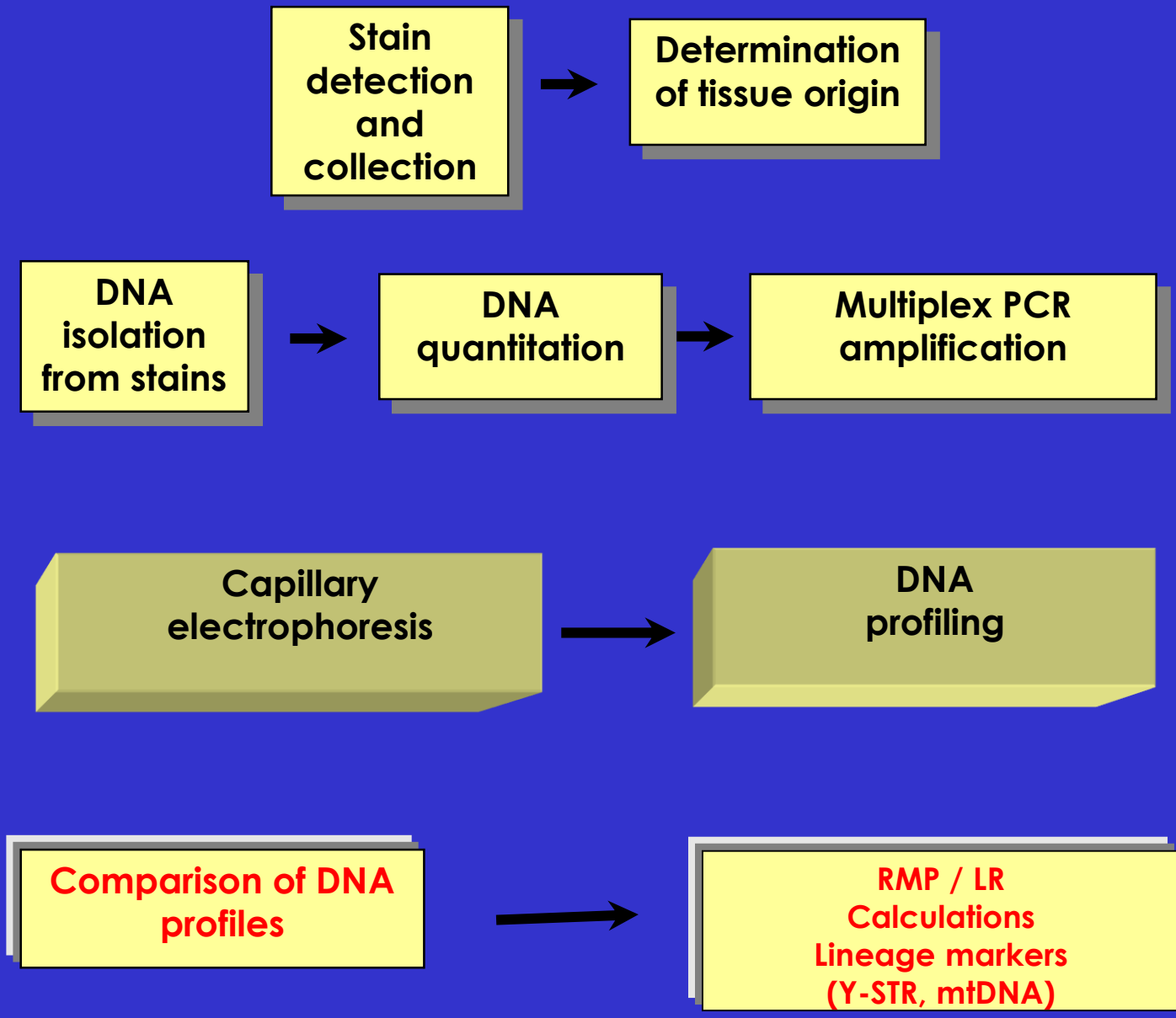
Forensic Genetics and Legal Medicine 2019-2020

29th April 2020

Basic match interpretation (Y-STR, mtDNA)


Stain detection and collection
Determination of tissue origin
DNA isolation from stains
DNA quantitation
Multiplex PCR amplification
Capillary electrophoresis
DNA profiling
Comparison of DNA profiles
RMP / LR
Calculations
Lineage markers
(Y-STR, mtDNA)

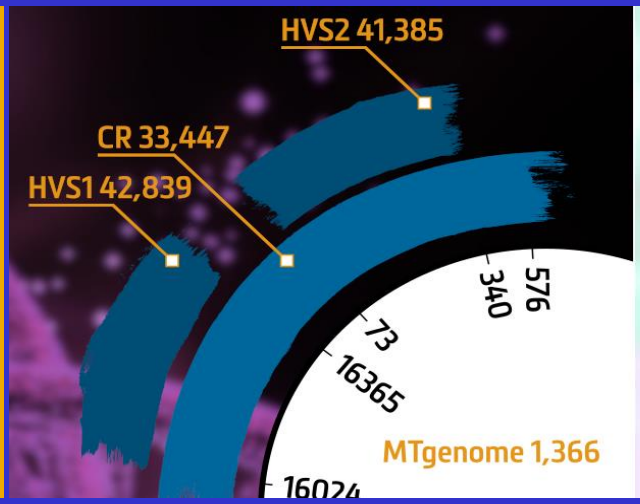
DNA polymorphisms on the Y chromosome and mtDNA are haploid lineage markers

✓Y chromosome is transmitted from the father to all male children in absence of recombination
✓mtDNA is transmitted from the mother to all children in absence of recombination
✓STRs/SNPs on Y/mtDNA are not independent, but completely linked to form "haplotypes"

- Allele frequencies of each STR/SNP cannot be multiplied as in RMP calculations of CODIS/ESS loci
- Haplotype frequencies are estimated through direct count in haplotype databases

Dataset	Y-STR loci	Number of haplotypes	Number of population samples	Number of national databases	Number of metapopulations
Minimal	DYS19 DYS389I DYS389II DYS390 DYS391 DYS392 DYS393 DYS385	269,383	1262	135	33
PowerPlex Y	DYS391 DYS389I DYS439 DYS389II DYS438 DYS437 DYS19 DYS392 DYS393 DYS390 DYS385	228,716	1036	126	32
Yfiler	DYS456 DYS389I DYS390 DYS389II DYS458 DYS19 DYS385 DYS393 DYS391 DYS439 DYS635 DYS392 YGATAH4 DYS437 DYS438 DYS448	209,111	951	119	32
PowerPlex Y23	DYS576 DYS389I DYS448 DYS389II DYS19 DYS391 DYS481 DYS549 DYS533 DYS438 DYS437 DYS570 DYS635 DYS390 DYS439 DYS392 DYS643 DYS393 DYS458 DYS385 DYS456 YGATAH4	52,312	310	67	28
Yfiler Plus	DYS576 DYS389I DYS635 DYS389II DYS627 DYS460 DYS458 DYS19 YGATAH4 DYS448 DYS391 DYS456 DYS390 DYS438 DYS392 DYS518 DYS570 DYS437 DYS385 DYS449 DYS393 DYS439 DYS481 DYF387S1 DYS533	45,892	222	50	30
Maximal	DYS19 DYS389I DYS389II DYS390 DYS391 DYS392 DYS393 DYS385 DYS438 DYS439 DYS437 DYS448 DYS456 DYS458 DYS635 YGATAH4 DYS481 DYS533 DYS549 DYS570 DYS576 DYS643 DYF387S1 DYS449 DYS460 DYS518 DYS627	6,471	42	14	19

yhrd.org

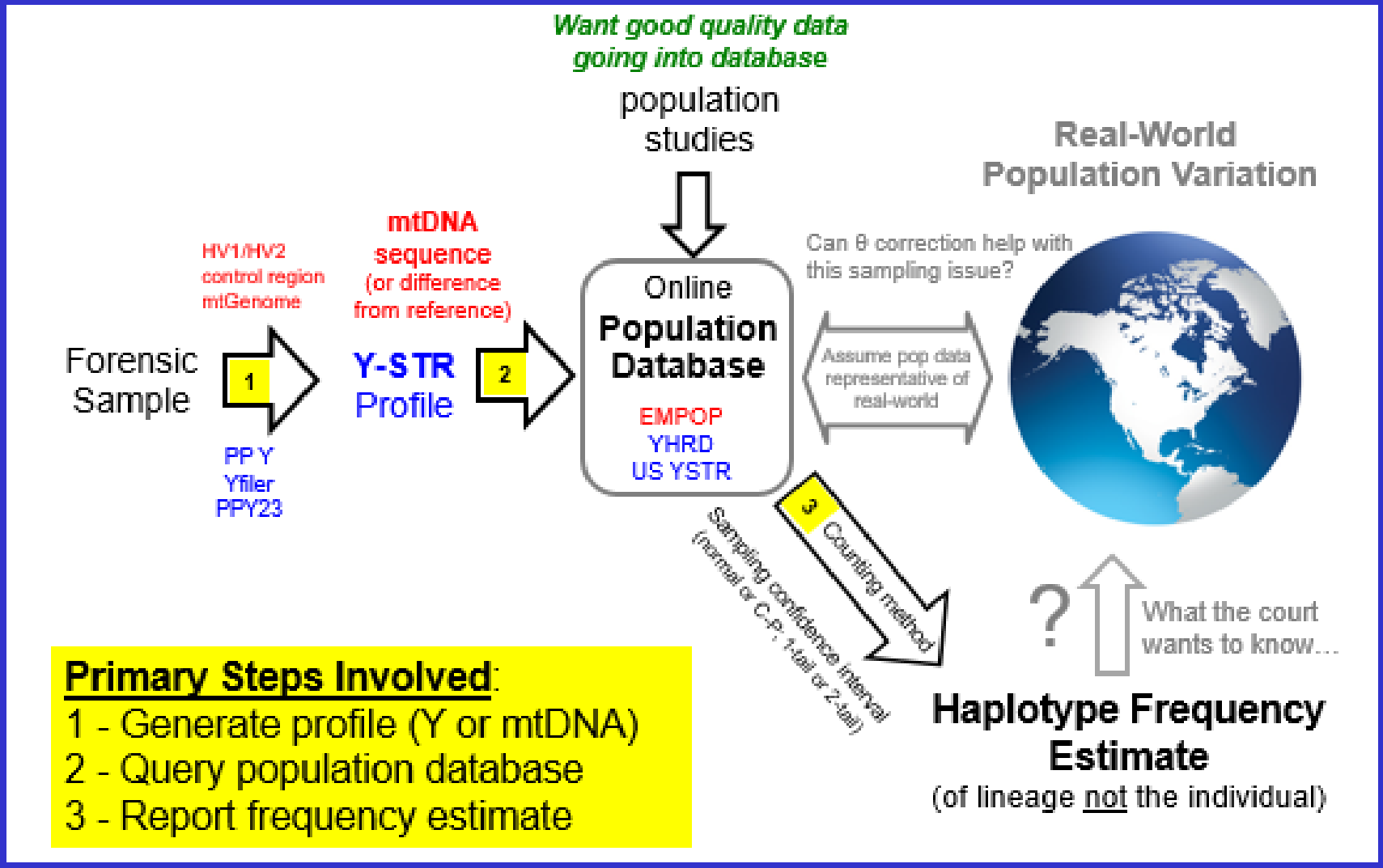


EMPOP Database

In total, the EMPOP Release 12 holds **42,839** quality-controlled mitotypes with at least HVS-I variation (16024-16365). Thereof,

- **41,385** cover HVS-I and HVS-II (16024-16365 73-340)
- **33,447** cover the Control Region (16024-576)
- **1,366** cover the entire mitogenome (ALL)

empop.online


Want good quality data going into database
population studies
Real-World Population Variation
HV1/HV2 control region mtGenome
mtDNA sequence (or difference from reference)
Forensic Sample
1
Y-STR Profile
2
Online Population Database
EMPOP
YHRD
US YSTR
Can θ correction help with this sampling issue?
Assume pop data representative of real-world
PP Y
Yfiler
PPY23
3
Counting method
Sampling confidence interval (normal or C-P; 1-tail or 2-tail)
?
What the court wants to know…
Haplotype Frequency Estimate
(of lineage not the individual)
Primary Steps Involved:
1 - Generate profile (Y or mtDNA)
2 - Query population database
3 - Report frequency estimate

yhrd.org/search/search

YHRD | Search the Database | Tools | Resources | Projects | News and Updates

Minimal | PowerPlex Y | Yfiler | PowerPlex Y23 | Yfiler Plus | Maximal

Report for Sample #1

Sample Name: Manual input

DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385
14	13	29	24	11	12	13	11, 16

Add feature to this Report

Worldwide

Observed

Found 13 matches in 307,169 Haplotypes. This is approx. 1 match in 23,628 Haplotypes (95% CI : 1 in 44,376 — 1 in 13,818).

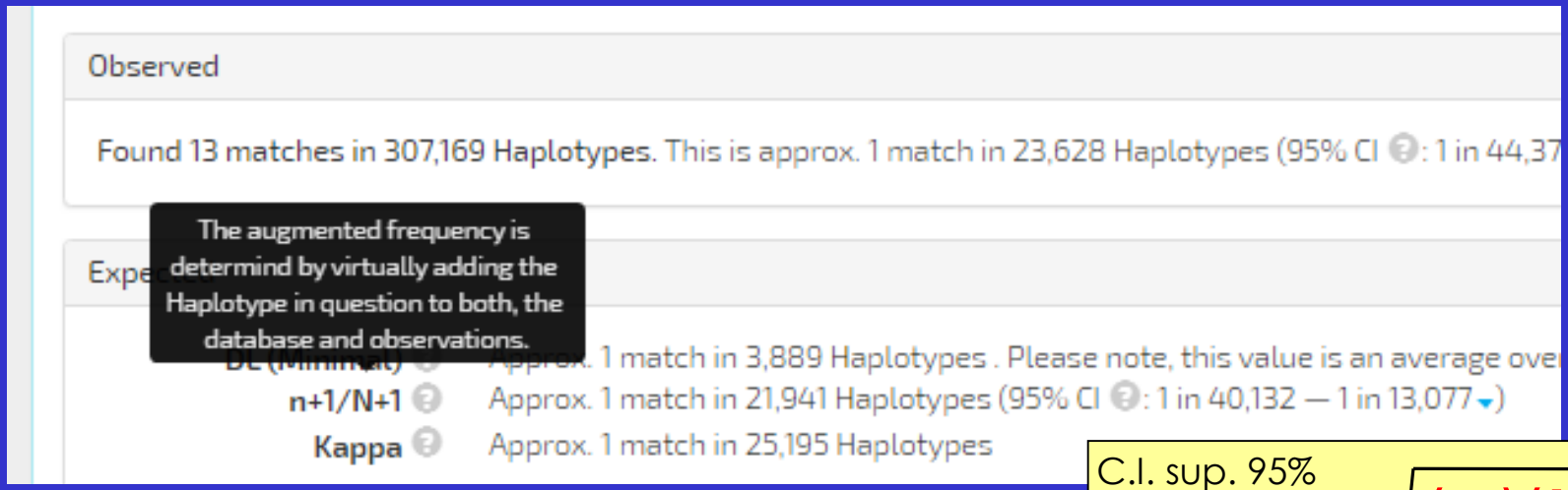
Expected

DL (Minimal) Approx. 1 match in 3,889 Haplotypes . Please note, this value is an average over the DL values of all nested feasible metapopulations.

n+1/N+1 Approx. 1 match in 21,941 Haplotypes (95% CI : 1 in 40,132 — 1 in 13,077)

Kappa Approx. 1 match in 25,195 Haplotypes

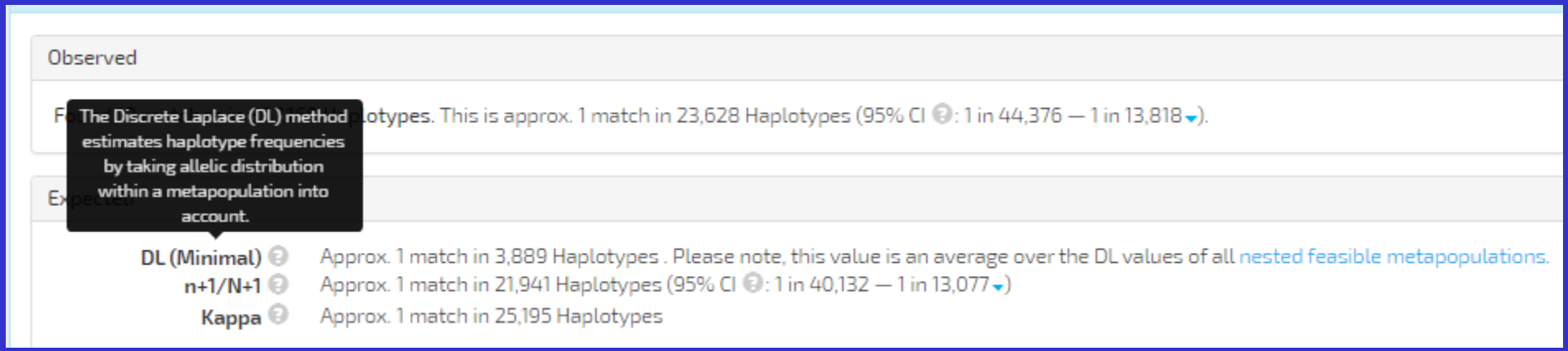
Augmented count method



C.I. sup. 95%

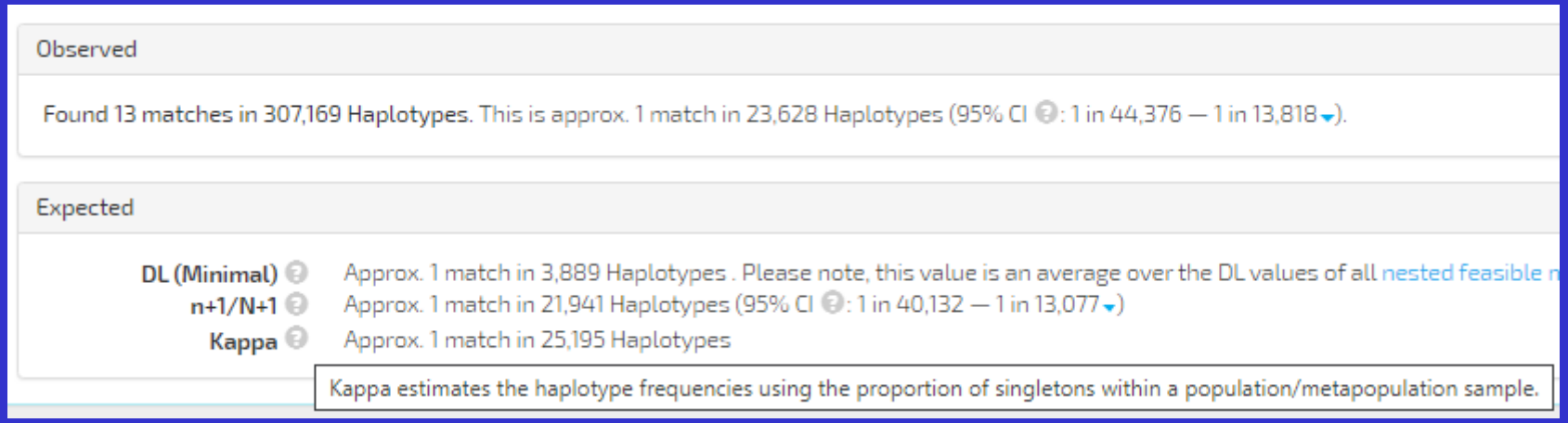
$$p + 1.96\sqrt{\frac{(p)(1-p)}{N}}$$

Discrete Laplace method

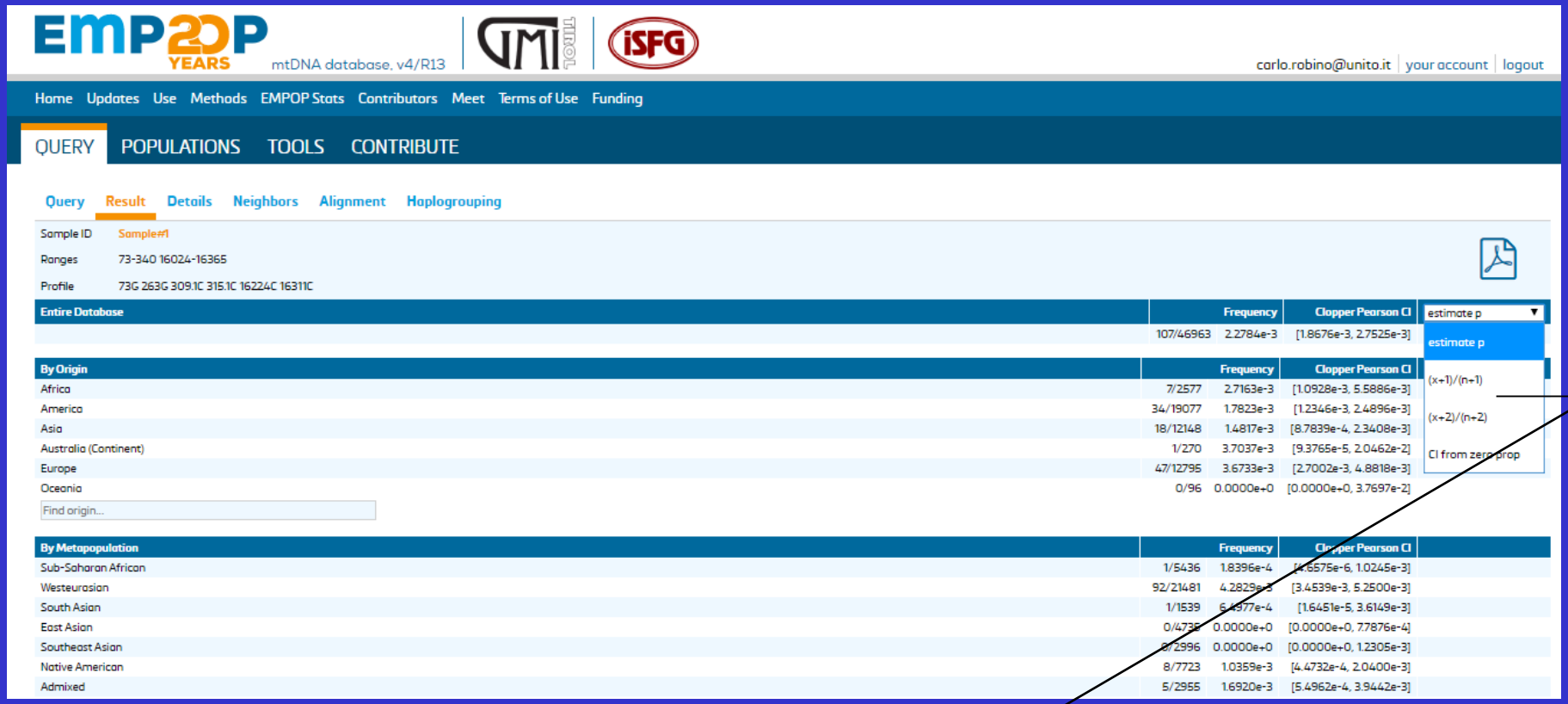


- DL method takes into account the distribution of observed haplotypes in metapopulations (populations grouped according to geographical, linguistic, demographic, previously known genetic data) and estimates haplotype frequencies according to a model of evolution by single-step mutation process.
- Haplotypes with close molecular "neighbors" will have higher frequency estimates than outlier haplotypes.

Kappa method



- Haplotype frequency is estimated as (n-k)/N, where k is the proportion of singletons (haplotypes that occur only once) in the N-sized dataset: k close to 0 would indicate that the database almost saturates all haplotypes present in the population and therefore a newly observed haplotype (n=1) will have frequency 1/N

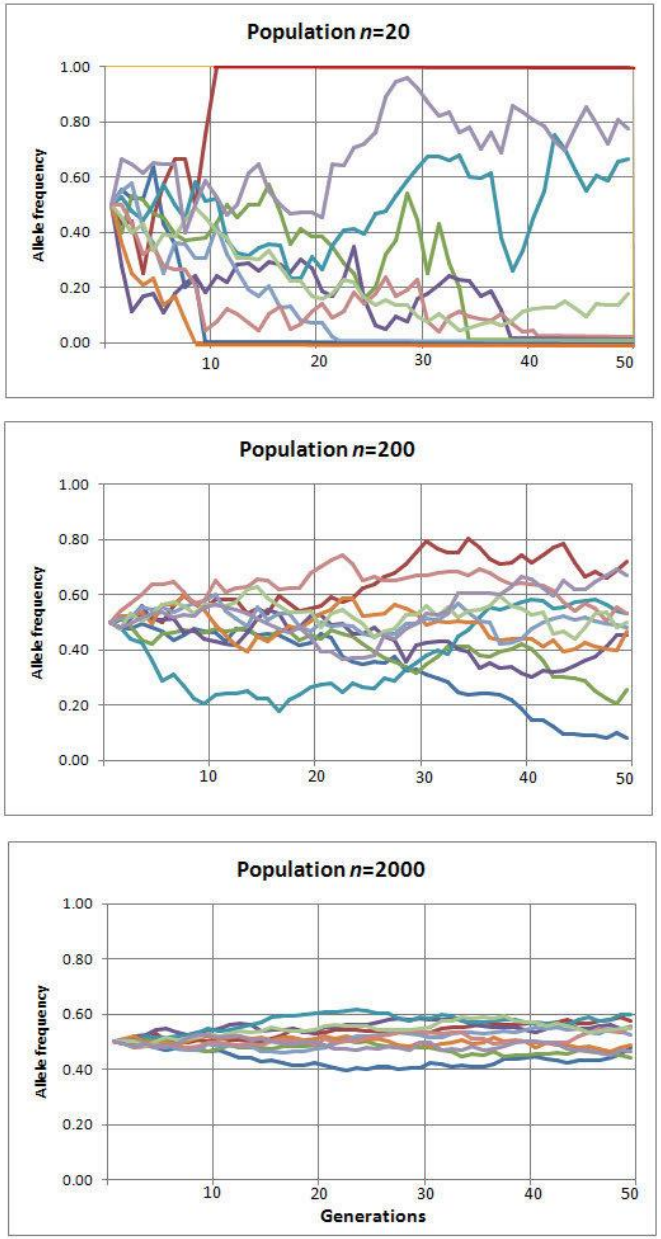


Two options for augmented counting method

- 1: observed mtDNA haplotype added to haplotype count and database size
- 2: matching mtDNA haplotypes observed in both suspect and stain added to haplotype count and database size

Why worldwide populations databases (covering all human "metapopulations") are needed?

- Assuming a 1:1 sex ratio, a human population can be represented in microcosm by one man and one woman. This couple carry four copies of each autosome, three X chromosomes, two mtDNAs (but only the female one is passed to the next generation) and one Y chromosome. The effective population size of the Y chromosome and mtDNA is therefore expected to be one-quarter of that of any autosome and one-third of that of the X chromosome.
- This makes haploid markers much more susceptible to genetic drift, which involves random changes in the frequency of haplotypes owing to sampling from one generation to the next and accelerates the differentiation between groups of Y chromosomes and mtDNAs in different populations.
- demographic events such as population bottlenecks and founder effect can create extremely uneven or peculiar distributions of Y-STR haplotypes



DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385
13	14	30	24	10	13	13	13, 17

+ Add feature to this Report

Worldwide

Observed

Found 4 matches in 307,169 Haplotypes. This is approx. 1 match in 76,792 Haplotypes (95% CI : 1 in 281,840 — 1 in 29,993).

Expected

DL (Minimal)	Approx. 1 match in 12,084 Haplotypes . Please note, this value is an average over the DL values of all ne
n+1/N+1	Approx. 1 match in 61,434 Haplotypes (95% CI : 1 in 189,203 — 1 in 26,325)
Kappa	Approx. 1 match in 70,548 Haplotypes

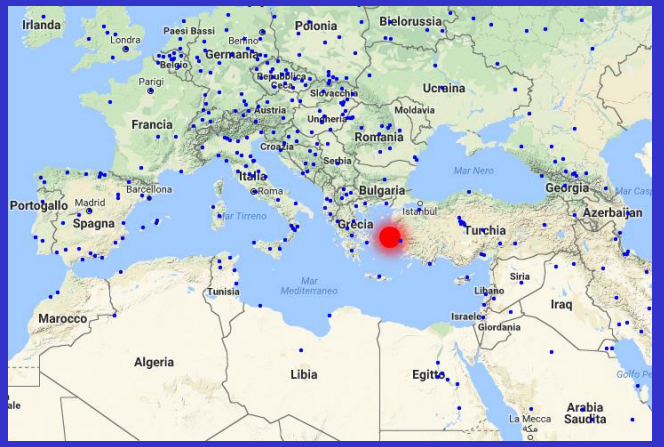
Eurasian - European - South-Eastern European (click to change)

Observed

Found 4 matches in 8,207 Haplotypes. This is approx. 1 match in 2,052 Haplotypes (95% CI : 1 in 7,529 — 1 in 802).

Expected

DL (Minimal)	Approx. 1 match in 9,265 Haplotypes
n+1/N+1	Approx. 1 match in 1,642 Haplotypes (95% CI : 1 in 5,055 — 1 in 704)
Kappa	Approx. 1 match in 2,230 Haplotypes



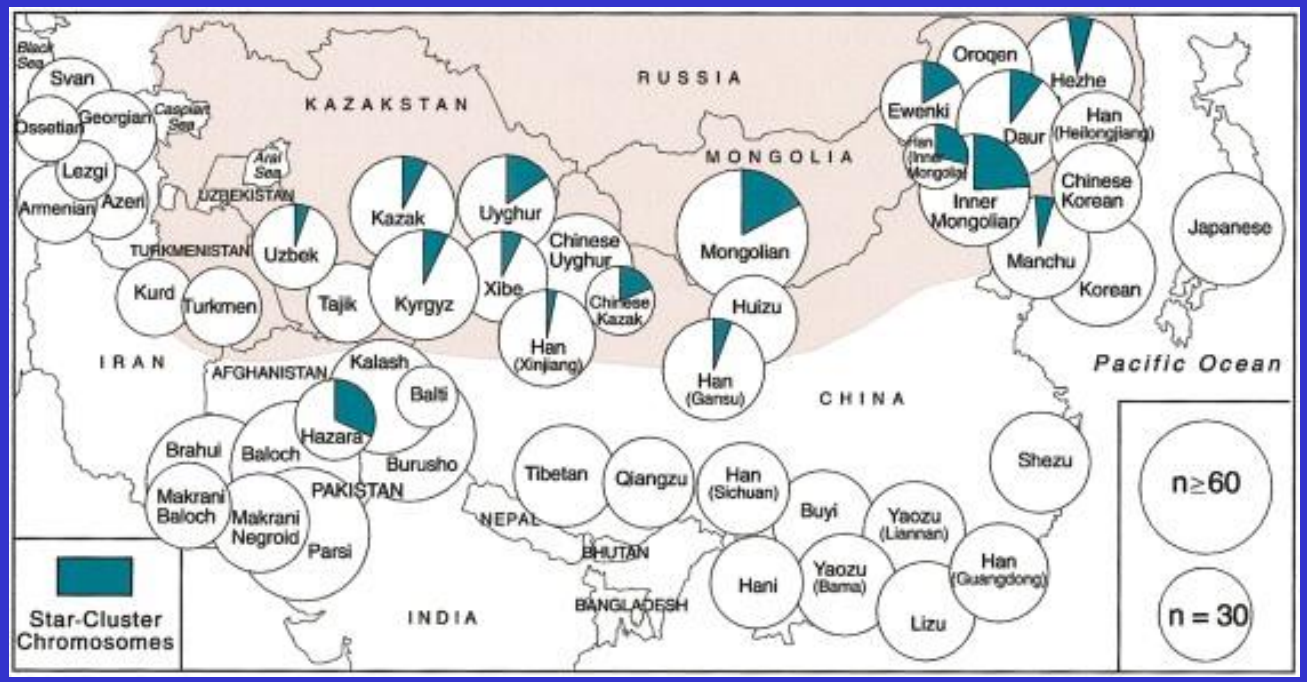
✓ Chios, Greece (observed) 4/16 (Robino et al Forensic Sci Int 2004)

We owe it all to superstud Genghis

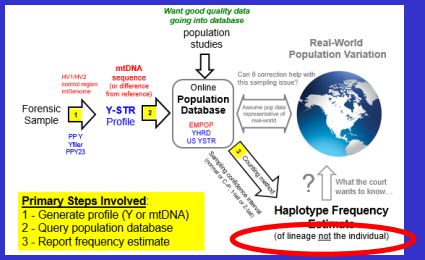
Warlord Khan has 16m male relatives alive now, says study

One in every 200 men alive today is a relative of Genghis Khan. An international team of geneticists has made the astonishing discovery that more than 16 million men in central Asia have the same male Y chromosome as the great Mongol leader.

It is a striking finding: a huge chunk of modern humanity can trace its origins to Khan's vigorous policy of claiming the most beautiful women captured during his merciless conquest.

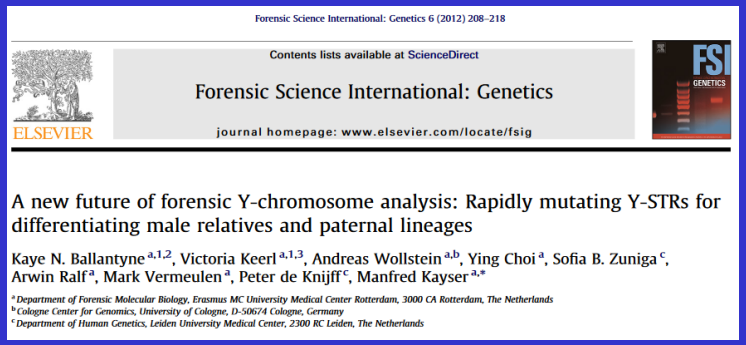


Zerjal T et al. Am J Hum Genet 2003



Variation within lineages is only caused by mutation

- Average mutation rates for standard STRs are ~ 1 every 1000 meiosis
- For a 17 Y-STR panel, therefore, the chance to observe a mutation in a father/son pair is <5% ~ 17/1000
- Mutation rates on single nucleotides in mtDNA control region is << 1/1000 so differentiation between close relatives is unlikely
- Recently a set of "rapidly mutating" Y-STR markers was described with mutation rate > 1/100, that increases the chance of discrimination between up to ~20% in father/son pairs (1 meiosis) and ~40% in sib pairs (2 meiosis)


Forensic Science International: Genetics 6 (2012) 208–218
Contents lists available at ScienceDirect
Forensic Science International: Genetics
journal homepage: www.elsevier.com/locate/fsig
ELSEVIER

A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages

Kaye N. Ballantyne [a,1,2], Victoria Keerl [a,1,3], Andreas Wollstein [a,b], Ying Choi [a], Sofia B. Zuniga [c], Arwin Ralf [a], Mark Vermeulen [a], Peter de Knijff [c], Manfred Kayser [a,*]

[a] Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, 3000 CA Rotterdam, The Netherlands
[b] Cologne Center for Genomics, University of Cologne, D-50674 Cologne, Germany
[c] Department of Human Genetics, Leiden University Medical Center, 2300 RC Leiden, The Netherlands