Is the Factor of 10 Still Applicable Today?

Simone Gittelson, PhD

John Buckleton, PhD

Disclaimer

Points of view in this presentation are ours and do not necessarily represent the official position or policies of the National Institute of Standards and Technology or the U.S. Department of Justice.

Certain commercial products are named in order to specify experimental procedures as completely as possible. In no case does such an identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that any of these products are necessarily the best available for the purpose.

Acknowledgements

We thank **Tamyra Moretti** and **Anthony Onorato** for their collaboration and sharing the FBI data.

We thank **Bruce Budowle** and **Bruce Weir** for their insightful comments on this study.

We thank **Alex Thiéry**, **James Curran**, **Jérôme Goudet** and **Bruce Weir** for data collection and analysis.

This work was supported in part by grant number 2011-DN-BX-K541 from the U.S. National Institute of Justice.

Contents

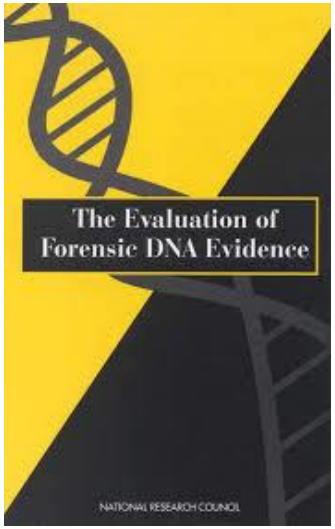
Factor of 10

Methods

Results

Conclusions

Factor of 10



NRC II, Overview, page 39:

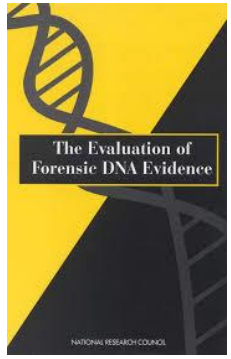
"(…) we examined empirical data from the comparison of different subpopulations and of subpopulations within the whole. The empirical studies show that the differences between the frequencies of the individual profiles estimated (…) from different adequate subpopulation databases (…) are **within a factor of about 10 of each other**,"

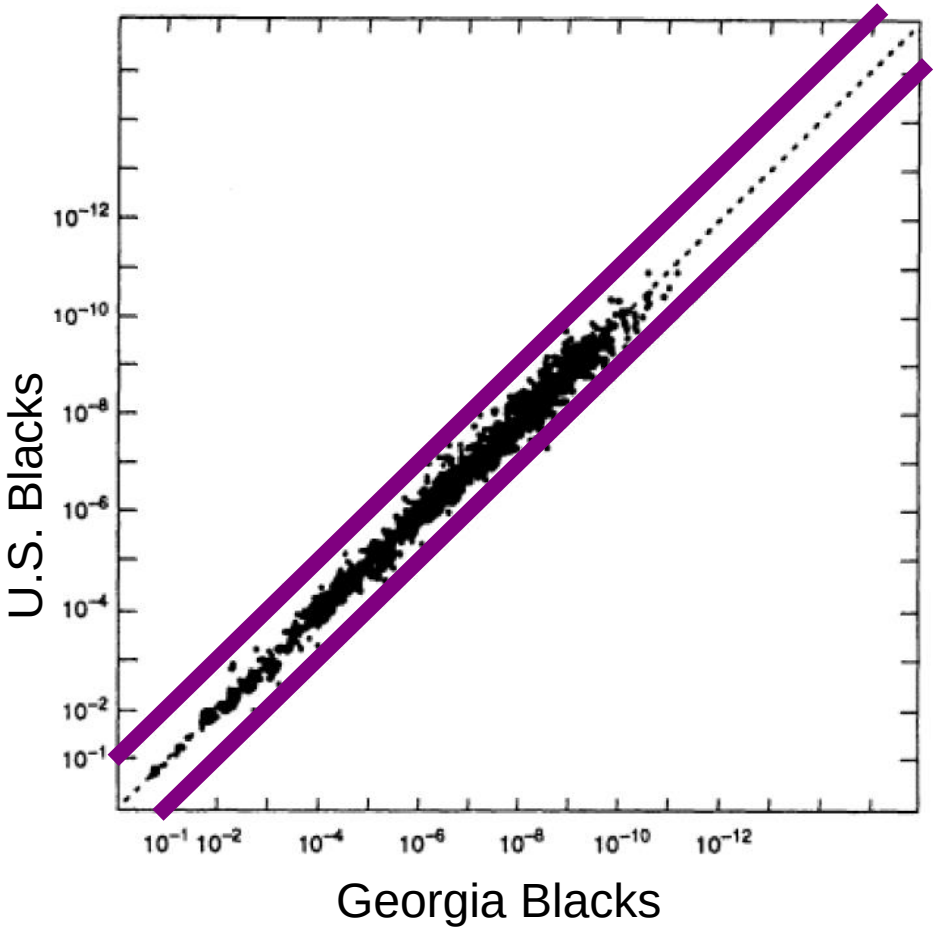
(emphasis added by me)

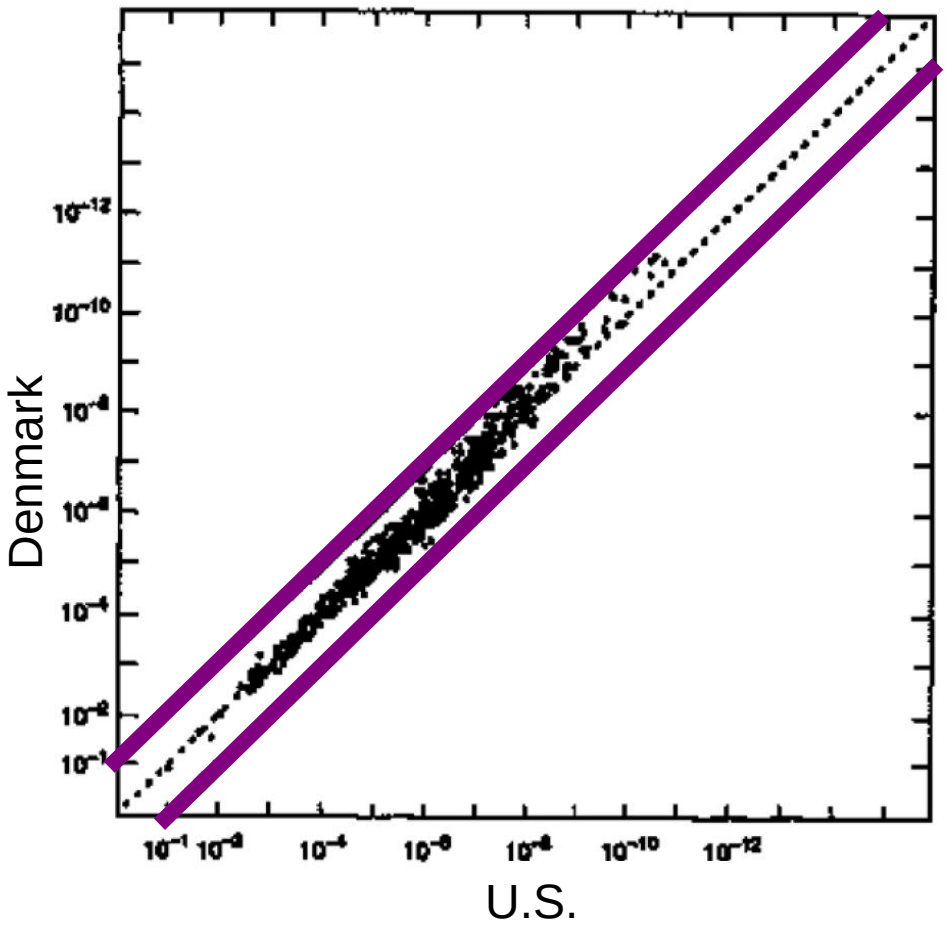
National Research Council Committee on DNA Forensic Science. The Evaluation of Forensic DNA Evidence. Washington, D.C.: National Academy Press; 1996.

Factor of 10

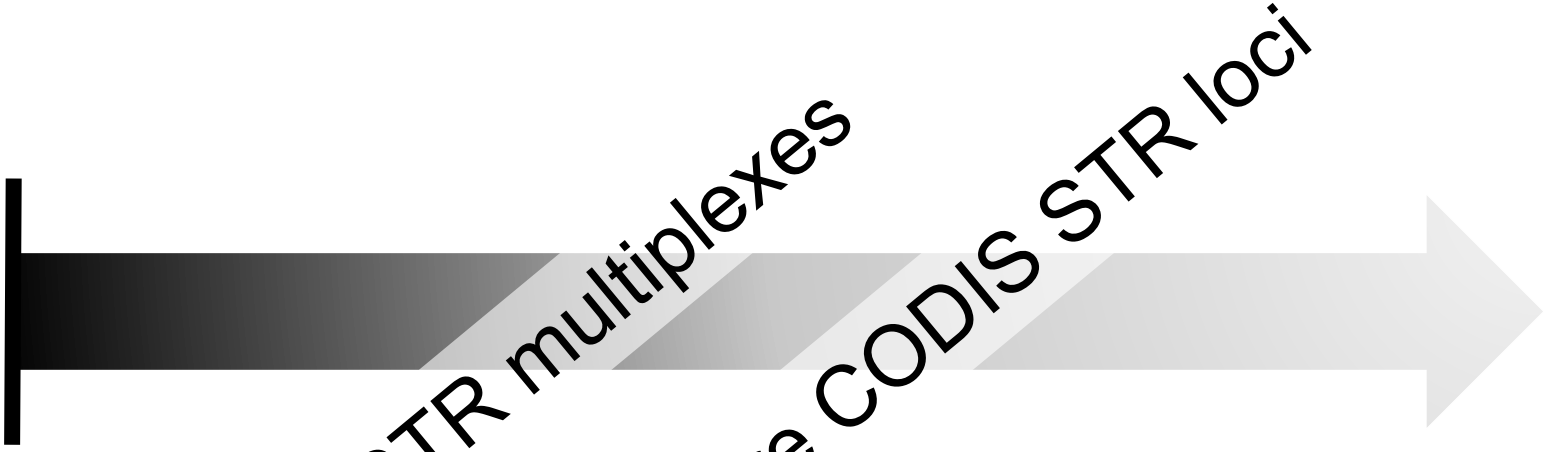
NRC II, Statistical Issues, pages 150 and 152:

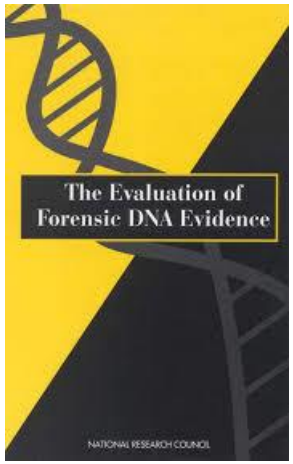






Factor of 10





Does the factor of 10 still apply?

Contents

Factor of 10

Methods

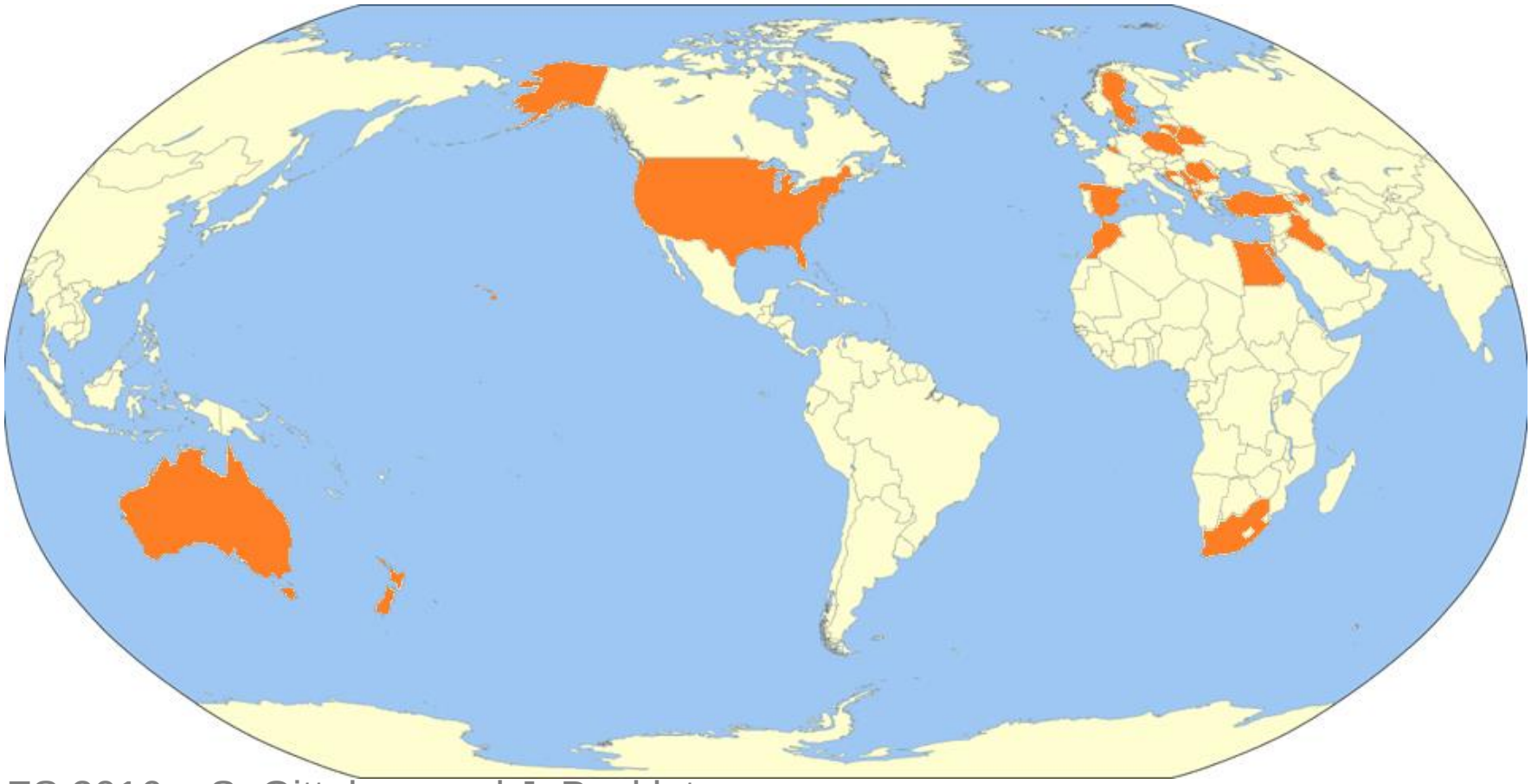
Results

Conclusions

Methods

AmpFℓSTR® Identifiler® PCR Amplification Kit
AmpFℓSTR® Identifiler® Plus PCR Amplification Kit

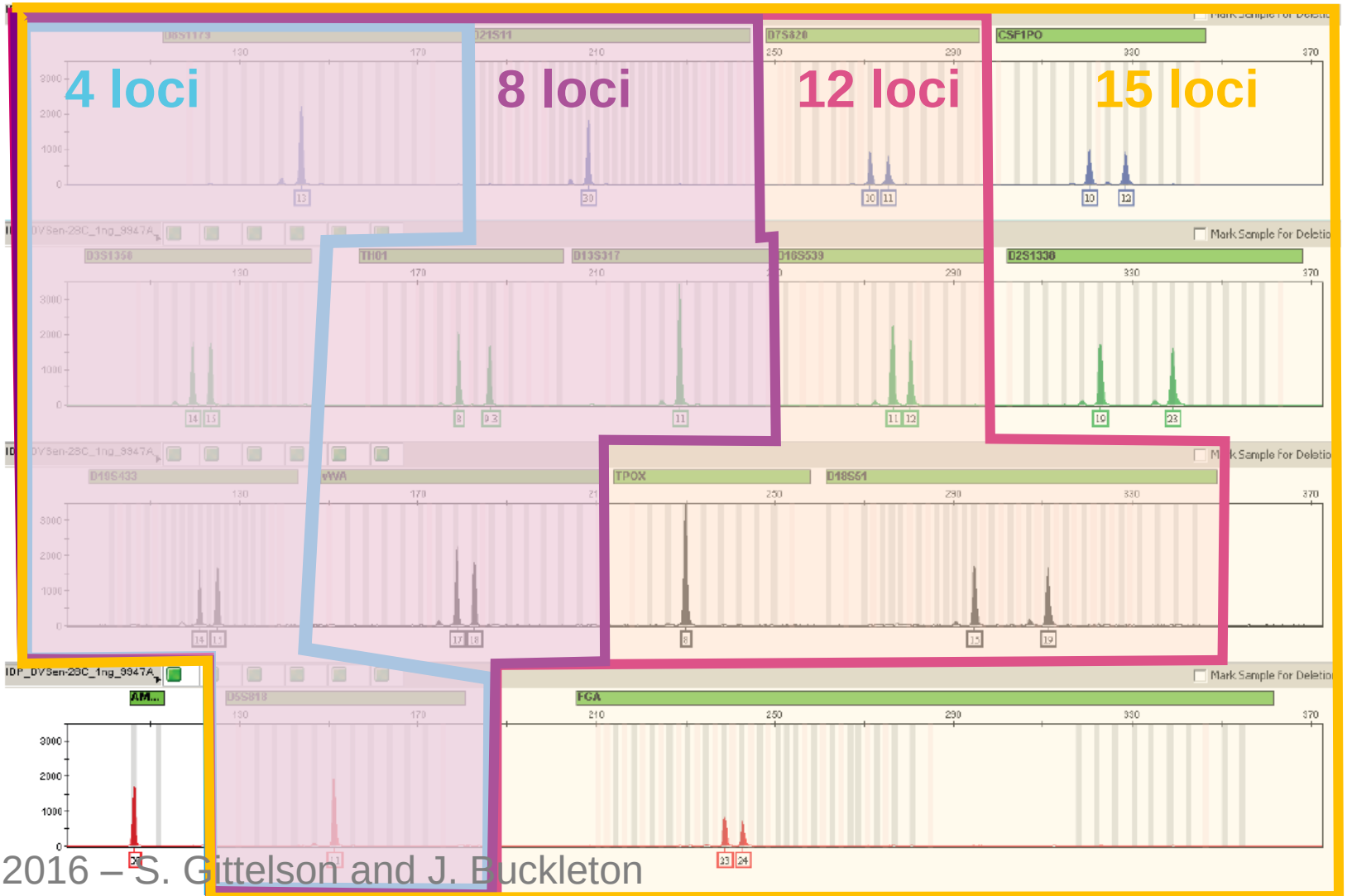
Caucasians:



Methods

AmpFℓSTR® Identifiler® PCR Amplification Kit

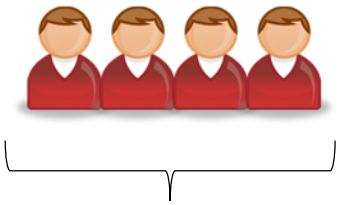
AmpFℓSTR® Identifiler® Plus PCR Amplification Kit



Methods

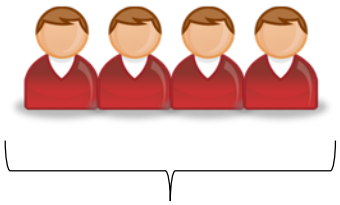
Simulations of:

4 loci



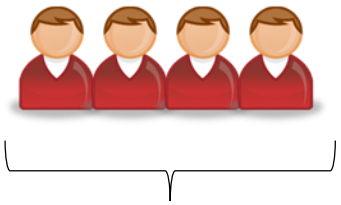
100 profiles

8 loci



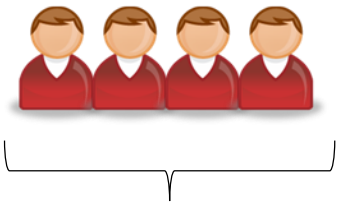
100 profiles

12 loci



100 profiles

15 loci



100 profiles

= 400 profiles per subpopulation

Profile Probability

NRC II, Recommendation 4.1.:

homozygote	heterozygote
$Fp_i + (1 - F)p_i^2$	$2p_i p_j$

→ takes into account inbreeding, but not co-ancestry

Contents

Factor of 10

Methods

Results

Conclusions

Results

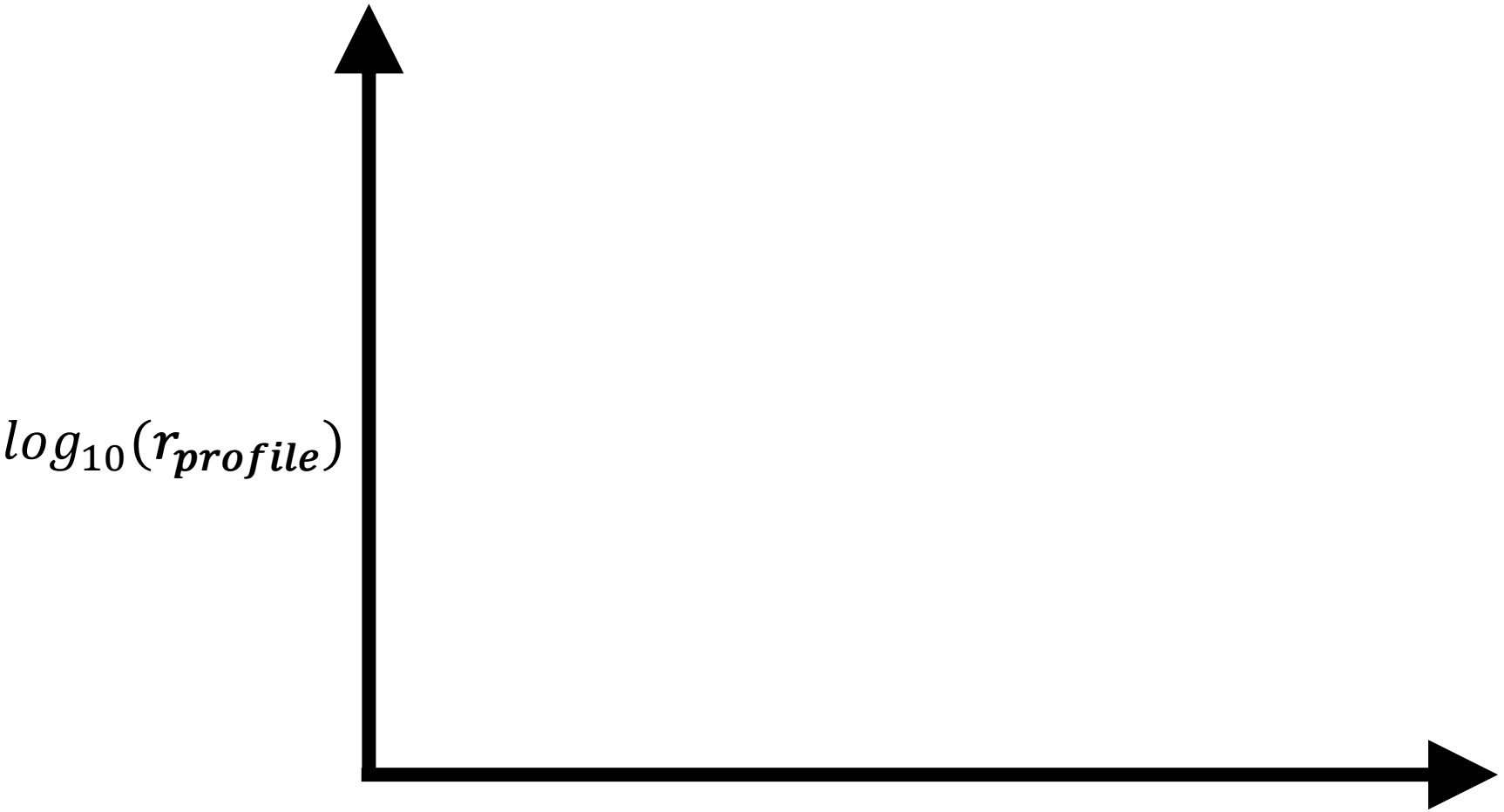
$$r_{profile} = \frac{\text{profile probability using non-cognate database}}{\text{profile probability using cognate database}}$$

if $r_{profile} = 1$: non-cognate profile probability is the same as the cognate profile probability

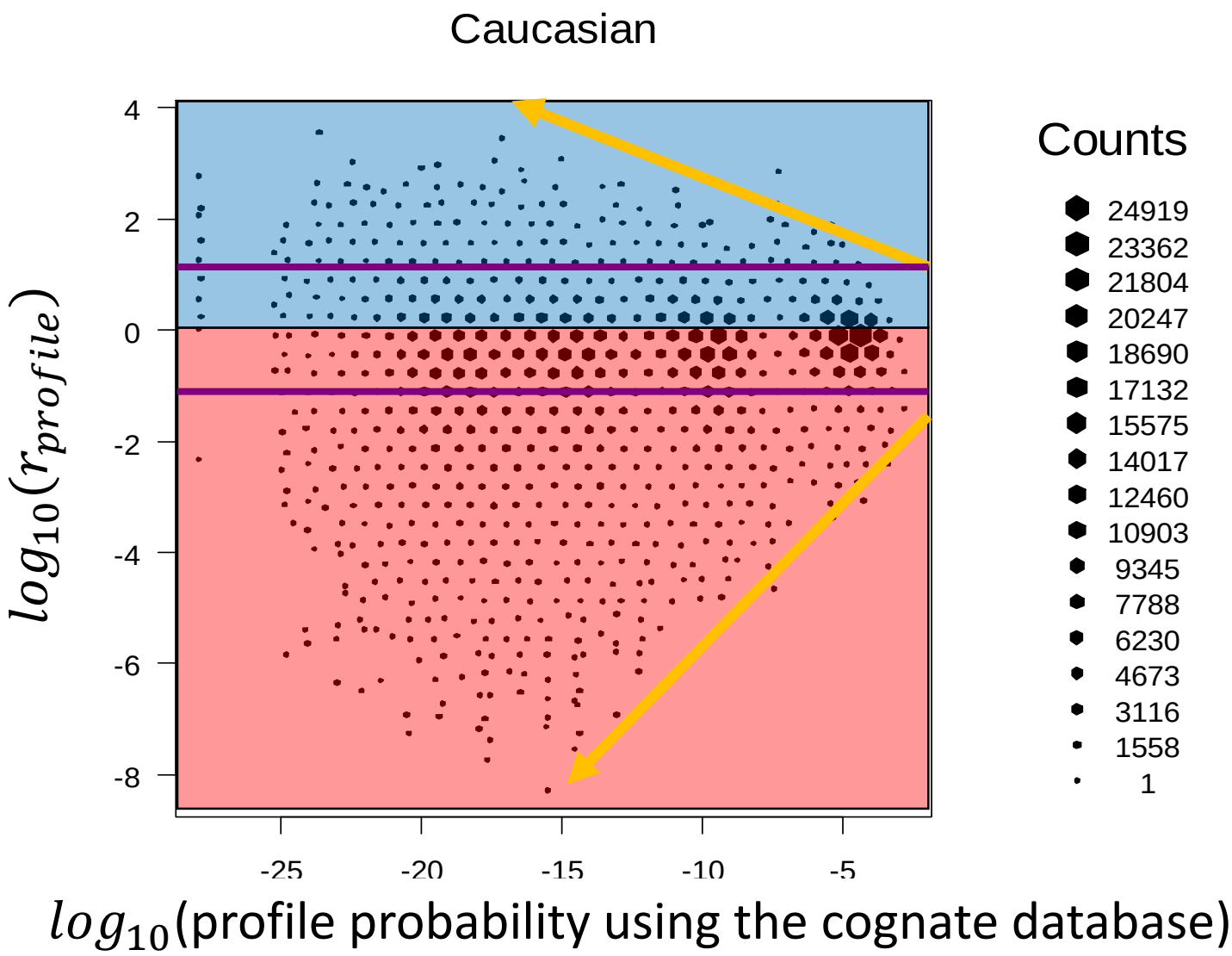
if $r_{profile} > 1$: the profile is **more common** in the non-cognate database than in the cognate database

if $r_{profile} < 1$: the profile is **rarer** in the non-cognate database than in the cognate database

Results

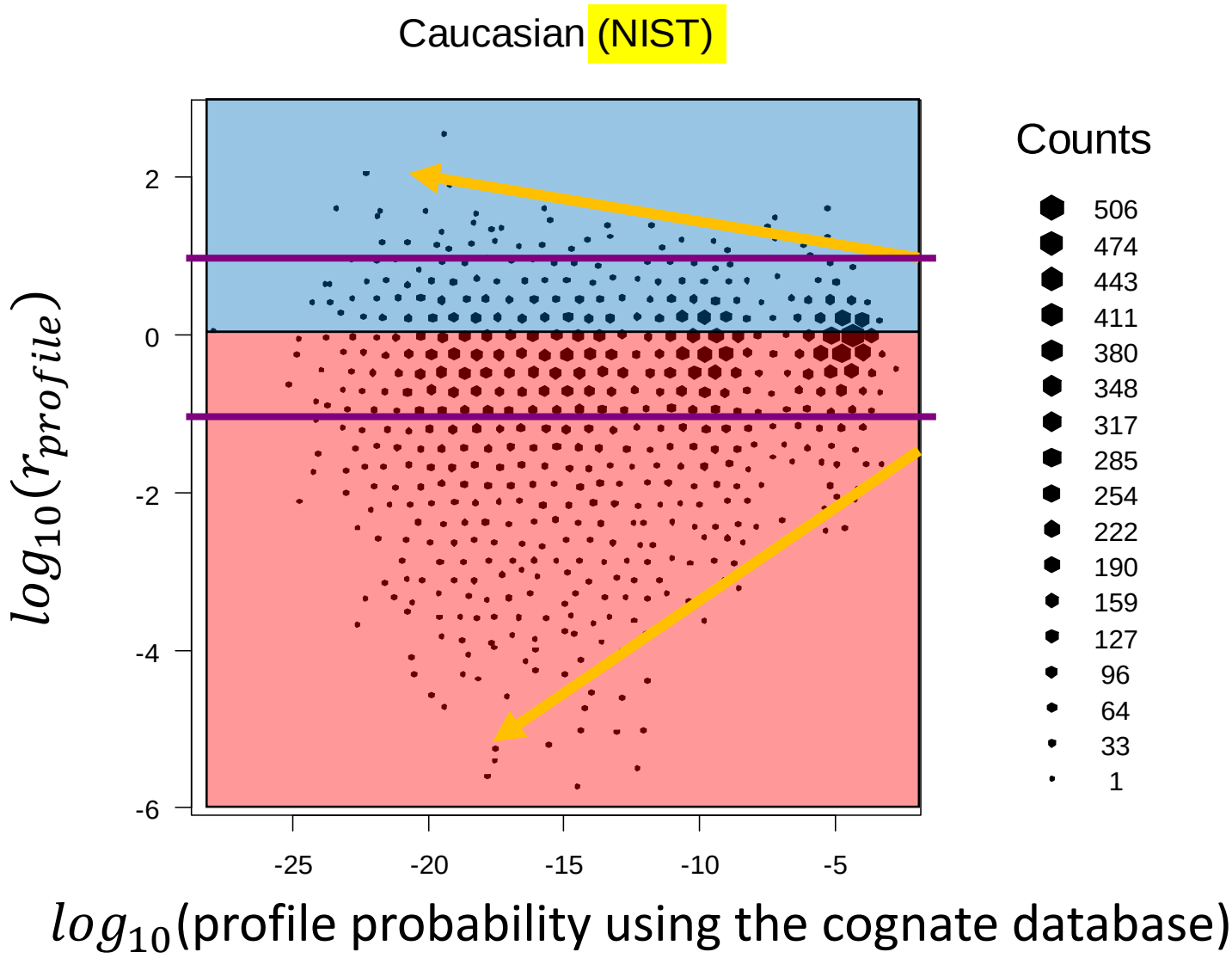

$log_{10}(r_{profile})$
log_{10}(profile probability using the cognate database)

Results


Caucasian
4
2
0
-2
-4
-6
-8
-25
-20
-15
-10
-5
$log_{10}(r_{profile})$
log_{10}(profile probability using the cognate database)
Counts
24919
23362
21804
20247
18690
17132
15575
14017
12460
10903
9345
7788
6230
4673
3116
1558
1

Results

Hill CR, Duewer DL, Kline MC, Coble MD, Butler JM. US population data for 29 autosomal STR loci. *Forensic Sci. Int.: Genet.* 2014; 7: e82-e83.



Match Probability

NRC II, Recommendation 4.2.:

homozygote

$$\frac{(2\theta+(1-\theta)p_i)(3\theta+(1-\theta)p_i)}{(1+\theta)(1+2\theta)}$$

heterozygote

$$2\,\frac{(\theta+(1-\theta)p_i)(\theta+(1-\theta)p_j)}{(1+\theta)(1+2\theta)}$$

→ takes into account inbreeding and co-ancestry

Balding DJ, Nichols RA. DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Sci. Int.* 1994; 64: 125-140.

Results

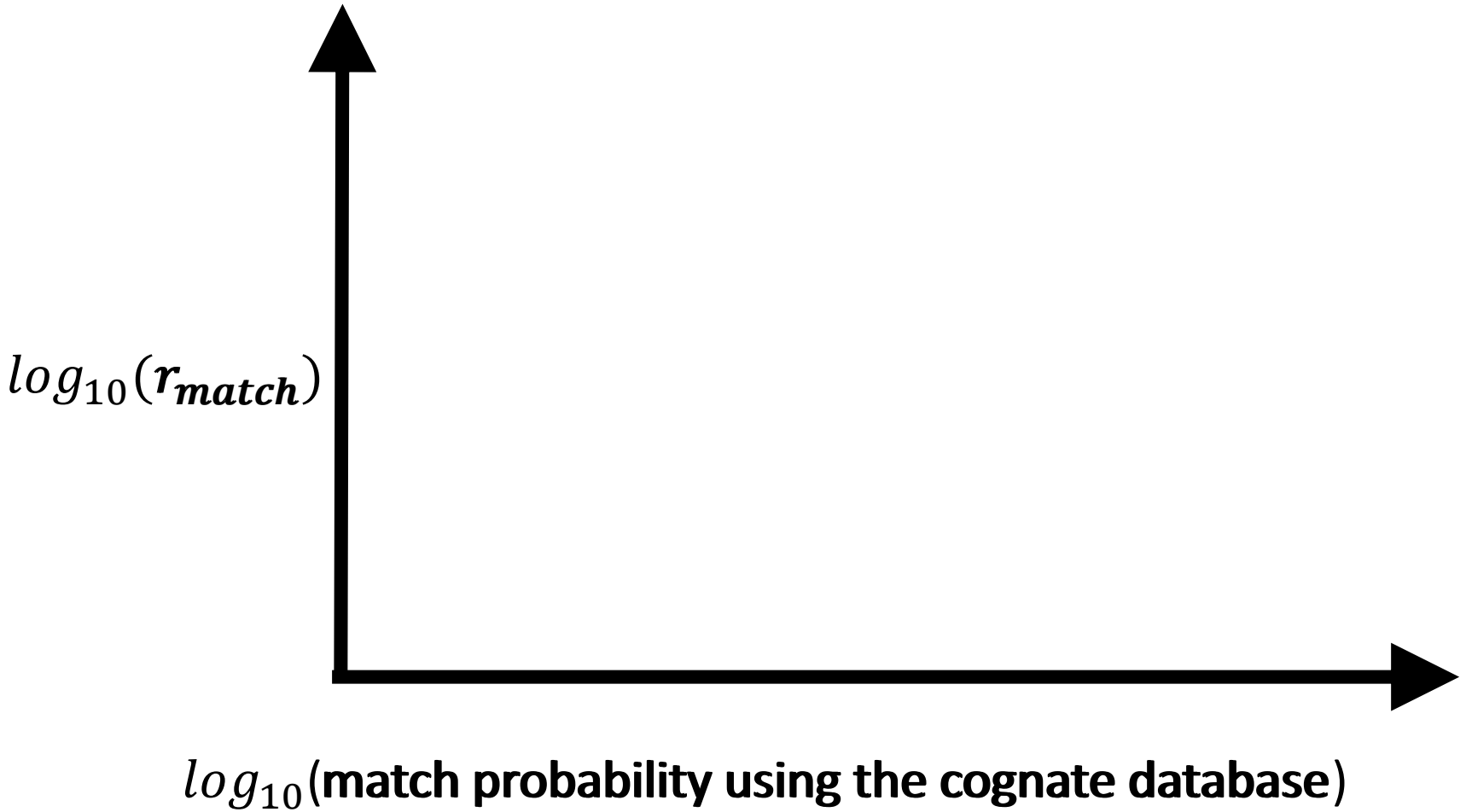
$$r_{match} = \frac{\text{match probability using non-cognate database}}{\text{match probability using cognate database}}$$

if $r_{match} = 1$: non-cognate match probability is the same as the cognate match probability

if $r_{match} > 1$: a match is **more common** in the non-cognate database than in the cognate database

if $r_{match} < 1$: a match is **rarer** in the non-cognate database than in the cognate database

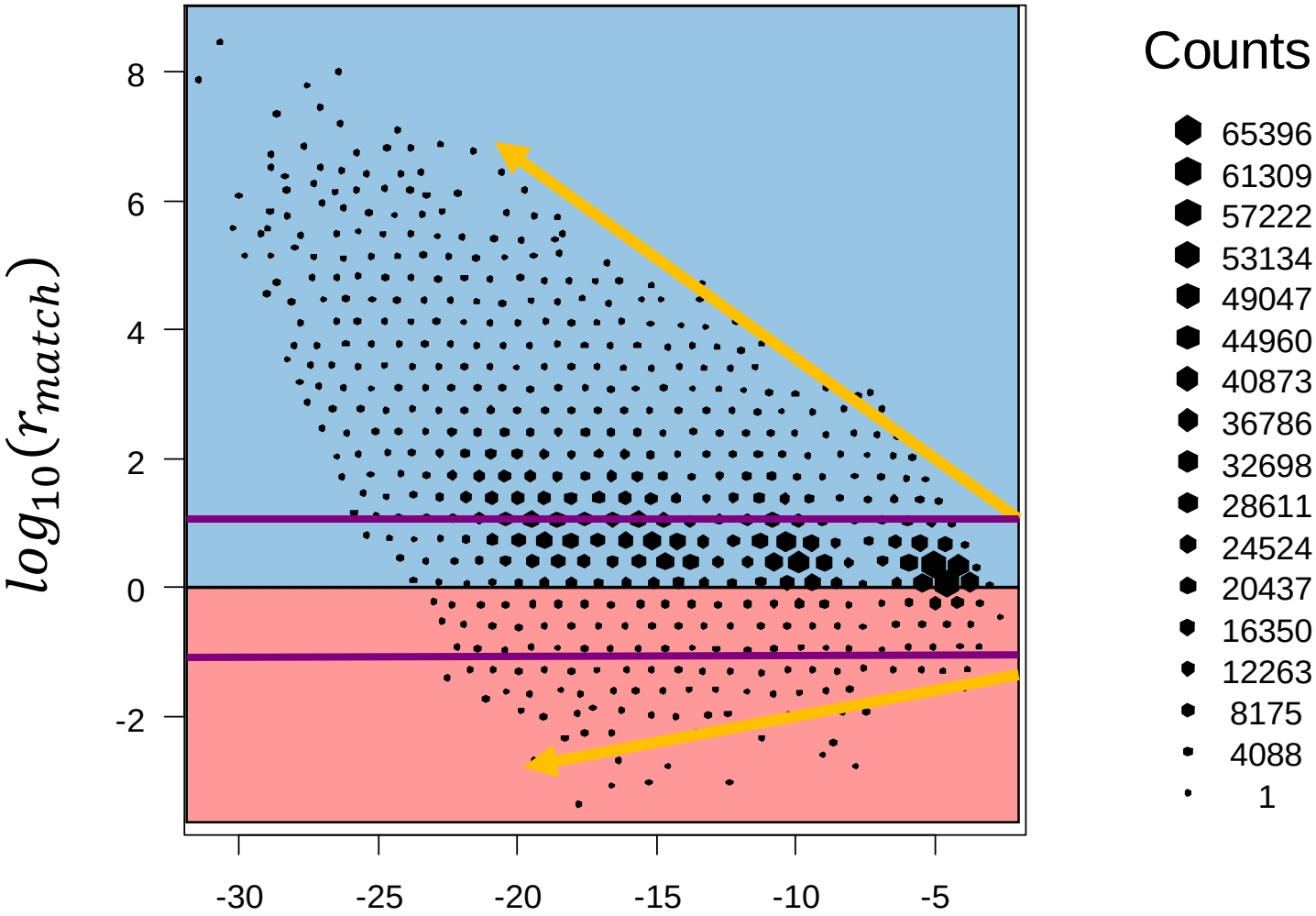
Results


$log_{10}(r_{match})$
log_{10}(match probability using the cognate database)

Results

Hill CR, Duewer DL, Kline MC, Coble MD, Butler JM. US population data for 29 autosomal STR loci. *Forensic Sci. Int.: Genet.* 2014; 7: e82-e83.

Caucasian (NIST)



Contents

Factor of 10

Methods

Results

Conclusions

Conclusions

Variation can be greater than a factor of 10.

Non-cognate **profile probabilities** tend to be rarer than the cognate profile probabilities.
→ **non-conservative**

Non-cognate **match probabilities** tend to be more common than the cognate match probabilities.
→ **conservative**

Thank you very much for your attention!