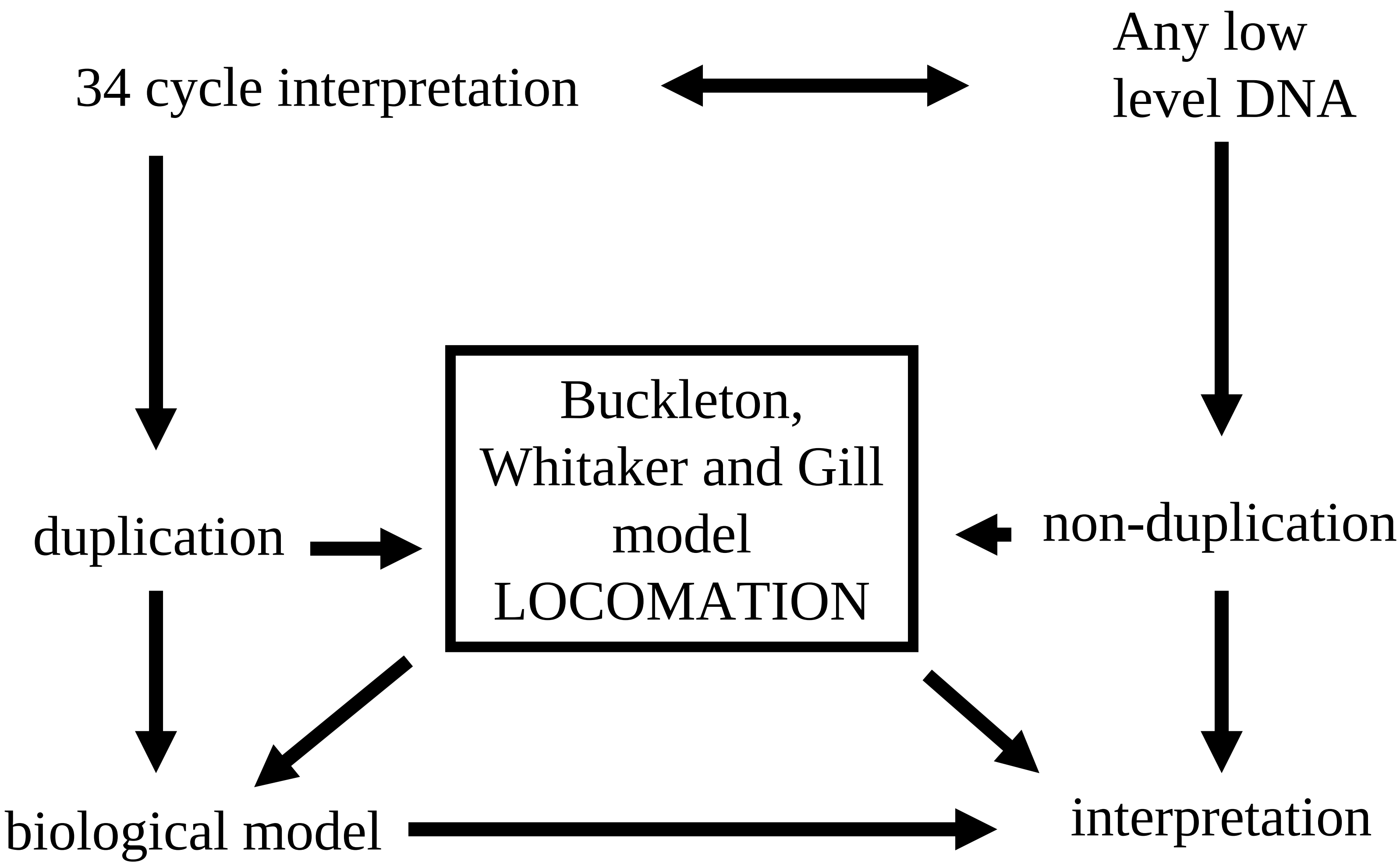
The Low Copy Number (LCN) STR interpretation strategy

Peter Gill and John Buckleton


34 cycle interpretation
Any low
level DNA
Buckleton,
Whitaker and Gill
model
LOCOMATION
duplication
non-duplication
biological model
interpretation

LCN STRATEGY

- The duplication strategy might be regarded as *ad hoc*.
- However, a statistical model was developed to underpin our rationale and to demonstrate the conservative nature of the method used.
- But - we *do* need to think in a very different way about interpretation

Models to interpret LCN profiles

- Probability of contamination
  - Calculated from probability of contaminant in a negative control
- Probability of allele drop-out
  - Predicted by the size of the associated allele
- Probability of stutter
  - Also predicted by the size of the associated allele

Procedure to estimate the LR

- Nomenclature:
- Replicates
- Say, R_1 = a R_2 = ab
- $Pr(E|H_p)$ is the probability of the evidence *if* the profile is the suspect's
- $Pr(E|H_d)$ is the probability of the evidence *if* the profile is from someone else

$$LR = \frac{p(R_1, R_2, ... | Hp)}{p(R_1, R_2, ... | Hd)}$$

$$= \frac{p(R_1, R_2, ... | Hp)}{\sum_j p(R_1, R_2, ... | M_j, Hd) p(M_j | Hd)}$$

Assume replicate 1 and replicate 2 etc are independent?

Once M_j is specified we don't need *Hd*.

$$= \frac{\prod_i p(R_i | H_p)}{\sum_j \prod_i p(R_i | M_j) p(M_j)}$$

Consider one replicate profile is *ab*
suspect is *ab*

Explanation of the evidence (H_p)

- We condition on the suspect who is *ab*
- If R_1 is really from the suspect how is the evidence explained?
- $R_1 = ab$ - explanation - no drop out of allele *a*, no drop out of allele *b*, no drop in

$$p(R_1 \mid Hp) = \overline{D}\,\overline{D}\,\overline{S}p$$

Explanation of the evidence (H_d)

- In LCN there are a lot of possible‘true offender” profiles. We call these *M*j.
- There is no need for restriction if you have a computer but there is a need if you do it by hand.
- I think in this case we could have *M*j=*ab*, *aa*, *bb*

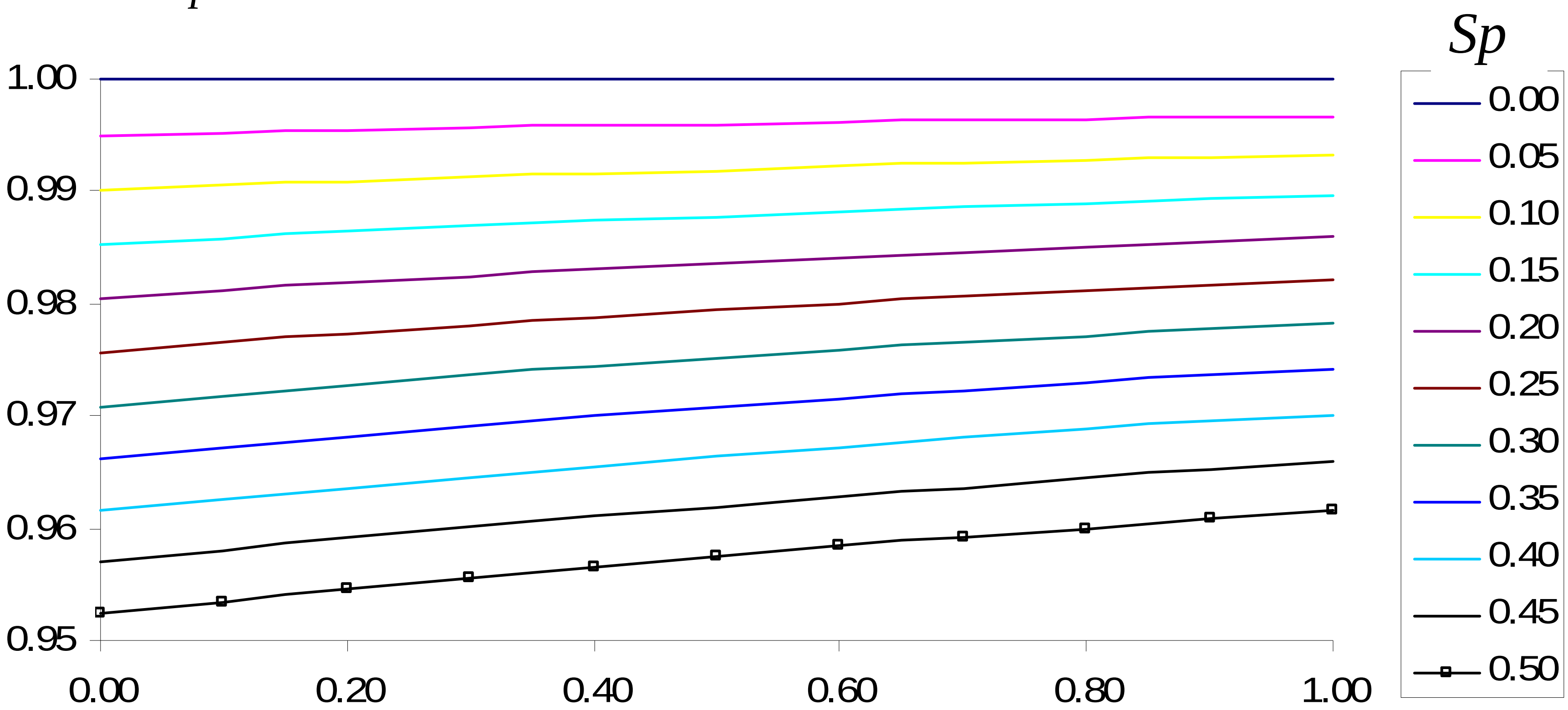
*M*j	*Pr*(*M*j)	$R_1 = ab$
ab	$2 f_a f_b$	$\overline{D}\,\overline{D}\,\overline{S}\,p$
aa	f_a^2	$\overline{D}\,?\,Spf_b$
bb	f_b^2	$\overline{D}Spf_a$
$2 f_a f_b \overline{D}^2 \overline{S} p + f_a^2 \overline{D} Spf_b + f_b^2 \overline{D} Spf_a$		

$$LR = \frac{\overline{D}^2 \overline{S} p}{2 f_a f_b \overline{D}^2 \overline{S} p + f_a^2 \overline{D} Spf_b + f_b^2 \overline{D} Spf_a}$$

$$= \frac{\overline{D}\overline{S} p}{2 f_a f_b \overline{D}\overline{S} p + f_a^2 Spf_b + f_b^2 Spf_a}$$

$$= \frac{1}{2 f_a f_b \left(1 + \frac{Sp\left(f_a + f_b \right)}{2\overline{D}\overline{S}p} \right)}$$

$$\frac{1}{1+\frac{Sp\left(f_a+f_b\right)}{2\overline{D}\,\overline{Sp}}}$$



Replicate profile is *a* suspect is *aa*

Explanation of the evidence (H_p)

- We condition on the suspect who is *aa*
- If R_1 is really from the suspect how is the evidence explained?
- R_1= *aa* - explanation - no drop out of allele *a*, a no drop in

$$p(R_1 \mid Hp) = \overline{D}\,\overline{Sp}$$

Explanation of the evidence (H_d)

- I think in this case we could have Mj=*ax*, *aa*

Mj	$Pr(M$j$)$	$R_1 = a$
ax	$2 f_a (1 - f_a)$	$\overline{D}D\overline{Sp}$
aa	f_a^2	$\overline{D}\overline{Sp}$
$2 f_a (1 - f_a) \overline{D}\overline{Sp} + f_a^2 \overline{D}$		

$$LR = \frac{\overline{D}\,\overline{Sp}}{2 f_a (1 - f_a)\overline{D} D \overline{Sp} + f_a^{2}\,\overline{D}\,\overline{Sp}}$$

$$= \frac{1}{2 f_a (1 - f_a) D + f_a^{2}}$$

$$= \frac{1}{2 f_a \left[(1 - f_a) D + \frac{f_a}{2} \right]}$$

$$= \frac{1}{2 f_a (1 - \frac{f_a}{2})}$$

Worst scenario

$D = 1$

$$LR = \frac{\overline{D}\overline{Sp}}{2f_a(1-f_a)\overline{D}D\overline{Sp} + f_a^2\,\overline{D}\,\overline{Sp}}$$

$$= \frac{1}{2f_a(1-f_a)D + f_a^2}$$

$$= \frac{1}{2\Pr(a\,|\,aa)(1-\theta)(1-f_a)D + \Pr(a\,|\,aa)\Pr(a\,|\,aaa)}$$

$$= \frac{(1+\theta)(1+2\theta)}{\left\{2\theta+(1-\theta)f_a\right\}\left[3\theta+(1-\theta)\left[f_a+2D(1-f_a)\right]\right]}$$

Eq 1

ESR

$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{2\theta+(1-\theta)f_a\}\left[3\theta+(1-\theta)[f_a+2D(1-f_a)]\right]}$$ Eq 1

$$LR \approx \frac{1}{2f_a}$$ Eq 2

$$LR \approx \frac{1}{2\Pr(a \mid aa)}$$ Eq 3

Set $D = 1$

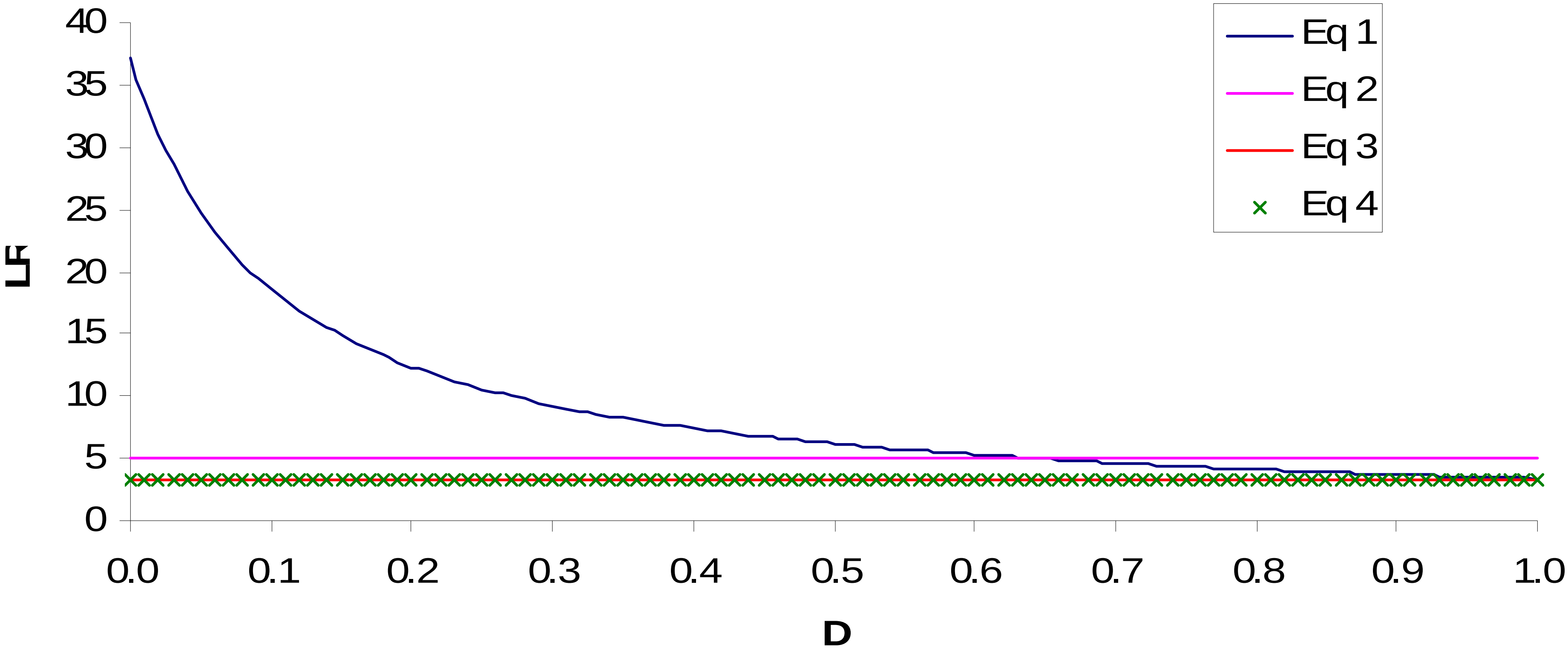
$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{2\theta+(1-\theta)f_a\}\left[3\theta+(1-\theta)[f_a+2(1-f_a)]\right]}$$ Eq 4

$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{2\theta+(1-\theta)f_a\}\left[3\theta+(1-\theta)\left[f_a+2D(1-f_a)\right]\right]} \qquad \text{Eq 1}$$

$$LR \approx \frac{1}{2f_a} \qquad \text{Eq 2}$$

$$LR \approx \frac{(1+\theta)}{2\{2\theta+(1-\theta)f_a\}} \qquad \text{Eq 3}$$

$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{2\theta+(1-\theta)f_a\}\left[3\theta+(1-\theta)\left[f_a+2(1-f_a)\right]\right]} \qquad \text{Eq 4}$$



Eq 1
Eq 2
Eq 3
Eq 4
LF
40
35
30
25
20
15
10
5
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
D

Suspect *ab* Stain $R_1 = a$

- Under *Hp*

$$\Pr(R_1 \mid Hp) = \bar{D}D\overline{Sp}$$

Suspect *ab* Stain R_1= a

- Under *Hd*
- Consider *Mj=aa or ax*

$$\Pr(R_1 \mid Hd) = \bar{D}\overline{Sp} f_a^{\,2} + D\bar{D}\overline{Sp} 2 f_a (1 - f_a)$$

Suspect *ab* Stain $R_1 = a$

$$LR = \frac{\bar{D} D \overline{Sp}}{\bar{D} \overline{Sp} f_a^2 + D \bar{D} \overline{Sp} 2 f_a (1 - f_a)}$$

$$= \frac{D}{f_a^2 + D 2 f_a (1 - f_a)}$$

$$= \frac{D}{f_a (f_a + 2D - 2D f_a)}$$

$$= \frac{D}{f_a (2D + (1 - 2D) f_a)}$$

$$LR \approx \frac{(1+\theta)(1+2\theta) D}{\{\theta + (1-\theta) f_a\}[2\theta + 2D + (1-2D)(1-\theta) f_a]} \quad \text{Eq 5}$$

$$LR \approx \frac{(1+\theta)(1+2\theta)D}{\{\theta+(1-\theta)f_a\}[2\theta+2D+(1-2D)(1-\theta)f_a]} \qquad \text{Eq 5}$$

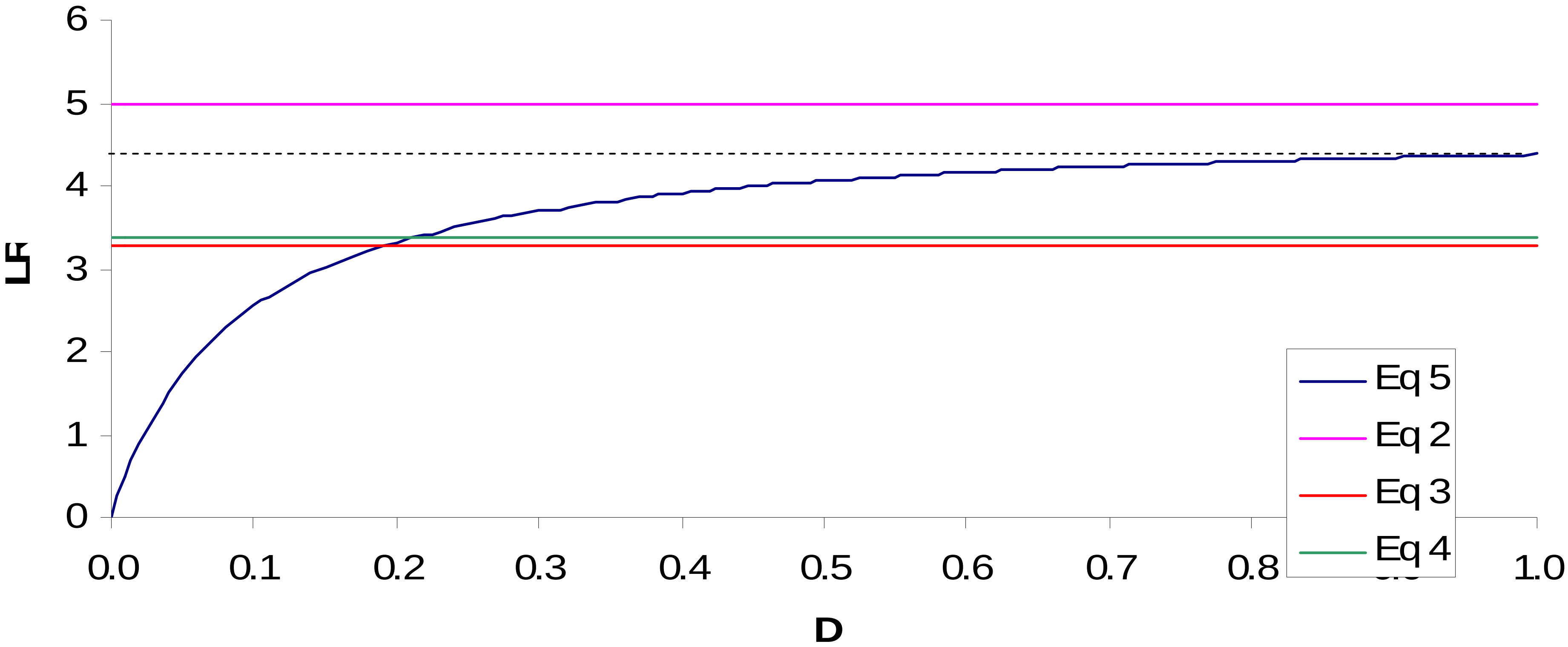
$$LR \approx \frac{1}{2f_a} \qquad \text{Eq 2}$$

$$LR \approx \frac{(1+\theta)}{2\{2\theta+(1-\theta)f_a\}} \qquad \text{Eq 3}$$

$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{2\theta+(1-\theta)f_a\}[3\theta+(1-\theta)[f_a+2(1-f_a)]]} \qquad \text{Eq 4}$$

$$LR \approx \frac{(1+\theta)(1+2\theta)}{\{\theta+(1-\theta)f_a\}[2\theta+2-(1-\theta)f_a]} \qquad \text{Eq 6}$$

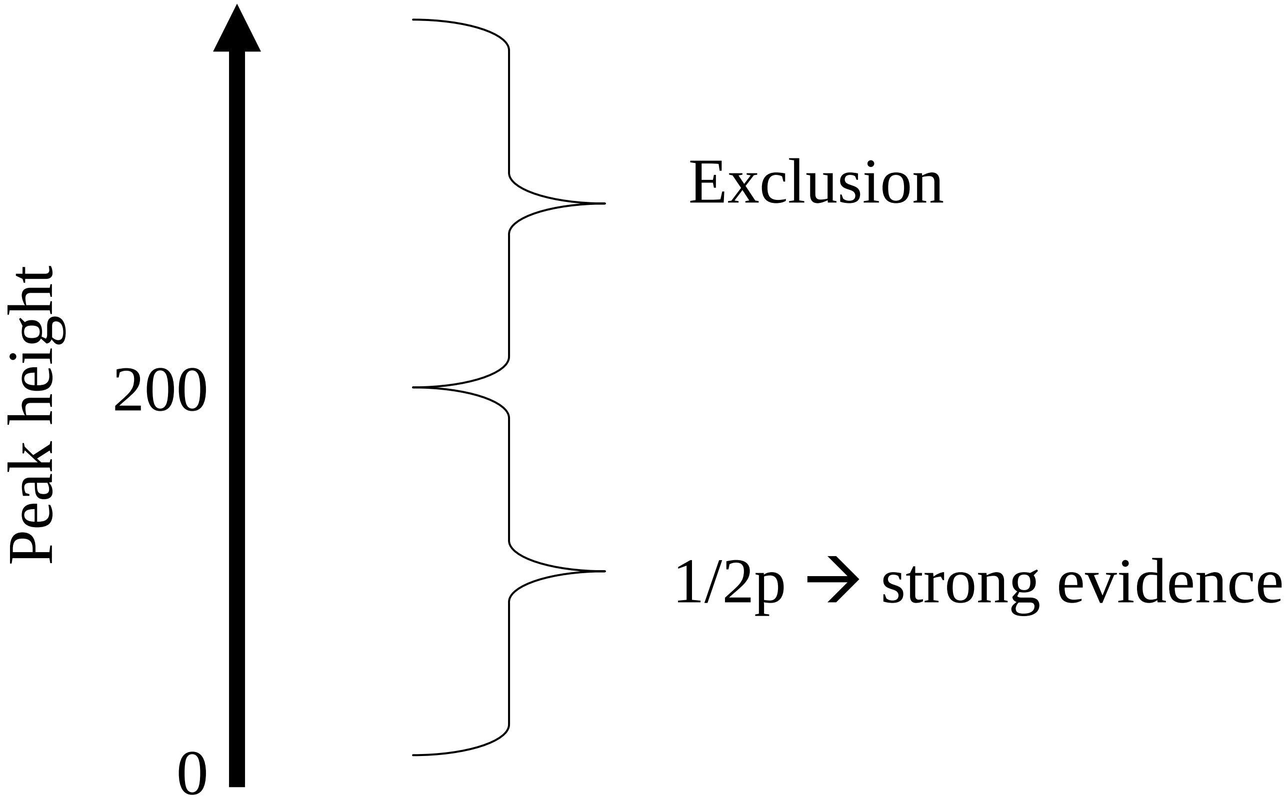
Eq 6 --------



Single replicate?

- Suspect *aa* Stain *a*F → no problem
- Suspect *ab* Stain *a*F → maybe a problem
- Suspect *ab* Stain *ab* caution needed!

Single replicate
Suspect *ab* Stain *a*



Two replicates

- Suspect is *ab*
- R_1 is *a*
- R_2 is *ab*

Mj	$Pr(Mj)$	$Pr(R_1=ab \mid Mj)$	$Pr(R_2=a \mid Mj)$
ab	$2 f_a f_b$	$\overline{D}\,\overline{D}\,\overline{Sp}$	$\overline{D} D \overline{Sp}$
aa	f_a^2	$\overline{D} Sp f_b$	$\overline{D}\,\overline{Sp}$

$$LR \approx \frac{\overline{D}\,\overline{D}\,\overline{Sp}\,\overline{D}\,D\,\overline{Sp}}{\overline{D}\,\overline{D}\,\overline{Sp}\,\overline{D}\,D\,\overline{Sp}\,2 f_a f_b + \overline{D}\,Sp\,\overline{D}\,\overline{Sp}\,f_b f_a^2}$$

$$= \frac{\overline{Sp}\,D\,\overline{D}}{\overline{Sp}\,D\,2 f_a f_b + Sp f_b f_a^2}$$

$$= \frac{1}{2 f_a f_b + \frac{Sp}{\overline{Sp}\,D\,\overline{D}} f_b f_a^2}$$

$$= \frac{1}{2 f_a f_b \left(1 + \frac{Sp f_a}{2\overline{Sp}\,D\,\overline{D}}\right)}$$

Two replicates

- Suspect is ab
- R_1 is a
- R_2 is a

Mj	Pr(Mj)	Pr(R_1=a|Mj)	Pr(R_2=a|Mj)
ax	$2f_a(1-f_a)$	$\overline{D}D\overline{Sp}$	$\overline{D}D\overline{Sp}$
aa	f_a^2	$\overline{D}\overline{Sp}$	$\overline{D}\overline{Sp}$

$$LR \approx \frac{\overline{D}D\overline{Sp}\,\overline{D}D\overline{Sp}}{\overline{D}D\overline{Sp}\,\overline{D}D\overline{Sp}\Pr(ax \mid ab) + \overline{DSpDSp}\Pr(aa \mid ab)}$$

$$\approx \frac{DD}{\Pr(aa \mid ab) + DD\Pr(ax \mid ab)}$$

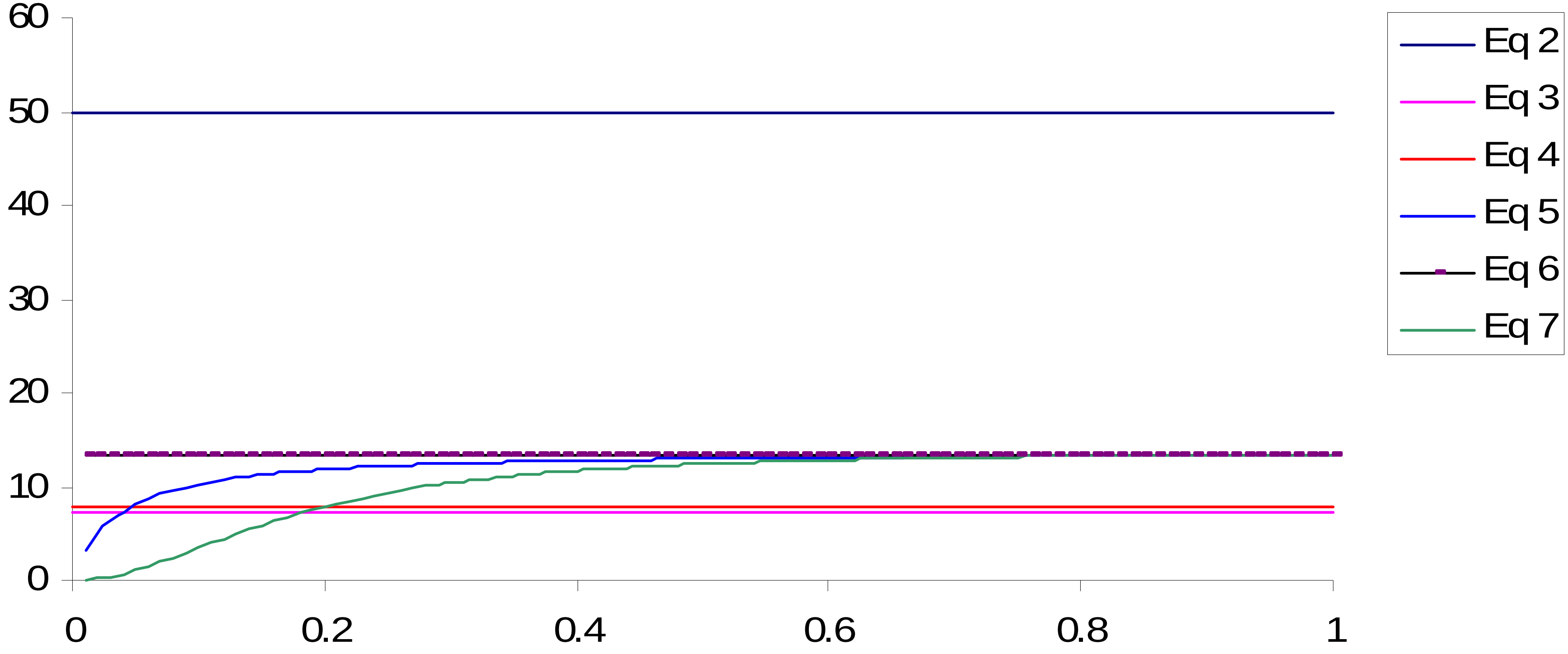
Eq 7

$$\approx \frac{DD}{\Pr(a \mid ab)\left[\Pr(a \mid aab) + DD\Pr(x \mid aab)\right]}$$

$$\approx \frac{(1+\theta)(1+2\theta)DD}{\left(\theta + (1-\theta)f_a\right)\left[2\theta + (1-\theta)f_a + 2DD\left[\theta + (1-\theta)(1-f_a)\right]\right]}$$

$$\approx \frac{(1+\theta)(1+2\theta)DD}{\left(\theta + (1-\theta)f_a\right)\left[2\theta + (1-\theta)f_a + 2DD\theta + 2DD(1-\theta)(1-f_a)\right]}$$

$$\approx \frac{(1+\theta)(1+2\theta)DD}{\left(\theta + (1-\theta)f_a\right)\left[2\theta + 2DD\theta + (1-\theta)(f_a + 2DD(1-f_a)\right]}$$


60
50
40
30
20
10
0
0
0.2
0.4
0.6
0.8
1
Eq 2
Eq 3
Eq 4
Eq 5
Eq 6
Eq 7

Example 8.8

- Suspect *ab*
- R_1 is *ac*
- R_2 is *a*

"*Inspection of the scaling function suggests…the biological model will be seriously non-conservative in this instance.*"

Scaling function

- Analysis of scaling functions enables critical levels of contamination to be estimated
- These feed back into quality assurance programs to ensure that limits are not exceeded

Acknowledgements

- Jonathan Whitaker
- Becky Sparkes
- Alex Lowe
- Peter Gill
- Amanda Kirkham

End