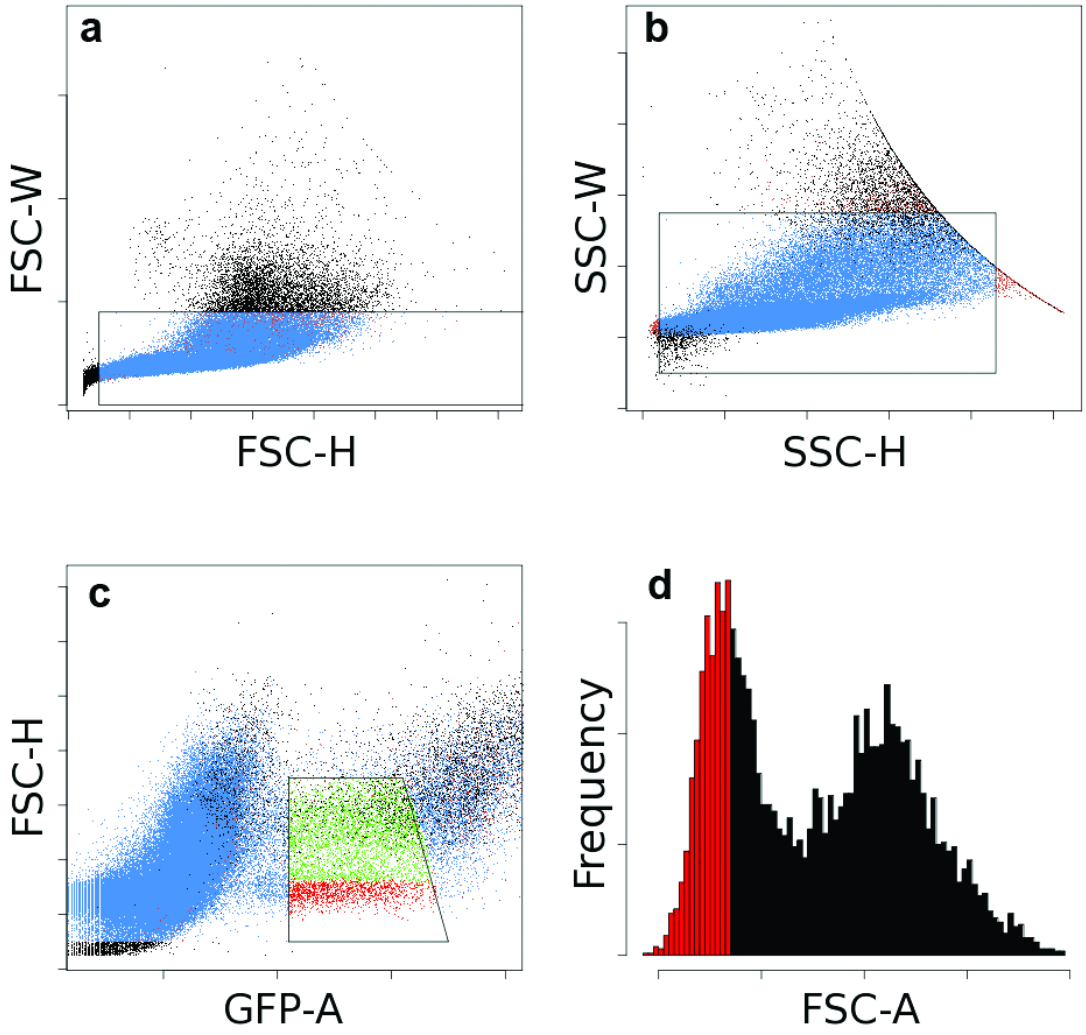
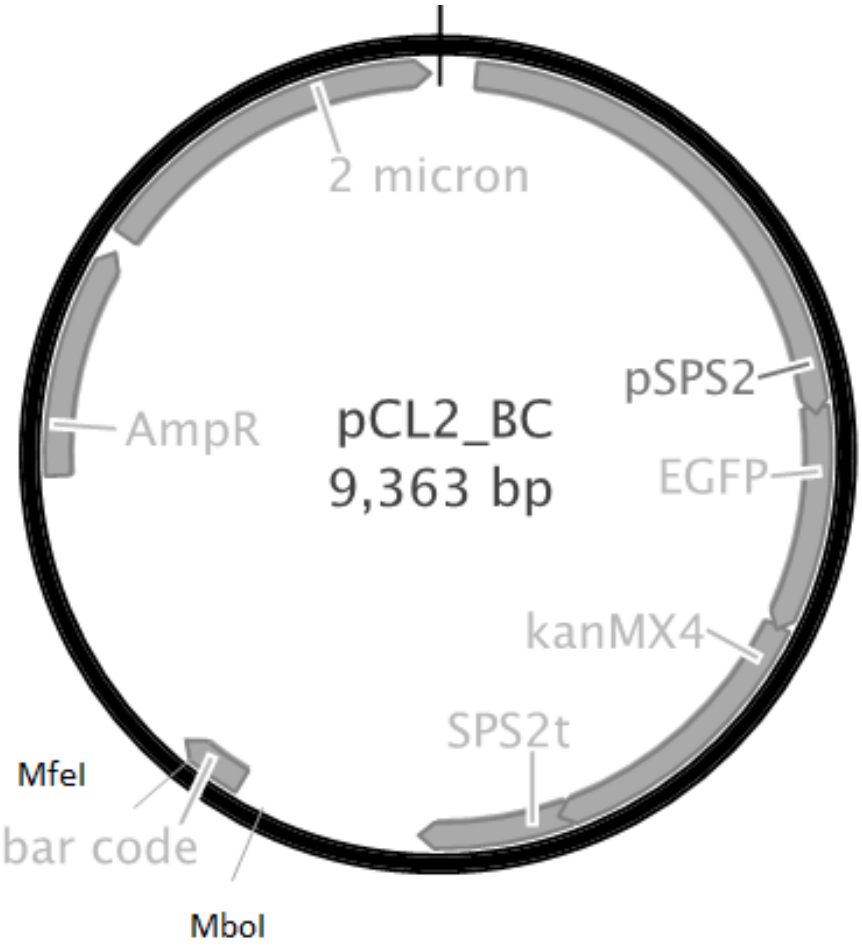
Supplementary Figure 1. FACS gates for tetrad sorting.



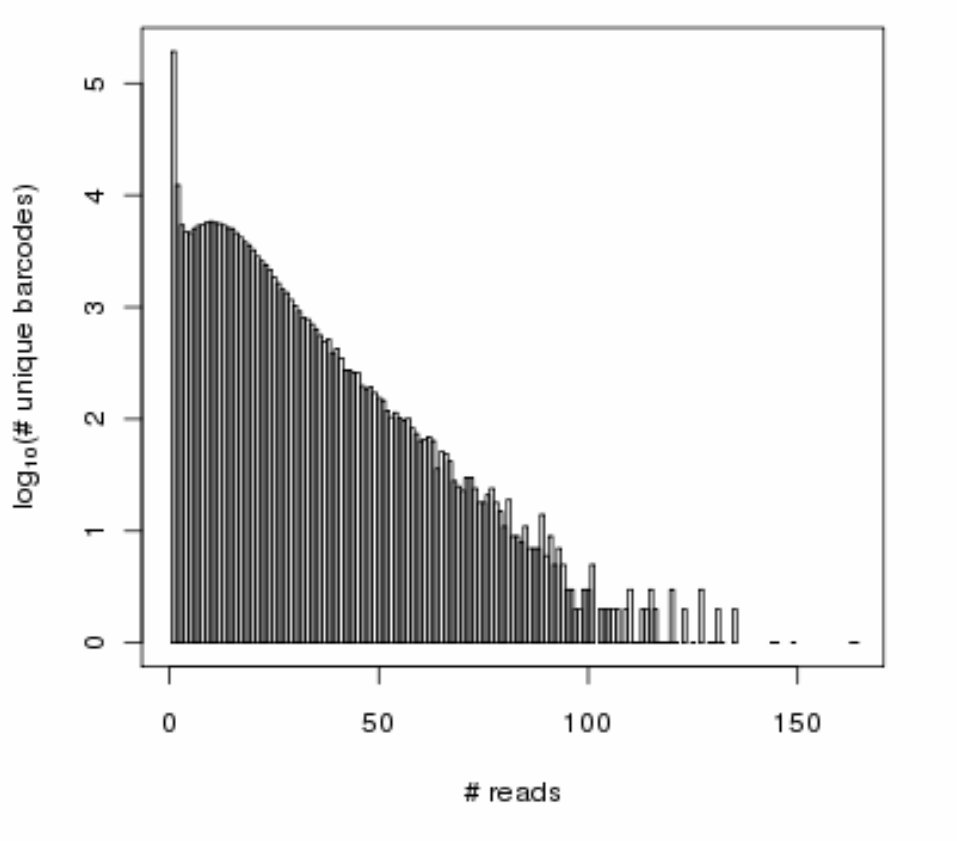
Supplementary Figure 1. FACS gates for tetrad sorting. The first two gates (**a**, **b**) are used to remove events which are either debris or have more than one cell by selecting a narrow width of the FSC and SSC signals, while permitting a large range of FSC and SSC heights. These combined gates remove ~7% of the total events. The next gate (**c**) is a GFP vs. FSC gate, which is used to identify fluorescent events. To choose this gate, we inspected the events for fluorescent and non-fluorescent populations, and initially selected all the fluorescent events. After checking the sorted cells from this population microscopically, we found many multi-tetrad clumps. We reduced the upper limit of fluorescence on this gate, which worked well for excluding clumps. The final gate (**d**) was on the FSC channel only; we found that within the fluorescently-gated population there were two peaks in the FSC histogram. Sorting events from each of these peaks revealed that the high-FSC peak was mostly composed of tetrads with an attached small bud. The low-FSC peak was enriched for isolated tetrads. Consequently, we set the gate to allow passage of events with an FSC measurement lower than that of the low-FSC peak. After applying these gates, the sorted events comprised 1% of the total events and were found by microscopy to be 95% tetrads.

Supplementary Figure 2. Plasmid map.



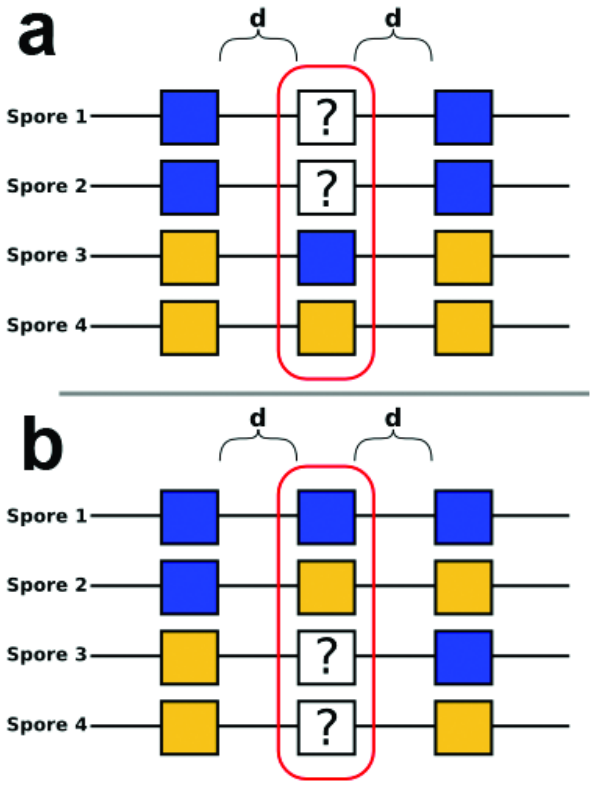
Supplementary Figure 2. Plasmid map. BEST utilizes a library plasmid that contains a copy of the sporulation-specific gene *SPS2* fused in frame to EGFP and a G418-resistance cassette (kanMX4). The plasmid also has a randomized 15-nucleotide barcode.

Supplementary Figure 3. **Barcode distribution.**



Supplementary Figure 3. **Barcode distribution**. We sequenced 1.5µg of our plasmid library (Online Methods) to determine the frequency of randomized 15-nucleotide barcodes. While many barcodes had more than one read in our sample (mean frequency of barcodes = 6.46 reads), the distribution is heavily skewed to rarely-seen barcode sequences (median frequency = 1 read).

Supplementary Figure 4. Examples in which full tetrad haplotyping improves the accuracy of linkage-based inference of missing genotypes.



Supplementary Figure 4. Examples in which full tetrad haplotyping improves the accuracy of linkage-based inference of missing genotypes. Two tetrads (A and B) with allele calls for the four spores within the tetrad at three consecutive chromosomal loci, separated by genetic distances of d in each interval are illustrated. Markers are colored blue and yellow to represent parent 1 (P1) and parent 2 (P2) alleles. For tetrad A, spore based inference suggests that both of the missing markers are parent 1 alleles (no crossovers versus crossovers in both intervals) with a probability of 0.9999 and 0.9977 for values of d of 1 cM and 5 cM respectively. Examination of information from the whole tetrad, however, reveals that either spore 1 or spore 2 must have undergone a double crossover with spore 3. Therefore each of the two missing markers has a 0.5 probability of having the P1 or P2 allele. For tetrad B, spore based inference suggests that spore 4 has the P2 allele (no crossovers versus crossovers in both intervals) with a probability of 0.9999 and 0.9977 for values of d of 1 cM and 5 cM respectively while spore 3 has a 0.5 probability of having the P1 or P2 allele (crossover equally likely in left or right interval). Examination of information from the whole tetrad, however, reveals that spore 2 and spore 3 are likely to have recombined with one another (single crossover in left interval of spore 3 versus crossover in right interval of spore 3 and crossover in both the left and right intervals of spore 4). Therefore, the full tetrad supports the P1 allele in spore 3 and the P2 allele in spore 4 with a probability of 0.9950 and 0.9750 for values of d of 1 cM and 5 cM respectively. For these calculations Haldane's Poisson-based mapping functions for single spores and tetrads[1,2] were used to calculate recombination pattern probabilities.

Supplementary Table 11. Plasmid barcode library sequencing results

Total reads	2,080,316
Unique reads	322,003
Single-read barcodes	192,646
Mean reads per barcode	6.46
Max reads per barcode	164

Additional FACS information

To isolate a population of yeast that is highly enriched for tetrads, we employed a series of FACS gating steps (**Supplementary Fig. 1**). The primary gate used for separating tetrads is a simple GFP fluorescence gate. Unsporulated cells formed a population with low fluorescence, while sporulated cells formed a population with high, but widely varying fluorescence. This population contained a high number of tetrads, including some tetrads that appeared to have an attached bud as well as attached pairs of tetrads that seem to have arisen from dividing cells that had not undergone a complete mitotic division before entering meiosis. To isolate individual tetrads from clumps, we set a gate that included events with moderate fluorescence, but excluded the highest intensity events.

To test the efficacy of our gating regime, we sorted 1000 events onto a microscope slide and counted tetrads and non-tetrads by phase contrast microscopy. In the FY x Σ1278b (pilot) cross, 262 tetrads and 13 non-tetrads were counted, all but one of these 13 "non-tetrad" events were actually tetrads with an attached bud. In the S288c x YPS163 (pilot) cross, 120 tetrads and 2 non-tetrads were counted, where one of the “non-tetrads” was a tetrad with an attached bud.

Additional plasmid library information

The complexity of our barcode (**Supplementary Fig. 2b**) is conferred by the presence of a randomized 15 nucleotides, which permits a theoretical 10^9 unique sequences. To determine the subset of that sequence complexity captured in our library, we sequenced a portion by Illumina sequencing. The results (**Supplementary Fig. 3**, **Supplementary Table 11**) demonstrate that the pCL2_BC library is of sufficient complexity for our experimental design. The ~2 million barcode sequences read included 322,003 different barcode sequences, of which 192,646 were present only once in the population. In our implementation of BEST, we start with ~10^4 transformants of the heterozygous diploid strain. Based on the empirically determined barcode distribution (**Supplementary Fig. 3**), 95% of the transformants should have a unique barcode. During the FACS isolation of tetrads, we select 25 random tetrads from this pool of transformants, which equates to a 99.8% chance that all 25 tetrads in a batch have a unique barcode.

Additional sporulation and tetrad disruption information

In our hands, the efficiency of spore disruption is the most variable part of the protocol, with different crosses/ strain backgrounds requiring different conditions for optimal spore separation. We have been able to achieve the levels of spore disruption reported by varying the sporulation conditions. In all cases strains were sporulated in liquid medium at room temperature. In some cases, optimal disruption was achieved using “fresh” spores sorted within a day of completing sporulation. In other cases, such as the two crosses presented here, optimal disruption was achieved by allowing cultures to age for 7-10 additional days at room temperature without agitation. While this part of the protocol might require some optimization, successful tetrad disruption is easily assayed by counting the number of single colonies that arise from the disrupted tetrads. Plates giving rise to a number of colonies significantly less than that expected for 4 times the number of tetrads sorted (corrected for the cross-specific spore viability) need not be processed further.

Sequence analysis of unrecovered tetrad barcodes

In ~5% of the strains in our two pilot crosses, no plasmid barcode sequence was detected, despite high read coverage of the genomic DNA sequence. To characterize these events we chose 89 such strains. In all cases, plasmid DNA could be isolated and capillary sequencing (Beckman Coulter Genomics) of the plasmid DNA identified two major classes of events accounting for the missing barcode information.

In the first class, a base change inactivated the *Mfe*I site. In the second class, a new *Mbo*I site was generated within the random tetrad barcode sequence.

References

1 Haldane, J. The combination of linkage values and the calculation of distance between loci of linked factors. *Journal of Genetics* **8**, 299-309 (1919).

2 Balding, D. J., Bishop, M. & Cannings, C. (Jon Wiley & Sons, West Sussex, England, 2007).