ESWG - Proficiency test 2018

Daniel Kling – Head Organizer

rmdakl@ous-hf.no

PAPER CHALLENGE

Disclaimer

All individuals and DNA data are fictional. Any resemblence to known persons is purely coincidental.

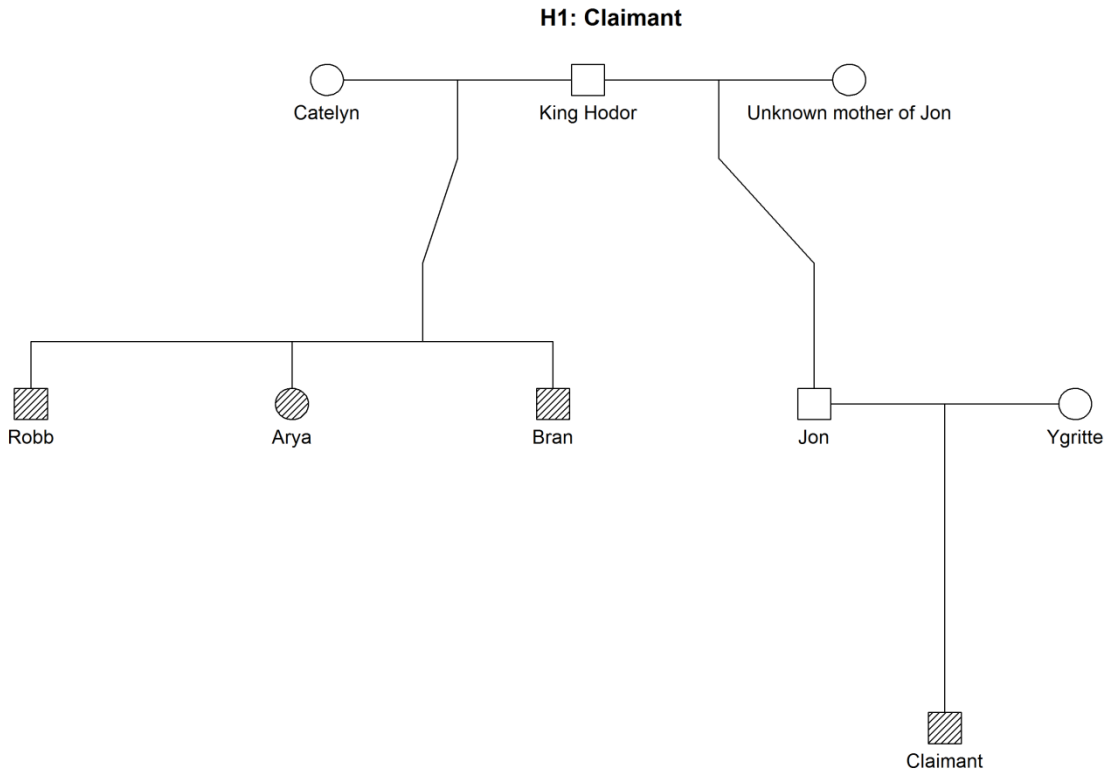
Paper challenge - Background

This year's paper challenge involves an ancient tale from a realm known as Westeros. The history has it that a claimant to the throne presents himself. He proposes that he is the grandchild of the deceased king. At the time, the allegations could not be ascertained and the claimant was regarded as a false pretender.

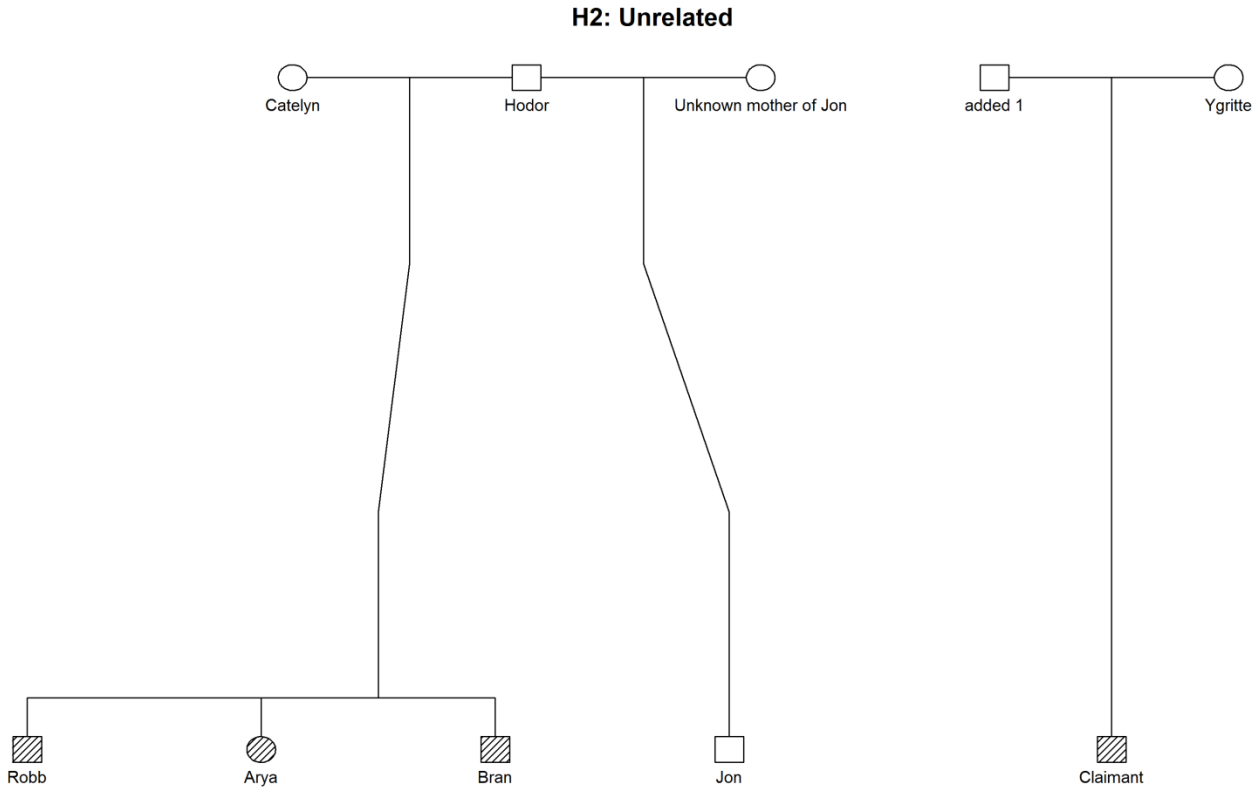
Using modern DNA techniques we are able to obtain samples and results from three known descendants of the King as well as from the Claimant.

Conduct a kinship analysis to uncover whether the Claimant is related to the three descendants of the king the way he claims. The pedigree corresponding to the claimant's allegations is depicted below. In the alternative hypothesis, the Claimant is unrelated to the descendants who are themselves still full siblings.

Paper challenge – Claimed relationship



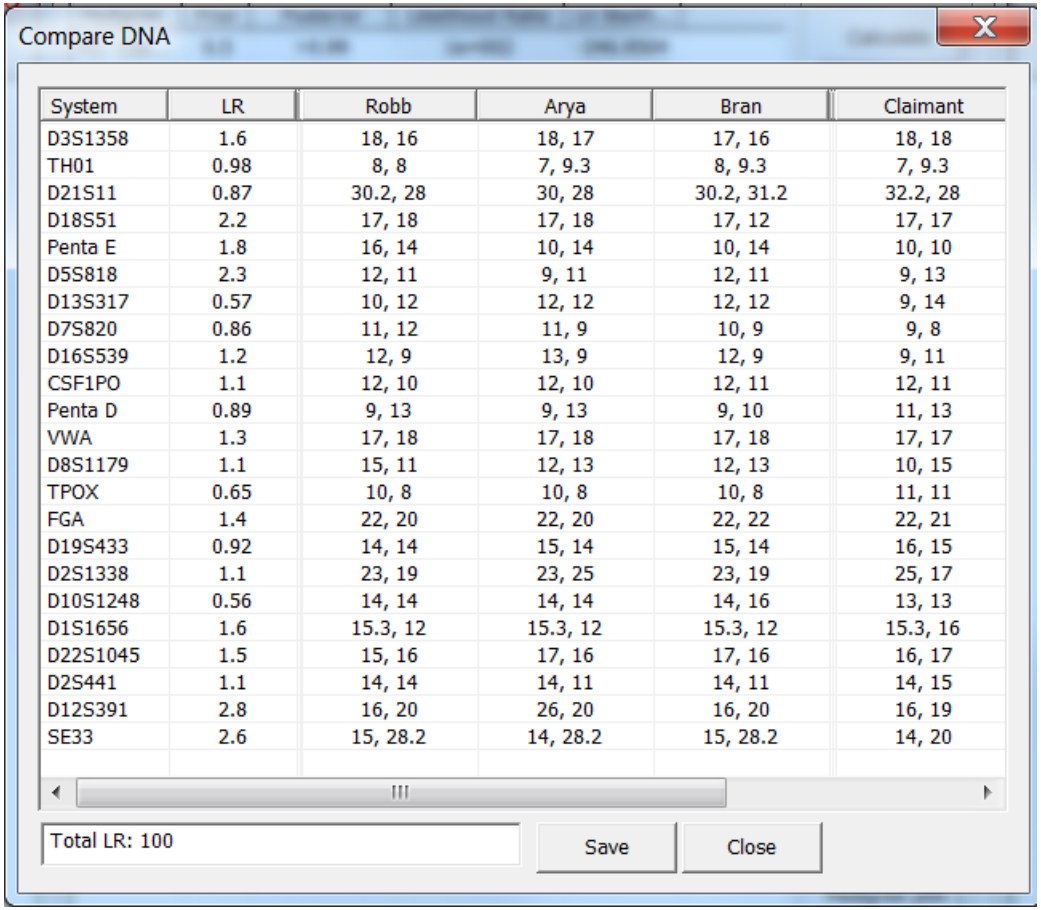
Paper challenge – Alternative



Paper challenge – Data

- Autosomal marker data for 23 STR markers
- Y-chromosomal data for 23 STR markers

Paper challenge – Autosomal results


Compare DNA

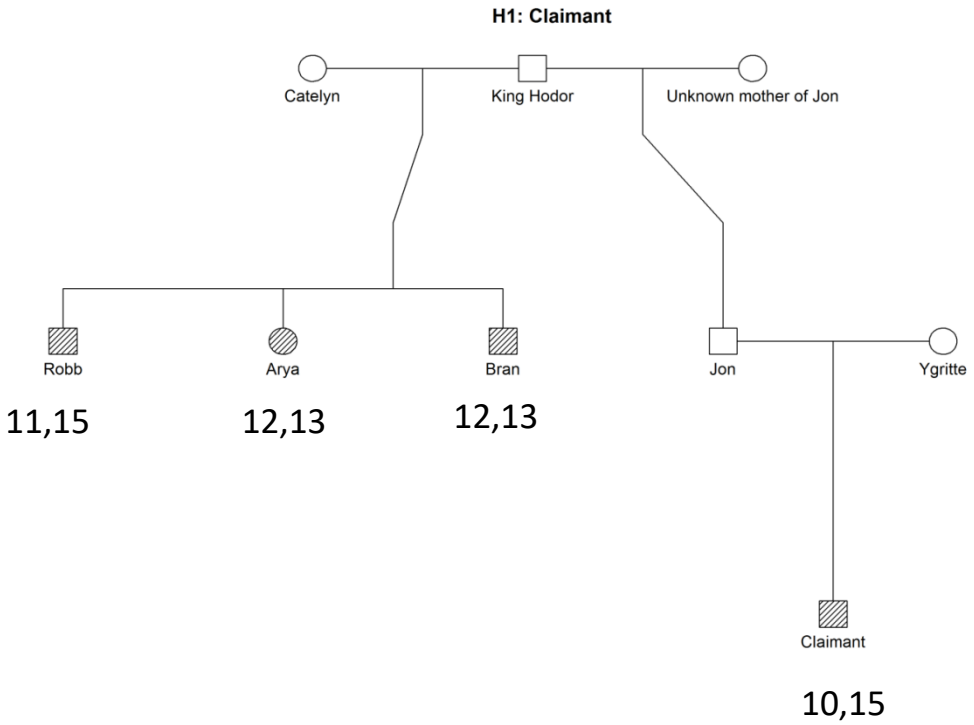
System	LR	Robb	Arya	Bran	Claimant
D3S1358	1.6	18, 16	18, 17	17, 16	18, 18
TH01	0.98	8, 8	7, 9.3	8, 9.3	7, 9.3
D21S11	0.87	30.2, 28	30, 28	30.2, 31.2	32.2, 28
D18S51	2.2	17, 18	17, 18	17, 12	17, 17
Penta E	1.8	16, 14	10, 14	10, 14	10, 10
D5S818	2.3	12, 11	9, 11	12, 11	9, 13
D13S317	0.57	10, 12	12, 12	12, 12	9, 14
D7S820	0.86	11, 12	11, 9	10, 9	9, 8
D16S539	1.2	12, 9	13, 9	12, 9	9, 11
CSF1PO	1.1	12, 10	12, 10	12, 11	12, 11
Penta D	0.89	9, 13	9, 13	9, 10	11, 13
VWA	1.3	17, 18	17, 18	17, 18	17, 17
D8S1179	1.1	15, 11	12, 13	12, 13	10, 15
TPOX	0.65	10, 8	10, 8	10, 8	11, 11
FGA	1.4	22, 20	22, 20	22, 22	22, 21
D19S433	0.92	14, 14	15, 14	15, 14	16, 15
D2S1338	1.1	23, 19	23, 25	23, 19	25, 17
D10S1248	0.56	14, 14	14, 14	14, 16	13, 13
D1S1656	1.6	15.3, 12	15.3, 12	15.3, 12	15.3, 16
D22S1045	1.5	15, 16	17, 16	17, 16	16, 17
D2S441	1.1	14, 14	14, 11	14, 11	14, 15
D12S391	2.8	16, 20	26, 20	16, 20	16, 19
SE33	2.6	15, 28.2	14, 28.2	15, 28.2	14, 20

Total LR: 100 Save Close

LR: 100
LR: 20-30 if you remove linked
LR: 130 if you account for linked

Paper challenge – Autosomal results

Manual calculations for a single marker, D8S1179

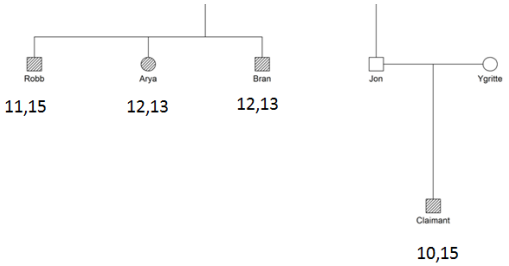


Founder genotypes

Hodor	Catelyn
11,12	13,15
12,15	11,13
11,13	12,15
13,15	11,12

Paper challenge – Autosomal results

Manual calculations for a single marker, D8S1179



Hodor	Catelyn
11,12	13,15
12,15	11,13
11,13	12,15
13,15	11,12

Founder probabilities

IBD probability between grandfather and grandson

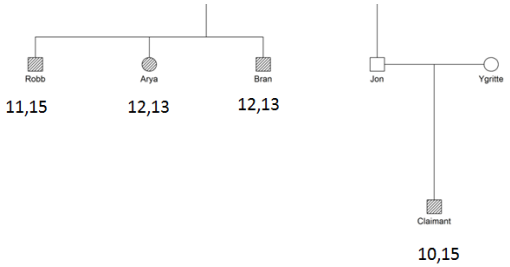
$$Likelihod(data|H1) = 2p_{11}p_{12} \cdot 2p_{13}p_{15} \cdot 0.5^6 \cdot (0.5 \cdot 2p_{10}p_{15}) + \cdots$$

Transmission probabilities to descendants

$$Likelihod(data|H2) = 2p_{10}p_{15} \cdot (2p_{11}p_{12} \cdot 2p_{13}p_{15} \cdot 0.5^6 + \cdots + 2p_{13}p_{15} \cdot 2p_{11}p_{12} \cdot 0.5^6)$$

Paper challenge – Autosomal results

Manual calculations for a single marker, D8S1179



Hodor	Catelyn
11,12	13,15
12,15	11,13
11,13	12,15
13,15	11,12

$$LR = \frac{0.5^6 \cdot (2p_{11}p_{12} \cdot 2p_{13}p_{15} \cdot 0.5 \cdot 2p_{10}p_{15} + \cdots + 2p_{13}p_{15} \cdot 2p_{11}p_{12} \cdot (0.5 \cdot 2p_{10}p_{15} + 0.5 \cdot 0.5p_{10}))}{{p_8}^2 \cdot 0.5^6 \cdot (2 \cdot 2p_{17}p_{18} \cdot 2p_{16}p_{17} + 2 \cdot 2p_{17}p_{18} \cdot 2p_{16}p_{18})}$$

$$= \frac{0.5 \cdot 2p_{10}p_{15} + 0.5 \cdot 2p_{10}p_{15} + (0.5 \cdot 2p_{10}p_{15} + 0.5 \cdot 0.5p_{10}) + (0.5 \cdot 2p_{10}p_{15} + 0.5 \cdot 0.5p_{10})}{4 \cdot 2p_{10}p_{15}}$$

$$= \frac{2p_{15} + 2p_{15} + (2p_{15} + 0.5) + (2p_{15} + 0.5)}{16p_{15}} = \frac{8p_{15} + 1}{16p_{15}} = \frac{8 \cdot 0.1 + 1}{16 \cdot 0.1} = 1.125$$

Paper challenge – Y results

Y STRs	Robb	Arya	Bran	Claimant
DYS576	18		18	18
DYS389I	13		13	13
DYS448	19		19	19
DYS389II	29		29	29
DYS19	14		14	14
DYS391	10		10	10
DYS481	21		21	21
DYS549	12		12	12
DYS533	13		13	13
DYS438	12		12	12
DYS437	14		14	14
DYS570	17		17	18
DYS635	23		23	23
DYS390	24		24	24
DYS439	11		11	11
DYS392	13		13	13
DYS643	10		10	10
DYS393	13		13	13
DYS458	18		18	18
DYS385	11,15		11,15	11,15
DYS456	16		16	16
YGATAH4	12		12	12

Inconcistency

$H_{R,B}$ H_C

Paper challenge – Y results

High mutation rate!

YHRD Search the Database Tools Resources Projects News and Updates 1 Help & Support

Locus Information on DYS570

Mutation rate

1.33e-02 (25 in 1873) based on Ballantyne2010, Maria Geppert, Josephine Purps, Patricia Entz, Carmen Krüger, Petra Otremba, Marion Nagy, Lutz Roewer, Chris Tyler-Smith, Wei Wei, Xue Y, Ayub Q, Mohyuddin A, Qamar R, Zerjal T, Helgason A, Mendi SQ, Sheila Marie Tabulina Angustia

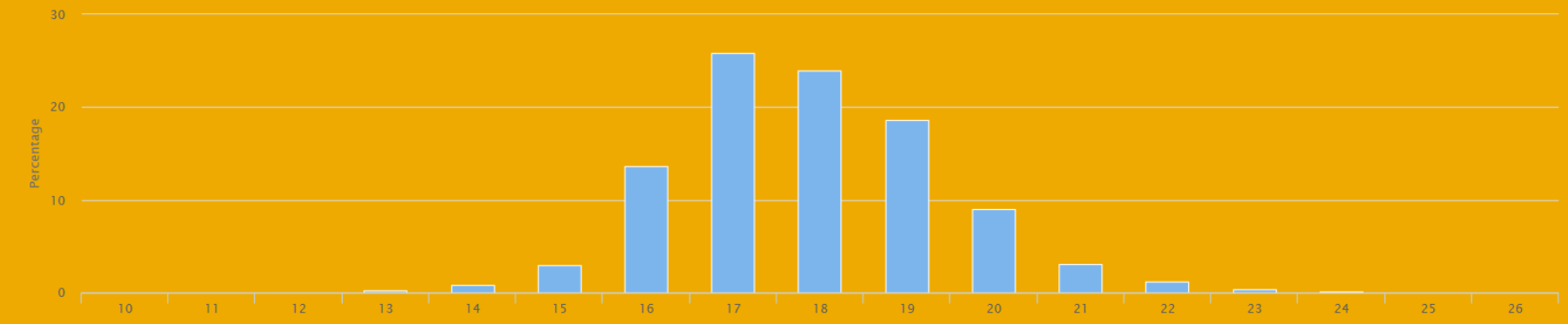
Observed alleles

10, 11, 12, 13, 13.3, 14, 15, 15.1, 15.2, 16, 16.2, 17, 17.2, 17.3, 18, 18.2, 18.3, 19, 19.1, 19.3, 20, 20.2, 20.3, 21, 21.2, 22, 22.2, 22.3, 23, 24, 25, 26

NULL alleles

24 NULL allele observations.

Allelic distribution

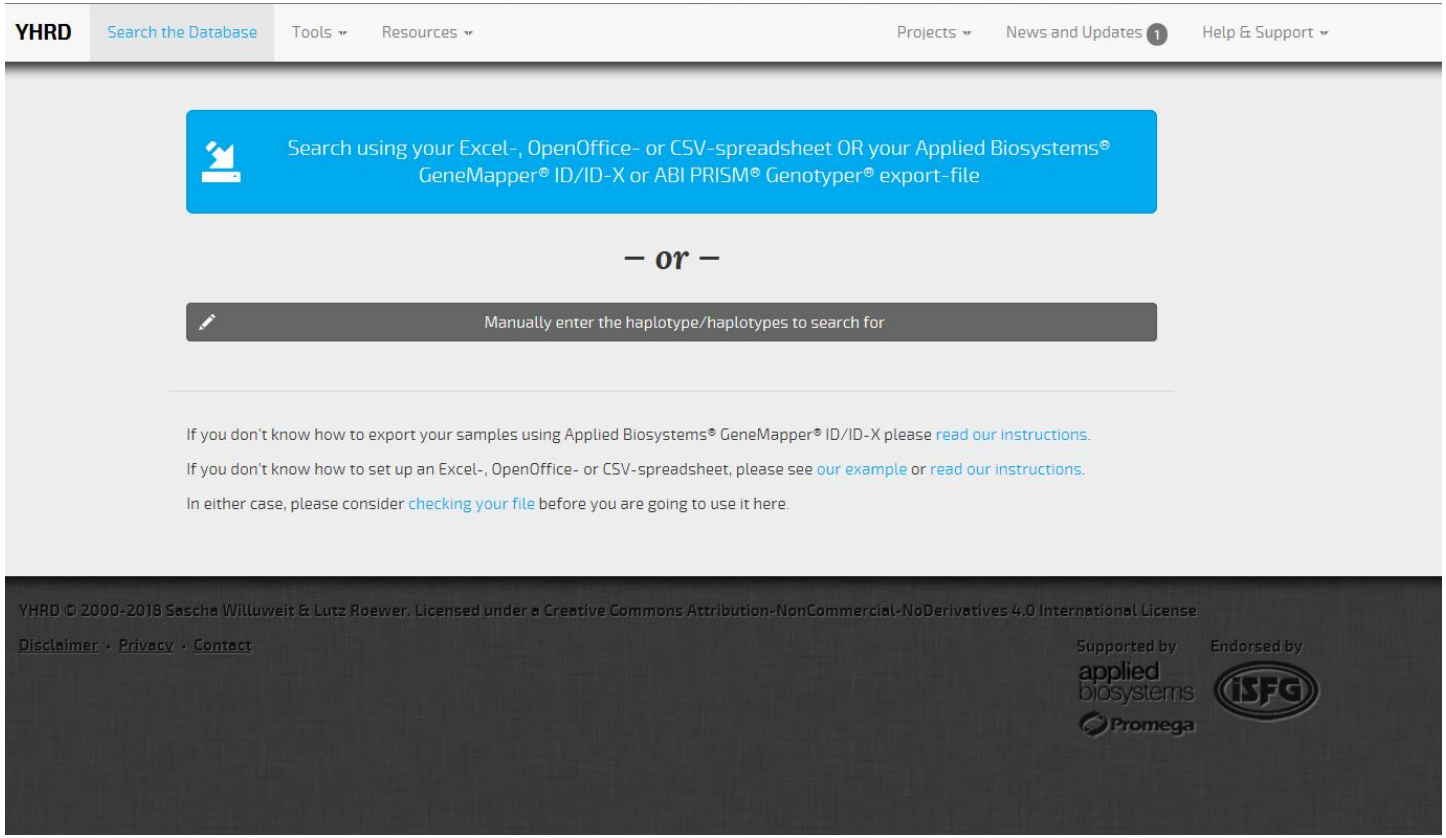


Linear, or Logarithmic y-axis. Hide intermediate alleles.

Treat duplicated alleles as two or more separate observations, or as one observation.

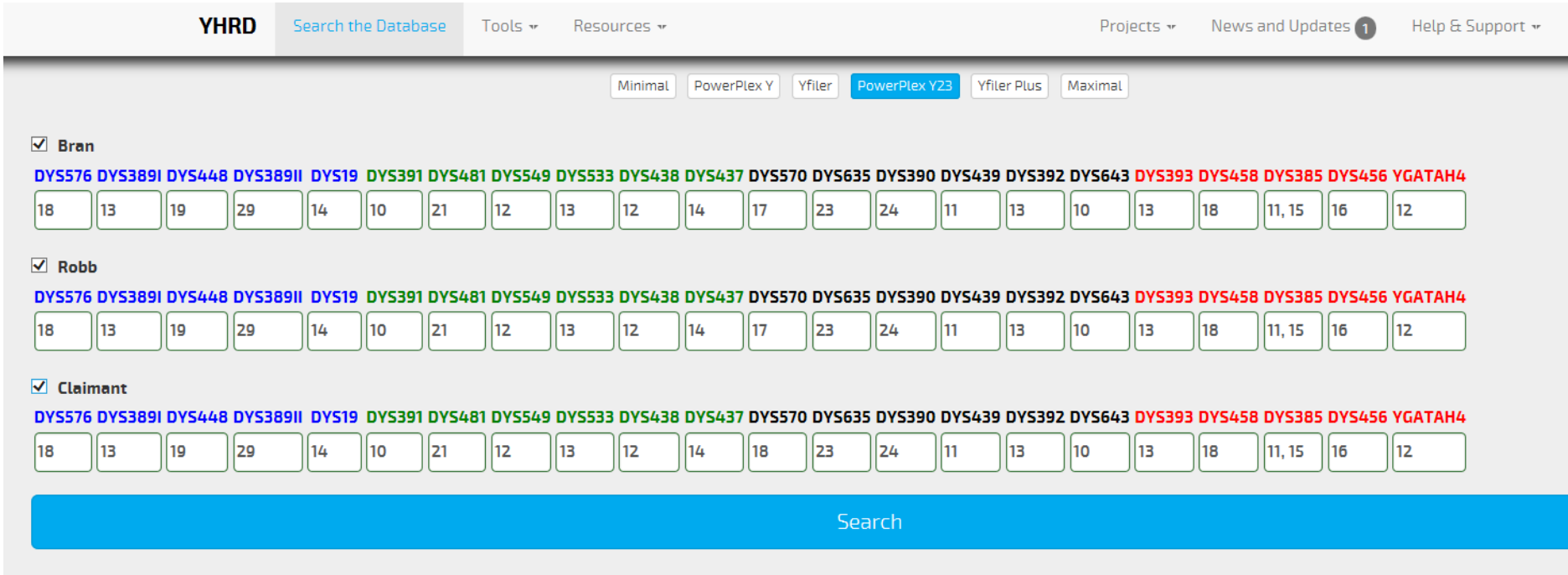
Paper challenge – Y results

Let's search YHRD to find if haplotypes have been observed!

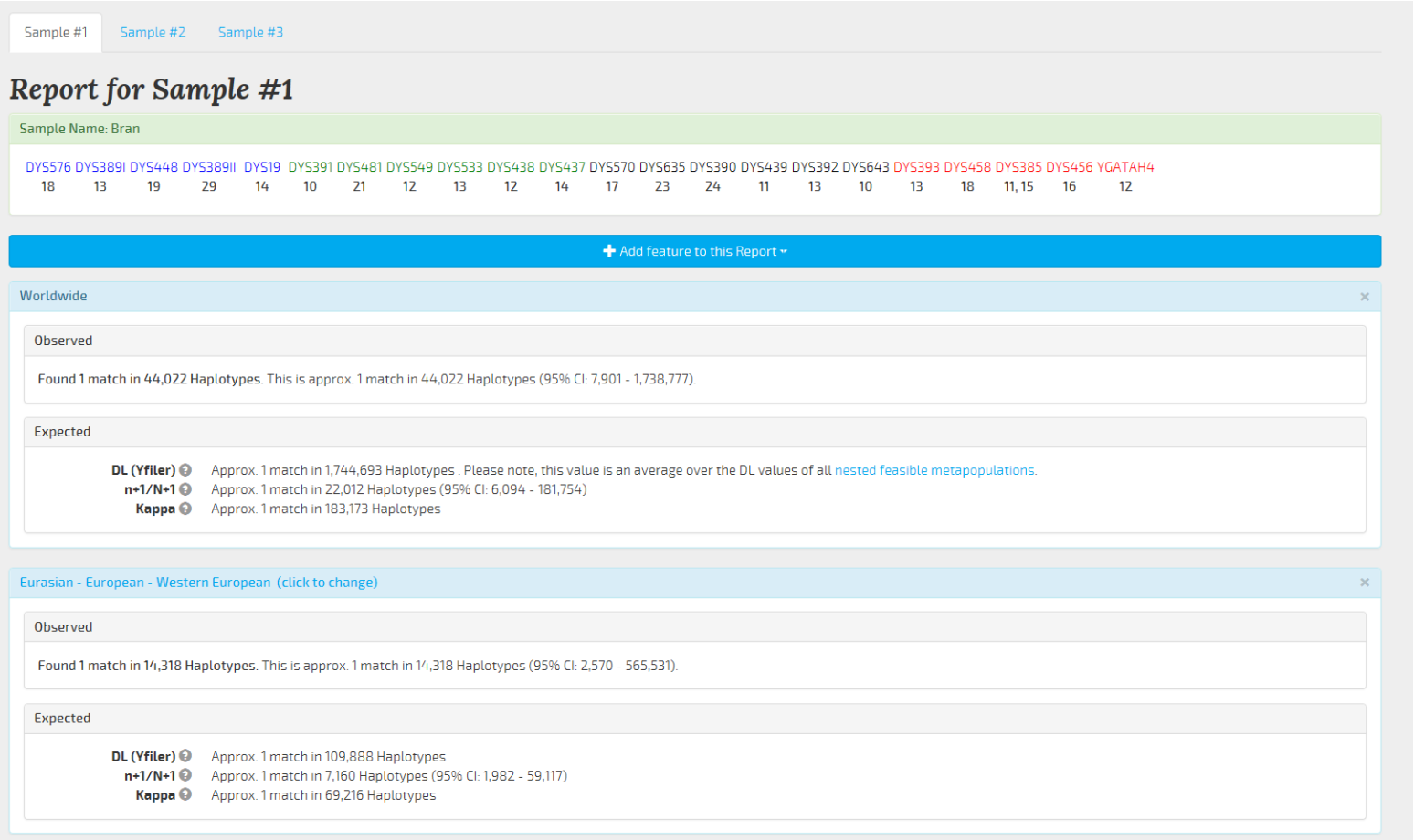


Paper challenge – Y results

Let’s search YHRD to find if haplotypes have been observed!



Paper challenge – Y results



Sample #1 | Sample #2 | Sample #3

Report for Sample #1

Sample Name: Bran

DYS576	DYS389I	DYS448	DYS389II	DYS19	DYS391	DYS481	DYS549	DYS533	DYS438	DYS437	DYS570	DYS635	DYS390	DYS439	DYS392	DYS643	DYS393	DYS458	DYS385	DYS456	YGATAH4
18	13	19	29	14	10	21	12	13	12	14	17	23	24	11	13	10	13	18	11, 15	16	12

+ Add feature to this Report

Worldwide

Observed

Found 1 match in 44,022 Haplotypes. This is approx. 1 match in 44,022 Haplotypes (95% CI: 7,901 - 1,738,777).

Expected

DL (Yfiler) Approx. 1 match in 1,744,693 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 22,012 Haplotypes (95% CI: 6,094 - 181,754)
Kappa Approx. 1 match in 183,173 Haplotypes

Eurasian - European - Western European (click to change)

Observed

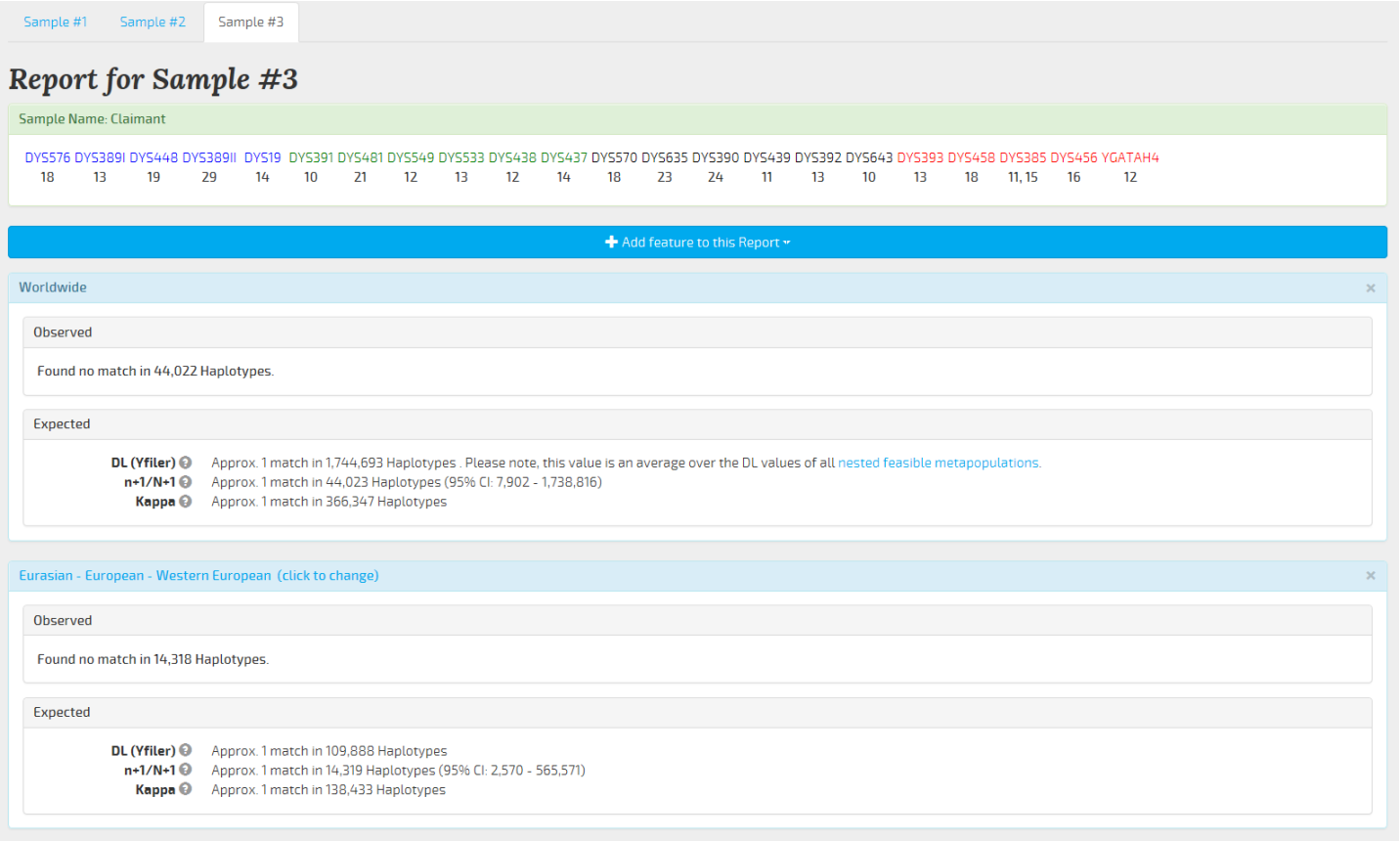
Found 1 match in 14,318 Haplotypes. This is approx. 1 match in 14,318 Haplotypes (95% CI: 2,570 - 565,531).

Expected

DL (Yfiler) Approx. 1 match in 109,888 Haplotypes
n+1/N+1 Approx. 1 match in 7,160 Haplotypes (95% CI: 1,982 - 59,117)
Kappa Approx. 1 match in 69,216 Haplotypes

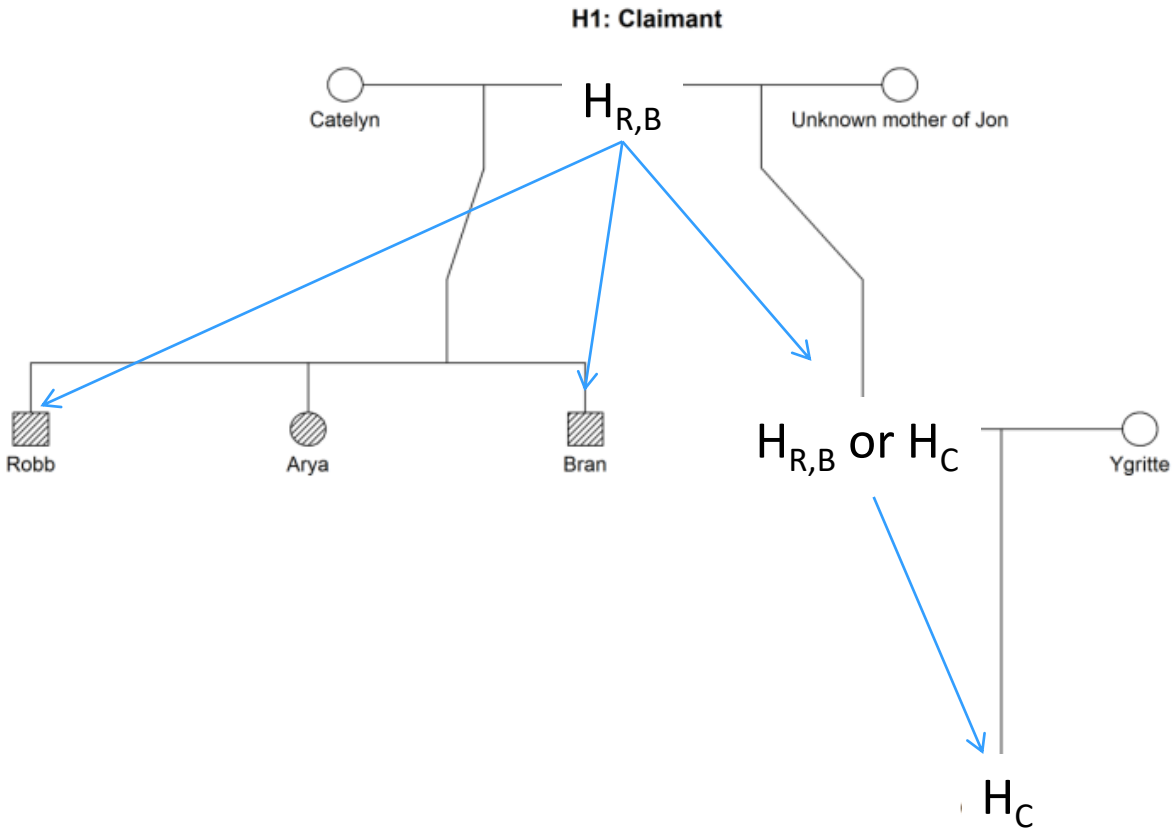
1 match in Western European metapopulation for $H_{R,B}$

Paper challenge – Y results



0 matches in Western European metapopulation for H_C

Paper challenge – Y results



Paper challenge – Y results

Let's get down to formulas!

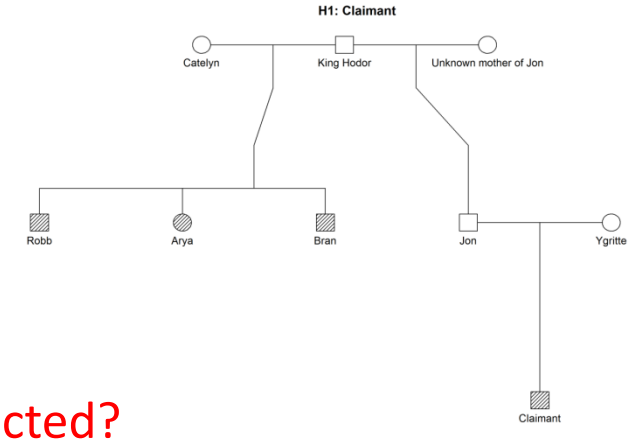
No mutation to Robb and Bran

Probability of $H_{B,R}$

$$Likelihood(Data|H_1) = \left[(1-\mu)^2 \cdot (2 \cdot \mu \cdot 0.5(1-\mu)) \cdot p_{H_{B,R}}\right] + \left[(1-\mu)^2 \cdot (\mu \cdot 0.5)^2 \cdot p_{H_C}\right]$$

Single mutation to Claimant

$$Likelihood(Data|H_2) = (1-\mu)^2 p_{H_{B,R}} \times p_{H_C}$$



Can be neglected?

$$LR = \frac{\left[(1-\mu)^2 \cdot (\mu(1-\mu)) \cdot p_{H_{B,R}}\right] + \left[(1-\mu)^2 \cdot (\mu \cdot 0.5)^2 \cdot p_{H_C}\right]}{(1-\mu)^2 p_{H_{B,R}} \times p_{H_C}} \approx \frac{\mu(1-\mu)}{p_{H_C}}$$

Paper challenge – Y results

We get LR=180 when using p_{H_C}=1/14,000. The frequency of $H_{R,B}$ does not affect results.

If we combine the automsomal and the Y LR:s we get

$LR_{combined}=100 \cdot 180=18,000$

A range of different LRs depending on frequency of haplotype!

Paper challenge – Simulations

Following the results, your lab decides to perform a simulation study to evaluate whether including further relatives could improve the results. We may assume we have already analyzed all available genetic markers. Evaluate the following scenarios,

a) Exhumation of the mother of the descendants (Catelyn).

b) Exhumation of the mother of the Claimant (Ygritte).

c) Exhumation of the King (Hodor).

d) Exhumation of all the above mentioned individuals.

Report the median of your simulations as well as the probability that the LR will exceed 1000 for each given scenario (assuming the Claimant is related to the descendants as proposed). Perform at least 1000 simulations.

What is your conclusion regarding the four possible scenarios? For instance, can we reach a more certain conclusion by only exhuming one extra individual (a-c) or do we need to exhume all three of them (d).

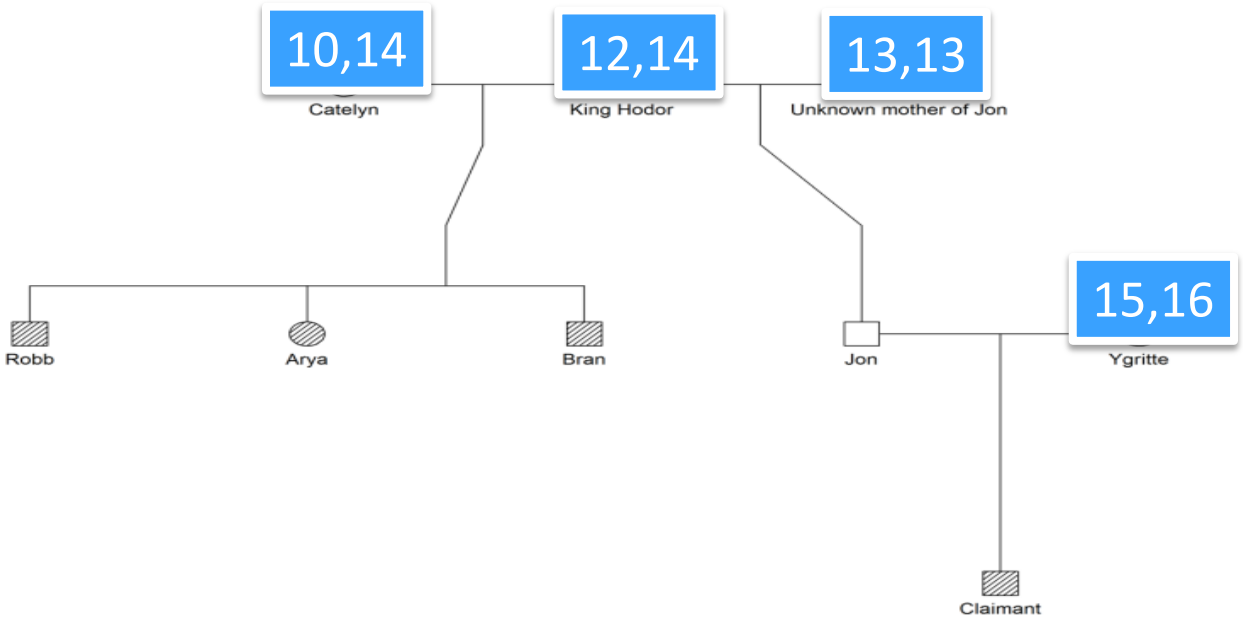
Paper challenge – Simulations

Simulations work by considering the pedigrees (H1 and H2); It starts by randomly drawing genotypes for the founders of the pedigree and subsequently randomly «creating» the children. In our case we were only interested in H1, i.e. in essence the true positive rate. We would like to simulate 1000 (or more) of H1 and count the number of times the LR exceeds 1000.

The next slides try to illustrate how a single simulation is conducted.

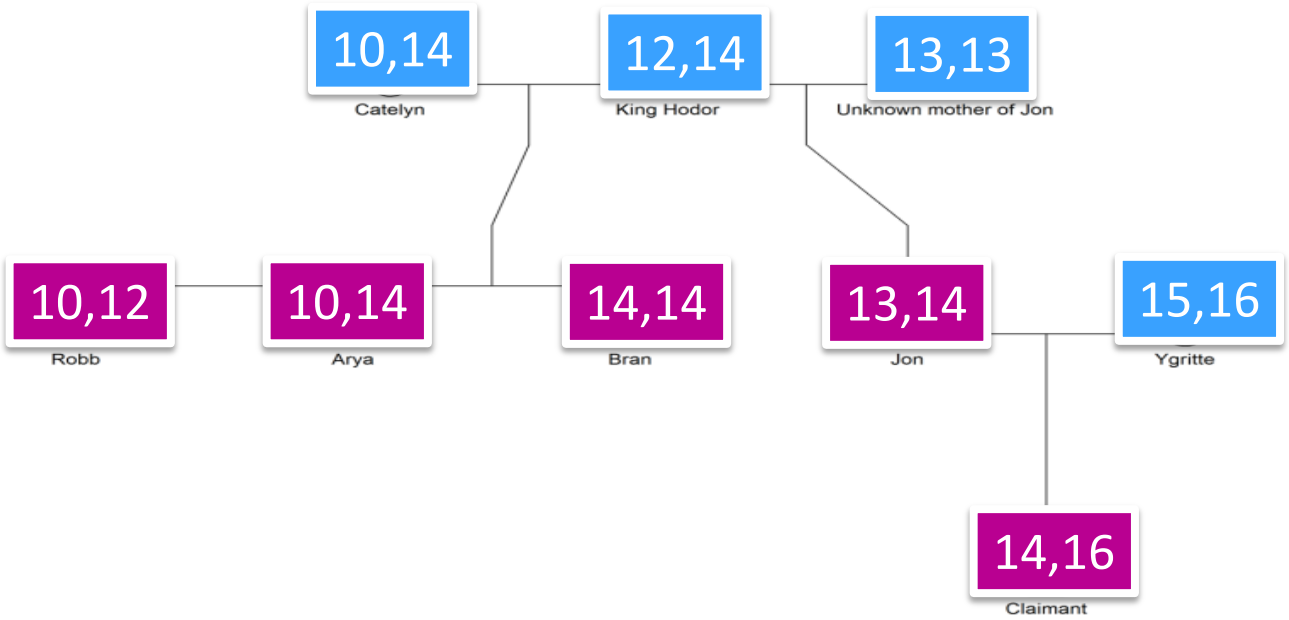
Paper challenge – Simulations

1. Simulate founders (randomly draw genotypes from the population frequencies). Here illustrated for a single marker.



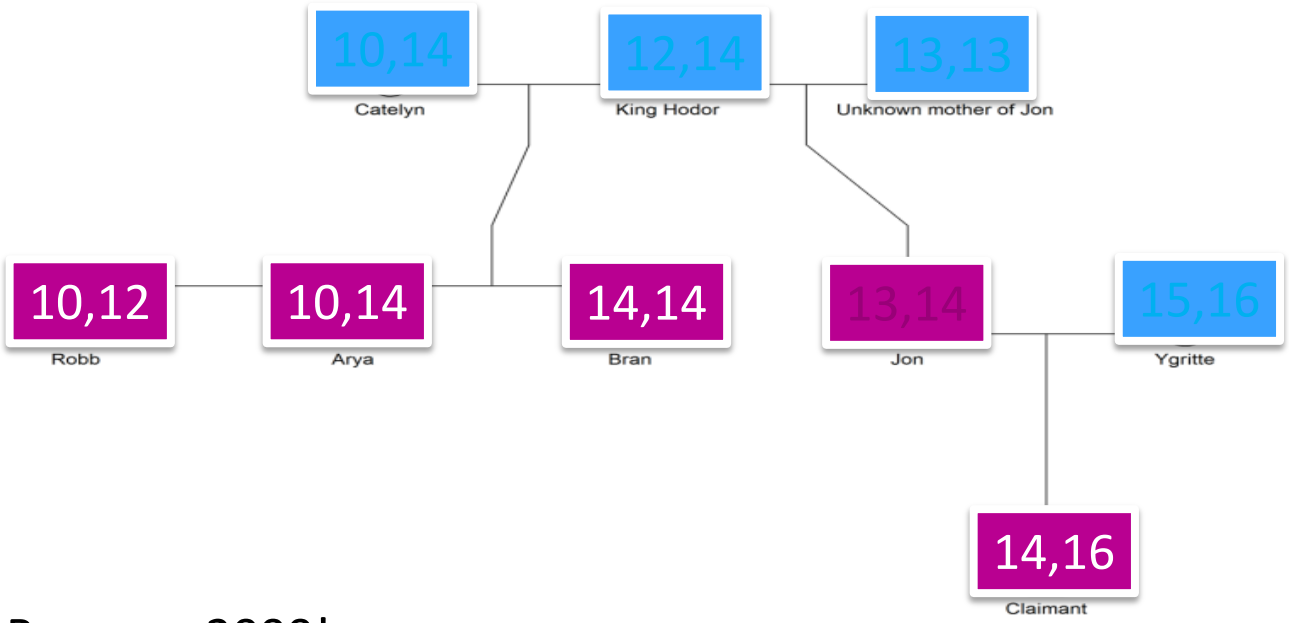
Paper challenge – Simulations

2. Simulate non-founders. Using laws of inheritance we know that there is a 50/50 chance for parents to transmit either of their alleles. This is basically just as flipping a coin and deciding what allele to pick for each child.



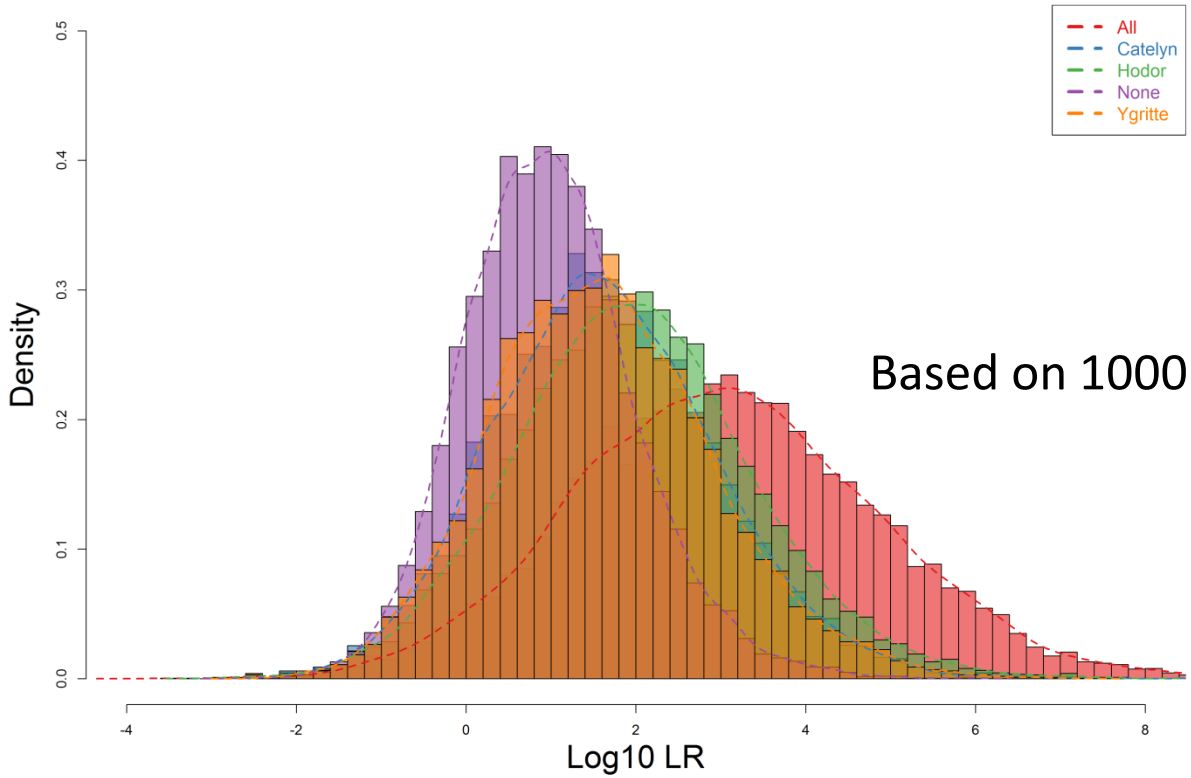
Paper challenge – Simulations

3. Finally, we compute the LR based on a subset of persons, starting with the once we have, i.e. Robb, Arya, Bran and the Claimant. We then add the exhumed persons one by one. We repeat steps 1-3 for all genetic markers and get a combined LR.



$LR_{combined}$=2000!

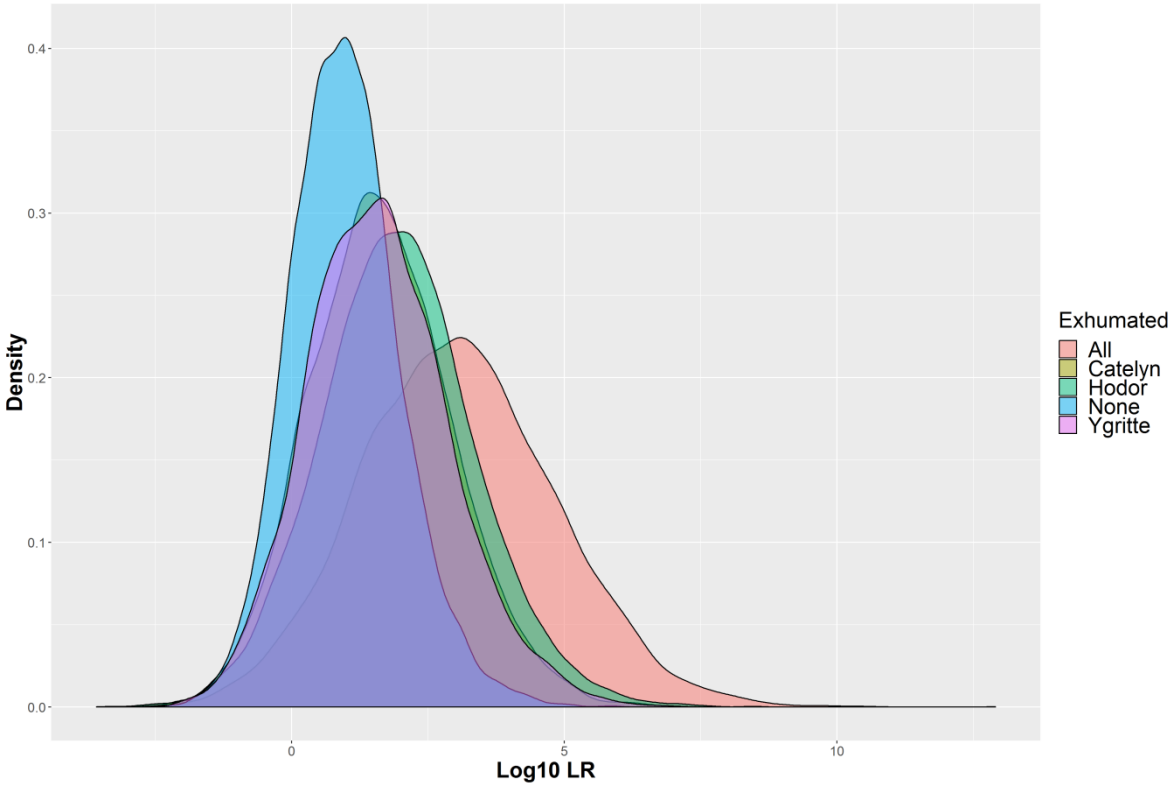
Paper challenge – Simulations



Based on 1000 simulations

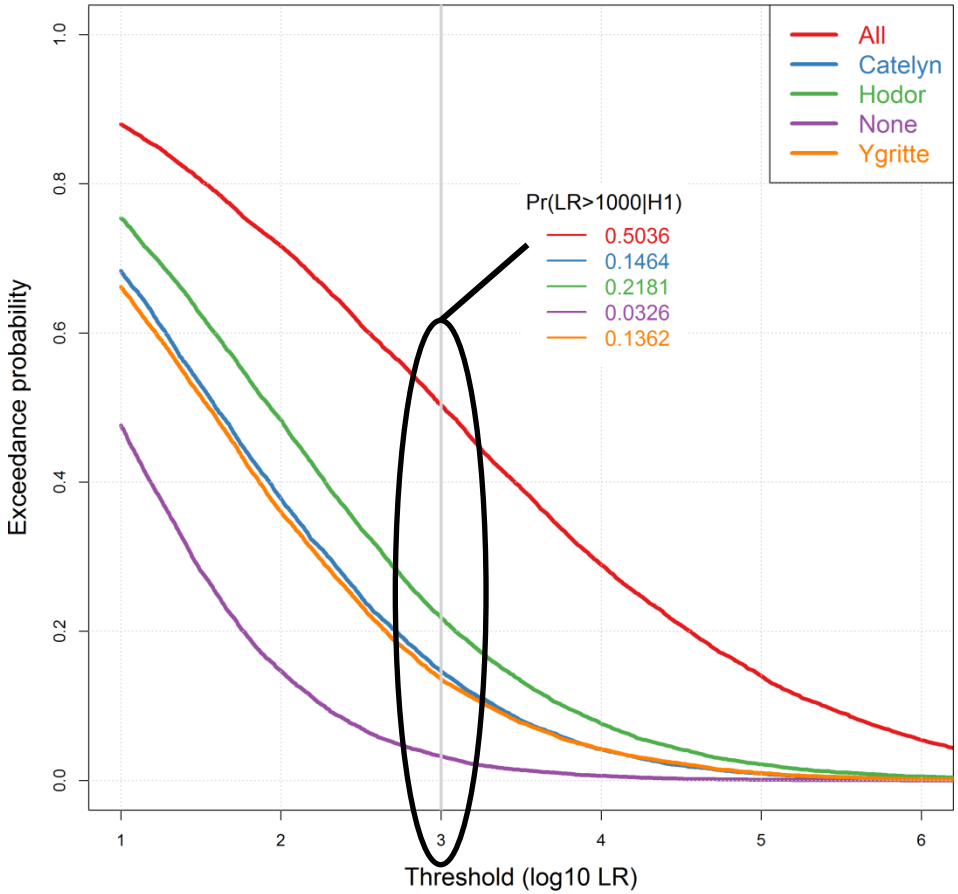
Interpretation: The distributions overlap, there is most information in exhuming all and, as expected, least information if none is exhumed.

Paper challenge – Simulations



Alternative representation (NB! Different colors)

Paper challenge – Simulations



16 labs did the second part!
10 labs got the results presented here
6 labs got the same deviating results

Interpretation: There is 50% chance to obtain LR>1000 if we exhume all

Paper challenge – Summary

- Complex kinship case.
- Combining autosomal and Y marker results.
- Accessing YHRD.
- Constructing pedigrees.
- Consistent results.
- Simulations.

PROFICIENCY TEST

Summary

- 47 participants
  - 36 did paper challenge
  - 43 did wet exercise
- 13 labs do sequencing
- 33 labs use Familias, 16 use Excel, 4 DNA-view
- Consistent results on the paper challenge (Autosomal part)
- More diverging results on the Y markers
  - Some labs did not realize the inconsistency
- Wet exercise results – Some typing differences have been addressed. Some labs report data also for SNP markers

Wet exercise

Your local police are investigating a case in which a young woman, Laura Dean, was found dead in her apartment. There are no signs of violence. Later, and in the same area, a lifeless baby is found in a container wrapped in a towel identical to towels in Laura Dean's apartment.

The police suspect that there is a connection between the two cases and ask for a DNA-test of the two bodies in order to decide whether Laura Dean is the mother of the child.

a) Perform a maternity test to clarify the maternal relationship between Laura Dean and the child (the samples are labelled Woman and Child).

Wet exercise

Given the DNA results from the DNA- tests, the police conclude that Laura Dean is the biological mother of the child. During the investigation, the police are approached by a woman claiming that her son, vanished since the case became public, very likely is the father of the baby. The police decide to test the claimed relation between this woman and the child.

b) Perform a test to clarify the claimed biological relationship between this woman (sample labelled Grandmother) and the child. Consider the maternity as confirmed and include the DNA-data from mother and child in task a).

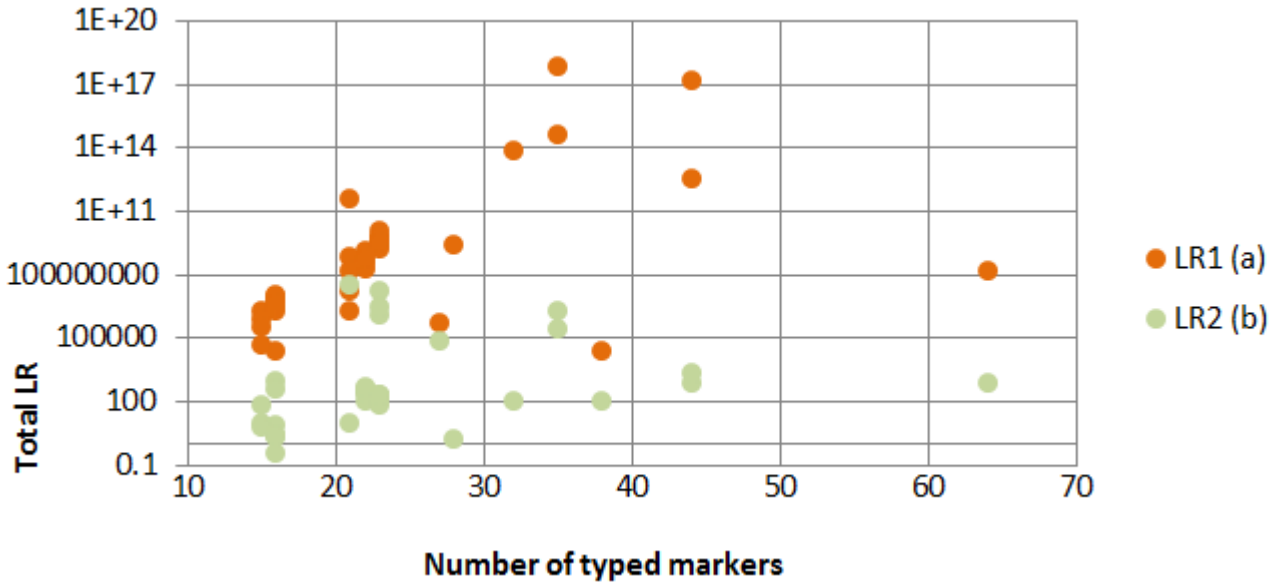
Report the likelihood ratios (LR) for the individual genetic markers included in the tests as well as the combined LR. State which frequency database you have used for the calculations. Part b) is optional while a) is mandatory to obtain the certificate.

Wet exercise

Samples consisted of diluted blood (EDTA) on FTA cards. Two labs reported problems with amplications and received undiluted blood instead.

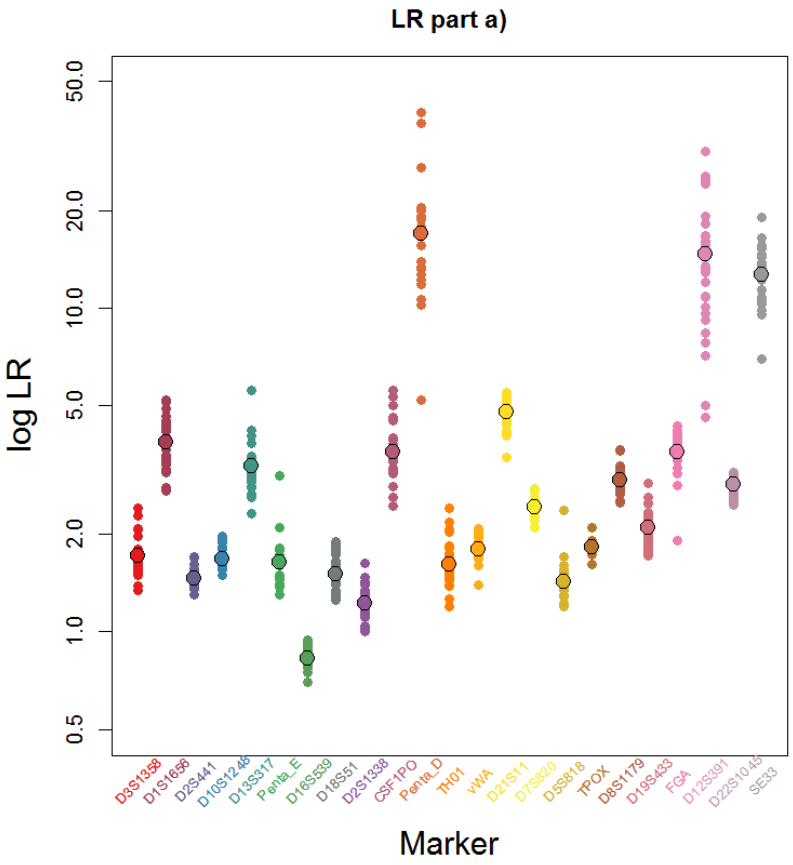
a) A plain maternity test (43 labs completed)

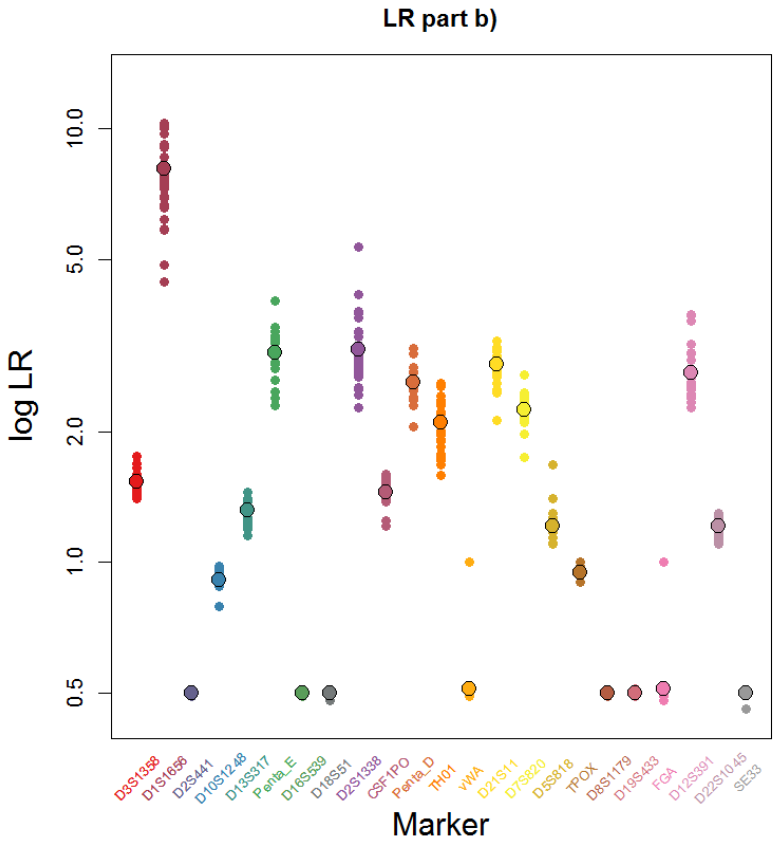
b) A deficient paternity case (alleged grandmother) (39 labs completed)



Wet exercise

Per marker results





Wet exercise

More in the Excel summary.

For the wet exercise, some lab's results have been highlighted (red) which indicates a result that deviates. Certificates will still be issued.

ESWG - Proficiency test 2018

Daniel Kling – Head Organizer

rmdakl@ous-hf.no