High-throughput tetrad analysis

Catherine L Ludlow1,4, Adrian C Scott1,4, Gareth A Cromie1, Eric W Jeffery1, Amy Sirr1, Patrick May1,2, Jake Lin1, Teresa L Gilbert3, Michelle Hays1 & Aimée M Dudley1

Tetrad analysis has been a gold-standard genetic technique for several decades. Unfortunately, the need to manually isolate, disrupt and space tetrads has relegated its application to small-scale studies and limited its integration with high-throughput DNA sequencing technologies. We have developed a rapid, high-throughput method, called barcode-enabled sequencing of tetrads (BEST), that uses (i) a meiosis-specific GFP fusion protein to isolate tetrads by FACS and (ii) molecular barcodes that are read during genotyping to identify spores derived from the same tetrad. Maintaining tetrad information allows accurate inference of missing genetic markers and full genotypes of missing (and presumably nonviable) individuals. An individual researcher was able to isolate over 3,000 yeast tetrads in 3 h, an output equivalent to that of almost 1 month of manual dissection. BEST is transferable to other microorganisms for which meiotic mapping is significantly more laborious.

Meiotic mapping, a linkage-based method for analyzing the recombinant progeny of a cross, has long been a cornerstone of genetics. The method is applicable to a wide range of eukaryotes, including yeasts and less tractable microorganisms such as the filamentous fungus Neurospora crassa and the unicellular green alga Chlamydomonas reinhardtii. The approach involves the isolation and cultivation of spores derived from individual asci (tetrads) and was first developed in Saccharomyces cerevisiae1. However, the manual nature of this process severely limits its throughput, even for experienced researchers. In an exceptional case of productivity, more than 12,000 crosses were dissected by the yeast geneticist C. Styles over a 30-year career (G. Fink, personal communication). A typical meiotic mapping experiment involving a single cross and several hundred tetrads might take several weeks—an efficiency that is too low for many genetic studies. Researchers trying to circumvent this bottleneck have used a variety of strategies. One strategy, random spore analysis, enriches for tetrads by selectively killing vegetative cells2. Spores are then randomly dispersed on solid medium to recover the recombinant progeny. A second strategy avoids much of the high variability and low specificity of random spore analysis by using a selectable reporter gene under the control of a mating-type–specific transcriptional promoter3. This approach has been applied with great success to generate specific classes of recombinant progeny needed to test synthetic growth defects3–5 and linkage between traits and gene deletions6. A third strategy, bulk segregant analysis7,8, has been used in organisms ranging from yeast to plants. Bulk segregant methods use a pooled genotyping strategy to identify genomic regions common to the majority of progeny selected under a specific condition, such as high drug concentrations. Although these strategies have been used effectively to address specific problems, they all have limitations that prevent them from achieving the broad applicability of conventional tetrad analysis. Chief among these limitations are the inability to recover all viable meiotic progeny (due to either the progeny isolation method or the phenotypic selection imposed for bulk comparisons) and the loss of the tetrad relationships (that is, knowledge of which sister spores were members of the same original tetrad). The limited throughput of manual tetrad dissection constrains two research areas. The first is the mapping of complex traits resulting from combinations of naturally occurring polymorphisms. The ability to detect only a small fraction of the genetic loci that underlie complex traits by meiotic mapping is at least partially due to the lack of sufficiently large numbers of individuals8. The second is the study of the molecular mechanisms of recombination, which often depends on capturing all the events from an individual meiosis. A striking example is the study of gene conversion, for which one seminal paper used over 19,000 tetrads10. In the absence of high-throughput alternatives, such fields are unable to effectively leverage the current revolution in DNA sequencing technology11.

We describe a high-throughput method, BEST, that enables the generation and genotyping of large numbers of progeny. With BEST, tetrads are isolated, genotyped and maintained as individuals in a manner that allows the sister-spore relationships of all four meiotic products to be recovered.

RESULTS

Tetrad dissection has two critical steps that are difficult to automate because they require microscopy-aided manipulation of individual tetrad spores. The first is the isolation of tetrads away from unsporulated cells in the culture, which outnumber tetrads 99:1 in the commonly used FY strain background (D. Swain-Lenz and J. Fay, unpublished data). The second is the physical separation of the tetrad spores and their arrangement in a grid.

1Institute for Systems Biology, Seattle, Washington, USA. 2Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg. 3Deceased. 4These authors contributed equally to this work. Correspondence should be addressed to A.M.D. (aimee.dudley@gmail.com).

RECEIVED 25 FEBRUARY; ACCEPTED 12 APRIL; PUBLISHED ONLINE 12 MAY 2013; DOI:10.1038/NMETH.2479
In *S. cerevisiae*, spores are held together by both an outer ascus, which is the remnant of the cell wall of the original diploid cell, and a set of interspore bridges. In conventional tetrad dissection, enzymatic digestion removes the ascus, and a researcher uses a micromanipulator to identify tetrads, break the interspore bridges and array the spores in a gridded pattern. The grid separates spores to prevent interspore mating, and it also preserves the knowledge of which spores came from the same tetrad. BEST (Fig. 1) overcomes both of these bottlenecks with a process in which tetrads are isolated and disrupted to give single spores, whose tetrad relationships are reconstructed using molecular barcodes.

**Tetrad isolation and disruption**

FACS permits rapid separation of four-spore tetrads from a mixed population that includes vegetative cells, dead cells, clumped cells and two-spore dyads (Fig. 2). Several reporter genes have been used to fluorescently label tetrads or individual spores; we chose the *SPS2-EGFP* fusion because it has been successfully used to quantitate sporulation in several genetically diverse, nonlaboratory strains. We introduced this construct into a 2-micron plasmid that, when transformed into yeast, is expressed in meiosis and fluorescently labels tetrads. We then established a series of FACS gating parameters (Online Methods and Supplementary Fig. 1) that reproducibly yielded 95% four-spore tetrads (Fig. 2), even from strains with sporulation efficiencies less than 1% (data not shown). The inclusion of this FACS step is how BEST achieves its largest gain in throughput: the identification and isolation of ten tetrads can be accomplished by FACS in less than 1 s, whereas the equivalent manual process takes an experienced researcher several minutes.

After tetrad isolation, the ascus must be digested and the spores digested and disrupted and physically separated to grow as individual clones (colonies). To avoid substantial spore loss during the liquid transfer steps, we developed a procedure for digesting, disrupting and distributing the spores on an agar plate (Online Methods), as follows. We sort GFP-positive tetrads directly into a pool of Zymolyase solution on the plate, where the ascus is enzymatically digested. Spores are then separated by agitation with glass beads, which provides the mechanical force necessary to break the interspore bridges connecting sister spores. This process also physically disperses spores randomly across the plate, far enough apart that most colonies are pure clonal isolates of each recombinant spore. Once cells have grown to single colonies, they are individually transferred to liquid medium in a 96-well plate in which they are stored as frozen stocks for subsequent genotype and phenotype analysis.

**Tetrad reconstruction**

The ease of random plating strategies and their amenability to automation come at the expense of losing information about which progeny originated from the same tetrad. To facilitate the reconstruction of sister-spore relationships, we devised a molecular barcoding strategy that satisfies four main criteria. First, the barcode must be reliably transmitted to all four tetrad spores. Second, the pool of barcodes must be complex enough to ensure that individuals share a common barcode because they were members of the same tetrad. Third, the presence of the barcode should be phenotypically neutral. Finally, the barcode should be compatible with the method used to determine the progeny genotypes so that the barcode can be read as part of the genotyping workflow. To our knowledge, no existing barcoding resource satisfies all of these criteria. For example, strategies that integrate barcodes at a neutral genomic location will be heterozygous in a diploid genome and thus present in only half of the tetrad's spores.

To satisfy these requirements, we constructed a barcoded 2-micron library. A random barcode sequence, flanked by restriction sites that ensure its representation in the sequencing reads of the chosen genotyping method, was inserted into the plasmid. 2-micron plasmids are maintained in high copy (10–40 copies per cell) and stably segregate during cell division, increasing the likelihood of plasmid transmission to all four spores. As with the native 2-micron circle, the presence of engineered 2-micron
plasmids should have a relatively neutral impact on most traits. However, because the plasmid is no longer required after the strain is genotyped, direct counterselection or simple failure to maintain selection could facilitate plasmid loss.

The plasmid library contains the SPS2-EGFP sporulation-specific fluorescent reporter, a complex DNA barcode flanked by restriction sites compatible with our genotyping protocol and a selectable drug-resistance marker for plasmid maintenance (Online Methods, Supplementary Fig. 2). This library is transformed into the heterozygous diploid cells of the cross (Fig. 1). The complexity of the library (Supplementary Fig. 3) is conferred by the presence of the randomized 15-nucleotide barcode, which permits a theoretical 10⁹ unique sequences. By pooling thousands of yeast transformants, we created a mixed population of barcoded diploid cells that fluorescence only when cells have undergone sporulation and that pass on a tetrad-specific barcode to each spore of the tetrad (Fig. 1). Because the pool of barcodes (Supplementary Fig. 3), which is determined by the number of initial yeast transformants, is much larger than the number of tetrads that will be mixed together on the same agar plate (Online Methods), the probability of isolating the same barcode in multiple tetrads is relatively low.

**Sequencing-based genotyping and reconstruction of tetrads**

Although BEST is compatible with numerous genotyping platforms, we favor sequencing-based approaches that permit the simultaneous determination of the strain’s genotype and the plasmid-borne tetrad barcode (Fig. 3a) in a manner that yields individual, rather than pooled, genotypes. Because recombination in yeast generates relatively few crossover events per chromosome per meiosis¹³, most of a recombinant genome’s sequence can be imputed from a relatively sparse set of genetic markers (Fig. 3b). We therefore chose a highly multiplexed genome-reduction strategy known as restriction site–associated DNA tag sequencing (RAD-seq)¹⁹, which directs genome sequencing to specific restriction sites (Fig. 3 and Online Methods). Our choice of restriction enzymes and 40-base Illuma single-end reads allowed us to sequence the same 2–3% of every yeast strain using a highly multiplexed sequencing strategy (Online Methods). For crosses between strains with a sufficient number of DNA polymorphisms, this provides a high-density set of genetic markers and drastically reduces costs relative to whole-genome sequencing of each progeny strain. If the complete genome sequences of both parental strains are known, RAD-seq markers permit the imputation of essentially the entire genome sequence of each recombinant individual.

Because the plasmid-borne tetrad barcode is flanked by the same restriction sites used in our RAD-seq method, it is captured by the genotyping reads. After RAD-seq, strains arising from the same FACS-sorted plate that share a common plasmid barcode sequence are assigned to the same tetrad (Fig. 3a), a hypothesis that is confirmed by a series of computational quality-control metrics (Online Methods). The small proportion of strains (<5%) that lack a tetrad barcode in their sequence reads can then be assigned to tetrads according to the expectation of 2:2 allele segregation of markers within tetrads (Online Methods). Thus, although the presence of a usable tetrad barcode in the sequencing data simplifies tetrad assignment, strains lacking this sequence can still be assigned to tetrads.

**Missing-marker and complete genotype inference**

Missing data arising from a stochastic lack of sequence coverage is a common problem in large-scale genetic analyses. In these cases, the expected 2:2 segregation of each allele in a tetrad can be leveraged to infer the values of markers that are not confidently assigned (Fig. 3b). Except in the case of rare gene-conversion events, it should always be possible to correctly infer a missing marker from a complete (four-spore) tetrad if the status of that marker in the other three members of the tetrad is known. Similarly, if a marker is missing in two members of a tetrad, it is possible to infer their values from the other two spores 50% of the time (Online Methods). Missing markers can also be inferred probabilistically on the basis of the genetic distances between markers because an untyped marker that is close to a typed marker has a high probability of being derived from the same parent as the typed marker. The use of both genetic distance and the known haplotypes of all spores in the tetrad (Online Methods) can improve the accuracy of inference, sometimes greatly (Supplementary Fig. 4), by incorporating the probability of all possible recombination patterns at the tetrad level. In the pilot crosses below, 5–10% of the final set of allele calls were made using these inference methods.

Tetrad-based genotype inference can be extended to infer the full genome sequence of nonviable spores. In the crosses below, we...
simulated the inference of a dead spore by removing all sequence information for one member of a four-spore tetrad and reconstructing its genotype by inference, and we were able to recover ~98% of the original allele calls with ~98% accuracy. Such an analysis could enable the discovery of synthetic interactions, such as those seen in synthetic lethal screens, resulting from combinations of naturally occurring polymorphisms and in a way that is less limited by strain background and the number of interacting genes than current methods. For example, it should be possible to uncover a synthetic interaction between four genes in two previously uncharacterized wild strains.

Pilot yeast crosses

We tested BEST by generating two crosses comparable in size to those commonly published for complex trait studies in yeast (~100 tetrads). The first cross, between two well-characterized and commonly used laboratory strains, FY4 (ref. 20) and S1278b (ref. 21), showed high (98%) spore viability by manual dissection. The second cross, between the laboratory strain S288c (ref. 22) and the wild oak isolate YPS163 (ref. 23), showed a spore viability (86%) more typical of crosses between genetically distant strains. In the FY4 × S1278b (pilot) cross, 77% of the progeny that passed our quality control filters could be assembled into three- or four-spore tetrads. With this percentage in tetrads’ metric used as a measure of efficiency, these results demonstrate that BEST can be successfully applied to crosses between strains over a range of genetic distances and beyond laboratory strain backgrounds (Table 1). For both crosses, we were also able to generate a dense set of genetic markers (Supplementary Tables 1–6); the larger number of markers in S288c × YPS163 reflects the greater sequence divergence of the two parental strains.

To assess the performance of the method in a large-scale implementation, we repeated the high-viability cross, FY4 × S1278b (full scale). In this experiment, a single researcher performed BEST in 3 h from the start of FACS through the completion of plating ~3,725 tetrads, a scale that would require almost a month of manual dissection. To assess efficiency, we sequenced strains from a subset of the agar plates generated in the experiment. The resulting 4,354 strains produced 1.7 billion high-quality sequencing reads that mapped to the genome sequence or to the tetrad barcode. Of the 3,652 progeny strains that passed our quality-control filters (Supplementary Tables 7 and 8), 63% could be assembled into three- or four-spore tetrads (Table 1). Taken together, these results demonstrate that BEST can be applied on a scale beyond that commonly performed for conventional studies.

**Table 1 | BEST results for three crosses**

| Cross                  | Colonies recovered | Strains passing QC | Strains in 3- or 4-spore tetrads (%) | Number of markers | Mean marker separation (kb) |
|------------------------|--------------------|--------------------|--------------------------------------|-------------------|-----------------------------|
| FY4 × Σ1278b (pilot)   | 385                | 325                | 77                                   | 481               | 21                          |
| S288c × YPS163 (pilot) | 378                | 347                | 70                                   | 831               | 14                          |
| FY4 × Σ1278b (full scale) | 4,354             | 3,652              | 63                                   | 579               | 18                          |

*Parental single-nucleotide polymorphism tables, progeny genotypes and barcode sequences are presented (Supplementary Tables 1–8).*

**DISCUSSION**

The ability to isolate large numbers of meiotically derived recombinant progeny in a manner that retains their sister-spore relationships advances a technique that has remained essentially unchanged for 75 years. With BEST we present an approach that marries the power of conventional genetics with ultrahigh-throughput DNA sequencing. As a result, an individual researcher can prepare separated spores for several thousand tetrads in a few hours, a vast improvement over manual dissection. BEST provides an even greater advantage in crosses with poor sporulation efficiency or low spore viability, for which the burden of manual dissection is increased.

In its current implementation, BEST has a relatively modest reduction in efficiency as compared to manual dissection, possibly due to the loss of spores caused by their adhesion to the glass beads during spreading, to plasmid loss, to increased spore death owing to the mechanical stresses of the process or to failure of sister spores to separate. These decreases in efficiency are easily overcome by the large advance in throughput that BEST affords. We have applied the method in *S. cerevisiae*, the most commonly used microorganism for meiotic mapping. However, minor substitutions of organism-specific reagents, such as alternative sporulation-specific proteins fused to GFP, should make the method readily transferrable to other microorganisms, including those in which meiotic mapping is significantly more labor intensive or currently intractable.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** RAD-seq reads of all progeny presented have been deposited in the NCBI Sequence Read Archive under accession codes ERP002316, ERP002317 and ERP002435.

*Note: Supplementary information is available in the online version of the paper.*

**ACKNOWLEDGMENTS**

We thank B. Cohen, A. Sherman and members of the Dudley lab for helpful discussions, and we thank B. Cohen (Washington University in St. Louis) for providing *S. cerevisiae* strain BC257. We thank D. Nickerson and B. Paeper for assistance with Illumina Hiseq 2000 sequencing and S. Bloom for assistance with Illumina GA IIx sequencing. This work was supported by a US National Institutes of Health–National Human Genome Research Institute (NIH/NHGRI) Genome Scholar/Faculty Transition Award (K22 HG002908) to A.M.D. and a strategic partnership between the Institute for Systems Biology and the University of Luxembourg. We note with sadness the passing of our friend and coauthor T.L. Gilbert.

**AUTHOR CONTRIBUTIONS**

C.L.L., A.C.S., A.S., E.W.J., T.L.G. and A.M.D. contributed to experimental design; A.C.S. and C.L.L. contributed strains and plasmids; A.C.S., A.S., M.H., C.L.L., E.W.J. and T.L.G. contributed to experiments; G.A.C., A.C.S., P.M. and J.L. contributed to data analysis and scripts; and A.C.S., C.L.L., G.A.C. and A.M.D. wrote the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

**Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.**

1. Wingé, Ő. & Laustsen, O. On two types of spore germination, and on genetic segregations in *Saccharomyces*, demonstrated through single-spore cultures. *C.R. Trav. Lab. Carlsberg Ser. Physiol.* 22, 99–116 (1937).
2. Amberg, D.C., Burke, D.J. & Strathern, J.N. Random spore analysis in yeast. CSH Protoc. 2006, pii: pdb.prot4162 (2006).
3. Tong, A.H. et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science 294, 2364–2368 (2001).
4. Pan, X. et al. A robust toolkit for functional profiling of the yeast genome. Mol. Cell 16, 487–496 (2004).
5. Schuldiner, M. et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell 123, 507–519 (2005).
6. Snitkin, E.S. et al. Model-driven analysis of experimentally determined growth phenotypes for 465 yeast gene deletion mutants under 16 different conditions. Genome Biol. 9, R140 (2008).
7. Michelmore, R.W., Paran, I. & Kesseli, R.V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88, 9828–9832 (1991).
8. Ehrenreich, I.M. et al. Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature 464, 1039–1042 (2010).
9. Mackay, T.F., Stone, E.A. & Ayroles, J.F. The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet. 10, 565–577 (2009).
10. Fogel, S., Mortimer, R., Lusnak, K. & Tavares, F. Meiotic gene conversion: a signal of the basic recombination event in yeast. Cold Spring Harb. Symp. Quant. Biol. 43, 1325–1341 (1979).
11. Lander, E.S. Initial impact of the sequencing of the human genome. Nature 470, 187–197 (2011).
12. Coluccio, A. & Neiman, A.M. Interspore bridges: a new feature of the Saccharomyces cerevisiae spore wall. Microbiology 150, 3189–3196 (2004).
13. Chin, B.L., Frizzell, M.A., Timberlake, W.E. & Fink, G.R. FASTER MT: isolation of pure populations of a and α ascospores from Saccharomyces cerevisiae. G3 (Bethesda) 2, 449–452 (2012).
14. Gerke, J.P., Chen, C.T. & Cohen, B.A. Natural isolates of Saccharomyces cerevisiae display complex genetic variation in sporulation efficiency. Genetics 174, 985–997 (2006).
15. Thacker, D., Lam, I., Knap, M. & Keeney, S. Exploiting spore-autonomous fluorescent protein expression to quantify meiotic chromosome behaviors in Saccharomyces cerevisiae. Genetics 189, 423–439 (2011).
16. Yan, Z. et al. Yeast Barcoders: a chemogenomic application of a universal donor-strain collection carrying bar-code identifiers. Nat. Methods 5, 719–725 (2008).
17. Armstrong, K.A., Som, T., Volkert, F.C., Rose, A. & Broach, J.R. Propagation and expression of genes in yeast using 2-micron circle vectors. Biotechnology 13, 165–192 (1995).
18. Mancera, E., Bourgon, R., Brozzi, A., Huber, W. & Steinmetz, L.M. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. Nature 454, 479–485 (2008).
19. Baird, N.A. et al. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3, e3376 (2008).
20. Winston, F., Dollard, C. & Ricupero-Hovasse, S.L. Construction of a set of convenient Saccharomyces cerevisiae strains that are isogenic to S288C. Yeast 11, 53–55 (1995).
21. Dowell, R.D. et al. Genotype to phenotype: a complex problem. Science 328, 469 (2010).
22. Mortimer, R.K. & Johnston, J.R. Genealogy of principal strains of the yeast genetic stock center. Genetics 113, 35–43 (1986).
23. Śniegowski, P.D., Dombrowski, P.G. & Fingerman, E. Saccharomyces cerevisiae and Saccharomyces paradoxus coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. FEBS Yeast Res. 1, 299–306 (2002).
ONLINE METHODS

Yeast strains, media and manipulation. Unless noted, standard media and methods were used for growth and genetic manipulation of yeast24. S. cerevisiae strains used in this study are listed in Supplementary Table 9.

pCL2_BC barcode library construction. The plasmid-based barcode library (pCL2_BC) was constructed in two steps.

First, the pCL2 plasmid backbone was constructed by gap repair in yeast as follows. The yeast 2-micron ADE2 plasmid, pRS422 (ref. 25), was cut with BglII. The ADE2-containing fragment was discarded, and the remaining plasmid backbone was treated with Antarctic phosphatase (New England BioLabs, NEB) to prevent re-ligation and was gel purified. An SPS2::EGFP::kanMX4 cassette14 was amplified from BC257 (gift of B. Cohen) using primers Gap1.1_F and Gap1.1_R (Supplementary Table 10) that bear homology to both the SPS2 genomic and plasmid DNA sequences. The resulting PCR product was cotransformed with the plasmid fragment into yeast. Transformants were selected on YPD-medium plates containing 200 µg/ml G418. G418-resistant clones were scraped and pooled; DNA was prepared and transformed into OneShot TOP10 chemically competent bacteria (Life Technologies). Bacterial transformants were selected on LB-carbenicillin plates and analyzed by restriction digest to identify the repaired plasmid.

Next, a complex library of random barcodes was inserted as follows: 20 nmol of a 200-mer oligo (Supplementary Table 10), including a high-complexity 15-base degenerate region, was amplified by 20 rounds of PCR using Phusion High-Fidelity DNA Polymerase (Thermo Fisher) with BC_F and BC_R primers (Supplementary Table 10) at a final concentration of 20 pM each. The DNA from 24 separate reactions was pooled and ligated to the linearized pCL2 at its unique SmaI site using the In-Fusion HD DNA Polymerase (Thermo Fisher) with BC_F and BC_R primers (Supplementary Table 10) at a final concentration of 20 pM each. The DNA from 24 separate reactions was pooled and ligated to the linearized pCL2 at its unique SmaI site using the In-Fusion HD Cloning System (Clontech). So that complexity was maintained, five ligation reactions were carried out and used for 18 independent bacterial transformations with LB-carbenicillin selection. Each transformation produced an average of 3.5 × 10^4 colonies. A pilot ligation, transformed and screened by X-gal blue-white, showed a low plasmid re-ligation background of ~5%. The transformants were scraped from the plates, resuspended and divided into five separate pools. Plasmid DNA from each pool was extracted and purified using a Qiagen Plasmid Maxi Purification kit (Qiagen).

pCL2_BC barcode complexity determination. The barcode complexity of the pCL2_BC library was assessed by Illumina DNA sequence analysis (Supplementary Table 11). Briefly, 1.5 µg of the plasmid library was fragmented by digestion with MfeI and Sau3A1 (a DAM-methylation insensitive isoschizomer of MboI). Digests were incubated for 2 h at 37 °C in a 20-µl reaction with 2 units of Sau3A1 and 10 units MfeI (NEB), which was followed by heat inactivation at 65 °C for 20 min. The annealed P2 adaptors and four sets of annealed barcoded P1 adaptors (Supplementary Table 10) were then ligated onto the plasmid fragments at room temperature for 20 min in a single 25-µl reaction containing 1 µg digested plasmid, 400 units T4 DNA ligase (NEB), 2.5 µl 10× T4 ligase buffer and 6 µl of a combined P1 (25 nM) and P2 (1 µM) adaptor mix (Supplementary Table 10). T4 ligase was heat inactivated for 20 min at 65 °C, and ligated plasmid DNA was concentrated to 10 µl using a MinElute PCR Purification Kit. The DNA was size selected and extracted, as below. Approximately 10 ng of the purified plasmid DNA library was enriched with a PCR reaction and sequenced in a single flow cell lane of a Genome Analyzer IIx (Illumina).

Generation of barcoded yeast tetrads. Heterozygous diploids resulting from crosses between two parental strains were grown to ~2 × 10^7 cells per ml and transformed with ~2 µg of the pCL2_BC barcoded plasmid library using a standard protocol26 modified to include 8% DMSO in the transformation mix. After the 30 min 42 °C heat-shock step, the transformed cells were gently washed with 1 ml ofYPD, resuspended in 1 ml ofYPD and allowed to recover by sitting at room temperature for 3 h. We then selected transformants by plating 200 µl of the recovered culture perYPD + 200 µg/ml G418 plate for a total of five plates per transformation. This protocol yielded a library of ~10^5 single colonies. Transformants were pooled by scraping the plates. A portion of the pool was saved as a frozen glycerol stock to set up sporulation cultures at a later date. All crosses described in this work were performed with frozen stocks revived by an overnight growth in liquidYPD + 200 µg/ml G418, washed and transferred to liquid sporulation medium3 containing 200 µg/ml G418. Sporulation was performed at room temperature with agitation and monitored daily. Cultures were deemed ready for sorting when sporulation had reached the point of completing well-formed tetrads without substantial numbers of dyads. In the crosses described here, spore separation was improved by allowing spores to sit at room temperature (without agitation) for an additional 7–10 d (Supplementary Note).

Tetrad isolation by FACS. Tetrads were isolated from the sporulation culture by FACS with a FACSAria II equipped with an automated cell deposition unit (BD Biosciences). GFP fluorescence was detected using the 488-nm laser and 530/30 filter. To achieve a reproducibly high proportion of tetrads, we implemented a series of gating steps. We selected a narrow width of the FSC and SSC signals, which, while permitting a large range of FSC and SSC heights, filtered out events containing cell or media debris as well as those containing multiple cells per droplet (Supplementary Fig. 1a,b). A GFP vs. FSC area gate was used to identify fluorescent (and therefore sporulated) cells (Supplementary Fig. 1c). The population selected by these steps consisted of two subpopulations: one subpopulation was composed of clumps of tetrads and tetrads with a small bud attached, and the other subpopulation was primarily composed of isolated tetrads. These subpopulations were distinguished from each other on the basis of their FSC signal. The clumps and budded tetrads had a higher FSC than the isolated tetrads, though the distribution of FSC in these two subpopulations did overlap as indicated by the overlapping peaks in Supplementary Figure 1d. To enrich for isolated tetrads, we set a final gate to include events with a low FSC (Supplementary Fig. 1d). During gate assignment, tetrad recovery was assessed by sorting 1,000 events onto a microscope slide and manually counting tetrads.

Spore separation by on-plate digestion and glass-bead spreading. To prevent spore loss during liquid handling, tetrads were sorted directly onto YPD + 200 µg/ml G418 agar plates with a 25-µl drop of 1 mg/ml Zymolyase in 0.7 M sorbitol on
samples in one sequencing library that was then enriched with a trated and size selected on agarose gels. Fragments between 150 and the MboI restriction enzyme–compatible overhang. After (a 6-nucleotide sequence based on Illumina TruSeq indexes 1–12) the Illumina PCR reverse primer sequence, a second barcode the P2-BC adaptors (Supplementary Note). After centrifugation, 250 µl of superna- 

**RAD-seq progeny genotyping.** Yeast genomic DNA was isolated for RAD-seq as follows. We used 96-well plates to seed 0.5-ml cultures in 2-ml deep-well plates of YPD with 200 µg/ml of G418. These were grown overnight at 30 °C on a Vibra Translator electro-magnetic shaker (Union Scientific). Cells were pelleted at 1,000g for 5 min and resuspended with 400 µl of lysis buffer (10 mM Tris (pH 8.0), 1 mM EDTA, 2% Triton X-100, 1% SDS, 100 mM NaCl), and the suspension was transferred to a lysis rack (Costar 4413, 1.2-ml tubes) containing 0.5-mm glass beads. The racks were processed at 1,300 r.p.m. for 2 min in a bead beater (Geno/Grinder 2010, Spex Sample Prep). After centrifugation, 250 µl of superna- 
thant were transferred to a 96–deep-well block, and 250 µl of 6 M guanidine HCl DNA–binding buffer were added. The mixture was then transferred to the 96-well DNA-binding plates (Pall AcroPrep #8032). Centrifugation, washing and elution procedures followed the manufacturer’s protocol.

RAD-seq was carried out as previously described. Briefly, yeast genomic DNA was fragmented by restriction enzyme digestion with MfeI and MboI. P1 and P2 adaptors were then ligated onto the fragments. The P1 adaptor contains the Illumina PCR forward sequencing primer sequence followed by 1 of 48 unique 4-nucleotide barcodes and then the MfeI overhang sequence. The two pilot crosses, FY4 × Σ1278B (pilot) and S288c × YPS163 (pilot), were sequenced in house on a Genome Analyzer IIx (Illumina) and used the P2 adaptor sequences (Supplementary Table 10), which allowed a multiplexing of 48 strains per lane. The larger cross, FY4 × Σ1278B (full scale), required a higher level of multiplexing on a HiSeq 2000 (Illumina) and therefore used the P2-BC adaptors (Supplementary Table 10), which contain the Illumina PCR reverse primer sequence, a second barcode (a 6-nucleotide sequence based on Illumina TruSeq indexes 1–12) and the MboI restriction enzyme–compatible overhang. After ligation, the barcoded ligation products were pooled, concen- 
trated and size selected on agarose gels. Fragments between 150 and 500 base pairs were extracted from the gel. Gel-extracted DNA was further pooled to multiplex 48 uniquely barcoded samples in one sequencing library that was then enriched with a PCR reaction using Illumina PCR forward and reverse primers (Supplementary Table 10). For the full-scale cross, 200 ng of a set of 12 libraries, each with a different P2 barcode, were combined and concentrated by QIAquick PCR Purification column (Qiagen). This combined, dual-indexed pool of 576 yeast strains was then sequenced in a single Illumina flow-cell lane. Dual-indexed sequencing runs were performed on the HiSeq 2000 (Northwest Genomics Center, University of Washington) using 50-base-pair single-end reads and the second index read only. Sequencing runs performed on the Genome Analyzer IIx for the pilot studies generated 40-base-pair single-end reads.

**Strain genotype and tetrad determination.** For each lane of sequencing, raw read sequences were split into pools on the basis of their P1 barcode sequences and, for the larger cross, also their P2 barcodes. Reads with unexpected strain barcodes or with bar- 
codes having Phred (−10 log10Perror) quality scores less than 20 or ambiguous (“N”) calls at any barcode base were discarded. Reads with more than 2 N calls in the body of the sequencing read were also discarded. The P1 barcodes were then removed from the read sequences. In each of the resulting strain-specific pools of reads, the sequences were searched for reads carrying the plasmid (tetrad) barcode. Tetrad barcodes were identified using the pattern ‘read start–NNNNNTGCCGACCC–barcode–GCAGG’, with the barcode restricted to a length of 11–19 nucleotides. A single mismatch or nucleotide deletion was allowed in the pattern match outside the barcode. The most frequent barcode sequence was then identified from the set of all plasmid barcode reads coming from each strain.

The strain-specific read pools were then used to infer the geno- 
types of the progeny strains. From each strain pool, the sequence reads were aligned to one parental genome. In both crosses, one parent was S288c/FY4, therefore the S288c genome sequence was used as a reference in both crosses. Read alignment was carried out using BWA (v.5.8) allowing six mismatches and using quality trimming (threshold of Phred = 20). The SAMtools (v.0.1.18) mpileup (–u –C 50) and bcftools view (–c) commands were then used to generate a variant call format (VCF) file for each strain. A table of SNPs (Supplementary Tables 1 and 2) between the two parental genomes was generated by aligning the parental genomes using Mummer (v.3.22). The bases called in the VCF file were then compared to the expected parent 1 (P1) and parent 2 (P2) alleles from this SNP file to convert the base calls into P1 or P2 allele calls. When the count of one parental allele was at least fivefold higher than the count of the other parental allele, a P1 or P2 allele assignment was made. Otherwise, the allele was defined as mixed/heterozygous.

Marker quality filtering was then carried out: markers were removed unless they were called at least 30% as frequently as the most called marker and unless the ratio P1/(P1 + P2) lay between 0.3 and 0.7. The strains were then grouped into tetrads on the basis of common tetrad barcode sequences, and strain quality-control filters were applied. Strains derived from each plate of 25 sorted tetrads were analyzed independently to reduce the risk of encountering more than one tetrad with the same plasmid barcode. Strains with too few reads (<50,000) or too many mixed/heterozygous sites (>20%) were flagged and removed from further processing and analysis. Duplicate strains were identified (>90% identical allele calls across >100 markers), and the lower-coverage strain
was removed. Strains with unique barcodes, for which the total number of tetrad barcode reads was <30 or the most common barcode sequence comprised less than 80% of all observed barcode sequences, had their tetrad barcode removed and were relabeled as unbarcoded. Among the remaining tetrads, any ‘tetrads’ with more than five spores, any four-sporo tetrads with >20% 3:1 or 4:0 segregation patterns and any three-sporo tetrads with >10% 3:0 segregation patterns were dissolved. All strains in dissolved tetrads were relabeled as unbarcoded. An attempt was then made to create three-sporo tetrads out of the group of unbarcoded strains. This was done by examining the frequency of allele missegregation (3:0) among each subset of three unbarcoded strains. Any set showing <10% 3:0 segregation over at least 100 sites was identified as a tetrad. After this process, an attempt was made to add the remaining unbarcoded strains to each two- or three-sporo tetrad. Again, this was done by examining the frequency of abnormal allele segregation in each existing tetrad when each single unbarcoded strain was added and requiring <20% missegregation over at least 100 informative sites. Finally, for each two- or three-sporo tetrad, the genotypes of the missing/dead spores were initialized to 0 (missing data) at all marker positions.

Missing data inference. After we generated an initial haplotype for each strain and assigned strains to tetrads, the tetrad organization was used to infer missing allele calls (including mixed calls). This included inferring alleles for any missing/dead spores. On the basis of the expected 2:2 allele segregation pattern in a tetrad, any marker called as P1 or P2 in two of the strains allows assignment of missing marker alleles in either of the other strains with almost complete confidence.

After the tetrad inference step, the data were then used to calculate a genetic map for each cross using R/qtl (v.1.21–1.22). This map was then used to carry out linkage-based inference of remaining missing values on a tetrad basis. Specifically, missing alleles were inferred from the relative probability of all possible local crossover patterns within the tetrad, anchored at flanking positions with allele calls in all four spores. Allele calls with probabilities greater than 0.99 were then accepted.

Linkage-based inference methods can suffer from two sources of error: gene conversion and markers with abnormal linkage patterns. To address the latter, haplotypes were first generated without linkage-based inference and then analyzed using R/qtl. Markers with abnormal linkage patterns (for example, markers not linked to any other marker, linked to different chromosomes or linked to a distant region of the same chromosome) were identified and flagged. These flagged markers were then removed, and the final haplotypes and genetic map were regenerated.

24. Rose, M.D., Winston, F.M. & Hieter, P. Methods in Yeast Genetics: A Laboratory Course Manual (Cold Spring Harbor, 1990).
25. Brachmann, C.B. et al. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14, 115–132 (1998).
26. Gietz, R.D. & Woods, R.A. Transformation of yeast by lithium acetate/ single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 350, 87–96 (2002).
27. Lorenz, K. & Cohen, B.A. Small- and large-effect quantitative trait locus interactions underlie variation in yeast sporulation efficiency. Genetics 192, 1123–1132 (2012).
28. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
29. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
30. Kurtz, S. et al. Versatile and open software for comparing large genomes. Genome Biol. 5, R12 (2004).