Supporting Text

Estimating the Strength of Interference Between Foci

The genetic and cytological coefficients of coincidence. In genetic experiments, the coefficient of coincidence (*CC*) is frequently used for the detection of crossover interference. If recombination is analyzed in two adjacent chromosomal segments (e.g., in a triple backcross), *CC* is defined as R_{12}/R_1R_2, with R_1 and R_2 representing the recombination fractions over the first and second segments and with R_{12} representing the frequency of recombination in both the first and the second segments simultaneously. Interference has been defined as $1 - CC$, so if $CC < 1$, there is (positive crossover) interference. Although it is possible to detect interference using *CC* as a metric, it is not possible to estimate the strength of interference from the *CC* value alone, because *CC* is a function of the recombination frequencies in the analyzed segments. As the length of the analyzed segments increases, R_1 and R_2 approach ½ and R_{12} approaches ¼, so that *CC* approaches 1, even if there is interference. This relation has recently been discussed in detail for the mouse (see Fig. 3*A* in ref. 1).

Fung *et al.* (2) used a metric similar to *CC* for the detection of cytological interference between Zip2 foci along SCs in yeast. They divided SCs into segments of equal length and determined for each segment the frequency of finding a focus. For two adjacent segments, they defined the cytological coefficient of coincidence (named *Z*) as F_{12}/F_1F_2, with F_1 and F_2 representing the frequencies of finding a focus in, respectively, the first and second segment, and with F_{12} representing the frequency of finding a focus in both segments on the same SC. The strength of cytological interference (Ic) was defined as $1 - Z$. However, like *CC*, *Z* has important drawbacks as a metric for interference.

Obviously, the assignment of foci to segments leads to loss of resolution. More importantly, the length of the chromosomal segments to which foci are assigned influences the estimate of *Z* in two ways. If one assigns foci to segments that are long compared with the average interfocus distance, many segments will have more than one focus, whereas F_1 and F_2 have not been defined for more than one focus in a segment. Fung *et al.* (2) assigned one of the two foci to an adjacent segment if they encountered

two foci in a segment. However, this changes the distribution of foci and thus affects the estimation of the strength of interference. Moreover, it leaves unsolved the problem of what to do if both adjacent segments have already a focus, or if there are several foci in a segment. If one redefines F_1 and F_2 as the frequencies of finding *at least* one focus in the relevant segments, then F_1, F_2, F_{12} and thus *Z* will all approach 1 as the segment size increases, so that little or no interference will be detected if the segments are about as long or longer than the average interfocus distance; this is illustrated by the simulations shown in Fig. 2. This effect is comparable to the effect of segment size on the estimate of *CC*. However, if one assigns foci to short chromosomal segments, other complications arise. If the segments are not much longer than the minimal interfocus distance that can be observed by light microscopy (~0.18 μm), a substantial proportion of segment pairs with a focus in each segment will go unnoticed because the two foci are too close; this leads to underestimation of *Z* and detection of interference even if there is none (Fig. 2).

Fig. 3 shows the *Z* values obtained for MSH4 foci on chromosome 2 of female *Sycp1* knockout mice for three different segment sizes. The dependency of *Z* on the segment size prohibits the estimation of the strength of interference between MSH4 foci from the *Z* value alone. It also precludes a comparison of the strength of interference between MLH1 foci with that between MSH4 foci on the same chromosome based on *Z* values: Because the average distance between MSH4 foci is far smaller than that between MLH1 foci, the segment size will influence the *Z* values for MSH4 and MLH1 foci differently.

Using the gamma model for estimating the strength of interference. In genetic studies, the precise positions of crossovers are usually not known; rather, only the segments (delimited by genetic markers) within which crossover(s) must have occurred are known. In such studies, crossover interference is necessarily analyzed in chromosomal segments. In cytological studies, in contrast, there is no need to study interference in chromosomal segments, because the precise positions of foci are known. Several models have been considered for estimating the strength of interference from the precise positions of chiasmata/crossovers/foci, and the gamma distribution has repeatedly emerged as a most useful model for this purpose (e.g., 3–5). The gamma distribution is commonly used for

the analysis of distances between events along a linear axis, and the properties of this distribution can be found in most textbooks on probability theory and/or stochastic processes.

The gamma model for interference of crossover events in meiosis can adequately be described as a modification of the Poisson process. In a Poisson process, events occur at random along an (imaginary) time or physical axis. The probability of an event in any interval does not depend on the occurrence or nonoccurrence in adjacent intervals. This property is sometimes referred to as “lack of memory” of the Poisson process.

Denoting the rate of occurrence (i.e., the mean number of events per unit of distance) by λ, the probability density function (p.d.f.) of interevent distances for a given value of λ reads

$$f(x;\lambda) = \lambda e^{-\lambda x}, \quad x > 0 .$$

The mean $E(x)$ and variance $\mathrm{var}(x)$ of the interevent distances read

$$E(x) = \frac{1}{\lambda}$$

and

$$\mathrm{var}(x) = \frac{1}{\lambda^2} .$$

Considering now the sum of n interevent distances, we obtain the gamma distribution with p.d.f. for interevent distances x

$$f(x \mid \lambda, n) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} .$$

Notice that the gamma distribution describes a Poisson process in which only every *n*th event is realized. One could also state that after every event in a Poisson process, the next $n - 1$ events are suppressed. As described by Stam (4) and McPeek and Speed (3), the concept of interference nicely corresponds with the "suppression" of events in a Poisson process. The result of (positive) interference is that events are more evenly distributed along the "axis" (i.e., the chromosome) than in a Poisson process.

Mathematically, the value of the parameter *n* is not restricted to integer values. Thus, a more general form of the gamma p.d.f. reads

$$f(x \mid \lambda, \nu) = \frac{\lambda^{\nu} x^{\nu-1} e^{-\lambda x}}{\Gamma(\nu)}, \ \nu > 0$$

where ν is equivalent to *n*, except that ν is not necessarily an integer, whereas $\Gamma()$ is the so-called gamma function. (The gamma function can be seen as a normalizing constant, such that the integral $\int_0^{\infty} f(x \mid \lambda, \nu) dx$ equals unity, as it should for a probability density function.)

When fitting a gamma distribution to a number of observed interevent distances, one obtains an estimate of both λ and ν.

The estimated value of ν can be taken as a measure of the level of interference [cf. Stam (4) and McPeek and Speed (3)]. The higher the value of ν, the more events are suppressed in the underlying Poisson process, and the stronger the interference is.

Thus, by comparing values of ν for different classes of foci, we may infer about different levels of interference among these classes.

The mean and variance of gamma-distributed interevent distances read

$$E(x \mid \lambda, \nu) = \frac{\nu}{\lambda}$$, and

$$var(x \mid \lambda, \nu) = \frac{\nu}{\lambda^2}$$.

Rescaling the distances such that the mean interevent distance equals unity, we observe that the ratio of mean and variance equals

$$\frac{E(x)}{var(x)} = \nu$$.

Thus, as is illustrated in Fig. 1*A* of the main text, as the level of interference (ν) increases, the variance (relative to the mean) gets smaller, meaning that events are becoming more and more equidistant. This again nicely captures the concept of (positive) interference.

1. Broman, K. W. (2005) *Genetics* **169,** 1133–1146.

2. Fung, J. C., Rockmill, B., Odell, M. & Roeder, G. S. (2004) *Cell* **116,** 795–802.

3. McPeek, M. S. & Speed, T. P. (1995) *Genetics* **139,** 1031–1044.

4. Stam, P. (1979) *Genetics* **92,** 573–594.

5. Foss, E., Lande, R., Stahl, F. W. & Steinberg, C. M. (1993) *Genetics* **133,** 681–691.