Role of *cis*, *trans*, and inbreeding effects on meiotic recombination in *Saccharomyces cerevisiae*

Xavier Raffoux, Mickael Bourge, Fabrice Dumas, Olivier C. Martin, and Matthieu Falque

Supplementary Methods

A-Segregation bias and fluorescence extinction

To check for possible experimental/technical biases in our experiment, we designed the following two tests for Mendelian segregation:

(1) *Single locus segregation in all hybrids:* we investigated to what extent our recombination phenotyping method is robust to different genetic backgrounds, particularly with hybrids from more or less distantly related parental strains. We observed single-locus segregation ratios of fluorescent *vs* non-fluorescent (NF) spores ranging from 0.47 (*sd*=0.013) to 0.51 (*sd*=0.008) among 22 hybrids (average of all markers), and ranging from 0.48 (*sd*=0.0128) to 0.50 (*sd*=0.011) among 15 markers (average of all hybrids). The lowest values of these ratios always corresponded to strains YIIc17_E5, UWOPS83_787_3, and marker CFP-XI-2 for which we observed a deficit of fluorescent spores.

(2) *Three-locus analysis*: from our data, we calculated the ratio between frequencies of spores with external markers only and spores with the central marker only. This ratio ranged from 0.54 (*sd*=0.078) to 0.64 (*sd*=0.135) among 22 hybrids (average of all markers) and from 0.49 (*sd*=0.040) to 0.81 (*sd*=0.069) among eight testers (average of all hybrids). Outliers showing an excess of spores with only the two external markers were again observed with hybrids from strains YIIc17_E5 and UWOPS83_787_3, and with tester XI-R1C2Y3 (containing marker CFP-XI-2 at middle position).

In a previous paper (Raffoux *et al.*, 2018. Yeast, doi:10.1002/yea.3315), we observed on strains SK1 and W303 that some parental-type tri-fluorescent spores lost the fluorescence phenotype for the central locus. Such fluorescence extinction resulted in some extent of segregation bias, mainly affecting multi-locus segregation and thus interference measurement. Our present results on many

more strains are fully consistent with these previous observations. To correct for this bias, we thus used the mathematical model described in our first paper to estimate fluorescent extinction frequencies as well as corrected recombination rates and interference strength (true coefficient of coincidence) from flow cytometry data (see Materials and Methods). Using this model with all testers and all hybrids, the average extinction frequencies estimated for the three consecutive markers of a tester were 0.5% (*sd*=0.8), 1.3% (*sd*=1.5), and 0.7% (*sd*=0.9). As expected, the highest extinction rates were obtained for strains YIIc17_E5, UWOPS83_787_3 which showed the strongest segregation distortion for all single markers (Supp Figure 2). These values are consistent with the rate of experimental fluorescence extinction measured in a SK1 homozygous fluorescent diploid as described in our previous paper, indicating that there is no bias in the recovery of all the expected classes of spores.

B-Experimental design and normalization

Each experiment was composed of one of the eight tri-fluorescent testers crossed with each of the 26 strains of the collection, plus one control sample: (Y12 Mat a X SK1-VI-Y3R4C5 Mat α). Four technical replicates were achieved for each cross and placed at different locations in the incubator during sporulation. As recombination rate values of the control sample didn't show significant variation between the eight experiments corresponding to the eight testers (ANOVA *p*-value = 0.99), we normalized the results as follows: every recombination rate obtained in experiment *E* was multiplied by the average recombination rate of the control across the eight experiments, and divided by the recombination rate of the control in experiment *E*.

C-Translating a coefficient of coincidence into an interference strength

When modeling interference, one usually assumes a process that generates COs on a continuous chromosome such that nearby COs are less likely to arise than by chance (*i.e.*, than when COs arise independently). In such models (Kleckner *et al.*, 2004, Proc Natl Acad Sci U S A 101: 12592-12597; McPeek and Speed, 1995, Genetics 139: 1031-1044), interference *strength* is quantified by a single parameter. Suppose the considered model is used to simulate a large number of meioses; interference in such data can be characterized either by that single model parameter or by the coefficients of coincidence (CoCs) produced. The « problem » with CoCs is that they require specifying the size of the intervals used. Since the CoCs will be very small for small intervals and will approach 1 for large intervals, CoCs do not give a direct measure of interference strength. It is thus wisest to rely on an underlying model if the objective is to quantify and compare interference strength. In the case of our dataset however, we do not have CO positions along a continuous

chromosome, we just know whether adjacent intervals are recombinant or not. In view of the previous remarks, it is best to « translate » the CoCs of these interval pairs into estimates of the model's interference strength parameter, thus overcoming the dependence of CoC on interval sizes. We do this for our dataset by asking which value of the model's interference parameter leads to the experimental CoC, given the recombination rates in each interval. For specificity, we use the Gamma model (McPeek and Speed, 1995, Genetics 139: 1031-1044) of CO interference but our method for mapping CoCs to interference strength is more general. The steps for a given pair of intervals are as follows.

(1) First it is necessary to relate genetic distance (« *d* » in Morgans) and recombination rate (« *r* », ratio without dimension). Indeed, Haldane's formula, namely $d = -0.5 \ln(1 - 2r)$, linking those two quantities, assumes that there is no interference; unfortunately, in the presence of interference, no such formula is available. We thus resort to simulation to determine the relation between *d* and *r*, assuming a given value of ν (the Gamma model interference parameter). This is done by simulating 10^5 meioses, considering two *ad hoc* markers at distance *d* and extracting the fraction of meioses leading to an odd number of COs between those two markers. In practice, we do this for regularly spaced values of *d*, *d*=0,2,4,6,8,… cM, thereby creating a table relating genetic distance and recombination rate. This table can then be used to interpolate and obtain the curve linking *d* and *r*. Supp Figure S3 illustrates this mapping for the specific values $\nu = 1.4$ and $\nu = 2.5$.

(2) Second, we extend the single interval analysis just described to study the CoC arising for the *two* intervals of interest. Specifically, the mapping between *d* and *r* allows us to infer each interval's genetic length, d_1 and d_2 (just search the table, with spline polynomial interpolation, to find the two distances that give rise to each interval's recombination r_1 and r_2.). Then as above we can introduce *ad hoc* markers delimiting the positions of the two intervals and extract the fraction of double recombinants generated by the simulations. This thereby provides a « theoretical » value for the CoC, and thus the translation between ν and CoC. Note that this translation requires knowing d_1 and d_2 (or equivalently, r_1 and r_2). In practice the experimental measurements provide the CoC and so it is necessary to perform a scan in ν to see which value of ν leads to a translation into the experimental value of the CoC. In Supp Figure S4 we illustrate the result of this scan for two of our pairs of intervals.

In summary, for each interval, we use these two steps to infer the value of the corresponding interference parameter ν, thereby overcoming the caveats of using the CoC directly as a measure of interference strength.