1. (a) *Haldane*: $r = (1 - e^{-2d})/2 \implies e^{-2d} = 1 - 2r \implies -2d = \log(1-2r) \implies$ $\boldsymbol{d = -\frac{1}{2}\log(1-2r)}$

   (b) *Kosambi*: $r = \tanh(2d)/2 = \frac{1}{2}[\exp(4d)-1]/[\exp(4d)+1] \implies (2r)[\exp(4d)+1] = [\exp(4d)-1] \implies \exp(4d)(1-2r) = 1+2r \implies \exp(4d) = (1+2r)/(1-2r) \implies$ $\boldsymbol{d = \frac{1}{4}\log[(1+2r)/(1-2r)]}$

	cM	
r	Haldane	Kosambi
0.01	1.0	1.0
0.05	5.3	5.0
0.10	11.2	10.1
0.20	25.5	21.2

2. Consider an interval of length d on a chromosome of length L.

   Let $A = \{\text{no chiasma in interval}\}$ and $N =$ total no. chiasmata on the chromosome.

$$\begin{aligned}\text{Map function: } M(d) &= [1 - \Pr(A)]/2 \\ &= [1 - \textstyle\sum_{n=0}^{\infty} \Pr(N = n \text{ and } A)]/2 \\ &= [1 - \textstyle\sum_{n=0}^{\infty} \Pr(N = n)\Pr(A|N = n)]/2 \\ &= [1 - \textstyle\sum_{n=0}^{\infty} p_n(1 - d/L)^n]/2\end{aligned}$$

   In the case $p_n = e^{-2L}(2L)^n/n!$, we obtain:

$$\begin{aligned}M(d) &= [1 - \textstyle\sum_{n=0}^{\infty} \frac{e^{-2L}(2L)^n}{n!}(1 - d/L)^n]/2 \\ &= \left[1 - \textstyle\sum_{n=0}^{\infty} \frac{1}{n!}e^{-2L}[2L(\frac{L-d}{L})]^n\right]/2 \\ &= \left[1 - e^{-2d}\textstyle\sum_{n=0}^{\infty} \frac{1}{n!}e^{-2(L-d)}[2(L-d)]^n\right]/2 \\ &= [1 - e^{-2d}]/2\end{aligned}$$

3. Let the chromosome be represented by the interval $[0, L]$.

   Consider any finite set of disjoint subintervals $I_1, I_2, \ldots, I_k$. Let $n_i =$ no. chiasmata in interval I_i (on the four-strand bundle) and $m_i =$ no. crossovers in I_i (on a random meiotic product).

   We wish to show that the $\{m_i\}$ are independent and $m_i \sim \text{Poisson}(|I_i|)$.

   Since the chiasma process is a Poisson process, the $\{n_i\}$ are independent with $n_i \sim \text{Poisson}(2|I_i|)$.

Under no chromatid interference (NCI), the chiasmata are "thinned" independently with probability 1/2 to get the crossover process. Since the $\{n_i\}$ are independent and since the thinning in the disjoint subintervals are independent, it should be clear that the $\{m_i\}$ are independent.

So, we have $m_i|n_i \sim \text{Binomial}(n_i, 1/2)$ and $n_i \sim \text{Poisson}(2|I_i|)$, and we need to show $m_i \sim \text{Poisson}(|I_i|)$. Let $d = |I_i|$ and drop the subscripts i, to save a few keystrokes.

$$\begin{aligned}
\Pr(m = j) &= \textstyle\sum_{k=j}^{\infty} \Pr(n = k \text{ and } m = j) \\
&= \textstyle\sum_{k=j}^{\infty} \Pr(n = k) \Pr(m = j | n = k) \\
&= \textstyle\sum_{k=j}^{\infty} \frac{e^{-2d}(2d)^k}{k!} \binom{k}{j} \left(\frac{1}{2}\right)^k \\
&= \textstyle\sum_{k=j}^{\infty} \frac{e^{-2d}d^k}{k!} \left(\frac{k!}{j!(k-j)!}\right) \\
&= \left(\frac{e^{-d}d^j}{j!}\right) \textstyle\sum_{k=j}^{\infty} \frac{e^{-d}d^{k-j}}{(k-j)!} \\
&= e^{-2d} d^j / j!
\end{aligned}$$

So $m_i \sim \text{Poisson}(|I_i|)$, and we're done.

4. Let $e_i(v) = \Pr(\text{observe } O_i | v_i = v)$.

	O					
v	A	H	B	C	D	–
AA	$1-\epsilon$	$\epsilon/2$	$\epsilon/2$	ϵ	$1-\epsilon/2$	1
AB	$\epsilon/2$	$1-\epsilon$	$\epsilon/2$	$1-\epsilon/2$	$1-\epsilon/2$	1
BA	$\epsilon/2$	$1-\epsilon$	$\epsilon/2$	$1-\epsilon/2$	$1-\epsilon/2$	1
BB	$\epsilon/2$	$\epsilon/2$	$1-\epsilon$	$1-\epsilon/2$	ϵ	1

Note: Think of C = H or B, D = A or H, – = A or H or B, so, for example, $\Pr(C|v) = \Pr(H|v) + \Pr(B|v)$.

5.

$$\begin{aligned}
& B_i(\delta) < B_{i-1}(\delta) \\
\iff\ & \ln \text{RSS}_i + i\delta\left(\frac{\ln n}{n}\right) < \ln \text{RSS}_{i-1} + (i-1)\delta\left(\frac{\ln n}{n}\right) \\
\iff\ & \ln\left(\frac{\text{RSS}_{i-1}}{\text{RSS}_i}\right) > \delta\frac{\ln n}{n} \\
\iff\ & \frac{n}{2}\log_{10}\left(\frac{\text{RSS}_{i-1}}{\text{RSS}_i}\right) > \frac{\delta}{2}\log_{10} n
\end{aligned}$$

(and similarly for i vs $i+1$). Thus, if $\frac{\text{RSS}_{i-1}}{\text{RSS}_i}$ is monotonically decreasing in i, then minimizing the BIC-δ criterion is equivalent to using $\frac{\delta}{2}\log_{10} n$ as a threshold for the conditional LOD score.