Haplotypes and databases

DANIEL KLING & ANDREAS TILLMAR

Daniel.kling@rmv.se

Andreas.tillmar@rmv.se

Teachers

Daniel Kling, PhD

- Forensic Expert
- National Board of Forensic Medicine, Sweden
- Worked in the field for almost 15 years
- Developer of Familias, FamLink and FamLinkX
- Applied biostatistics, relationship inference, genetic genealogy

Andreas Tillmar, PhD

- Forensic geneticist & Associate professor
- National Board of Forensic Medicine, Sweden and Linköping University, Sweden
- Worked in the field for almost 20 years
- Technical leadership mixed with R&D
- Applied biostatistics, relationship inference, population genetics, genetic genealogy

Disclaimer!

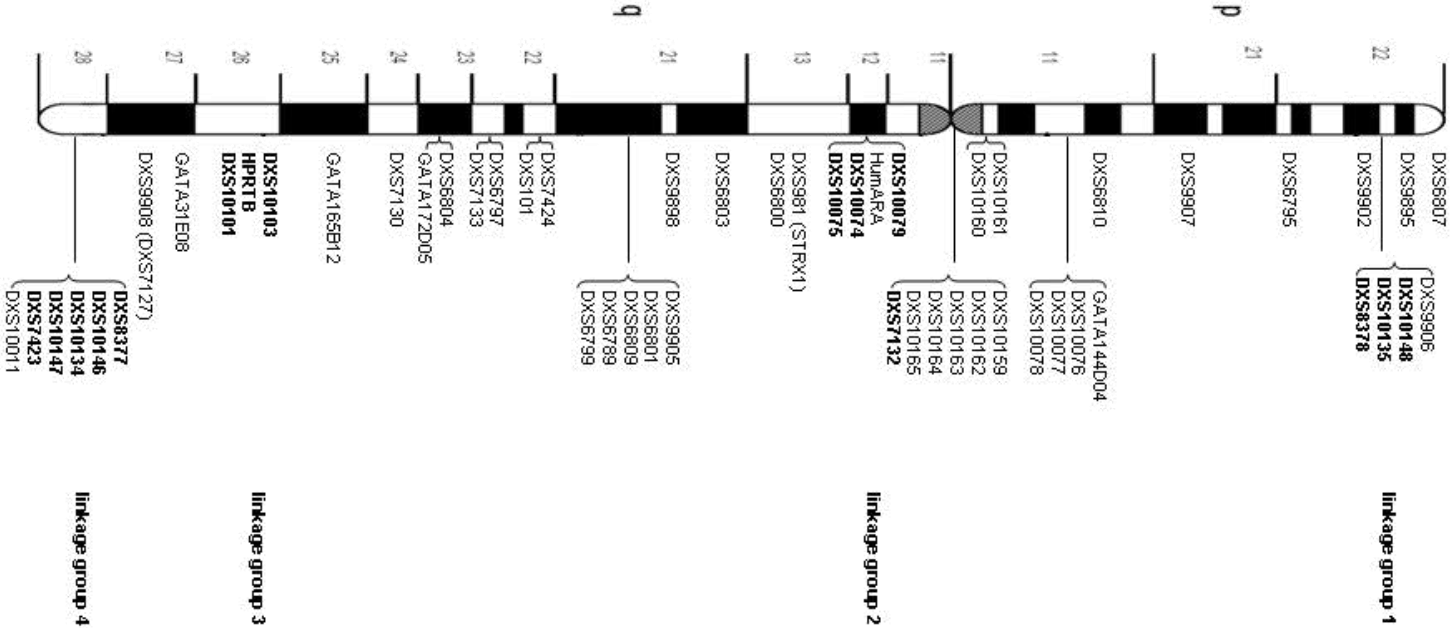
Points of view are those of the presenters and do not necessarily represent the official position or policies of the National Board of Forensic Medicine or ISFG. Certain commercial software, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Topics

- Markers
- Haplotypes
- Exploring population genetics of X-STRs

Markers

Markers


p
22
21
11
11
q
21
22
23
24
25
26
27
28
DXS6807
DXS9695
DXS9902
DXS6795
DXS9907
DXS6810
DXS10161
DXS10160
DXS10079
HumARA
DXS10074
DXS10075
DXS981 (STRX1)
DXS6800
DXS6803
DXS9898
DXS7424
DXS101
DXS6797
DXS7133
DXS6804
GATA172D05
DXS7130
GATA165B12
DXS10103
HPRTB
DXS10101
GATA31E08
DXS9908 (DXS7127)
DXS8377
DXS10146
DXS10134
DXS10147
DXS7423
DXS10011
DXS9906
DXS10148
DXS10135
DXS8378
GATA144D04
DXS10076
DXS10077
DXS10078
DXS10159
DXS10162
DXS10163
DXS10164
DXS10165
DXS7132
DXS9905
DXS6801
DXS6809
DXS6789
DXS6799
linkage group 1
linkage group 2
linkage group 3
linkage group 4

Markers

Argus X12

Cytogenetic localisation	Marker	Linkage group	Physical localisation [Mb] NCBI 36	Rutgers Map v.2 [cM Kosambi]
p 22.31	DXS10148	X1	9.198	19.84
p 22.31	DXS10135	X1	9.199	20.03
p 22.31	DXS8378	X1	9.33	20.21
centromere	DXS7132	X2	64.572	90.75
q 12	DXS10079	X2	66.632	90.82
q 12	DXS10074	X2	66.894	90.83
q 26.2	DXS10103	X3	133.246	149.37
q 26.2	HPRTB	X3	133.443	149.66
q 26.3	DXS10101	X3	133.482	149.75
q 28	DXS10146	X4	149.335	183.72
q 28	DXS10134	X4	149.401	183.96
q 28	DXS7423	X4	149.46	184.19

https://pubmed.ncbi.nlm.nih.gov/24632058/

Decaplex

Cytogenetic localisation	Marker	Physical localisation [Mb] NCBI 36	Rutgers Map v.2 [cM Kosambi]
p 22.31	DXS8378	9.33	20.21
p 22.2	DXS9902	15.234	32.32
centromere	DXS7132	64.572	90.75
q 21.31	DXS9898	87.682	101.29
q 21.33	DXS6809	94.825	108.12
q 21.33	DXS6789	95.336	108.47
q 22.3	DXS7133	108.928	118.18
q 23	GATA172D05	113.061	124.36
q 27.1	GATA31E08	140.062	160.54
q 28	DXS7423	149.46	184.19

<0.5cm

https://pubmed.ncbi.nlm.nih.gov/19082839/

Markers

- Different kits

STRs

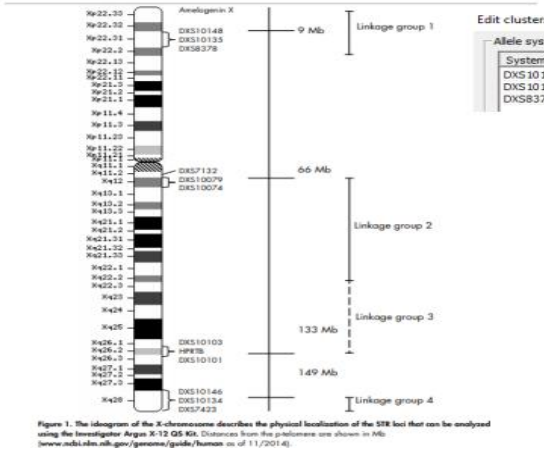
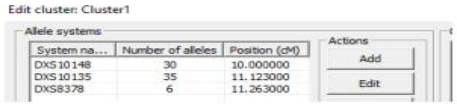


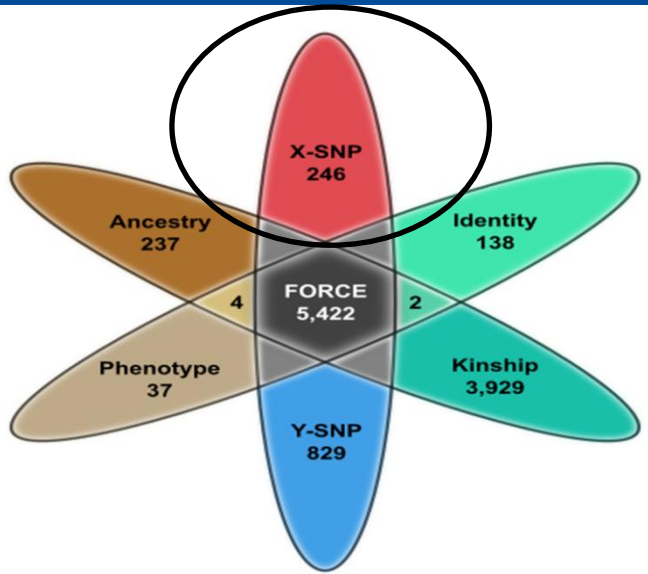
Figure 1. The ideogram of the X-chromosome describes the physical localization of the STR loci that can be analyzed using the Investigator Argus X-12 QS Kit. Distances from the p-telomere are shown in Mb (www.ncbi.nlm.nih.gov/genome/guide/human as of 11/2014).



1. Many alleles
2. Mutations
3. LD

Markers

	Marker	Genetic position (cM)
2	rs4892897	7.582
3	rs1637781	7.805
4	rs5983084	8.5874
5	rs6642174	8.8134
6	rs6641753	9.1834
7	rs6641574	9.4251
8	rs2058865	9.7825
9	rs5962087	11.3138
10	rs5915796	12.2585
11	rs5916138	12.71
12	rs5915672	13.1805
13	rs6529997	13.9103
14	rs1637788	15.1505
15	rs4240138	15.6932
16	rs2108400	16.1319
17	rs5933710	17.0167
18	rs845444	17.6324
19	rs768568	17.9229
20	rs929217	18.1746



Open Access Article

The FORCE Panel: An All-in-One SNP Marker Set for Confirming Investigative Genetic Genealogy Leads and for General Forensic Applications

by Andreas Tillmar [1,2,*], Kimberly Sturk-Andreaggi [3,4,5], Jennifer Daniels-Higginbotham [3,4], Jacqueline Tyler Thomas [3,4] and Charla Marshall [3,4,6,*]

[1] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, SE-587 58 Linköping, Sweden
[2] Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, SE-582 25 Linköping, Sweden
[3] Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL), Dover Air Force Base, Dover, DE 19902, USA
[4] SNA International, LLC, Contractor Supporting the AFMES-AFDIL, Alexandria, VA 22314, USA
[5] Department of Immunology, Genetics and Pathology, Uppsala University, SE-751 08 Uppsala, Sweden

X-SNPs are chosen to be have no significant LD

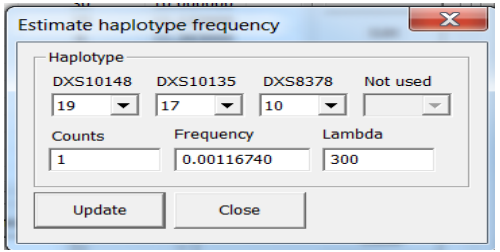
Haplotypes

Haplotypes

- What is a haplotype? Sequence of alleles on the same chromosome
- What is special with the X-chromosome? Males only have one
- Phase uncertainty

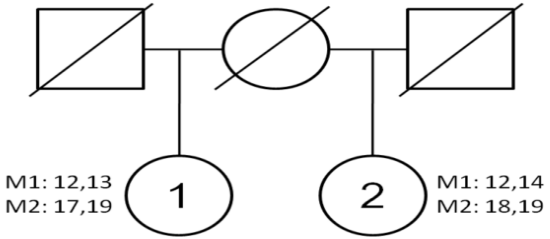
Marker	Child	Sibling
DXS10148	19,20	19 19
DXS10135	16,17	16 17
DXS8378	10,10	10 10

Haplotype



Haplotypes

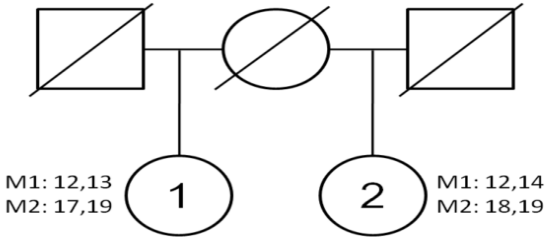
Ambigious haplotypes



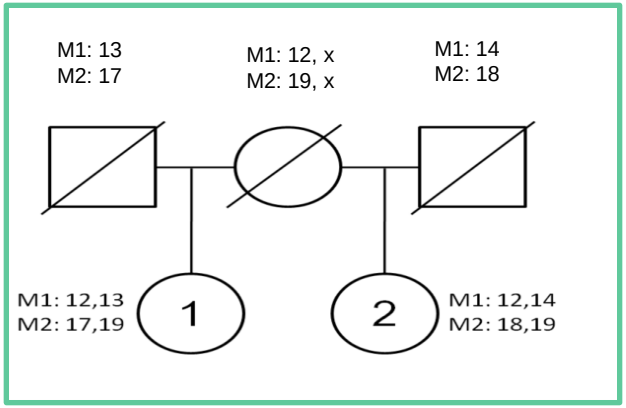
12_19 can be from mother or from different fathers

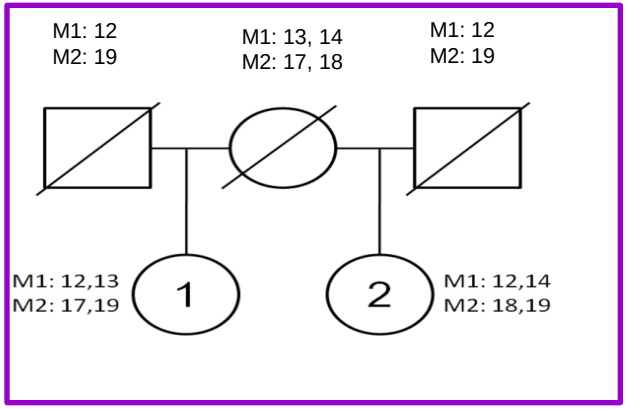
Haplotypes

Ambigious haplotypes



12_19 can be from mother or from different fathers

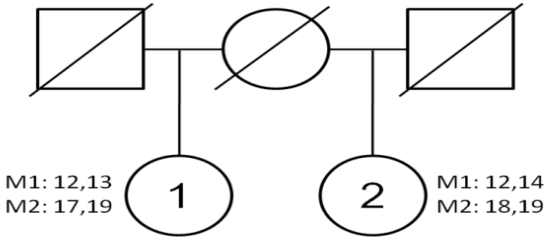




Other possibilities?

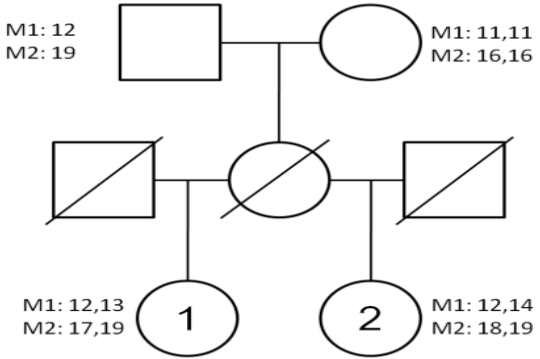
Haplotypes

Ambigious haplotypes



12_19 can be from mother or from different fathers

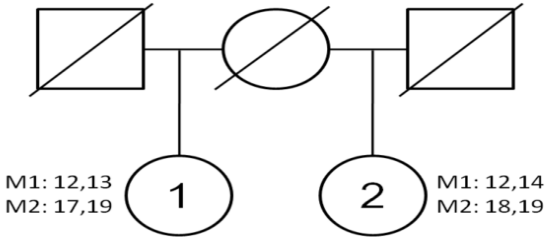
Haplotypes ”certain”



12_19 is from the grandfather

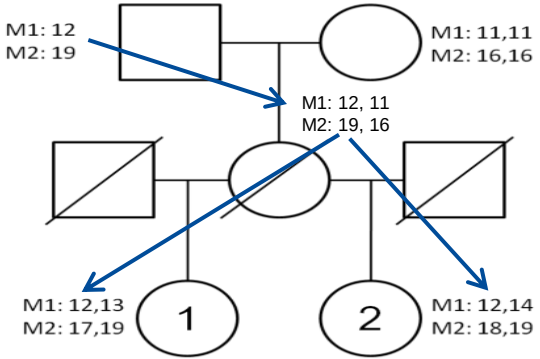
Haplotypes

Ambigious haplotypes



12_19 can be from mother or from different fathers

Haplotypes ”certain”



12_19 is from the grandfather

Haplotypes

Ambigious haplotypes

	DXS10103	DXS8378	DXS7132	DXS10134	DXS10074	DXS10101	DXS10135	DXS7423	DXS10146	DXS10079	HPRTB	DXS10148
Mother	19	10	13	37	16	28.2	18	14	28	20	12	24.1
	21	11	14	38	17	30.2	23	16	30	21	12	24.1
Son 1	19	10	13	37	16	28.2	18	14	28	20	12	24.1
Son 2	21	11	14	37	16	28.2	18	14	28	20	12	24.1

One crossover

Haplotypes

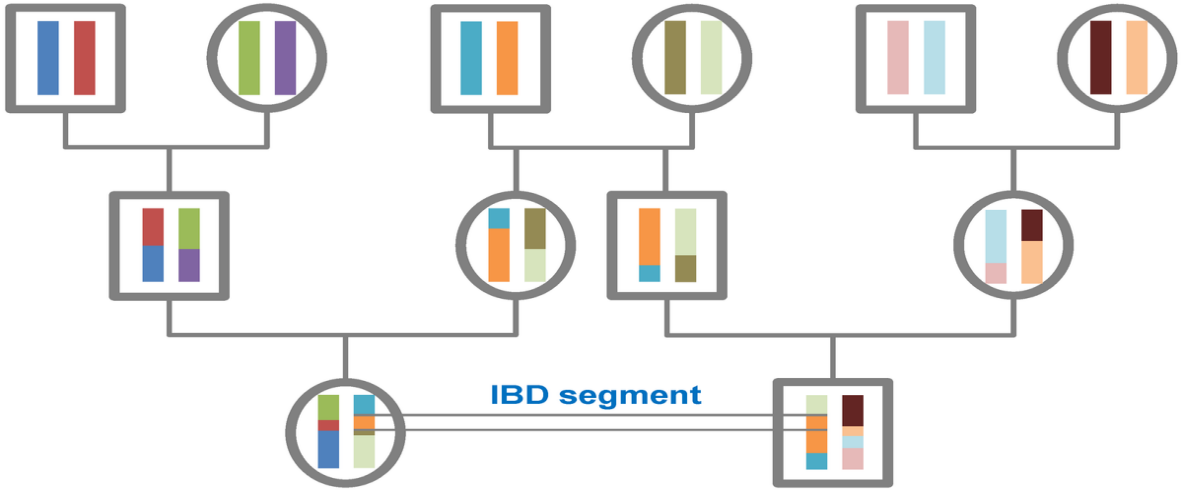
Ambigious haplotypes

	DXS10103	DXS8378	DXS7132	DXS10134	DXS10074	DXS10101	DXS10135	DXS7423	DXS10146	DXS10079	HPRTB	DXS10148
Grandfather	19	10	13	37	17	28.2	23	16	30	21	12	24.1
Mother	19	10	13	37	16	28.2	23	16	30	21	12	24.1
	21	11	14	38	17	30.2	18	14	28	20	12	24.1
Son 1	19	10	13	37	16	28.2	18	14	28	20	12	24.1
Son 2	21	11	14	37	16	28.2	18	14	28	20	12	24.1

Three crossovers!

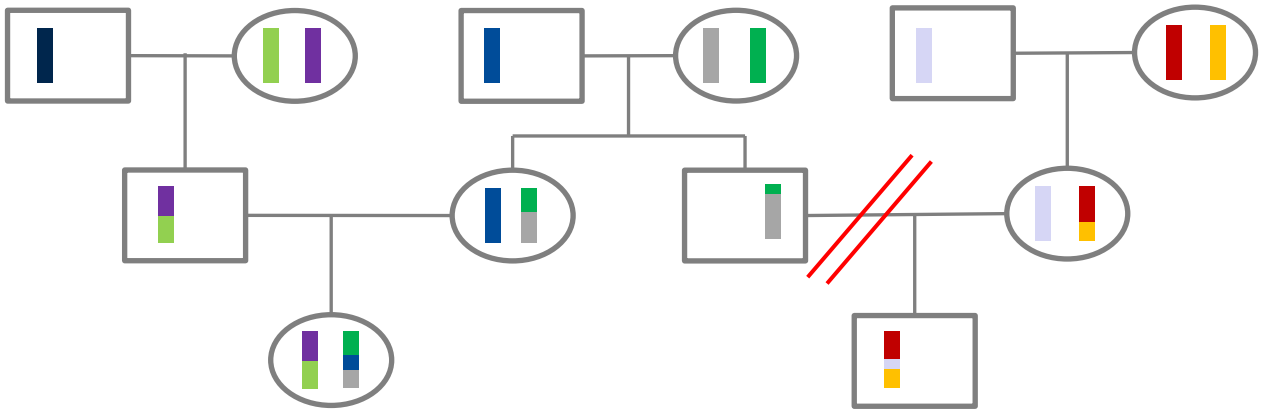
Haplotypes

Tracing X-haplotypes through pedigrees



Haplotypes

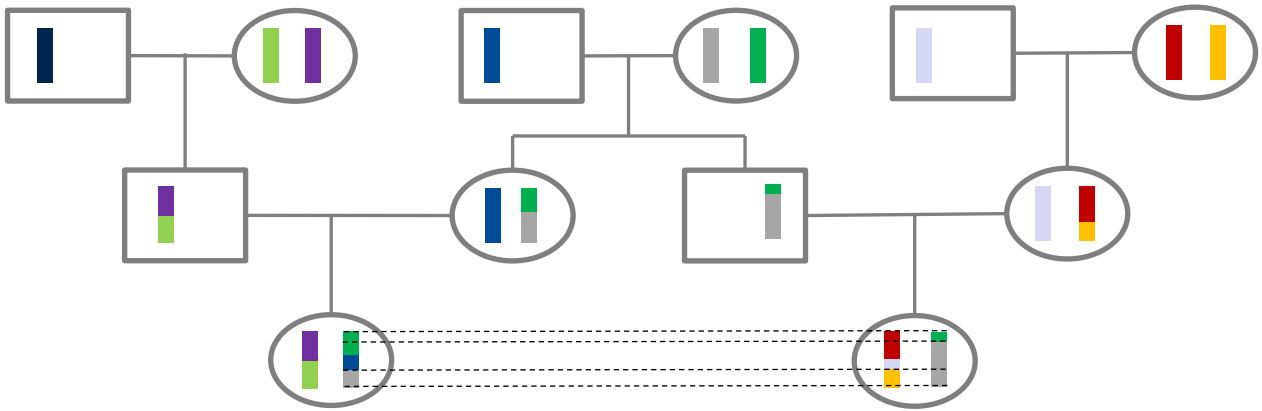
Tracing haplotypes through pedigrees



No IBD possible!

Haplotypes

Tracing X-haplotypes through pedigrees



IBD

Haplotypes

- For X-chromosomal databasing – males are prefered
- Exact inference of haplotypes

Marker	Male1	Male2	Male3	Male4	…	Male*n*
DXS10148	19	19	20	20	…	23
DXS10135	16	16	16	17	…	17
DXS8378	10	10	11	11	…	11

Haplotypes

ELSEVIER

Forensic Science International: Genetics
Volume 15, March 2015, Pages 127-130

Expanding X-chromosomal forensic haplotype frequencies database: Italian population data of four linkage groups

Carla Bini [a], Laura Natalia Riccardi [a], Stefania Ceccardi [a], Francesco Carano [a], Stefania Sarno [b], Donata Luiselli [b], Susi Pelotti [a]

Show more

https://doi.org/10.1016/j.fsigen.2014.11.008

Highlights

- X-chromosome STRs analysis has a potential usefulness in paternity testing.
- Haplotype frequencies in 200 unrelated Italian males have been estimated.
- Comparison with European populations showed some minor differences.
- Contribute to the statistical evaluation of the X-STRs polymorphism.

Bini et al., Forensic Science International: Genetics, Vol. 15, p127–130

LG I				LG II				LG III			
DXS8378	DXS10135	DXS10148	Number	DXS7132	DXS10074	DXS10079	Number	DXS10103	DXS10101	HPRTB	Number
8	29	27,1	1	10	17	19	1	15	33	12	1
9	22	24,1	1	11	8	20	1	16	26	14	1
10	19	23,1	1	11	17	20	1	16	28,2	12	3
10	19	24,1	2	11	17	21	1	16	28,2	14	1
10	19	25,1	2	11	19	20	1	16	30	12	1
10	19	26,1	2	12	7	19	2	16	31	14	1
10	19,1	18	1	12	8	18	2	16	31,2	14	1
10	20	24,1	1	12	8	20	1	16	32	12	1
10	20,1	25,1	1	12	8	21	1	16	34	14	1
10	20,1	26,1	1	12	9	21	1	17	26,2	12	1
10	20,1	27,1	1	12	16	18	1	17	28	12	1
10	20,1	28,1	1	12	16	20	1	17	29	13	1
10	21	20	1	12	17	18	2	17	30	13	2
10	21	23	1	12	18	16	1	17	30	14	2
10	21	26,1	2	12	18	18	1	17	30	15	2
10	21,1	18	1	12	18	19	3	17	30,2	16	1
10	21,1	26,1	2	12	18	22	1	17	31	13	4
10	22	18	1	12	19	19	1	17	31	14	3
10	22	21	1	12	20	20	1	17	31	15	3
10	22	24,1	2	13	7	20	2	17	31,2	12	1

Haplotypes

famlink.se/fx_databases.html

FamLinkX – A software for X-chromosomal markers

Home Download **Databases** Getting started Manual **Contact** Help

Databases

The files below are prepared based on data given in the references. Please keep in mind that we have not performed quality control of the submitted data.

Send us a request on new databases or merging of several of the databases below.

All data for the Decaplex kit is obtained via Supplementary data in Gusmão et al.

Frequency databases (haplotypes and allele frequencies)

Databases may be found here or downloaded in the tables below. Size refers to the number of male haplotypes in the publication.

Please note that several populations can appear twice (different sources) and in continental databases as well.

Argus X12

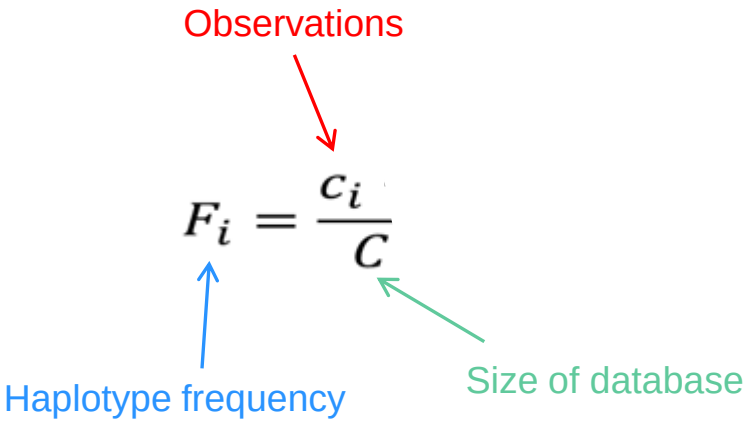
Population	Size	Database	Publication	Comment	Kit
Sweden	652	SWE	Tillmar A.		Argus X12
Norway	631	NOR	Bergseth et al.		Argus X12
Czech Repubic	307	CZE	Zidkova et al.		Argus X12
Germany	1037	GER	Edelmann et al.		Argus X12
Greece	121	GRE	Tomas et al.		Argus X12
Italian	200	ITA	Bini et al.		Argus X12
Sardinia	316	SAR	Robino et al.		Argus X12
Serbia	220	SER	Veselinovic et al.		Argus X12

Haplotypes

- Estimating haplotype frequencies
  - A long history of different methods
    - Expectation maximization (EM)
    - MCMC - PHASE
    - IMPUTE2, SHAPEIT2
- For X-chromosomal STR markers we consider a Dirichlet model

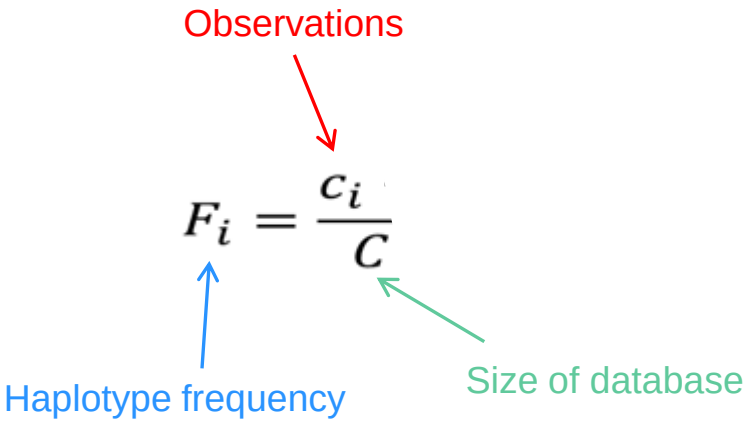
Haplotypes

The Lambda model



Haplotypes

The Lambda model



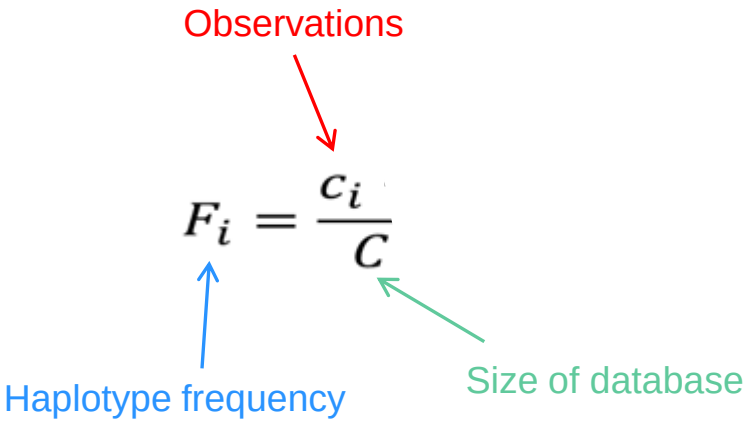
Example

Haplotype 12_20 is observed 10 times in a database of size 652

F_{12_20}=10/652

Haplotypes

The Lambda model



Example

Haplotype 12_20 is observed 0 times in a database of size 652

F_{12_20}=0/652

Haplotypes

The Lambda model

Observations

$$F_i = \frac{c_i}{C}$$

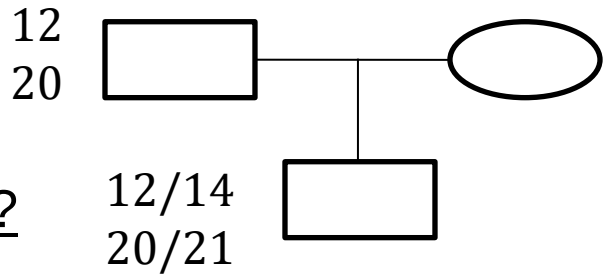
Haplotype frequency

Size of database

Example

Haplotype 12_20 is observed 0 times in a database of size 652

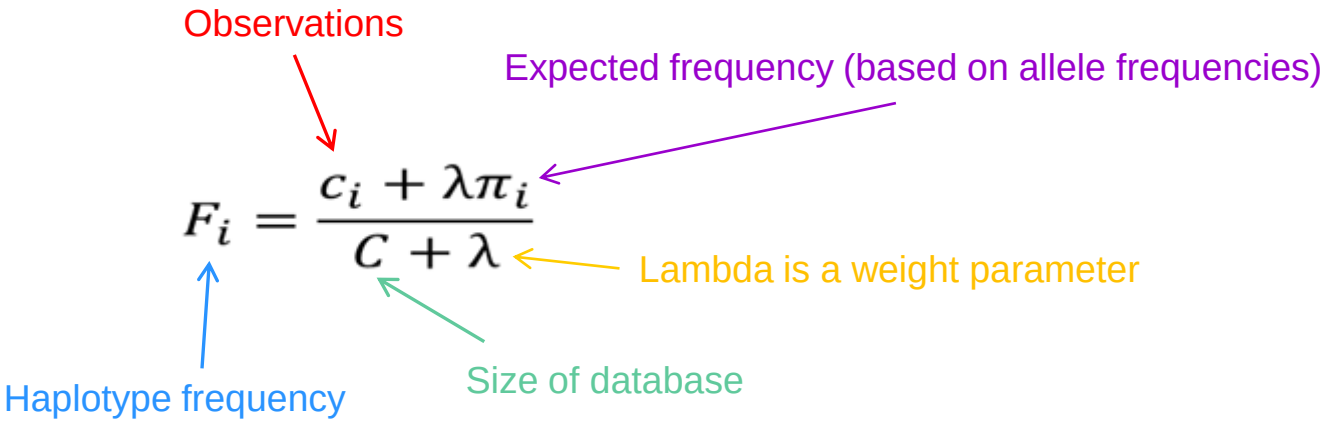
F_{12_20}=0/652



What about this case?

Haplotypes

The Lambda model



Haplotypes

The Lambda model

Observations

Expected frequency (based on allele frequencies)

$$F_i = \frac{c_i + \lambda\pi_i}{C + \lambda}$$

Lambda is a weight parameter

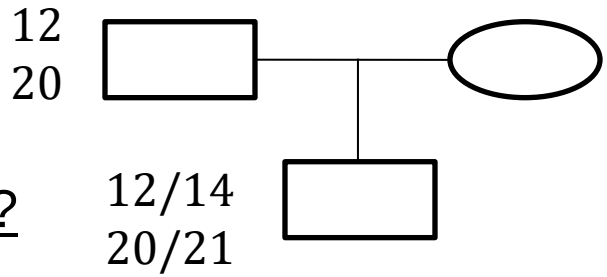
Haplotype frequency

Size of database

Example

Haplotype 12_20 is observed 0 times in a database of size 652

F_{12_20}=?



What about this case?

Haplotypes

The Lambda model

Observations

Expected frequency (based on allele frequencies)

$$F_i = \frac{c_i + \lambda\pi_i}{C + \lambda}$$

Lambda is a weight parameter

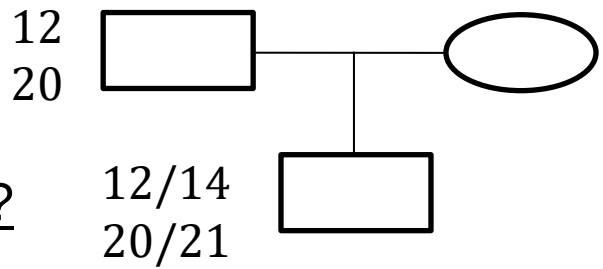
Haplotype frequency

Size of database

Example

Haplotype 12_20 is observed 0 times in a database of size 652. Lambda=100, p12=0.2, p20=0.3

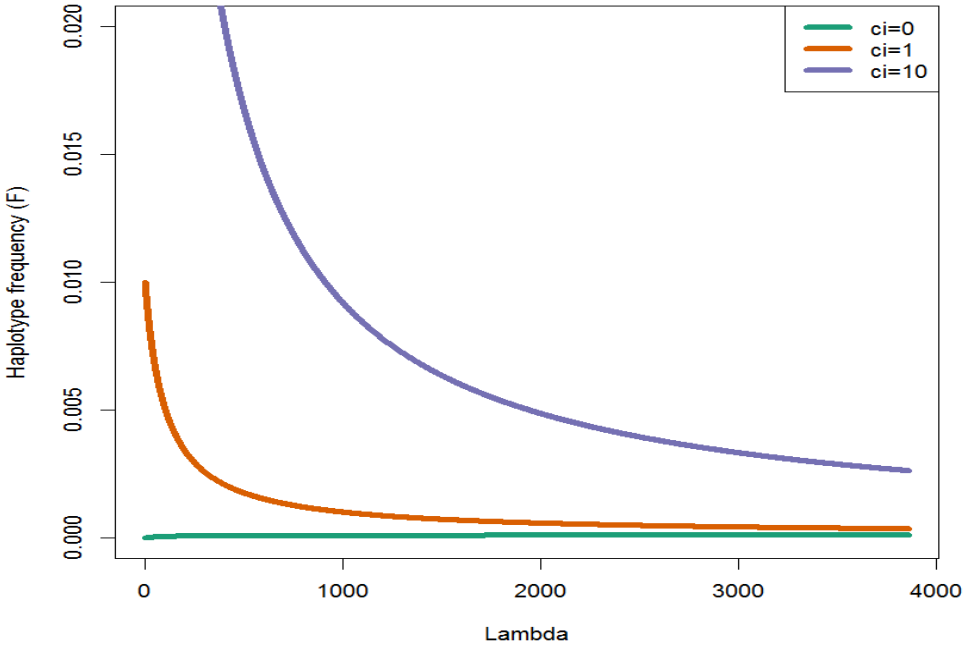
F_{12_20}=(0+100*0.2*0.3))/(652+100)= 6/752



What about this case?

Haplotypes

$$F_i = \frac{c_i + \lambda\pi_i}{C + \lambda}$$



Marker	Haplotype
DXS10148	19
DXS10135	16
DXS8378	10

c_i	$F_i(\lambda=0)$	$F_i(\lambda=4000)$	$F_i(\lambda=\infty)$
0	0	0.0001	0.0001
1	1/100	0.00035	0.0001
10	10/100	0.0025	0.0001

C=652

Difficulties

- In computations of likelihoods we,
  - Model population haplotype diversity – Estimate frequencies
  - Model haplotype transmission within a pedigree – Recombination
- Requires iteration over possible haplotype «states» - Phasing
- Mutations are highly relevant for STR markers (less so for SNPs)

More next week!

Haplotypes are not needed for statistical calculation when using markers from Decaplex (or FORCE)!

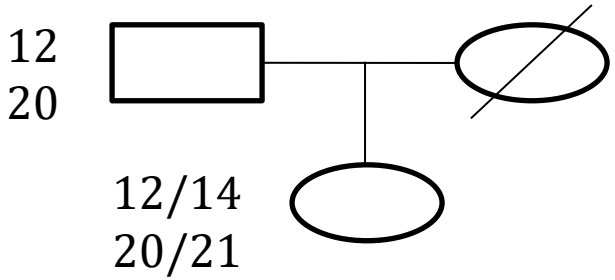
Instead, allele frequencies are sufficient

FamLinkX – Demonstration 4

Demonstration

Marker	Position (cM)	Alleles	Frequencies
DXS10148	19.84	12, 14	0.1, 0.9
DXS10135	20.03	20, 21	0.1, 0.9

Haplotype	Observations	Freq
12 20	1	1/100
12 21	9	9/100
14 20	1	1/100
14 21	89	89/100



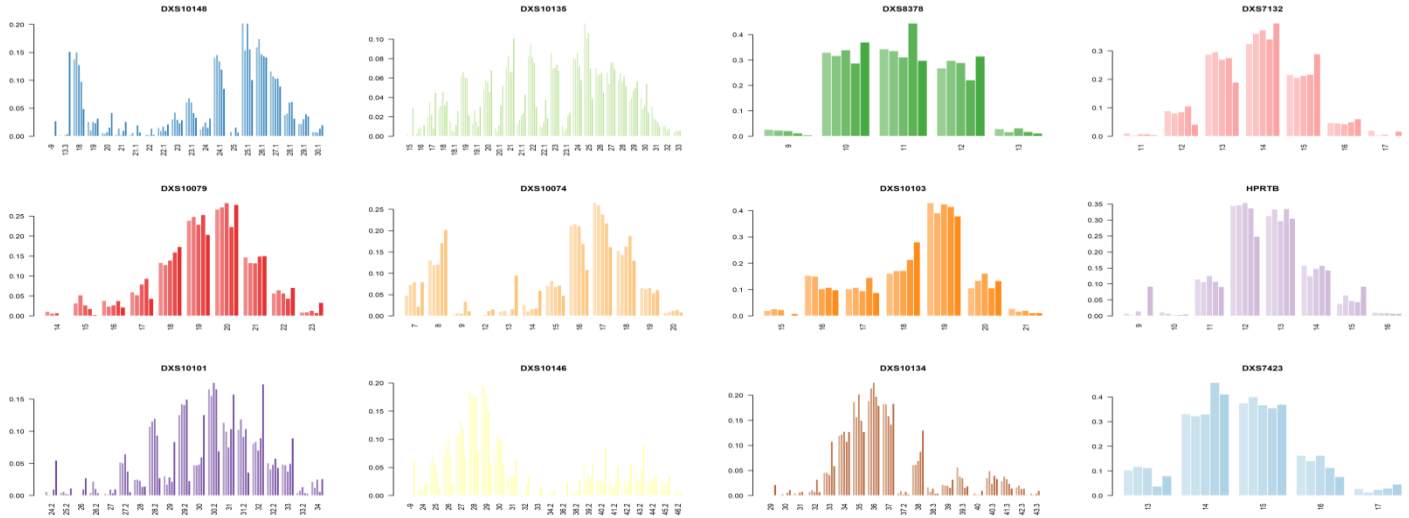
$LR_{af}=1/2p_{12} * 1/2p_{20}=5*5=\underline{25}$

$LR_{hf}=p_{14,21}/(2p_{12,20}p_{14,21} + 2p_{12,21}p_{14,20})=0.89/ 2(0.01*0.89 + 0.09*0.01)=\underline{45}$

Databases and populations

Population genetics

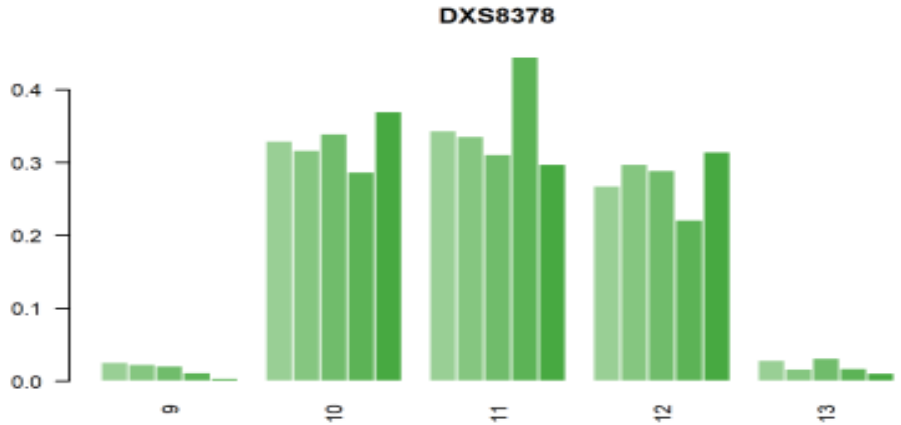
Allele frequencies for 5 different populations



Interpretation: Little information in allele frequencies alone?

Population genetics

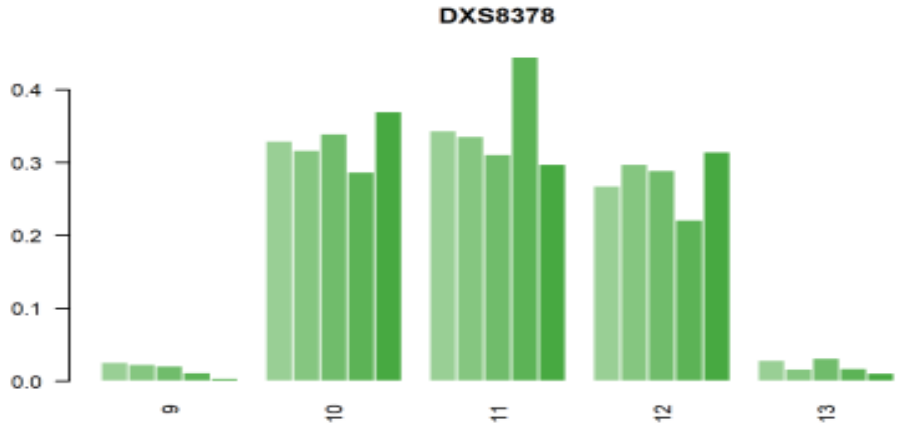
Allele frequencies for 5 different populations



Interpretation: Little information in allele frequencies alone?

Population genetics

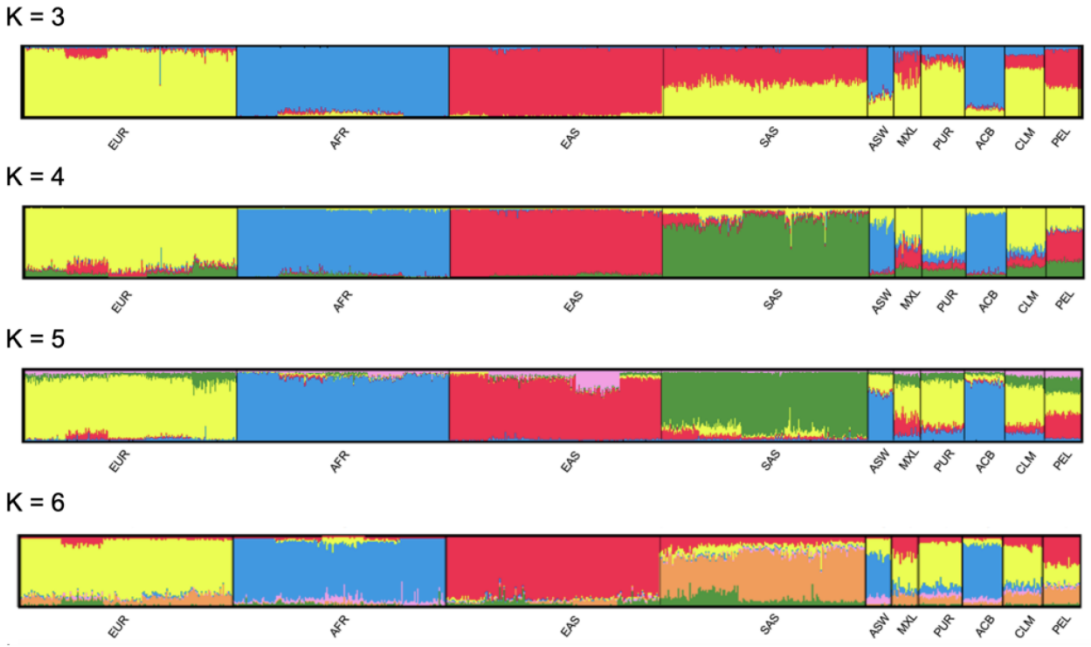
Allele frequencies for 5 different populations



Interpretation: Little information in allele frequencies alone?

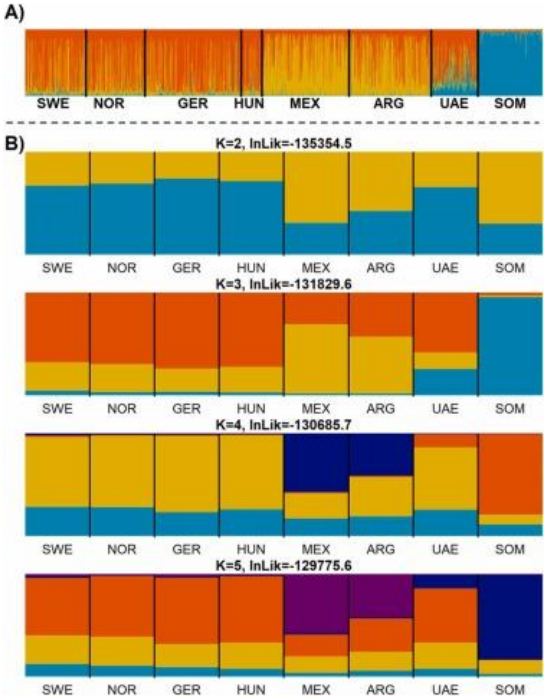
STRUCTURE

Each sample is assinged a mosaic of populations assuming K populations

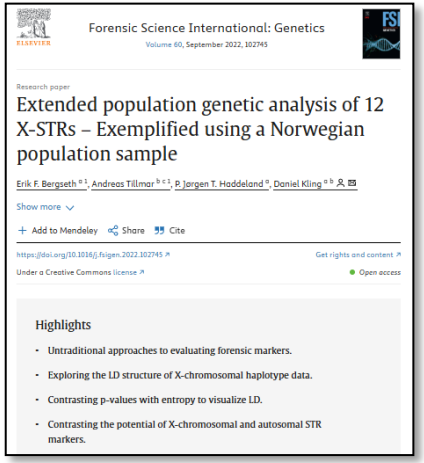


E.g. yellow represents a European ancestry

STRUCTURE



Interpretation: A vague trend where Somalia is represented by blue-ish. Too few markers to draw strong conclusions about population differences using this method.



Forensic Science International: Genetics
Volume 60, September 2022, 102745

Research paper

Extended population genetic analysis of 12 X-STRs – Exemplified using a Norwegian population sample

Erik F. Bergseth, Andreas Tillmar, P. Jørgen T. Haddeland, Daniel Kling

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.fsigen.2022.102745

Get rights and content

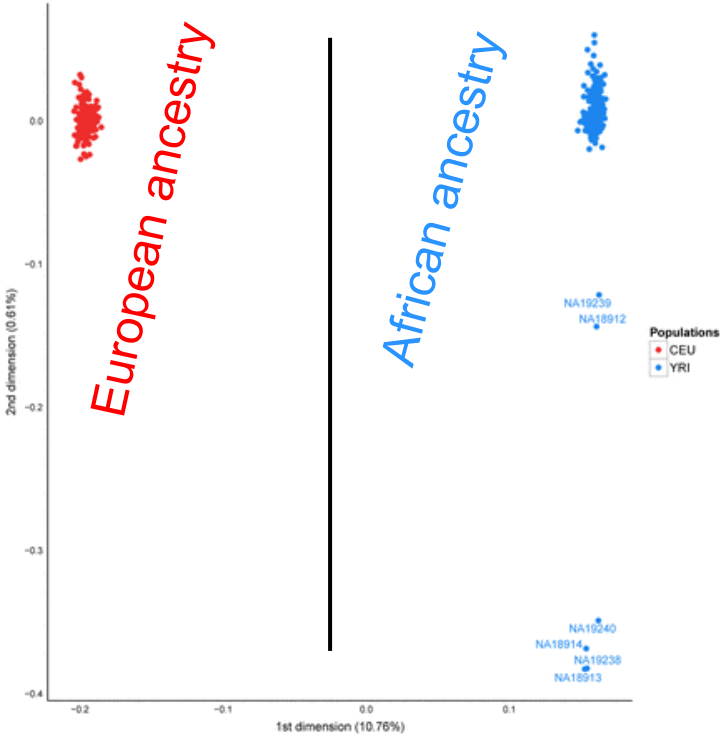
Under a Creative Commons license

Open access

Highlights

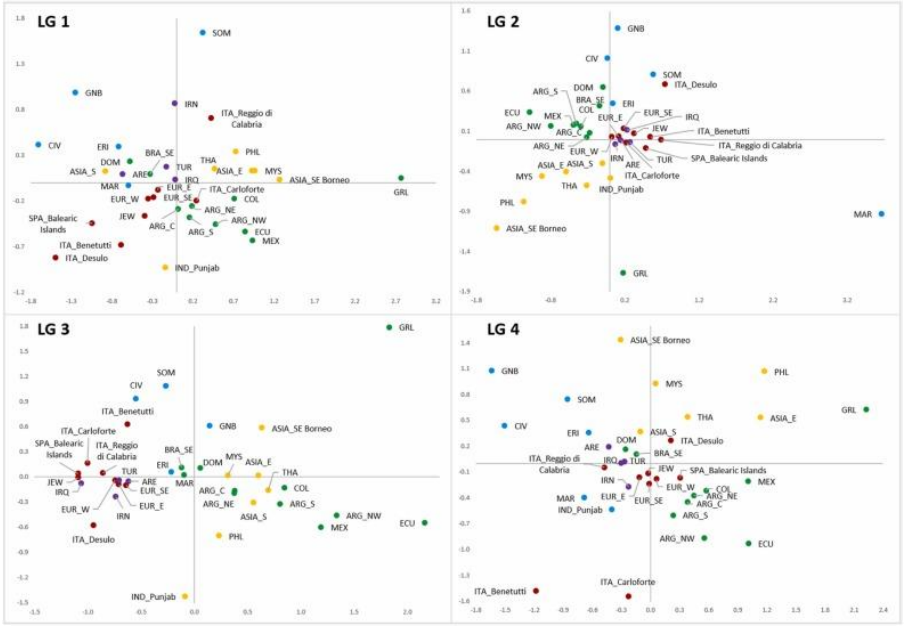
- Untraditional approaches to evaluating forensic markers.
- Exploring the LD structure of X-chromosomal haplotype data.
- Contrasting p-values with entropy to visualize LD.
- Contrasting the potential of X-chromosomal and autosomal STR markers.

Differentiation



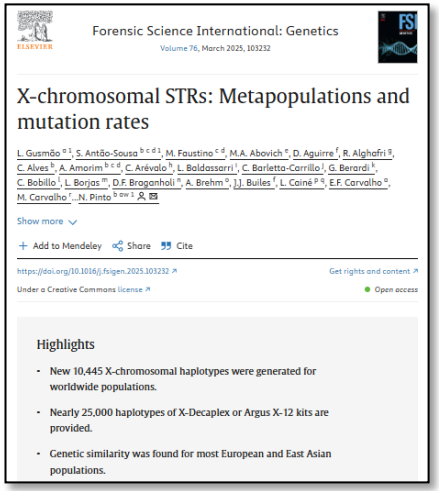
Separates multi-dimensional data into two Dimensions (MDS or PCA)

Differentiation



From Gusmao et al. 2025

Interpretation: A vague separation with outliers



Forensic Science International: Genetics
Volume 76, March 2025, 103232

X-chromosomal STRs: Metapopulations and mutation rates

L. Gusmão [a 1], S. Antão-Sousa [b c d 1], M. Faustino [c d], M.A. Abovich [e], D. Aguirre [f], R. Alghafri [g], C. Alves [h], A. Amorim [b c d], C. Arévalo [h], L. Baldassarri [i], C. Barletta-Carrillo [j], G. Berardi [k], C. Bobillo [l], L. Borjas [m], D.F. Braganholi [n], A. Brehm [o], J.J. Builes [f], L. Cainé P [q], E.F. Carvalho [o], M. Carvalho [r], N. Pinto [b ow 1]

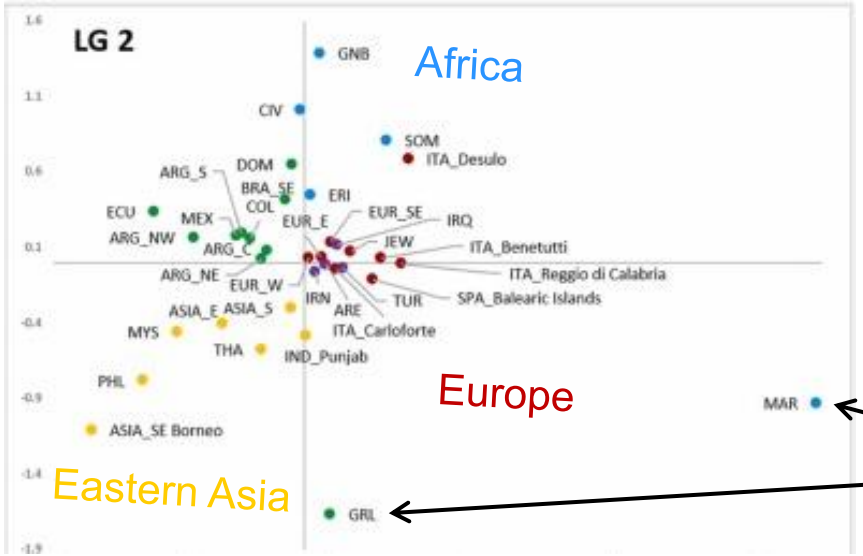
Show more
Add to Mendeley Share Cite
https://doi.org/10.1016/j.fsigen.2025.103232
Get rights and content
Under a Creative Commons license
Open access

Highlights

- New 10,445 X-chromosomal haplotypes were generated for worldwide populations.
- Nearly 25,000 haplotypes of X-Decaplex or Argus X-12 kits are provided.
- Genetic similarity was found for most European and East Asian populations.

Differentiation

South America

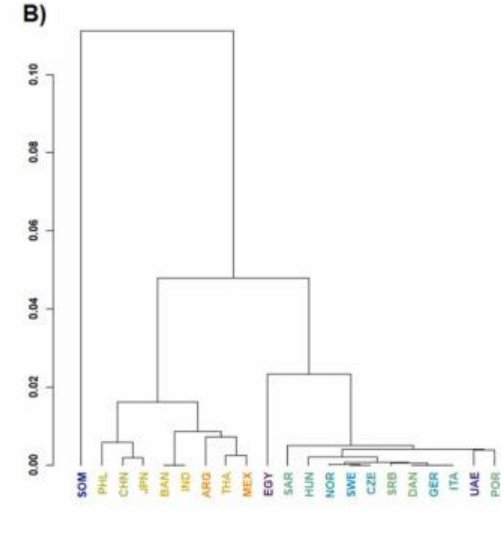


Interpretation: A vague separation with outliers

Outliers

From Gusmao et al. 2025

Differentiation



From Bergseth et al. 2023

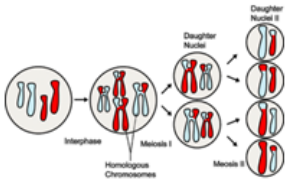
Interpretation: Somali population is outlier. Asian populations cluster and European/Middle East

Linkage disequilibrium (LD)

Recall from first presentation

Linkage and Linkage disequilibrium

- Linkage (or genetic linkage)
  - Can be described as the co-segregation of closely located loci within a family or pedigree.
  - **Effects the transmission probabilities!**



- Linkage disequilibrium (LD)
  - Allelic association.
  - Two alleles (at two different markers) which is observed more often/less often than can be expected.
  - **Effects the founder genotype probabilities, not the transmission probabilities!**
  - Haplotype frequencies rather than allele frequencies must be used

Linkage disequilibrium (LD)

- Measures, e.g. D and r2

$$D = P_{A,B} - P_A P_B$$

$$r^2 = \frac{(P_{A,B} - P_A P_B)^2}{P_A P_B (1 - P_A)(1 - P_B)}$$

Example

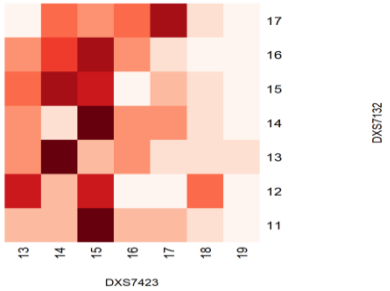
Haplotype	Observations	Ho	He
12 20	9	9/20	0.5*0.5
12 21	1	1/20	0.5*0.5
14 20	1	1/20	0.5*0.5
14 21	9	9/20	0.5*0.5

D=0.45-0.25=0.2

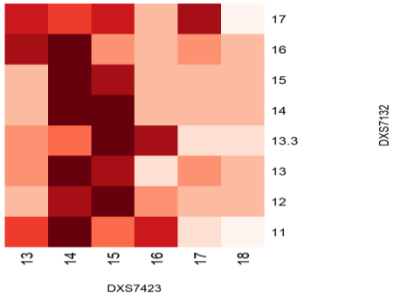
r^2=(0.45-0.25)^2/(0.5*0.5*0.5*0.5)=0.64

Population genetics

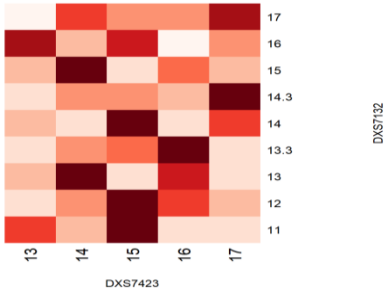
LD measured by D between pairs of alleles at two markers



Somalia



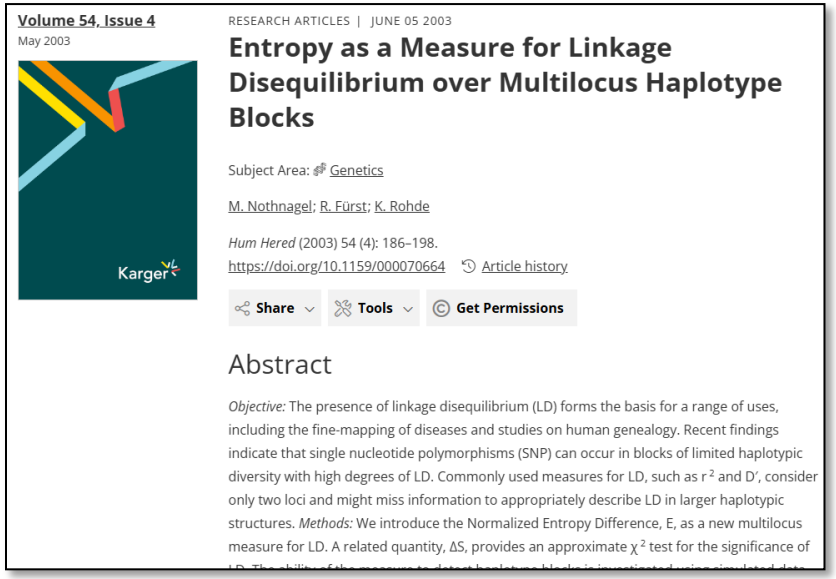
Norway



Sweden

Difficult to assess the overall LD

Linkage disequilibrium (LD)



Volume 54, Issue 4
May 2003

RESEARCH ARTICLES | JUNE 05 2003

Entropy as a Measure for Linkage Disequilibrium over Multilocus Haplotype Blocks

Subject Area: Genetics

M. Nothnagel; R. Fürst; K. Rohde

Hum Hered (2003) 54 (4): 186–198.
https://doi.org/10.1159/000070664 Article history

Abstract

Objective: The presence of linkage disequilibrium (LD) forms the basis for a range of uses, including the fine-mapping of diseases and studies on human genealogy. Recent findings indicate that single nucleotide polymorphisms (SNP) can occur in blocks of limited haplotypic diversity with high degrees of LD. Commonly used measures for LD, such as r^2 and D′, consider only two loci and might miss information to appropriately describe LD in larger haplotypic structures. *Methods:* We introduce the Normalized Entropy Difference, E, as a new multilocus measure for LD. A related quantity, ΔS, provides an approximate χ^2 test for the significance of

Entropy

Linkage disequilibrium (LD)

- Measures, e.g. D and r2

$$D = P_{A,B} - P_A P_B$$

$$r^2 = \frac{(P_{A,B} - P_A P_B)^2}{P_A P_B (1 - P_A)(1 - P_B)}$$

- What about multiallelic markers?
  - We use entropy!

$$H = -\sum_s p_s \times \ln(p_s)$$

H is entropy and p_s is allele frequency

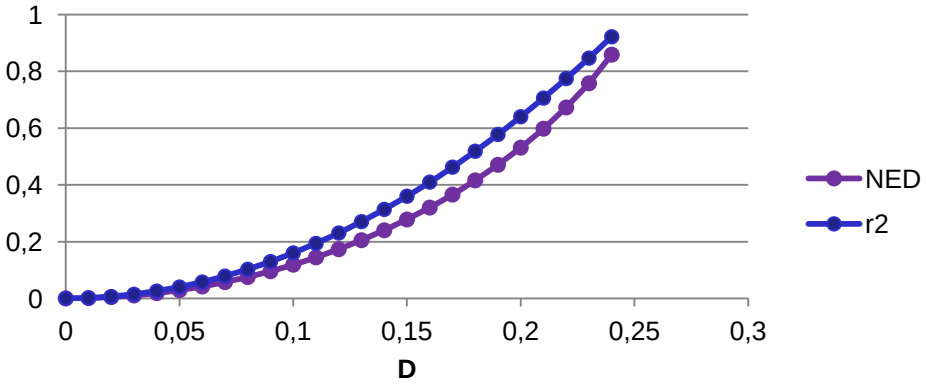
Linkage disequilibrium

Normalized Entropy Difference (NED)

Single marker entropies

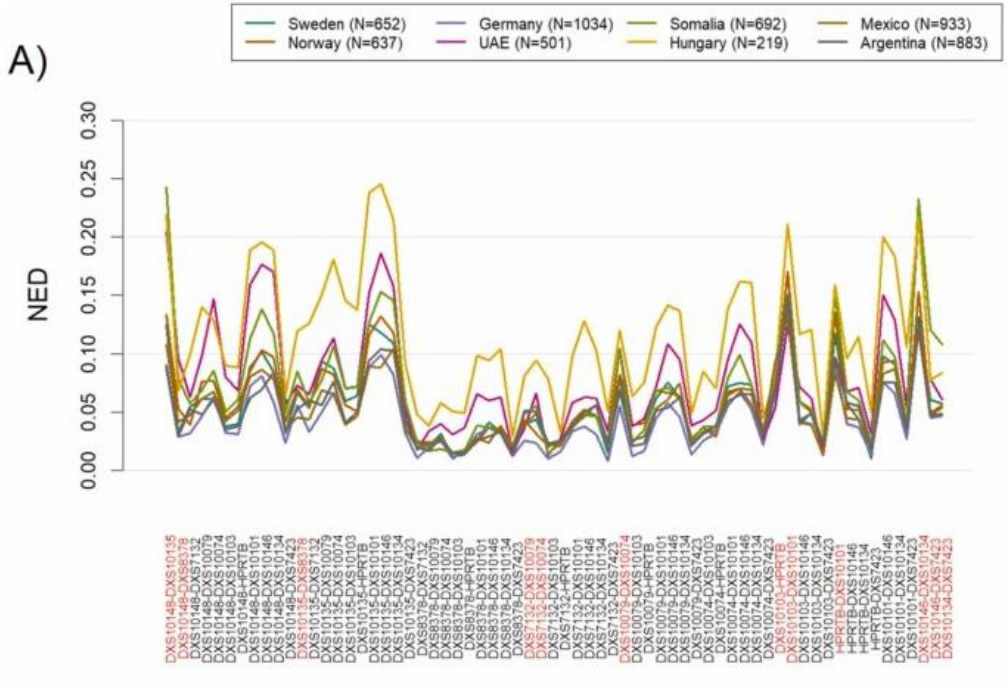
Joint entropy

$$NED_{i,j} = 2 \times \frac{(H_i + H_j) - H_{i,j}}{H_i + H_j}$$



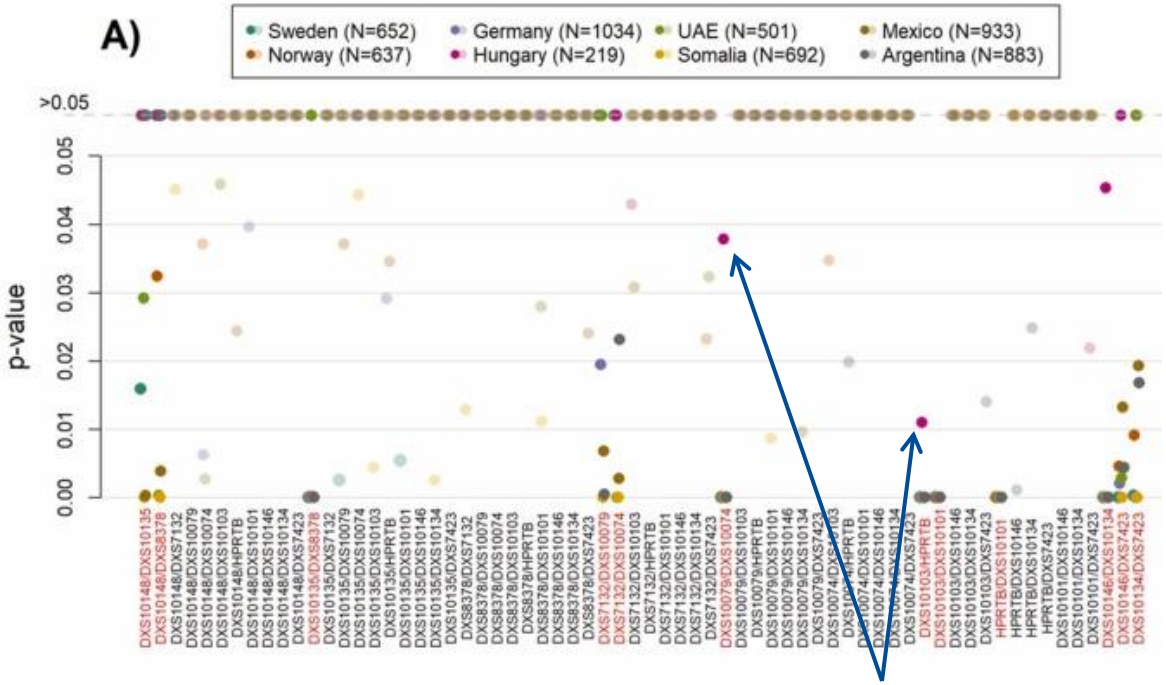
Two SNPs with minor allele frequencies of 0.5. NED as computed based on the degree of LD (D)

Population genetics



Interpretation: Similar patterns of NED for the populations

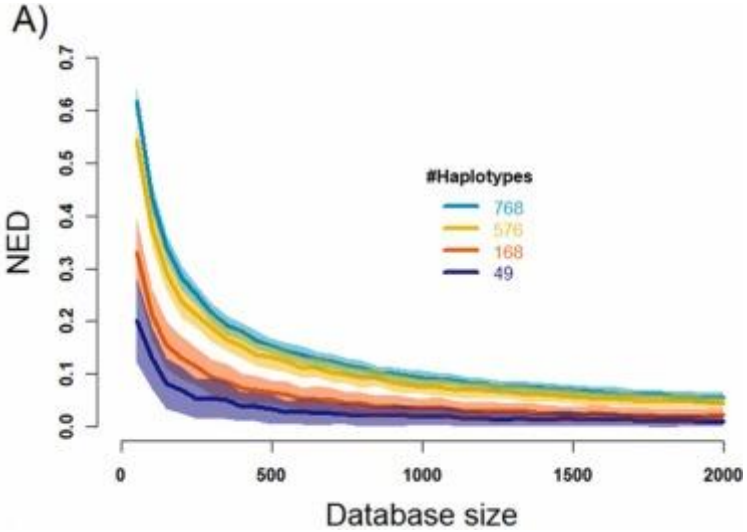
Population genetics



Interpretation: Smaller databases may fail to detect the presence of LD

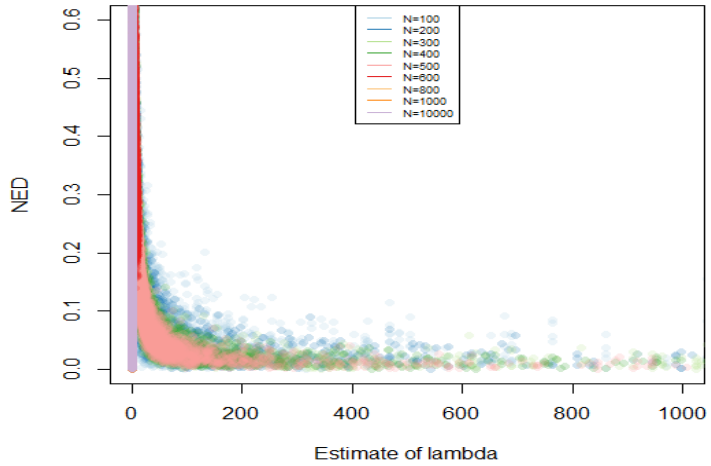
Population genetics

Data based on linkage equilibrium (LE)



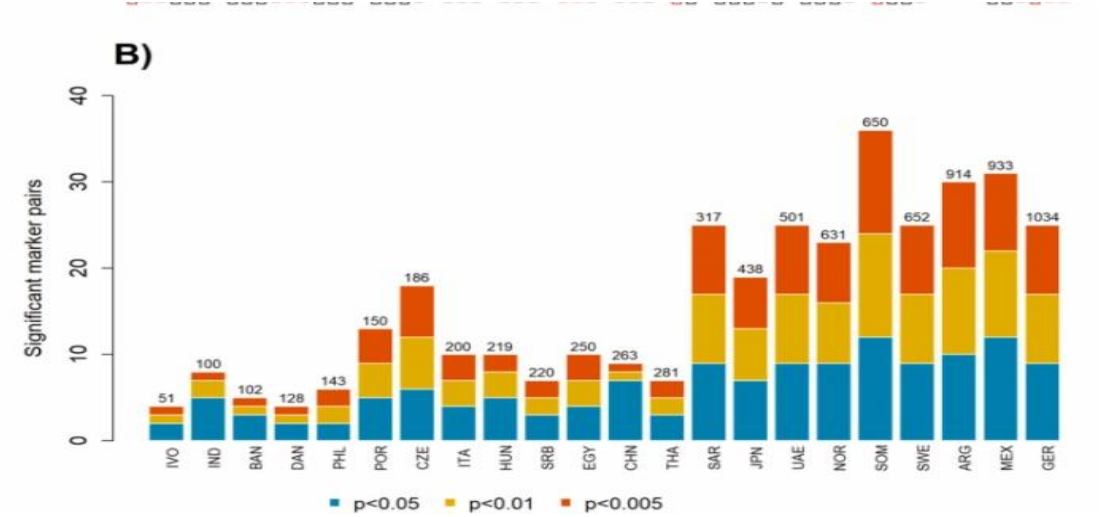
Interpretation: The smaller the sampled database, the higher the NED

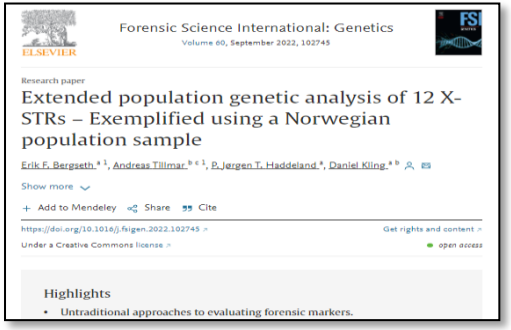
Population genetics



Interpretation: Large databases leads to low (zero) estimates of lambda

On the importance of size




Forensic Science International: Genetics
Volume 60, September 2022, 102745
ELSEVIER
Research paper
Extended population genetic analysis of 12 X-STRs – Exemplified using a Norwegian population sample
Erik F. Bergseth, Andreas Tillmar, P. Jørgen T. Haddeland, Daniel Kling
Show more
Add to Mendeley Share Cite
https://doi.org/10.1016/j.fsigen.2022.102745
Get rights and content
Under a Creative Commons license
open access
Highlights
• Untraditional approaches to evaluating forensic markers.

From Bergseth et al. 2022

In small sized databases there is less LD??

On the importance of size

Example

Two SNPs with 2 alleles each => Total 4 possible haplotypes
Two STRs with 15 alleles each => Total of 225 possible haplotypes

On the importance of size

Table 2. Population genetic data for the Norwegian population sample (N = 631), with relevant forensic efficiency parameters for four triads of X-chromosomal STR markers in the Argus X-12 QS kit. MEC = Mean exclusion chance, PE=Power of Exclusion. PDF/PDM/MEC Desmarais Trio and Duo are calculated using formulas in Desmarais et al. [11] whereas PE (Paternal half sisters) and PE (Paternal grandmother/granddaughter) are calculated using formulas in Pinto et al. [12].

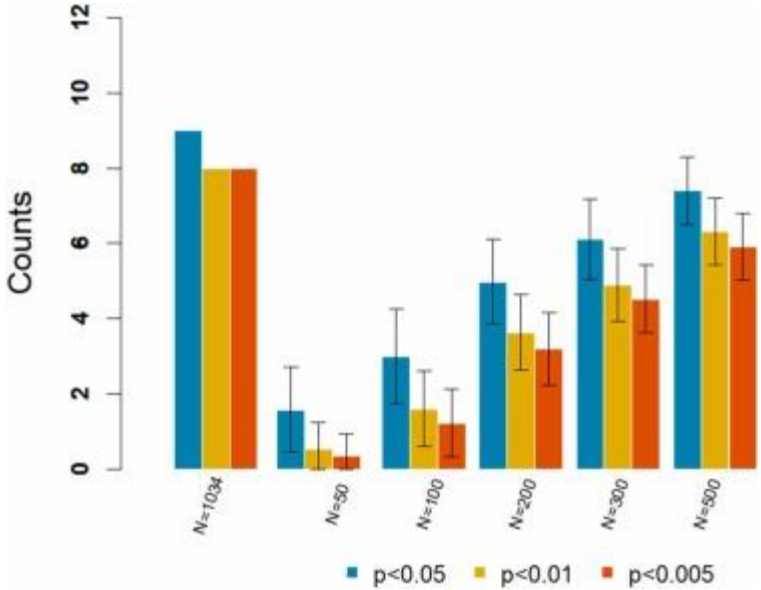
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Theoretical haplotypes	5376	1560	1197	3456
Observed haplotypes	362	204	193	275
Most common haplotype	7	20	24	13
Singletons	218	91	91	142
Power of Discrimination in Females (PDF)	0.999969	0.999791	0.99972	0.99991
Power of Discrimination in Males (PDM)	0.996074	0.989685	0.988032	0.993256
MEC Desmarais Trio	0.996059	0.989582	0.987896	0.993212
MEC Desmarais Duo	0.99217	0.979544	0.976307	0.986584
PE (Paternal half sisters)	0.984411	0.959637	0.953365	0.973399
PE (Paternal grandmother/granddaughter)[a]	0.992139	0.979344	0.976042	0.986496

5376 possible haplotypes!

Only 362 observed

From Bergseth et al. 2022

On the importance of size



From Bergseth et al. 2022

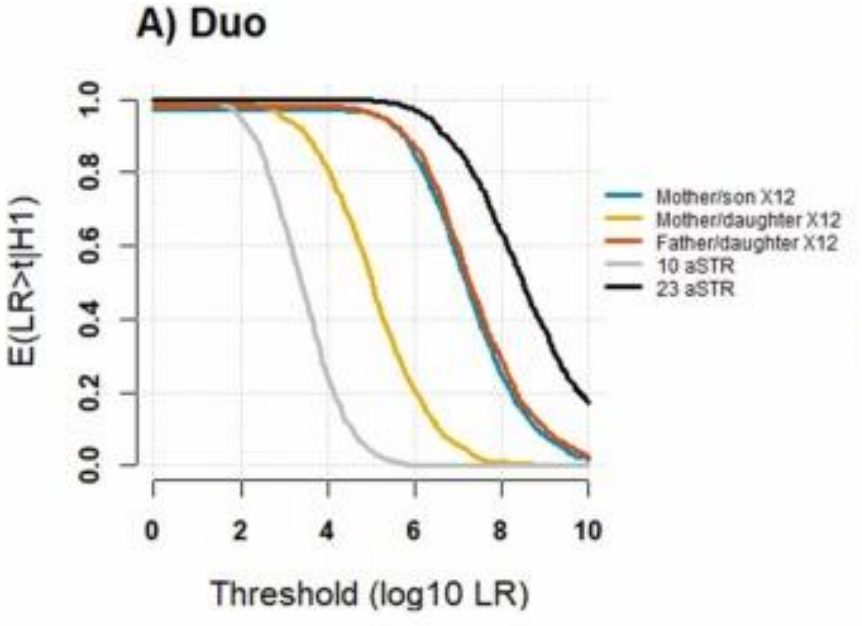
Simulation as a mean to evaluate markers

Table 2. Population genetic data for the Norwegian population sample (N = 631), with relevant forensic efficiency parameters for four triads of X-chromosomal STR markers in the Argus X-12 QS kit. MEC = Mean exclusion chance, PE=Power of Exclusion. PDF/PDM/MEC Desmarais Trio and Duo are calculated using formulas in Desmarais et al. [11] whereas PE (Paternal half sisters) and PE (Paternal grandmother/granddaughter) are calculated using formulas in Pinto et al. [12].

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Theoretical haplotypes	5376	1560	1197	3456
Observed haplotypes	362	204	193	275
Most common haplotype	7	20	24	13
Singletons	218	91	91	142
Power of Discrimination in Females (PDF)	0.999969	0.999791	0.99972	0.99991
Power of Discrimination in Males (PDM)	0.996074	0.989685	0.988032	0.993256
MEC Desmarais Trio	0.996059	0.989582	0.987896	0.993212
MEC Desmarais Duo	0.99217	0.979544	0.976307	0.986584
PE (Paternal half sisters)	0.984411	0.959637	0.953365	0.973399
PE (Paternal grandmother/granddaughter)[a]	0.992139	0.979344	0.976042	0.986496

Are these markers useful?

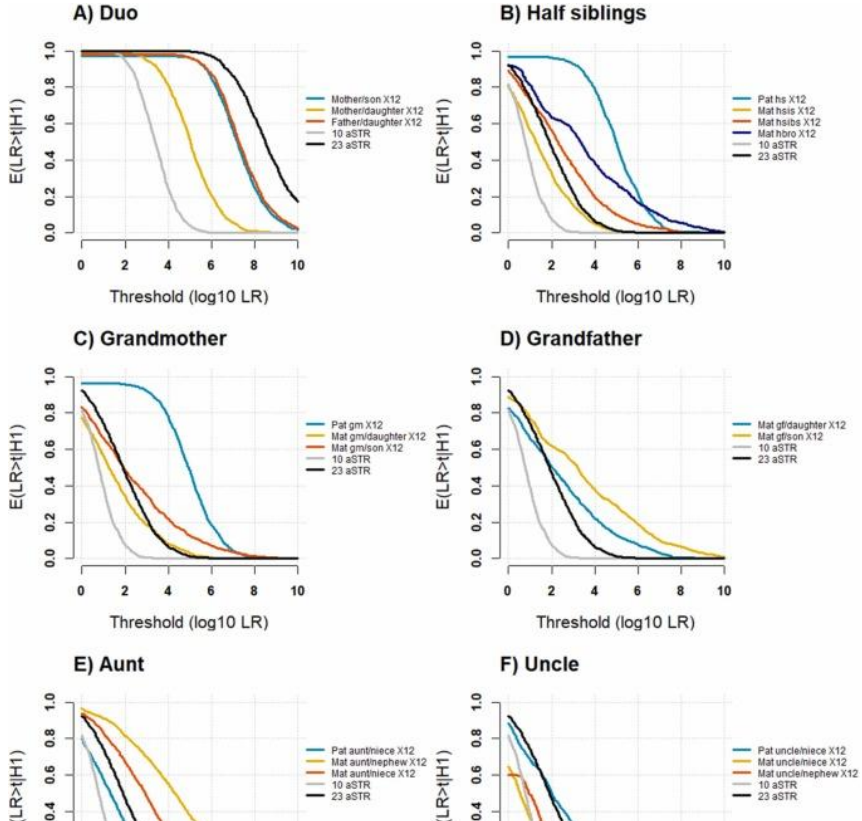
Simulation as a mean to evaluate markers



Conclusion: 12 X-STRs is more informative then 10 extra aSTRs

From Bergseth et al. 2022

Simulation as a mean to evaluate markers



From Bergseth et al. 2022

Haplotypes and databases

DANIEL KLING & ANDREAS TILLMAR

Daniel.kling@rmv.se

Andreas.tillmar@rmv.se