology,  
ICMS, 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 19). Nearly 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 59 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  $5.0 \times 10^{-19}$  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

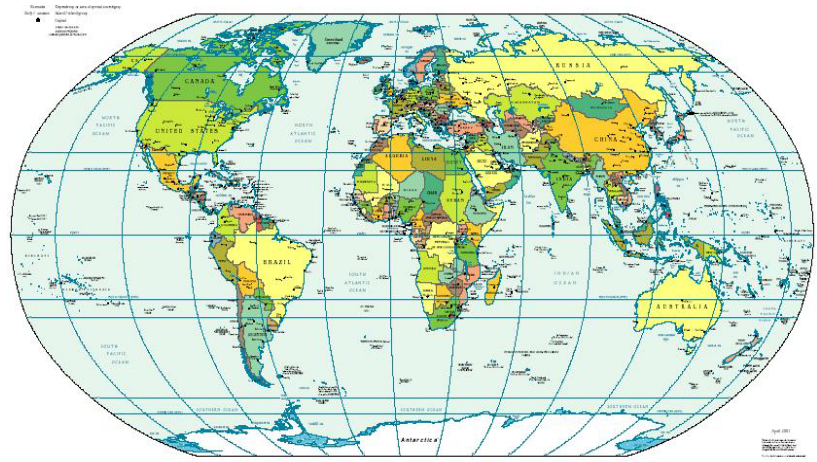


# SNP for ID

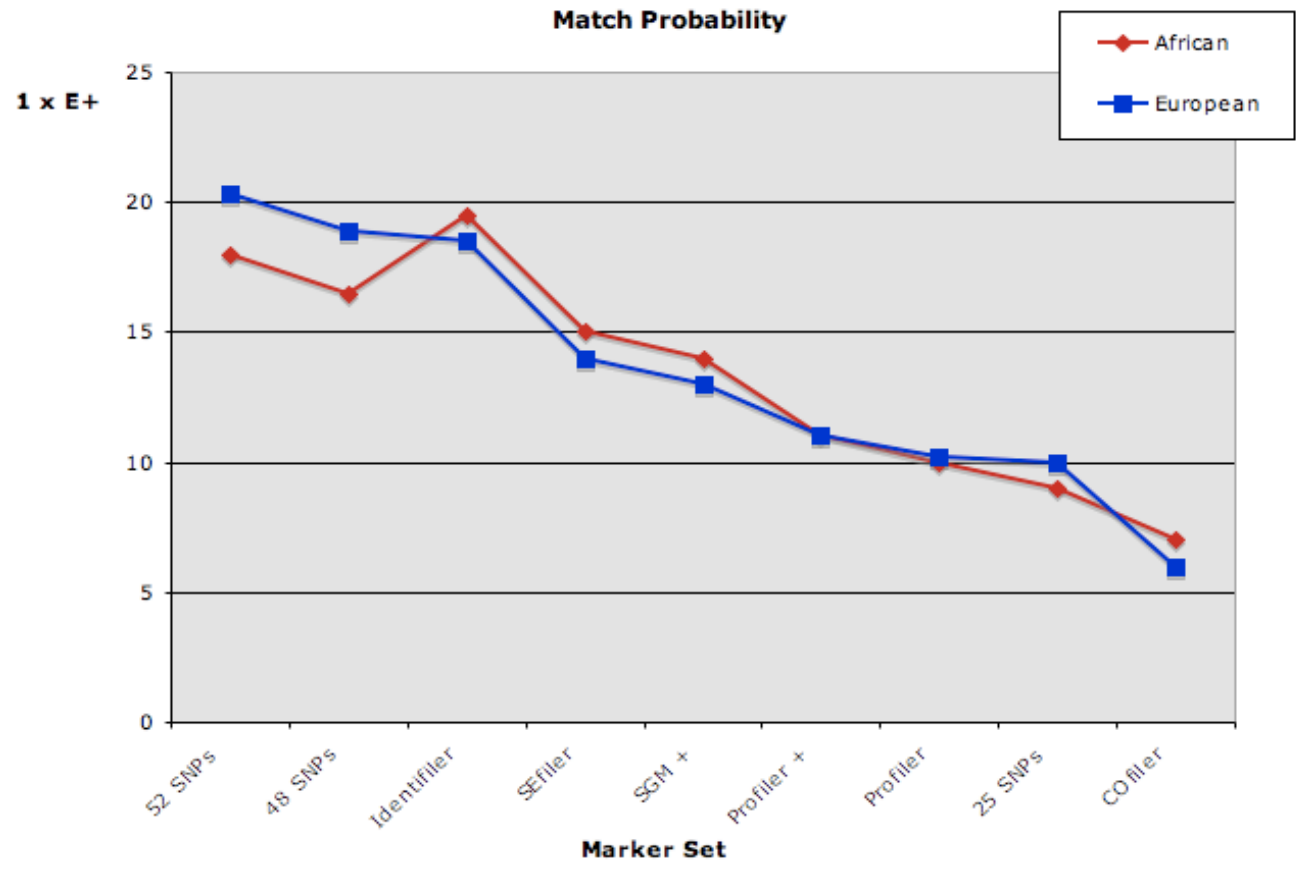


# Selecting SNPs for the SNPforID multiplex

- Selection criteria used
  - ✿ avoiding coding regions and STRs - ensuring sufficient variability - pre-screening candidates on sequence quality - genomic distribution to maximize segregation
- A single multiplex PCR serves multiple platforms
  - two simple genotyping systems:
    - ✿ SNaPshot      Genplex
- SNPs as supplements to STRs: casework applications that require some of the SNP advantages such as the importance of reduced amplicon size



1350 samples + CEPH panel



# Forensic validation in critical samples

Saliva stains degraded for 147 days

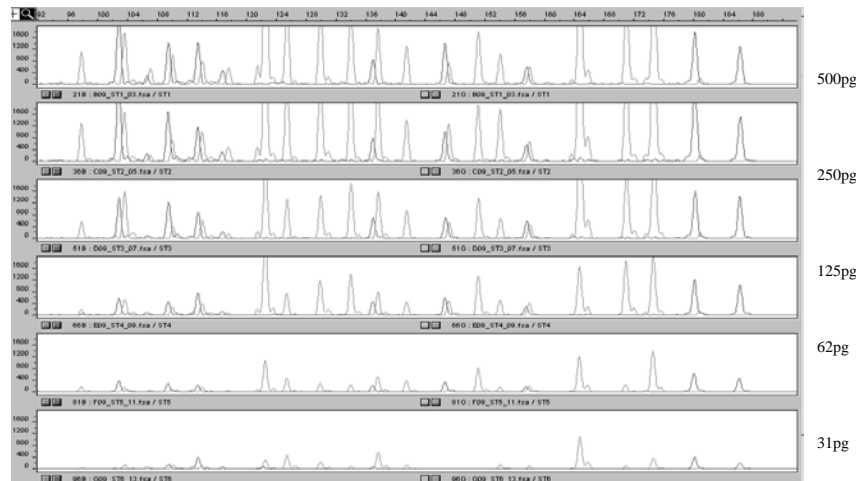
81% complete SNP profile /STRs were only 18% complete

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%



ISFG'05 Azores

# SNP typing applications

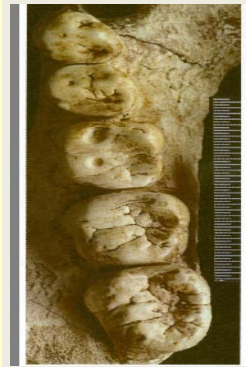
## 11-M MADRID



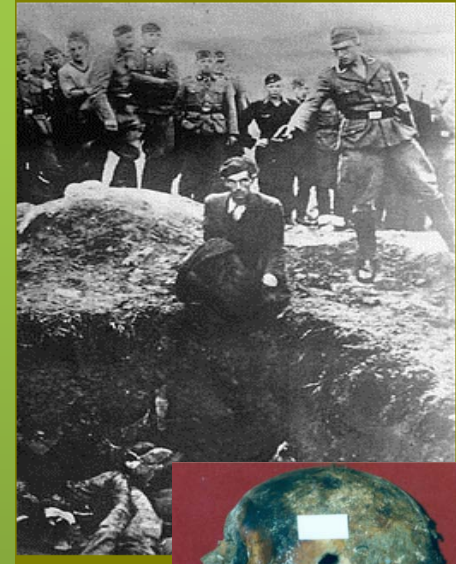
Identification



WTC (New York)

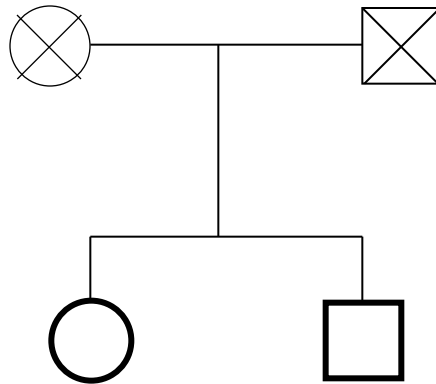


Archeogenetics



# Half brothers?

H<sub>1</sub>



21 STRs

$$\text{LR: } \frac{1}{3} \begin{matrix} H_2 \\ H_1 \end{matrix}$$

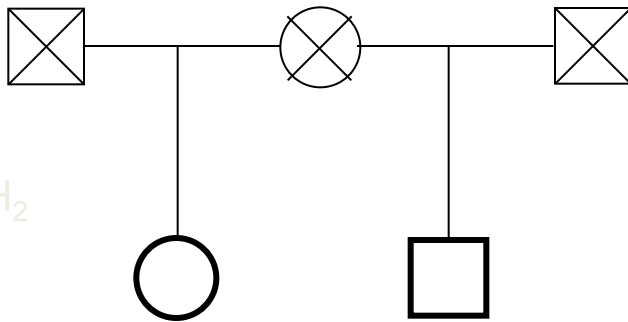
P: 75%

21 STRs + 52 SNPs

$$\text{LR: } \frac{1}{1,193}$$

P: 99.91% H<sub>1</sub>

H<sub>2</sub>



21 STRs

$$\text{LR: } \frac{1}{897} \begin{matrix} H_2 \\ H_1 \end{matrix}$$

P: 99.89%

21 STRs + 52 SNPs

$$\text{LR: } \frac{1}{12,140,628,977}$$

P: 99.999999% H<sub>1</sub>

# 34plex AIM panel for the prediction of geographic origin of samples

Population specific AIMS



15

Skewed frequency AIMS



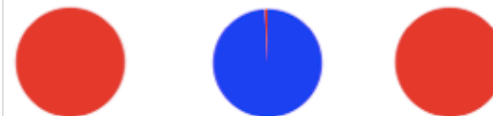
14

Tri-allelic SNPs



2

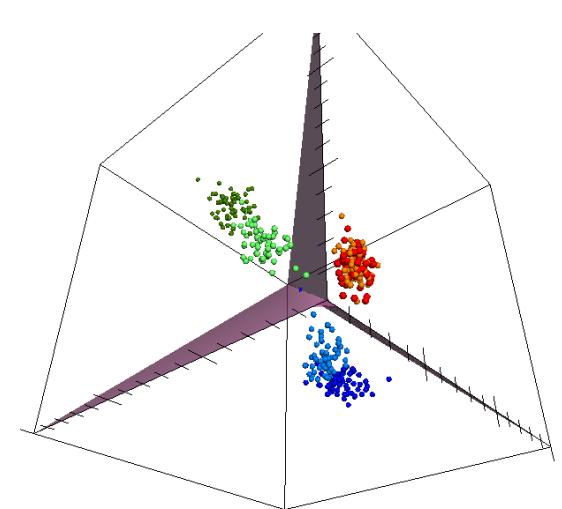
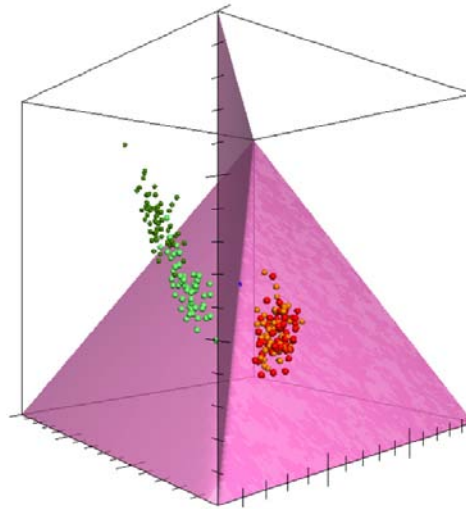
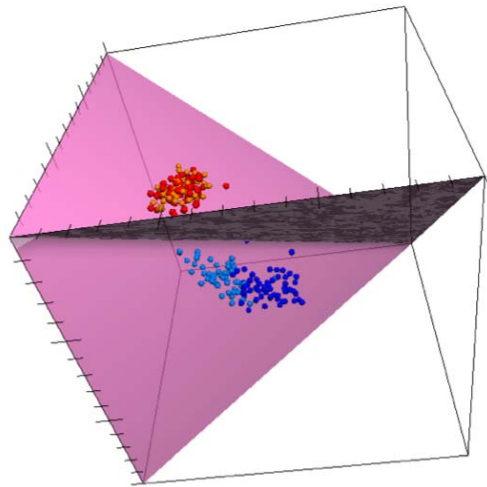
Fixed difference SNPs



3



# 3-D plots provide better depth in plotting divergence of population groups



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



# 11-M ancestry analysis

# casework application



- 7 STR profiles unmatched to suspects
- Presiding udge asked: are these profiles North African or European ? i.e. *a directed or closed population comparison*
- Contact trace extracts meant that DNA was very limited



# Calculation results

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

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

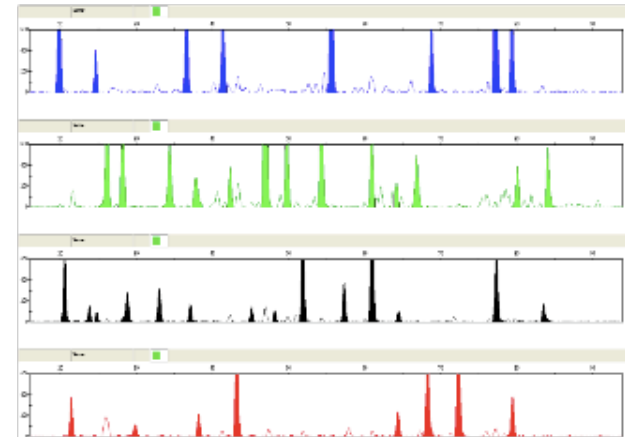


training set:	-log likelihood	percentile	exp	"times more likely to be:	verbal predicate
Gal-Dani	41.109466	2.50%	1.40083E-18	EUR not AFR	164 billion times more likely to be European than African
Moz-Som	66.935324	0.00%	8.5184E-30	1.64448E+11	
Gal-Dani	41.109466	2.50%	1.40083E-18	EUR not ASN	44 billion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

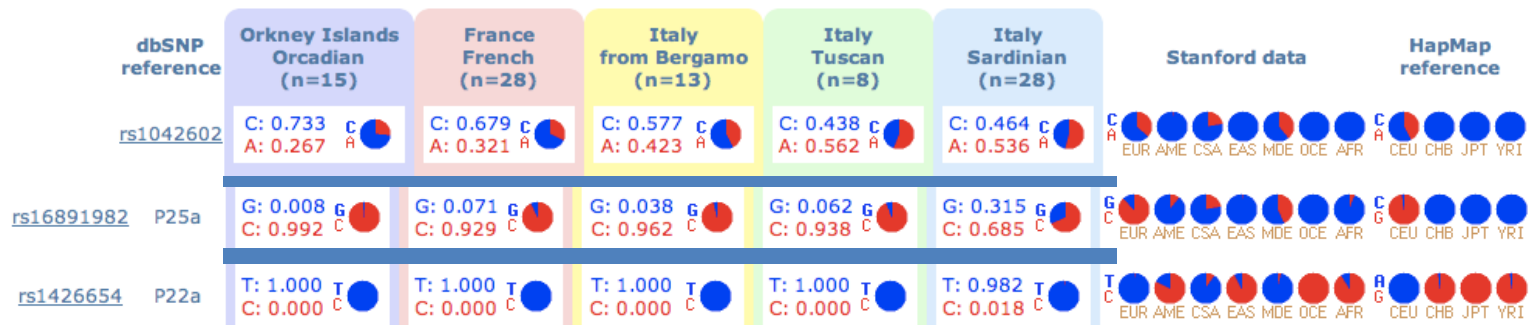


# Physical traits

- Hair color (red hair is a single gene)
- Skin pigmentation
- Eye color
- Facial features

## Skin pigmentation GWAs

- Phenotypic extremes in S ASNs, 1.6M SNPs
- 3 critical coding SNPs in genes SLC24A5, SLC45A2 & TYR
- Additive: both within SNPs (alleles) and between SNPs



# Calculation results

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

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

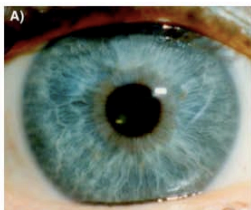
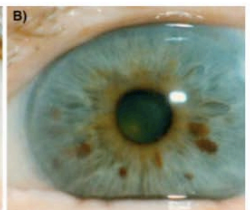

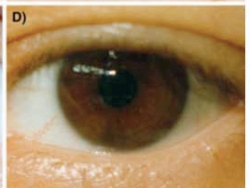
Gal-Dani	41.109466	2.50%
Moz-Som	66.935324	0.00%
Chi-Taiw	65.629561	0.00%

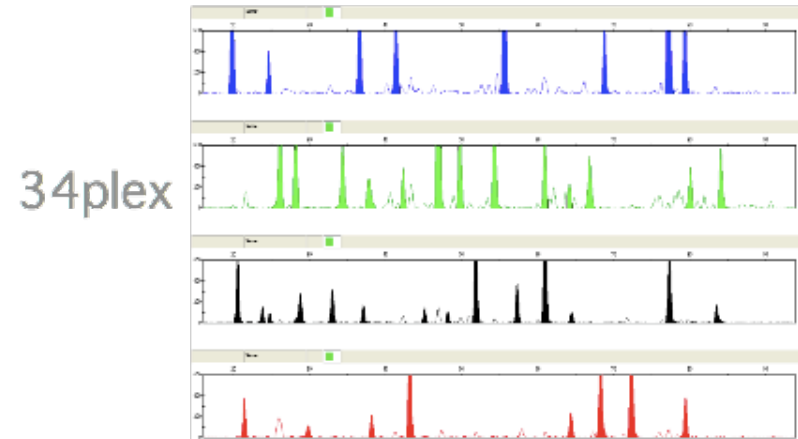
This person was European

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	164 billion times more likely to be European than African
Moz-Som	66.935324	0.00%	8.5184E-30	1.64448E+11	
Gal-Dani	41.109466	2.50%	1.40083E-18	EUR not ASN	44 billion times more likely to be European than Asian
Chi-Taiw	65.629561	0.00%	3.14372E-29	44559668806	
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

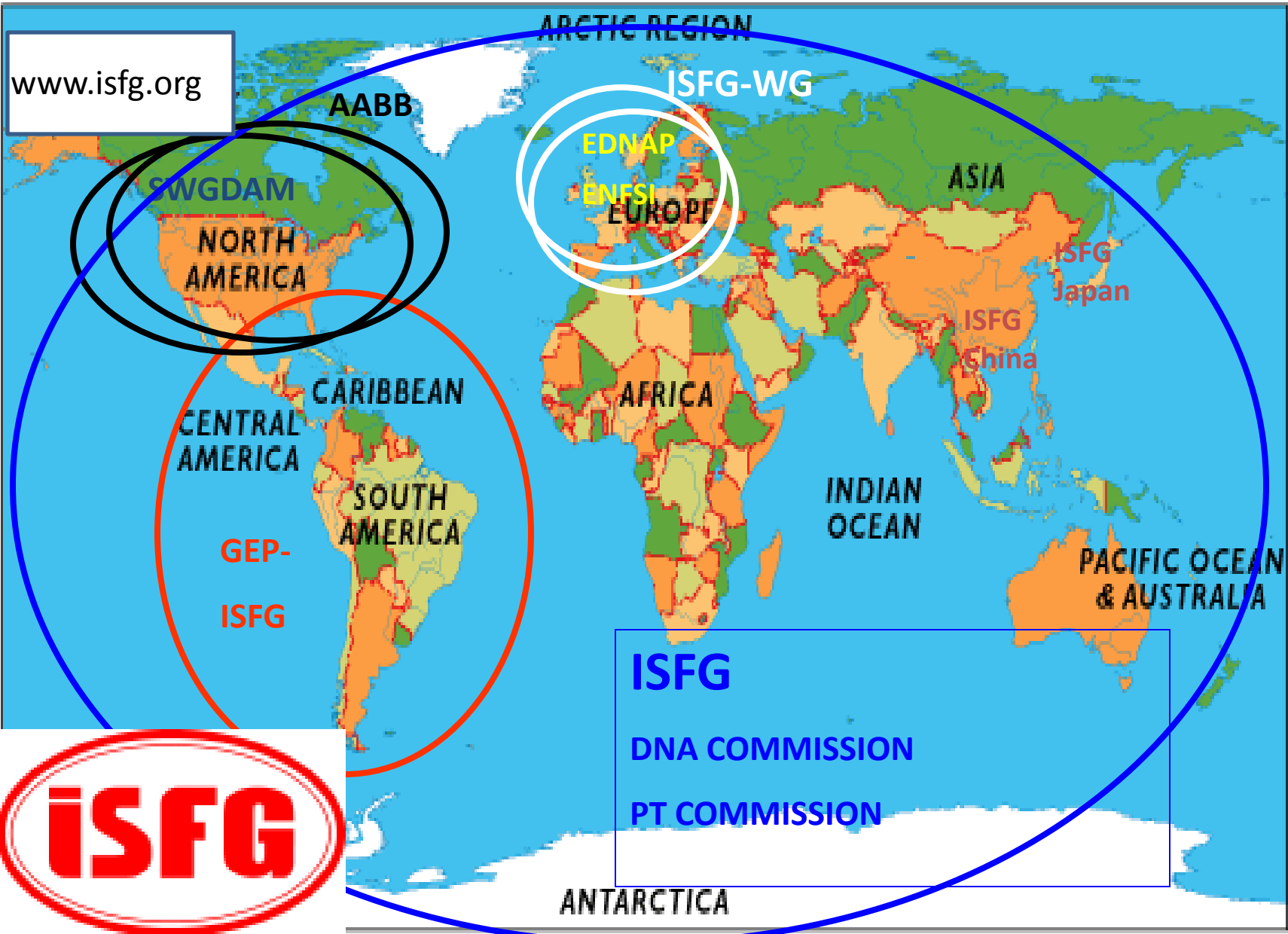
... and had blue eyes: rs12913832 = GG

GG		AG
Blue		Green (Unknown)
BEY1		AG / (AA)
Green Hazel		Brown
AG		BEY2



# DNA Databases

- 1995                    England
- 1996                    New Zealand, N. Ireland, Scotland
- 1997                    Netherlands, Austria
- 1998                    USA, Germany, Slovenia
- 1999                    Finland, Norway
- 2000                    Denmark, Switzerland, Sweden, Croatia,  
Bulgaria, Canada, Australia
- 2001                    France, Czech Republic
- 2002                    Belgium, Estonia, Lithuania, Slovakia
- From 2002            Rest of the UE countries, others



# 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

*The forensic craftsman*

Experience

Heuristics

Intuition

Absolute value of the  
opinion



Ian Evett, 2002



## *The forensic craftsman*



Experience

Heuristics

Intuition

Absolute value of the  
opinion

## *The forensic scientist*

Knowledge base

Data

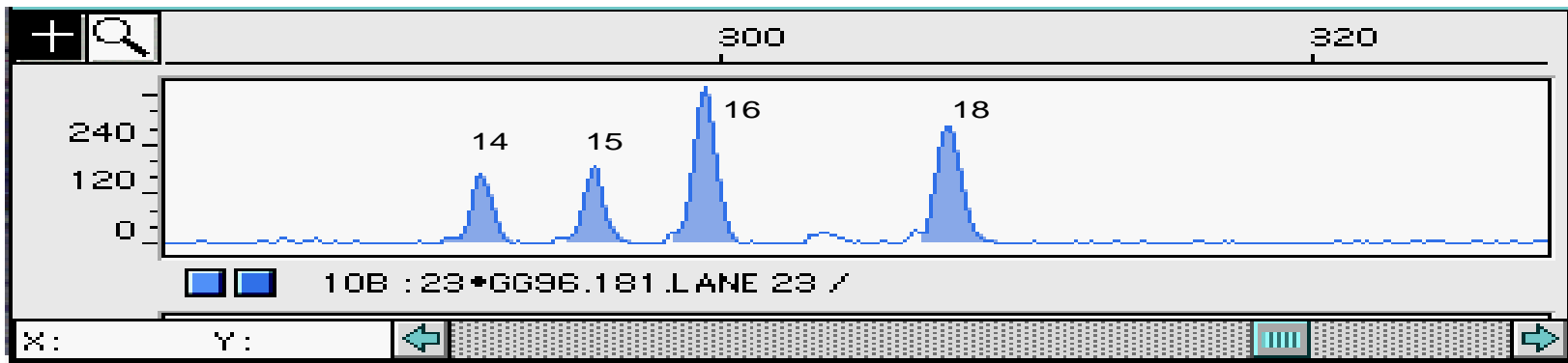
Understanding

Reasoning

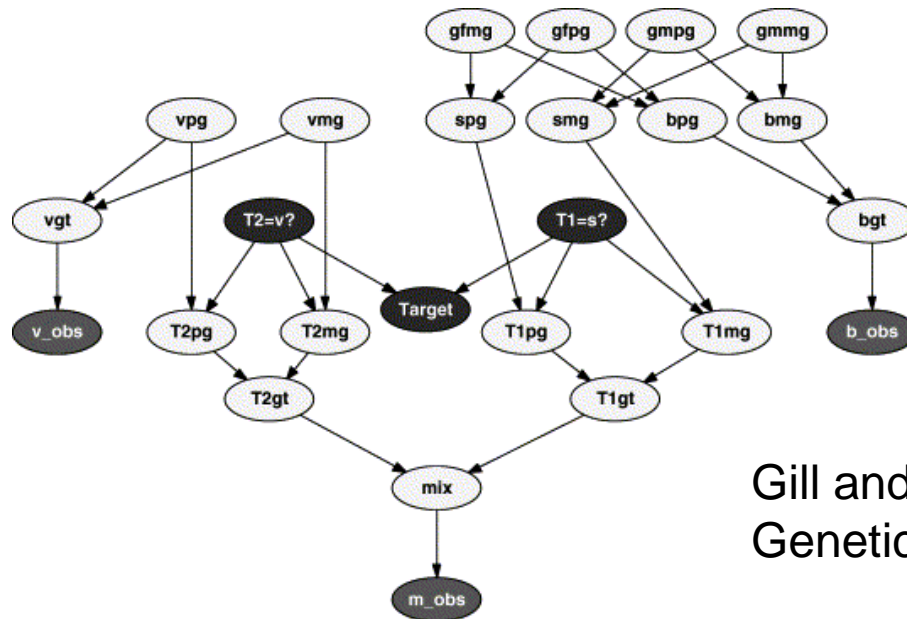
Probability of the  
**evidence given both**  
propositions



Ian Evett, 2002



Dye / Sample Peak	Minutes	Size	Peak Height	Peak Area
10B, 23	270.80	291.87	171	1375
10B, 24	274.30	295.71	192	1465
10B, 25	277.70	299.45	390	2867
10B, 26	281.10	303.49	31	226
10B, 27	283.70	306.60	33	142
10B, 28	284.60	307.68	291	2281



Probabilistic expert systems

Bayesian networks

Gill and Buckleton, FSI: Genetics, 2009

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



**FIN**

Urbano  
y  
Molinar