

Polyploidy can drive rapid adaptation in yeast

Anna M. Selmecki^{1,2,3,†}, Yosef E. Maruvka^{4,5}, Phillip A. Richmond^{6,7}, Marie Guillet^{1,2,3}, Noam Shores⁸, Amber L. Sorenson^{6,7}, Subhajyoti De^{9,10,11}, Roy Kishony^{12,13}, Franziska Michor^{4,5}, Robin Dowell^{6,7} & David Pellman^{1,2,3,14}

Polyploidy is observed across the tree of life, yet its influence on evolution remains incompletely understood^{1–4}. Polyploidy, usually whole-genome duplication, is proposed to alter the rate of evolutionary adaptation. This could occur through complex effects on the frequency or fitness of beneficial mutations^{2,5–7}. For example, in diverse cell types and organisms, immediately after a whole-genome duplication, newly formed polyploids missegregate chromosomes and undergo genetic instability^{8–13}. The instability following whole-genome duplications is thought to provide adaptive mutations in microorganisms^{13,14} and can promote tumorigenesis in mammalian cells^{11,15}. Polyploidy may also affect adaptation independently of beneficial mutations through ploidy-specific changes in cell physiology¹⁶. Here we perform *in vitro* evolution experiments to test directly whether polyploidy can accelerate evolutionary adaptation. Compared with haploids and diploids, tetraploids undergo significantly faster adaptation. Mathematical modelling suggests that rapid adaptation of tetraploids is driven by higher rates of beneficial mutations with

stronger fitness effects, which is supported by whole-genome sequencing and phenotypic analyses of evolved clones. Chromosome aneuploidy, concerted chromosome loss, and point mutations all provide large fitness gains. We identify several mutations whose beneficial effects are manifest specifically in the tetraploid strains. Together, these results provide direct quantitative evidence that in some environments polyploidy can accelerate evolutionary adaptation.

To determine how polyploidy affects the rate of adaptation, we performed hundreds of independent passaging experiments in a poor carbon-source medium (raffinose, Fig. 1a), comparing isogenic haploid (1N), diploid (2N), and tetraploid (4N) strains (Extended Data Fig. 1 and Extended Data Table 1). The evolution experiments were performed as competitions between equal numbers of cells of the same ploidy expressing cyan fluorescent protein (CFP) or yellow fluorescent protein (YFP)^{17,18}, where the acquisition and spread of beneficial mutations is visualized by divergence from a 50:50 ratio of CFP- and YFP-expressing cells (Fig. 1b). The rate of adaptation was determined by measuring the change in

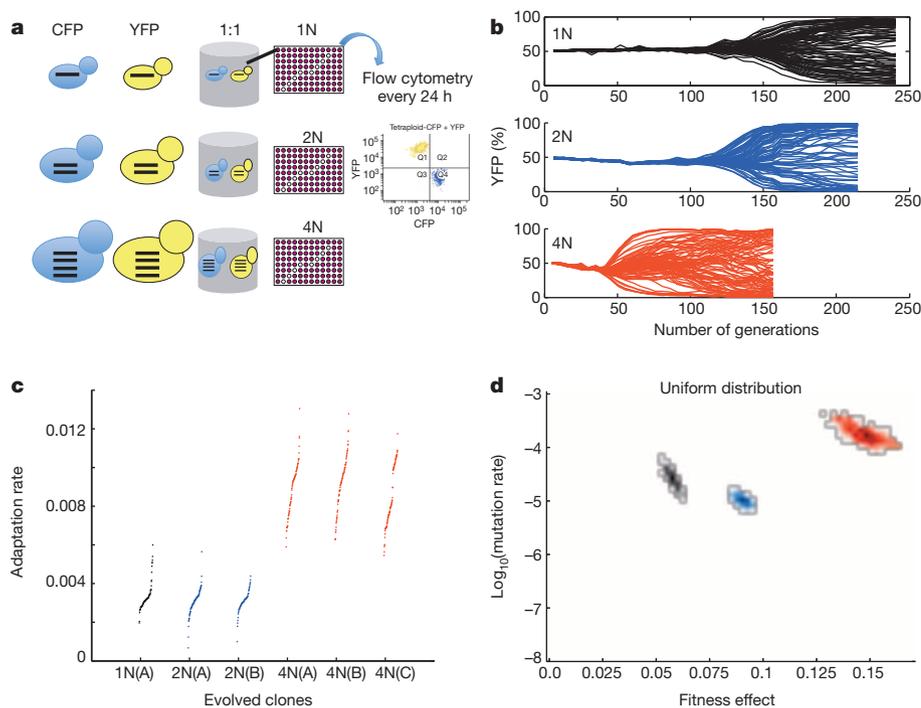


Figure 1 | Rapid spread of beneficial mutations in tetraploid yeast. **a**, Schematic diagram of the evolution experiment. **b**, Flow cytometry analysis of isogenic haploid (black), diploid (blue), and tetraploid (red) populations during adaptation to raffinose medium. Each line is the percentage of YFP cells in an independent population of YFP and CFP cells. Here and below, data from haploids is black, from diploids is blue, and from tetraploids is red. **c**, The adaptation rate of the evolved clones relative to the diploid ancestor after 250 generations. Data points are the average rate of adaptation (change in fitness between generation 250 and generation zero, divided by 250 generations) of two replicate fitness measurements for the evolved clones. Clones from replicate evolution experiments (A, B, and C) are indicated. The tetraploids acquired significantly more fitness in the same number of generations compared with the haploids and diploids (*t*-test, $P < 1 \times 10^{-10}$). **d**, Estimates from the branching evolution model of the best-fit value of the selection coefficient and beneficial mutation rate of each ploidy experiment, and their error range, determined using a uniform distribution of acquired mutations (other distributions are analysed in Extended Data Fig. 2c, d, and the equivalence principle model is analysed in Extended Data Fig. 2e). Error ranges were obtained by parametric bootstrap of 1,000 independent realizations (Methods).

¹Department of Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. ²Department of Cell Biology, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02215, USA. ³Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA. ⁴Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue Boston, Massachusetts 02215, USA. ⁵Department of Biostatistics, Harvard School of Public Health, 158 Longwood Avenue, Boston, Massachusetts 02215, USA. ⁶BioFrontiers Institute, University of Colorado at Boulder, 3415 Colorado Avenue, Boulder, Colorado 80303, USA. ⁷Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, 347 UCB, Boulder, Colorado 80309, USA. ⁸Broad Institute, 415 Main Street, Cambridge, Massachusetts 02142, USA. ⁹Department of Medicine, University of Colorado School of Medicine, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹⁰Department of Biostatistics and Informatics, Colorado School of Public Health, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹¹Molecular Oncology Program, University of Colorado Cancer Center, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹²Department of Systems Biology, Harvard Medical School, 200 Longwood Ave, Boston, Massachusetts 02115, USA. ¹³Department of Biology, Technion - Israel Institute of Technology, Haifa, 32000, Israel. ¹⁴Department of Pediatric Hematology/Oncology, Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA. †Present address: Department of Medical Microbiology and Immunology, Creighton University School of Medicine, 2500 California Plaza, Omaha, Nebraska 68178, USA.

fitness relative to the diploid ancestor over time (Methods). Over 250 generations, the tetraploids adapted at a rate that was significantly faster than haploids or diploids (Fig. 1c; t -test, $P < 1 \times 10^{-10}$, Methods). This faster rate of adaptation in tetraploids may be due to a higher rate of beneficial mutations, higher fitness effects of the acquired mutations, or both.

To gain insight into the rapid adaptation of tetraploids, we applied two complementary mathematical modelling approaches (see Methods). First, we use a model based on a branching evolutionary process¹⁹, designed to mimic closely the divergence experiments. At each time-step, a cell is chosen at random to die or to divide, with a probability corresponding to its fitness. Mutations arise with rate μ . If a mutation occurs, the fitness of the daughter cell may change and the fitness increase is then chosen from a fitness distribution. Second, we use the 'equivalence principle' model¹⁷, which focuses on beneficial mutations that establish in the population, and estimates that these mutations confer a single effective fitness advantage. Proliferation of clonal subpopulations under this model is deterministic. These simplifications are relevant to examples of high clonal interference, and are therefore only appropriate when the population size is large or when the beneficial mutation rate is high²⁰. In both models, we assume no epistasis; the fitness change is independent of whether the cell already had one or more mutations. Furthermore, there is no restriction on the number of cells that acquire beneficial mutations, thus allowing clonal interference to occur^{18,21}.

Both modelling approaches led to the same general conclusion: the rapid adaptation of tetraploids results from both more frequent beneficial mutations and stronger fitness effects (Extended Data Fig. 2 and Methods). For the branching evolution model, these conclusions are independent of the assumed distribution of beneficial mutations, although there are differences in the magnitude of the best-fit values that are expected from the shape of the chosen distribution (Fig. 1d and Extended Data Fig. 2). Moreover, the conclusions are insensitive to the inclusion of deleterious mutations in the model (Extended Data Fig. 3 and Methods).

To evaluate these conclusions experimentally, we performed whole-genome sequencing (WGS) to compare the frequency of mutations in the 1N and 4N ancestors with 74 evolved clones. In total, we identified 240 *de novo* sequence variants (single nucleotide polymorphisms (SNPs) and small insertions/deletions): 45 from the 1N-, 69 from the 2N-, and 126 from the 4N-evolved clones, an average of 2.05, 2.87, and 4.5 variants respectively per cell type (Supplementary Table 1). We observed significantly more variants per 4N clone than per 1N- and 2N-evolved clones (Fig. 2a; t -test, $P < 1 \times 10^{-4}$ and $P = 0.0040$, respectively). Note that these results are not a direct measurement of the mutation rate or beneficial mutation rate (μ), but rather the total number of mutations acquired during the experiment (see Supplementary Discussion).

Sequence variants frequently occurred in genes encoding proteins in the *Snf3/Rgt2* glucose-signalling pathway (*SNF3*, *RGT2*, *MTH1*, *RGT1*), as expected from previous yeast evolution experiments under carbon-source limitation^{22–24}. Several independent mutations in these genes resulted in either identical base-pair changes or altered the same amino acid (Supplementary Table 1). Non-synonymous *SNF3* mutations were identified in all ploidy types, whereas loss-of-function mutations in *MTH1* were observed most frequently in the 1N-evolved clones.

In addition to WGS, we used a combination of flow cytometry, microarray comparative genome hybridization (aCGH), and quantitative PCR (qPCR) to measure the frequency of DNA copy number variations (CNVs) in the evolved clones. The only CNVs that arose in all three ploidy types was amplification of two adjacent genes encoding the high-affinity hexose transporters, *HXT6* and *HXT7*, a frequently identified beneficial mutation in low glucose environments^{18,22,23}. The *HXT6/7* amplification was significantly more common in 2N- and 4N-evolved clones than in 1N clones (t -test, $P = 0.005$ and $P = 1 \times 10^{-4}$, respectively, Methods), which may be due to negative epistasis between *HXT6/7* amplification and mutations in 1N cells, such as those in *MTH1* (ref. 24).

Additional CNVs, including recurrent chromosome aneuploidy, were detected only in the 4N-evolved clones. With the exception of a small

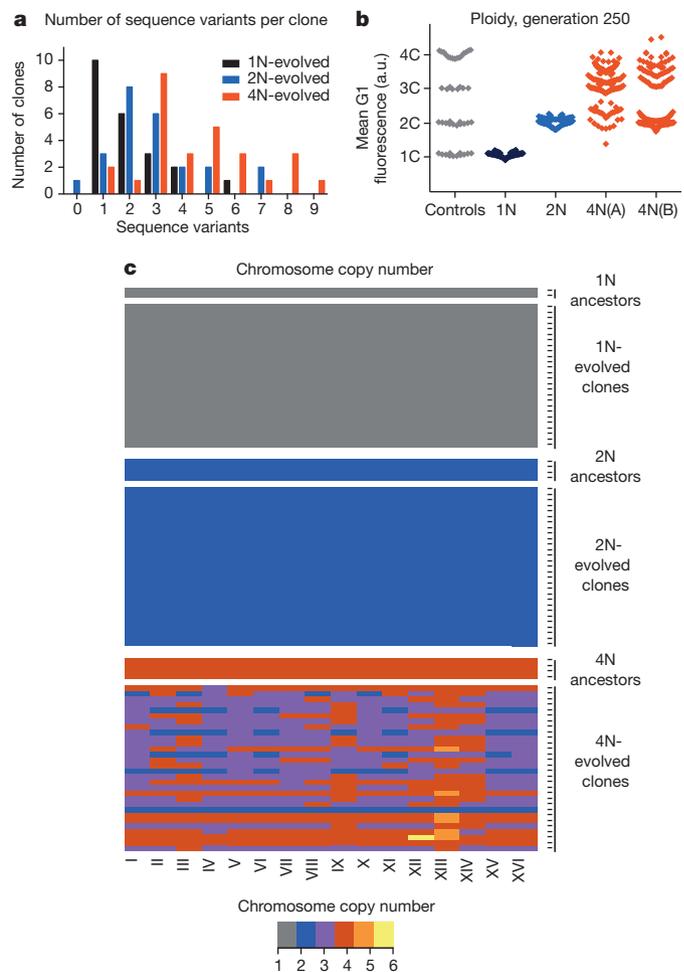


Figure 2 | Tetraploid clones acquire frequent sequence variants, large-scale ploidy shifts and recurrent whole chromosome aneuploidy during adaptation.

a, The number of sequence variants per clone was determined with WGS of 74 evolved clones (22 haploid, 24 diploid, and 28 tetraploid clones; Supplementary Table 1). The difference between tetraploids and haploids or diploids was significant (t -test, $P < 1 \times 10^{-4}$ and $P = 0.004$, respectively). **b**, DNA content (1C–4C) of evolved clones at generation 250, measured as the mean G1 propidium iodide fluorescence for each evolved clone ($n = 88$ haploid and diploid evolved clones, $n = 176$ tetraploid evolved clones) in arbitrary units (a.u.). For reference, the DNA content of ancestral, control strains (1N, 2N, 3N, and 4N) is shown in grey. **c**, Heat map of chromosome copy number data obtained from aCGH and WGS for the ancestral and evolved 1N, 2N, and 4N clones at generation 250; colour key at bottom. See Extended Data Figs 4–6, and Supplementary Table 2 for all individual clones.

segmental amplification in one 2N-evolved clone, there were no CNVs or aneuploidy in the ancestral strains or the 1N- and 2N-evolved clones (Fig. 2b, c and Extended Data Figs 4 and 5). By contrast, many 4N clones underwent large shifts in ploidy (Fig. 2b) and all but two of the 4N-evolved clones were aneuploid at generation 250 ($n = 30$; Fig. 2c, Extended Data Fig. 6 and Supplementary Table 2). These alterations included large segmental aneuploidies with breakpoints at loci of transposable elements (Extended Data Fig. 7a). Pairwise patterns of chromosome copy number alterations were observed, indicating that there is a strong copy number relationship between certain pairs of chromosomes (Extended Data Fig. 7b, c). Notably, increased copy number of chromosome XIII was significantly more common than all other aneuploidies (Extended Data Fig. 7d; Cochran–Armitage test, $P < 1 \times 10^{-7}$). These chromosome-level alterations were present early, at the time of CFP/YFP marker divergence in the 4N populations (\sim generation 45; Extended Data Fig. 8). Therefore, 4N-evolved clones had a higher frequency and greater diversity of mutations, supporting the inference from our mathematical

model that 4N-evolved clones have a relatively higher beneficial mutation rate.

Next, we determined the effects of specific mutations on the fitness of the ancestral cells of differing ploidy. We first determined whether chromosome XIII gain contributed directly to the rapid adaptation of 4N cells. Isogenic 2N and 4N strains, with and without an extra copy of chromosome XIII, were generated (Methods and Extended Data Fig. 9). The increased copy number of chromosome XIII provided a significant fitness increase to 4N strains specifically in raffinose medium relative to the 2N ancestor (Fig. 3a; *t*-test, $P < 1 \times 10^{-4}$), not in glucose (Fig. 3b). This was not a general effect of aneuploidy because the gain of a different chromosome, chromosome XII, had the opposite effect on fitness (Fig. 3a). In striking contrast to 4N cells, chromosome XIII trisomy was not beneficial to 2N strains in raffinose medium and decreased fitness of 2N cells in glucose. Although increased fitness due to whole and segmental chromosome gain is known to occur during adaptation^{13,14,23}, to our knowledge this is the first observation of a ploidy-specific fitness advantage for an aneuploid chromosome. Thus, aneuploidy, acquired through high rates of mitotic errors, is one way that 4N cells can acquire more beneficial mutations with higher fitness effects.

We also characterized how ploidy impacts the fitness effect of recurrently isolated mutations in *SNF3*, a gene encoding a plasma-membrane glucose sensor²⁵. We identified SNPs that changed the codon for the same amino acid in the ninth transmembrane domain of Snf3p (G439E, G439V, G439R; Supplementary Table 1), and increased *HXT* expression in raffinose (Methods)²⁵. By analysing the fitness of isogenic *SNF3*-G439E strains differing only by ploidy, we found that *SNF3*-G439E had a dominant, raffinose-specific, beneficial effect that was relatively stronger in the 4N strain (Fig. 3c, d; *t*-test, $P < 1 \times 10^{-4}$).

Ploidy-specific effect size of mutations could be an intrinsic property of polyploidy, as was recently suggested in plants¹⁶, or it could be related

to the fitness of the 4N ancestor relative to the 1N and 2N ancestors^{26,27}. To address the impact of initial fitness generally, we isolated 48 clones from the 4N evolution experiments at generation 250 (4N₂₅₀) with fitness values equal to the 2N ancestor (competitive fitness difference < 0.05), and determined the speed of their next adaptive step. We compared the fitness acquired by the selected 4N₂₅₀ clones after an additional 250 generations with that of 2N clones evolved for 250 generations (2N₂₅₀, $n = 160$). Despite comparable starting fitness, the 4N-derived clones still underwent more rapid adaptation and achieved significantly higher fitness. This occurred irrespective of large-scale shifts in ploidy: 29% of the 4N₅₀₀ clones maintained a ploidy of 3N–4N and acquired higher fitness than the 2N₂₅₀ clones (Fig. 4, Kolmogorov–Smirnov test, $P < 1 \times 10^{-6}$); 71% of the 4N₅₀₀ clones underwent chromosome loss to become near-diploid and acquired even higher fitness relative to 2N₂₅₀ clones (Fig. 4, Kolmogorov–Smirnov test, $P < 1 \times 10^{-8}$). Thus, the rapid adaptation of tetraploid cells was at least partly independent of their initial fitness.

Here, we measured the acquisition and spread of beneficial mutations in isogenic yeast populations that differed only by ploidy. Mathematical modelling enabled us to infer parameters driving the evolutionary dynamics of these strains and indicated that in a poor carbon-source environment, polyploidy increases the rate and fitness effects of the acquired mutations. Polyploidy increased the genetic diversity of the population. We identified examples of mutations that are selectively beneficial in polyploid strains, including whole chromosome aneuploidy. Because aneuploidy itself is mutagenic²⁸, the high rates of aneuploidy induced by whole-genome duplication may further increase the rate at which beneficial mutations are acquired. If these mutations are beneficial at lower ploidy states, then the long-term benefit of polyploidy will be preserved, even if polyploidy is transient during adaptation. Indeed, 4N-evolved clones that became near-diploid had higher fitness than the 2N-evolved clones. Moreover, although we only studied one environmental condition, polyploidy buffers the effects of partly recessive deleterious mutations^{12,29}, which in principle can then accumulate², providing a reservoir of mutations that might be adaptive in a new environment. Interestingly, the evolved tetraploid karyotypes closely resemble the polyploid and aneuploid karyotypes of fermentation, industrial, baking, natural desert isolates³⁰, and antifungal drug-resistant yeasts¹⁴, consistent with a role for polyploidization events during adaptation to

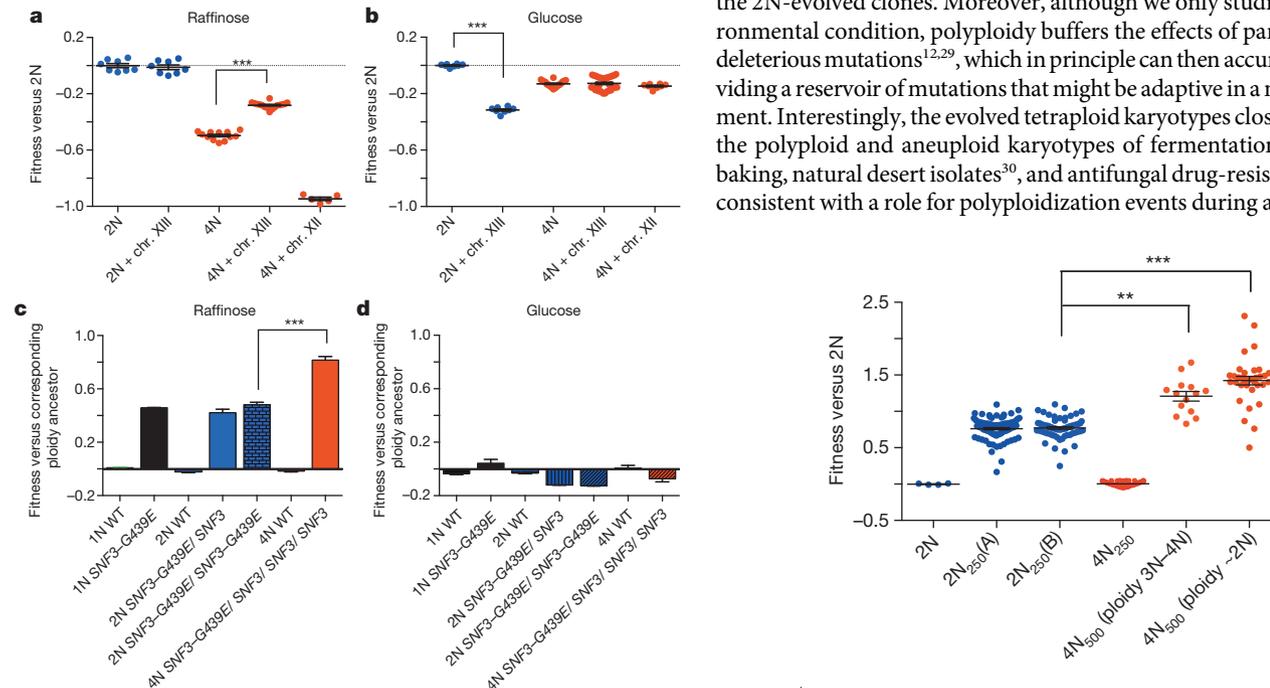


Figure 3 | Ploidy-specific fitness effects for certain beneficial mutations. Gain of chromosome (chr.) XIII is beneficial to tetraploid cells grown in raffinose medium but not for diploids. Shown is the fitness of isogenic wild-type 2N and 4N strains, with or without chromosome XIII gain, relative to the 2N ancestor in raffinose (a) or glucose (b) medium. Error bars, mean \pm s.e.m. of four individual clones and two or three technical replicates. c, Competitive fitness of engineered isogenic strains of the indicated ploidy and genotype, relative to the corresponding ploidy ancestor, in raffinose and d, glucose medium. Error bars, mean \pm s.e.m. of three independent *SNF3*-G439E transformants of each ploidy type, *t*-test $***P < 1 \times 10^{-4}$.

Figure 4 | Rapid adaptation of tetraploids normalized for initial fitness. Fitness of 2N and 4N clones relative to the 2N ancestor. Evolved tetraploids (4N₂₅₀) with fitness equivalent to the diploid ancestors were identified and passed for another 250 generations to generate 4N₅₀₀ clones ($n = 48$). The fitness of these 4N₅₀₀ clones was then compared with the fitness of evolved diploids after 250 generations (2N₂₅₀, $n = 160$, replicate experiments A and B). 4N₅₀₀ clones reached a higher fitness than 2N₂₅₀ clones, irrespective of whether the 4N₅₀₀ clones maintained a 3N–4N DNA content ($n = 14$, Kolmogorov–Smirnov test, $**P < 1 \times 10^{-6}$) or underwent large-scale chromosome loss to a near diploid chromosome content ($n = 34$, Kolmogorov–Smirnov test, $***P < 1 \times 10^{-8}$). Error bars, mean \pm s.e.m.

these stressful environments. Thus, the genetic plasticity of polyploid cells, together with ploidy-specific beneficial effects, can facilitate rapid adaptation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 January; accepted 29 December 2014.

Published online 2 March 2015.

- Ohno, S., Wolf, U. & Atkin, N. B. Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169–187 (1968).
- Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
- Semon, M. & Wolfe, K. H. Consequences of genome duplication. *Curr. Opin. Genet. Dev.* **17**, 505–512 (2007).
- Hufton, A. L. & Panopoulou, G. Polyploidy and genome restructuring: a variety of outcomes. *Curr. Opin. Genet. Dev.* **19**, 600–606 (2009).
- Paquin, C. & Adams, J. Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature* **302**, 495–500 (1983).
- Anderson, J. B., Sirjusingh, C. & Ricker, N. Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* **168**, 1915–1923 (2004).
- Zorgo, E. *et al.* Ancient evolutionary trade-offs between yeast ploidy states. *PLoS Genet.* **9**, e1003388 (2013).
- Mayer, V. W. & Aguilera, A. High levels of chromosome instability in polyploids of *Saccharomyces cerevisiae*. *Mutat. Res.* **231**, 177–186 (1990).
- Bennett, R. J. & Johnson, A. D. Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *EMBO J.* **22**, 2505–2515 (2003).
- Gerstein, A. C., Chun, H. J., Grant, A. & Otto, S. P. Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. *PLoS Genet.* **2**, e145 (2006).
- Fujiwara, T. *et al.* Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
- Storchova, Z. *et al.* Genome-wide genetic analysis of polyploidy in yeast. *Nature* **443**, 541–547 (2006).
- Rancati, G. *et al.* Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell* **135**, 879–893 (2008).
- Selmecki, A., Forche, A. & Berman, J. Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* **313**, 367–370 (2006).
- Zack, T. I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nature Genet.* **45**, 1134–1140 (2013).
- Chao, D. Y. *et al.* Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. *Science* **341**, 658–659 (2013).
- Hegreness, M., Shores, N., Hartl, D. & Kishony, R. An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* **311**, 1615–1617 (2006).
- Kao, K. C. & Sherlock, G. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nature Genet.* **40**, 1499–1504 (2008).
- Haccou, P., Jagers, P., Vatutin, V. A. & International Institute for Applied Systems Analysis. *Branching Processes: Variation, Growth, and Extinction of Populations* 82–106 (Cambridge Univ. Press, 2005).
- Barrett, R. D., M'Gonigle, L. K. & Otto, S. P. The distribution of beneficial mutant effects under strong selection. *Genetics* **174**, 2071–2079 (2006).
- Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571–574 (2013).
- Brown, C. J., Todd, K. M. & Rosenzweig, R. F. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol. Biol. Evol.* **15**, 931–942 (1998).
- Gresham, D. *et al.* The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303 (2008).
- Kvitek, D. J. & Sherlock, G. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**, e1002056 (2011).
- Ozcan, S., Dover, J., Rosenwald, A. G., Wolf, S. & Johnston, M. Two glucose transporters in *Saccharomyces cerevisiae* are glucose sensors that generate a signal for induction of gene expression. *Proc. Natl Acad. Sci. USA* **93**, 12428–12432 (1996).
- Barrick, J. E., Kauth, M. R., Strelieff, C. C. & Lenski, R. E. *Escherichia coli* rpoB mutants have increased evolvability in proportion to their fitness defects. *Mol. Biol. Evol.* **27**, 1338–1347 (2010).
- Kryazhimskiy, S., Rice, D. P., Jerison, E. R. & Desai, M. M. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* **344**, 1519–1522 (2014).
- Sheltzer, J. M. *et al.* Aneuploidy drives genomic instability in yeast. *Science* **333**, 1026–1030 (2011).
- Dewhurst, S. M. *et al.* Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Disc.* **4**, 175–185 (2014).
- Ezov, T. K. *et al.* Molecular-genetic biodiversity in a natural population of the yeast *Saccharomyces cerevisiae* from “Evolution Canyon”: microsatellite polymorphism, ploidy and controversial sexual status. *Genetics* **174**, 1455–1468 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the Howard Hughes Medical Institute, the National Institutes of Health (R37 GM61345), the G. Harold & Leila Y. Mathers Charitable Foundation, the Dana-Farber Cancer Institute Physical Sciences-Oncology Center (U54CA143798), the Boettcher Foundation's Webb-Waring Biomedical Research Program, the National Science Foundation (NSF 1350915), the National Institutes of Health (R01 GM081617), and an American Cancer Society Postdoctoral Fellowship.

Author Contributions A.M.S., M.G., N.S., R.K., and D.P. contributed to the overall study design. A.M.S. and M.G. performed the experiments. Y.E.M. implemented the mathematical modelling with contributions from N.S., R.K., and F.M. A.M.S., P.A.R., and A.L.S. generated WGS libraries; P.A.R. developed the sequencing pipeline and analysed the WGS data with help from A.M.S. and A.L.S. under the supervision of R.D.D. Data were analysed by A.M.S., Y.E.M., P.A.R., M.G., N.S., A.L.S., S.D., F.M., and D.P. The manuscript was written primarily by A.M.S., Y.E.M., and D.P. with contributions from the other authors.

Author Information All aCGH data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus database under accession number GSE51017 and all WGS data have been deposited in the National Center for Biotechnology Information Sequence Read Archive database under accession number SRP047435. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.S. (AnnaSelmecki@creighton.edu) or D.P. (david_pellman@dfci.harvard.edu).

METHODS

Batch culture evolution experiment. All *Saccharomyces cerevisiae* strains used in this study were in the S288c background (detailed information on strain construction is provided in the section entitled 'Yeast strain construction'). Briefly, the isogenic ploidy series was generated in a *matA ste4A* background to eliminate mating and meiosis during the course of the experiment. Either a *pGAL-CFP* or a *pGAL-YFP* construct was integrated at the *TRP1* locus near the chromosome IV centromere in a haploid strain (PY5998 and PY5999, respectively). These haploid strains were used to generate isogenic diploids, from which isogenic tetraploids were then derived (Extended Data Fig. 1). This procedure ensured that all copies of chromosome IV had the capacity to express the inducible fluorescent marker even if the strains became aneuploid. Mating-competent haploids were generated from the *matA ste4A* ancestor, PY5998, by transformation with either plasmid PB2647 (*CEN-LEU2-STE4*) or PB2648 (*CEN-URA3-STE4-Matz*). Zygotes from mating-competent haploids were isolated by micromanipulation to obtain diploid CFP ancestors (PY6008 and PY6022). Similarly, zygotes from mating-competent diploids were isolated by micro-manipulation to obtain tetraploid CFP ancestors (PY6031 and PY6032). The same mating scheme was performed for the YFP lineage starting with PY5999 to generate diploid YFP (PY6006 and PY6014) and tetraploid YFP (PY6040 and PY6045) ancestors.

The ancestor strains were grown to saturation from the -80°C stock, in Synthetic Complete + 2% glucose. The cell density of each ancestor was determined using a haemocytometer and an automated cell counter (Vi-Cell-XR from Beckman Coulter). An equal number of YFP and CFP cells of the same ploidy were diluted into fresh Synthetic Complete + 2% raffinose medium, and combined into a single tube for an initial concentration of 1×10^5 cells per millilitre. The 50:50 YFP:CFP culture was distributed equally into the wells of a 96-deep-well plate (1 ml per well, U-bottom block plate from Qiagen). Seven or eight wells were not inoculated, to detect cross-well contamination during the experiment. The plates were covered with 'breathe-EASIER' tape (Electron Microscopy Science) and incubated at 30°C on a 96-well plate shaker (Union Scientific). Two plates of haploid and three plates of diploid and tetraploid cells were analysed, representing 173 parallel haploid evolutions, 264 parallel diploid evolutions, and 265 parallel tetraploid evolutions.

At 24 h intervals, the cells were resuspended (by pipetting) and diluted into fresh Synthetic Complete + 2% raffinose medium. The dilution factor was determined for each ploidy type based on the initial strain fitness to maintain an equivalent population size, as reported previously¹⁷. The number of cells transferred each day was calculated by counting the number of cells in ten replicate wells of each ploidy before and after dilution with an automated cell counter (Vi-Cell-XR from Beckman Coulter), and averaged across 3 consecutive days. The dilution factor for the haploid, diploid, and tetraploid experiments was 1/100, 1/50, and 1/33, respectively. This corresponds to 6.64, 5.64, and 5.04 generations per day¹⁷. The tetraploid evolution experiment from generation 250 to 500 (Fig. 4) was performed with the same dilution factor as the diploid experiments (1/50).

The number of CFP and YFP cells in each population was measured at the same time each day. First, expression of the fluorescent proteins was induced by transferring 10 μl of the overnight culture into 200 μl Synthetic Complete + 2% galactose medium for 4 h at 30°C . The number of CFP- or YFP-expressing cells was determined using a BD LSRII flow cytometer high-throughput plate reader (10,000 cells were analysed from each well). Pacific Blue and FITC filters were used to detect CFP and YFP, respectively. All experiments were passaged for 250 generations, but daily acquisition of CFP:YFP ratios was not always continued to the 250th generation.

To ensure that the flow cytometer measurement and the galactose induction of CFP and YFP was an accurate reflection of the size of these populations, the ratio of CFP:YFP cells was determined by both flow cytometry and microscopy, and the ratio was determined both before and after galactose induction. To do this, we combined overnight cultures of the 1N, 2N, and 4N ancestor CFP and YFP strains at three different ratios (nine populations in total) and analysed the ratios in two ways. First, for an aliquot of the mixture, we induced the expression of the fluorescent proteins with 2% galactose for 4 h and analysed 10,000 cells using flow cytometry. In parallel, we also added 2% galactose for 4 h and then counted about 300 cells by fluorescence microscopy. Finally, to ensure that the induction with 2% galactose did not alter the CFP:YFP ratio, a portion of the population was used to determine the number of CFP and YFP cells in the population before adding galactose to the medium. To do this, cells from each population were struck for single colonies on YPD plates for 2 days. Ninety-six colonies were chosen randomly from each plate and added to a single well of a 96-well plate containing Synthetic Complete + 2% galactose. The fluorescence of each colony was determined by flow cytometry, and the percentage YFP of the initial population was determined. There was a strong correlation between the percentage YFP-expressing cells obtained from all three measurements (Extended Data Fig. 3a): including the flow cytometer and fluorescence microscopy (Pearson correlation coefficient = 0.979), and both before and after galactose induction (Pearson correlation coefficient = 0.985).

Finally, frozen stocks of the evolution experiments were made at 3- to 4-day intervals throughout the experiment. At the end of each experiment, single colony clones were isolated and used for competitive fitness assays, flow cytometry analysis of ploidy, and preparation of DNA for aCGH.

We isolated 48 clones from the 4N evolution experiments at generation 250 ($4N_{250}$) with fitness values equal to the 2N ancestor (competitive fitness difference < 0.05), and determined the rate of adaptation after an additional 250 generations (Fig. 4). Each $4N_{250}$ clone was grown to saturation from the -80°C stock in Synthetic Complete + 2% raffinose medium. Cell counts were performed as above, and each population was diluted to an initial concentration of 1×10^5 cells per millilitre. At 24 h intervals, the cells were resuspended (by pipetting) and diluted 1/50 into fresh Synthetic Complete + 2% raffinose medium (the same dilution factor as the diploid experiments). These evolution experiments were not performed as CFP:YFP competitions, so daily flow cytometry was not necessary. After 250 generations, single colony clones ($4N_{500}$) were isolated on Synthetic Complete + 2% raffinose plates. Each $4N_{500}$ clone was cultured overnight in 1 ml Synthetic Complete + 2% raffinose medium and aliquots of this culture were immediately used for competitive fitness assays, flow cytometry analysis of ploidy, preparation of DNA for aCGH, and frozen stocks.

Measuring the variation in the flow cytometer measurements. We determined the amount of noise in our flow cytometer measurements by calculating the mean and standard deviation of the percentage YFP obtained from 48 independent populations at six different ratios of CFP:YFP, for each ploidy type. The 1N, 2N, and 4N ancestor strains were cultured separately overnight in 2% raffinose medium and transferred to 2% galactose for 4 h to induce expression of CFP and YFP. Next, CFP and YFP cells of the same ploidy were combined at ratios of 100:0, 85:15, 75:25, 50:50, 25:75, and 0:100 to reach the same final volume (200 μl). Ten thousand cells from each population were analysed by the LSRII (BD) flow cytometer using the same parameters (for example gating and flow rate) that we used for the evolution experiments, the total numbers of CFP and YFP cells were obtained, and the percentage YFP was calculated ((number of YFP cells)/(number of CFP + number of YFP cells)) $\times 100$. The standard deviations are presented in Extended Data Fig. 3b, and indicate that there was little well-to-well variability for the same CFP:YFP ratio (across 48 wells), and that this variability changed only slightly across different CFP:YFP ratios of all three ploidy types. Importantly, it was never greater than 0.66% of the measurement. This small variability had a minimal effect on our analysis because the fitting procedure we used to measure the deviation from an equal percentage of CFP and YFP cells used bins of 5% deviation to combine the number of wells that had deviations between 0 and 5%, 5 and 10%, and so on, for each experiment.

Statistical analysis of the experimental data. No statistical methods were used to predetermine sample size.

Adaptation rate. The CFP versus YFP evolution experiments were designed to analyse the dynamics of the adaptation process at the population level within and across wells. The adaptation rate was determined with additional competition experiments that were designed to measure the change in fitness from generation zero to generation 250 for cells of each ploidy. To this end, we isolated single colony clones from each evolved well of different ploidy types at generation 250 and measured their competitive fitness relative to the 2N ancestor (see detailed methods below). Fitness was defined as the slope of the log of the ratio of the evolved clone to the reference strain (2N ancestor) over time; more precisely, we performed a linear least-squares fit of $\log(N_{t1}/N_{t0})$ over multiple dilution cycles (where N_{t1} is the number of cells from the evolved clone and N_{t0} is the number of ancestor cells). The fitness relative to the ancestor is defined as $s = d/dt [\log_2(N_{t1}/N_{t0})]$, where t is measured in days¹⁷. The rate of adaptation was the relative fitness at generation 250 minus the relative fitness of generation zero, divided by 250 generations. The rate of adaptation for each ploidy is shown in Fig. 1c. We found that the tetraploid populations had a significantly larger rate of adaptation (0.009 [0.0062, 0.011]) than haploids (0.0031 [0.0022, 0.0052]) or diploids (0.0031 [0.0018, 0.0041]) during the 250 generations in raffinose medium (values indicate the median rate followed by the 95% confidence intervals in square brackets, Fig. 1c, t -test, $P < 1 \times 10^{-10}$).

Mathematical modelling of population dynamics. Branching evolution model. We formulated a mathematical model of the population dynamics of cells that was then used to infer evolutionary parameters using the experimental data. This branching evolution model was designed to mimic closely the divergence experiments containing two equally fit populations (CFP or YFP), each initially consisting of 50,000 cells. The model is based on a stochastic birth and death process called a branching process³¹. In this process, at each time-step, a cell is chosen to die at random, or to divide with a probability corresponding to its fitness. During each cell division, a mutation arises with mutation rate μ . If no mutation occurs, the fitness of the daughter cell is equal to the fitness of the mother cell. If a mutation does occur, the fitness may change; the additive fitness of the daughter cell is then chosen from a fitness distribution. The fitness change is independent of whether the ancestors of this clone had already obtained one or more mutations. Furthermore, there

is no restriction on the number of cells that acquire beneficial mutations, thus allowing clonal interference to occur^{18,21}.

We compared results assuming either a uniform, exponential, or delta distribution of fitness values (Extended Data Fig. 2a–d). For the initial formulation of the model, we considered half of the newly arising mutations to be beneficial and half deleterious, and their fitness effects were considered to be additive to the fitness value of the mother cell. Because complete simulations of the branching process would be prohibitively slow, we approximated this branching process with a Wright–Fisher process¹⁹ with non-overlapping generations. This process was implemented as a Monte Carlo simulation in C++ and the code is provided as a Supplementary Software file.

In the simulation, each competition experiment was initiated with 1×10^5 cells (5×10^4 of each CFP and YFP cell type). At every generation, each cell reproduces and gives birth to a random number of surviving offspring distributed according to a Poisson distribution. The initial population of each ploidy had a different growth rate per day: the haploid population increased approximately 100-fold, the diploids about 50-fold, and the tetraploids about 30-fold in a 24 h time interval (described previously in the section entitled ‘Batch culture evolution experiment’). Assuming a population doubling every generation, the average number of generations per day was 6.64, 5.65, and 5.04 for haploids, diploids, and tetraploids, respectively. In the Wright–Fisher model, the number of generations is discrete; thus we rounded these numbers to the closest integers (7, 6, and 5, respectively). The initial fitness (f) of each ploidy was chosen to satisfy the growth rate of that ploidy with this number of generations: that is, $(f_{1N})^7 = 100$, $(f_{2N})^6 = 50$, $(f_{4N})^5 = 30$. The fitness of a cell is the average number of its surviving offspring in the next generation.

During each cell division, a new mutation might arise. In haploid cells, the impact of a mutation is given by its additive fitness value s drawn from the fitness distribution. In diploid and tetraploid cells, however, a mutation might have a degree of dominance, be recessive or have a different fitness effect than it would have in a haploid cell. Following ref. 2, we assumed that the effect of a given mutation in a diploid or tetraploid cell compared with that in a haploid cell is scaled by a certain factor h . Therefore, the fitness effect of a given mutation in one allele will be $\tilde{s} = hs$.

In our model, we considered a beneficial mutation to arise independently at rate μ (per whole genome). Because our populations proliferated asexually, we could not differentiate between the two components of \tilde{s} , namely the haploid fitness effect s and the dominance coefficient h , and therefore we could only infer the combined value of \tilde{s} . Our simulations were done in exactly the same way for all ploidy types, and the selection coefficient of a new mutation was taken from a given distribution. However, the meaning of the added fitness is different for the haploids, where it is s , than for the diploids and tetraploids, where it is \tilde{s} . The assumption of independence of mutations is justified by the low point mutation rate, which is of the order of magnitude of 2×10^{-10} per base per generation in yeast³². The low per-base mutation rate means that the probability of independently obtaining a second identical mutation is vanishingly small (probability of μ^2 , or 4×10^{-20}), and therefore we did not consider such events in the model. Similarly, the probability of obtaining a given mutation that is then copied by a recombination-based mechanism such as gene conversion³³ is low ($(2 \times 10^{-10}) \times (4 \times 10^{-5})$), and was not included. These assumptions are validated by our WGS data demonstrating single copies of all point mutations.

Our identification of specific mutations that have a larger fitness effect in the tetraploid strains than in the haploid or diploid strains (Fig. 3) is consistent with the overall larger fitness effect of the tetraploids in our experiment. Note, however, that our parameters describe the entire distribution of mutations, rather than any specific mutation. Thus, it is also possible that some mutations could have a larger fitness effect in the haploids or diploids relative to the tetraploids. Note also that in our model clonal interference can occur, as there is no limitation on the number of independent mutations that can arise within a population. Thus multiple clones can emerge and compete with each other. As shown previously, clonal interference is an important aspect of microbial population dynamics and thus cannot be ignored^{17,18,21}. Additionally, multiple mutations can exist in the same cell.

As a sensitivity analysis, we also varied the ratio between advantageous and deleterious mutations, and found that even when the fraction of the deleterious mutation is very large, their inclusion has only a negligible effect on the rate and dynamics of adaptation. This is consistent with a large body of prior literature^{17,34–36}. To accomplish this, we fixed the rate of the advantageous mutations and generated data sets with different ratios of beneficial to deleterious mutations: 100:0, 90:10, 50:50, 10:90, and 1:99. For each ratio we generated 100 data sets (each data set with 264 single deviation experiments) using a beneficial mutation rate of 1.2×10^{-6} and a fitness effect of $s = 0.16$ (the best-fit values of the delta function for the diploid experiments, as a typical example, we used the same distribution as in ref. 17), and fitted them against the simulations that we used to fit the empirical experiments (see below for the fitting procedure). The means of the fitted values are presented in Extended Data Fig. 3c. We found no significant difference between the different

ratios of beneficial to deleterious mutations (t -test, $P > 0.2$ between all ratios tested). Therefore, in our model, we assumed that half of all non-neutral mutations were deleterious and half were beneficial, and that the deleterious and beneficial mutations had the same fitness effect distribution with the same parameters. Note that we also obtained similar results using the equivalence principle model¹⁷ that did not include deleterious mutations (see below).

The upper bound of cellular fitness. For biological plausibility, it was necessary to set an upper bound for acquired fitness. On the basis of the well-described growth rates achievable for *S. cerevisiae* in optimal conditions, we set this boundary at a doubling time of 1 h (2^{24} per day). Given that g is the number of generations a strain experiences in a day, the upper bound of the fitness f of any ploidy type was therefore set to satisfy the equation $f^g = 2^{24}$.

The initial growth rate of the tetraploid cells was lower than the growth rate of the haploids and diploids (see previous section entitled ‘Batch culture evolution experiment’). In the simulations we rounded the number of generations to satisfy the assumption of non-overlapping generations in the Wright–Fisher model: the tetraploid cells underwent about five generations per day, whereas the haploid cells underwent around seven, and diploid cells around six generations per day. In the simulations, as in the experiments, we diluted the populations every day by choosing at random 1% of the haploid, 2% of the diploid, and 3.3% of the tetraploid cells. This dilution was done by using a hypergeometric random generator for populations smaller than 100 million (<http://www.agner.org/random/>); for populations larger than 100 million this method is not applicable and we used a direct Bernoulli sampling of cells³⁷, forcing the total sampled cells to be 1%, 2%, or 3.3% for haploid, diploid, and tetraploid populations, respectively. The concentration of each of the cell types was recorded in the simulation output at the end of every day, after dilution.

In the experiments with the diploid and tetraploid cells, we observed an initial small bias against the YFP-labelled population: in the diploid cells there was a decline in the YFP-labelled population of 0.5% per day, and in the tetraploid cells there was a 1% decline per day (Fig. 1b). We included these biases in the simulations by including this initial small difference in the fitness (that is, a reduced average number of offspring) of CFP- versus YFP-labelled cells^{17,26}.

Note that drift could not feasibly play a role in our experiment, as the timescale (in generations) for fixation of an allele due to drift is approximately equal to the effective population size³⁸. This time was of the order of 10^6 generations in our experiments¹⁷, much larger than the timeframe of our experiment (250 generations). Thus the simulations continued until one of the cell types overtook the whole population or until the end of the time that the CFP:YFP data were collected on the 30th day; the first of either event terminated a simulation. To increase the efficiency of the simulations, extinction of a certain colour (and thus fixation of the other) was defined not by a value of zero frequency, but as a frequency of less than or equal to 1%, as the probability was negligible that a sub-clone present at a frequency of 1% would overtake another sub-clone present at a frequency of 99% in the timeframe of our experiment²¹. Furthermore, while the variability between flow cytometer measurements was never greater than 1% of the measurement (see above, and ref. 26), our ability to detect changes in CFP or YFP populations below a frequency of 1% was a limitation of the flow cytometer.

Fitting procedure. The empirical data for each ploidy type were combined to represent the average deviation from an equal percentage of YFP and CFP cells. We then used the combined ploidy data to estimate the best-fit values of mutation rate (μ) and selection coefficient (s), using least squares fitting.

To compare data from the experiment and the simulations, we investigated several different summary statistics and used the one that performed best. The summary statistics evaluated were as follows.

1. Mean deviation from an equal percentage of CFP and YFP cells: for every experiment corresponding to an individual well in a 96-well plate, we calculated the deviation from equal percentage every day. Then, for every data set generated (experiment and simulation) we calculated the average deviation per day. This procedure created a vector of the mean deviations per day for every data set.

2. Mean and standard deviation of the deviation from an equal percentage of CFP and YFP: as in 1, except that we also calculated the standard deviation of the deviation from equal percentage for every day. This procedure generated two vectors for every data set.

3. A distribution of the deviation from an equal percentage of CFP and YFP with ten bins: we calculated the absolute value of the deviation from equal percentage for each deviation experiment for each day. Then for every day, we binned the deviation values into ten bins, each with a size of 5%. That is, we counted the number of wells that had a deviation between 0–5%, 5–10%, ..., 45–50%. This procedure generated a matrix of ten bins \times 30 days for each data set.

4. A distribution of the deviation from equal percentage CFP and YFP with three bins: similar to (3), except we used only three bins, one for the non-deviated wells (defined as 0–0.1), one for fixated wells (defined as 0.4–0.5) and one for those in the middle (0.1–0.4).

The comparison between the experiments and the simulations was done by calculating the sum of squares (SOS) between the summary statistic (SS) of the experiment and the summary statistic of the simulations:

$$\text{SOS}_{(\mu,s)} = \sum_{\Omega} \left[\left(\text{SS}_{(\mu,s)}^{\text{Experiment}}(\Omega) - \text{SS}_{(\mu,s)}^{\text{Simulation}}(\Omega) \right)^2 \right]$$

where Ω spans the day and the number of values that each summary statistic has each day. The best-fit pair was the set of μ and s values with the smallest sum of squares. To determine the performance of the different summary statistic, we generated 1,000 data sets with the same parameters (1N, uniform distribution with $s = 0.05$, and $\mu = 8 \times 10^{-5}$), each containing 264 single deviation experiments, and inferred their values by scanning the parameter space for each of the 1,000 artificial data sets. The expected values from the simulations were then generated by 1,000 single deviation experiments for each set of parameters.

The distribution of inferred μ values is shown in Extended Data Fig. 3d. Whereas all of the summary statistics had the mutation rate used in the simulations as their mode, the observed means varied modestly. The ten-bin summary statistic had the narrowest range, and therefore this was used for further analysis. A similar summary statistic was also used previously³⁹.

The scanned range of mutation rates and fitness effects. The best-fit value was found by scanning a range of mutation rates and fitness effects²⁶; to this end, we scanned the parameter space of the fitness effect in linear steps from 0.005 to 0.35, with increments of 0.005. The mutation rates were scanned in logarithmic steps from $\log_{10}(\mu) = -8$ to $\log_{10}(\mu) = -4$, with increments of 0.1. Thus, per ploidy per fitness distribution, we scanned 2,460 parameter regimes, analysed 1,000 divergence simulations per μ and s combination, and scanned three ploidy types and three different fitness distributions in total.

For the exponential fitness distribution, larger mutation rates and smaller selection coefficients needed to be investigated to find the best-fitting pair: we scanned mutation rates from $\log_{10}(\mu) = -8$ to $\log_{10}(\mu) = -3$, and selection coefficients from 0.002 to 0.1, with increments of 0.002. Additionally, in the exponential distribution, the majority of mutations had a fitness value that was too small to contribute to the competition between the two cell populations (YFP and CFP). However, increasing the mutation rate would increase the probability that mutations with larger fitness value were obtained. We found that for mutation rates higher than 10^{-4} per genome per cell division, the computational time constraints were substantial; thus we excluded all mutations that had an s value smaller than 10% of the average fitness effect s . Support for this choice was provided by the finding that for large mutation rates (that is, $N\mu > 1$), the mutations chosen from an exponential distribution that eventually reach fixation within the population had a larger fitness effect than the distribution average (ref. 20 and Fig. 3). The above criteria for large mutation rates were within our experimental regime ($N\mu = 10^6 \times 10^{-4} > 1$). Furthermore, in Extended Data Fig. 3e we present simulation results of the average deviation from equal percentages of CFP and YFP for a given mutation rate and fitness effect drawn from an exponential distribution. We found that the results were robust to including or excluding those mutations with fitness effects smaller than the distribution average. Thus, for high mutation rates ($N\mu > 1$), we could exclude weak mutations, as was shown in ref. 20.

The fitting of the empirical experiments and the simulations was done as described above. The results of the best-fit values are presented in Extended Data Fig. 2a–d.

The equivalence principle model. As a complementary approach we evaluated the equivalence principle model developed in ref. 17 to analyse a similar competition experiment in *Escherichia coli*. This study concluded that the mutations that eventually reach fixation in large microbial populations have a very narrow range of fitness values. Very weak mutations are unlikely to lead to a takeover of one population (CFP or YFP) by the other, whereas very strong mutations are very infrequent. On the basis of this prediction, the competition experiments were described by assuming that all beneficial mutations have exactly the same fitness value; that is, assuming that the distribution of fitness effects is a delta function. On the basis of this assumption, the authors developed a method for inferring the two key parameters of the dynamics of the adaptation process: the mutation rate of a beneficial mutation and the single fitness effect value. Note that the assumption of a single fitness value for beneficial mutations was shown to be true only if the mutation rate is very high²⁰, which cannot be known *a priori*. This is why the branching evolution model was implemented first and then compared with the equivalence principle model. Note also that the equivalence principle model analyses only the initial CFP:YFP divergence phase, rather than using the entire data set as in the branching evolution model. Although this excludes some data, it has the advantage of not necessitating the assumption that the distribution of fitness effects is constant⁴⁰.

We followed the procedure outlined in ref. 17, including the following assumptions and modelling steps (see the Supplementary Information in ref. 17 for a more complete description).

1. Rather than the computationally intensive simulation of growth and dilution, we used the effective population size⁴¹. The effective population size is $N_e = N_0 \log(r)T$, where N_0 is the initial population size, r is the growth rate of the population, and T is the number of generations between dilutions. In addition to reducing computation time, the use of the effective population size allows for an analytical approximation of certain quantities, such as the probability of escaping drift.

2. Even large mutations may be eliminated by drift, and only those that escape drift can contribute to the competition between the two labelled cell populations. The probability of emergence of a new mutation that will escape drift can be calculated analytically for a fixed population, making it possible to generate in simulations only those mutations that escaped drift, and contribute to the competitiveness of the population. Under the above simplifying approximations of (1) a fixed population size and (2) a delta distribution of fitness effects, the probability that a mutation will emerge and escape drift (P_{escape}) was calculated analytically, as a function of the mutation rate, μ (the rate for a new mutation to appear in a single cell division), the selection coefficient, s , and the effective population size, N_e .

3. Given that P_{escape} can be calculated, simulations can be performed rapidly by generating the time in the experiment when such mutations occur, rather than randomly assigning mutations at each cell division.

We applied the equivalence principle model as follows. To compare the experiments with the simulations, as in ref. 17, we used time and initial slope of CFP:YFP ratio divergence as summary statistics for the divergence curves. Each CFP:YFP divergence experiment (one single well from a 96-well plate) was fit to an exponential growth model, with slope α that starts at a given time τ , using the expression $-\log_{10}(1 + 0.5\exp(\alpha(t - \tau)))$. An exponential model is a good description of the initial divergence phase. The end of the exponential growth phase was defined by the time that had the maximum likelihood to be described by exponential growth (see the Supplementary Information in ref. 17 for more details). The initial CFP:YFP bias that we detected in some 2N and 4N populations was included in cases where it occurred.

Each divergence experiment is thus described by its values τ and α , whether it is a real experiment or a simulated one. Because a single divergence experiment is subject to significant stochasticity, we compared the empirical experiments and the simulations by combining all the wells of each ploidy type. This was done by collecting all of the α values of a given experiment or a given set of simulations (that is, those simulations that we generated by the same pair of μ and s values) into a distribution of α values. The experimental distribution was compared with the distribution of the simulations by a Kolmogorov–Smirnov test between the two. The same was done for the τ values. For each pair of s and μ values, we calculated the sum of the Kolmogorov–Smirnov test of its α distribution and the Kolmogorov–Smirnov test of its τ distribution. The pair of values (μ, s) that had the smallest sum was declared the best fit (Extended Data Fig. 2a). This was done independently for each ploidy type (Extended Data Fig. 2e).

Modelling results. The fitting of the empirical experiments and the simulations was done as described above. The results of the best-fit values from the branching evolution model and the equivalence principle model are presented in Extended Data Fig. 2a. To generate the error range for each distribution and ploidy type, we used the parametric bootstrapping method⁴². For each estimated set of values, we generated 1,000 simulated data sets and compared those with the empirical data sets (data set sizes were 172, 264, and 265 independent deviation experiments for the haploids, diploids, and tetraploids, respectively). We then inferred the best-fit values of those data sets. The 95% confidence intervals of μ and s from those 1,000 data sets were defined as the error ranges (Extended Data Fig. 2a–e).

We infer from these results that the tetraploids have a higher mutation rate and that these mutations have, on average, a stronger fitness effect compared with haploids or diploids. The trend for tetraploids to have higher μ and s values occurs independently of the different assumed distributions of fitness effects. However, the different distributions lead to significantly different absolute values. This is expected from the characteristic shape of these distributions. To illustrate why this result is expected, we show a schematic diagram of the three distributions of fitness effects (Extended Data Fig. 2f).

The mutations that mainly govern the adaptation process are hypothesized to come from a relatively narrow range of fitness effect values that are sufficiently strong, but not extremely rare (Extended Data Fig. 2f, double arrow region from the idealized narrow Gaussian curve). The three assumed distributions that we used in our modelling approximate the true distribution but with the following differences. The delta distribution is located in the centre of the double arrow region. Thus, by definition, every mutation from the delta distribution will have a fitness effect that is strong enough to have a significant probability of promoting adaptation. If every mutation has the possibility of contributing, then a lower mutation rate suffices for adaptation. By contrast, the exponential distribution is dominated by small-effect or near-neutral mutations, with a relatively low fraction of mutations near the value for the delta distribution. Thus, the assumption of an exponential

distribution is accompanied by a requirement for a compensatory higher mutation rate to achieve numbers of equivalently beneficial mutations. By the same line of reasoning, the uniform distribution necessitates a rate of beneficial mutations that is intermediate between the exponential and delta distributions. Indeed, our results match all of these expectations (Extended Data Fig. 2a–d). A similar difference in the estimation of μ was also observed in another recent study that modelled this value based on these three assumptions about the distribution of beneficial mutations⁴³.

In terms of the fitness effect values of the different distributions, the uniform distribution is governed by its upper limit, which, by definition, is the strongest mutation allowed. Therefore, it is expected that the uniform distribution mean, which is half of the upper limit, will be half of the fitness effect value of the delta function. Again, this is what we observed (Extended Data Fig. 2a–d). The exponential distribution mean, is much smaller than the strongest mutation that can be generated by the exponential distribution, and the best-fit parameters are expected to be much smaller than the other two distributions, which is also observed (Extended Data Fig. 2a–d). Furthermore, because the exponential distribution has no upper bound on the fitness effects, a larger mutation rate can lead to the emergence of mutations with much larger fitness effects. In this way, under the exponential distribution assumption, the mutation rate affects the range of possible fitness effects, which results in a larger error range for the exponential distribution compared with the other distributions (Extended Data Fig. 2a, brackets).

In summary, despite expected differences in absolute values, we reach the same overall conclusions with either the branching evolution model with varied assumptions about the distribution of beneficial mutations or the equivalence principle model¹⁷.

Computer code availability. The computer code is available as a Supplementary Software file. The code was compiled by g++ (version 4.2.1) and was tested on Unix (CentOS5 operating system) and Mac (OS X version 10.9.2) machines.

Plasmid construction. All plasmids used in this study are listed in Extended Data Table 1. To construct plasmids for the inducible expression of either *CFP* or *YFP*, the galactose-inducible *GAL1* promoter was subcloned into the YFP plasmid PB1500 and the CFP plasmid PB2452. These plasmids were derived from the green fluorescent protein (GFP) tagging plasmid generated in ref. 44. Both plasmids contained the *ADH* gene terminator (*tADH*) after the *YFP* or *CFP* gene and the sequence of the *SpHIS5* gene of *Schizosaccharomyces pombe* as a selectable marker. Plasmids PB1500 and PB2452 were digested with BamHI and PacI to introduce the pGAL promoter, 461 base pairs (bp) upstream of the start codon of *GAL1* (ref. 45), which was amplified using the primers pGAL1 BamHI 5' (5'-ACGGATCCCCGGGTTGAA GTACGGATTAGAAGCCGCCGAG-3') and pGAL1 PacI 3' (5'-CGTTAATTA ATATAGTTTTTCTCCTTGACGTTAAAG-3'). Site-directed mutagenesis (Quick Change Mutagenesis Kit, Stratagene) was used to introduce an ATG translation start codon to the *YFP* and *CFP* genes (using the GAPATGpFA6 primer, 5'-CAA TCAATCAATCAATCATCACATAAATTAATTAATGAGTAAAGGAGAAG AACTTTTCACTGGAGTTGTC-3'). The resulting plasmids PB2694 and PB2697 contained the cassette pGAL1-*CFP*-*tADH*-*SpHIS5* and pGAL1-*YFP*-*tADH*-*SpHIS5*, respectively.

PB2314 was used to delete the *MAT* locus as previously described¹². PB1308 was used to perform a *URA3* to *TRP1* marker swap, as previously described⁴⁶. PB1640 (hphMX4, ref. 47) was used for PCR-mediated deletion of *STE4*. PB2647 (*STE4*-*LEU2*) was used to restore mating competency and was constructed by amplifying the *STE4* gene with primers *STE4*P BamHI 5' (5'-CCGGATTCTGTAGCCCTG TTAGGTTTACC-3') and *STE4*T BamHI 3' (5'-CCGATTCCAATACATAAG GACGAGCCAGTG-3'), and cloning it into pRS315. PB2649 (*STE4* *URA3* *CEN* *MAT α*) was also used to restore mating competency, and was constructed by sub cloning the *STE4* fragment from PB2647 (digested with SmaI and NotI) into PB2577 (*MAT α* *URA3* *CEN*, digested with SmaI and NotI).

Yeast strain construction. All *S. cerevisiae* strains used in this study were isogenic to PY3295 (BY4741, S288c genetic background *MATa his3A leu2 Δ met15A ura3 Δ*) and are listed in Extended Data Table 1. The strategy used to generate the isogenic ploidy series is illustrated in Extended Data Fig. 1 and genotypes of key intermediates are indicated. The CFP and YFP ancestors were derived from the haploid strain PY5997 (*matA::pSTE5-ura3::TRP1, ste4A::Hygro^R, trp1::Nat^R*, strain construction details available upon request). Isogenic strains with either the *CFP* or *YFP* cassettes at the *TRP1* locus (chromosome IV) were generated (PY5998 and PY5999) as follows: the pGAL1-*CFP*-*tADH*-*SpHIS5* or pGAL1-*YFP*-*tADH*-*SpHIS5* cassette was PCR amplified from plasmid PB2694 (*CFP*) or PB2697 (*YFP*), respectively, with primers delTRPGFP5' (5'-TATTGACACGTGAGTATACGTGATTAAG CACACAAAGGCAGCTTGGAGTGCAGGTCGACGGATCCCCGGG-3') and delTRPGFP3' (5'-GAACGTGCACTGAGTATGTTGCAGTCTTTGGAA ATACGAGTCAATTCGAGCTCGTTAAAC-3') and transformed into PY5997 at the *TRP1* locus. The haploid ancestor strains expressing *CFP* (PY5998) or *YFP* (PY5999) were confirmed by PCR and fluorescence microscopy. The haploid ancestors

were modified to become mating competent by transformation with plasmids PB2647 (*LEU2-STE4*) and PB2649 (*URA3-STE4-MAT α*). Diploid zygotes were selected on -Ura -Leu plates, and then colony purified on YPD plates to allow plasmid loss. Diploid chromosome content was confirmed by flow cytometry and aCGH, and strains PY6006, PY6008, PY6014, and PY6022 were selected. The diploid ancestors were made mating competent by transformation with plasmids PB2647 and PB2649. Tetraploid zygotes were pulled onto YPD plates using a micromanipulator, and after 2 days growth at 30 °C the ploidy of each zygote was determined by flow cytometry and aCGH.

The *SNF3-G439E* mutation was constructed in the haploid YFP strain background (PY5999) using the pCORE counter-selectable reporter system⁴⁸, a gift from M. Resnick. Primers *SNF3_pCORE_KAN* (5'-TGTTGGGGGTGTTATCATGA CTATAGCCAACCTTTATGTGGCCATGTGGGAGCTCGTTTCGCAC TGG-3') and *SNF3_pCORE_URA* (5'-TATAAATGCTATCATAAAGTTTGGCG GCCGCTACAGTCTTTAAGGAACACTCCTTACCATTAAAGTTGATC-3') were designed to integrate the *CORE* sequence at the *SNF3* locus; PCR amplification and transformation procedures were followed as detailed previously⁴⁹. Sanger sequencing was used to identify clones with the desired mutation (chromosome IV: 112,896 G>A). Diploid *SNF3-G439E* mutants (heterozygous *SNF3-G439E/SNF3* and homozygous *SNF3-G439E/SNF3-G439E* clones) were constructed by mating after introduction of plasmids to confer mating competence (PB2649 or PB2647), as described above for the construction of the CFP- and YFP-marked strains. An analogous strategy was used to generate tetraploid *SNF3-G439E* strains (heterozygous *SNF3-G439E/SNF3/SNF3/SNF3*).

The chromosome XIII aneuploid strain series was constructed in the S288c background from the diploid strain PY7295 (RL4737) and the diploid PY7296 (RL4888), which is trisomic for chromosome XIII⁵⁰. PY7296 was isolated from a triploid meiosis and a minimal number of cell divisions⁵⁰. We confirmed the chromosome XIII trisomy by aCGH (Extended Data Fig. 9). We generated tetraploid clones by mating PY7295 to PY7296, with changes in mating-type accomplished as described previously¹². Tetraploid clones were isolated on selective media and analysed by flow cytometry and aCGH (representative clones Extended Data Fig. 9). Additional details for all yeast strain constructions are available upon request.

Relative fitness assays. Competitive fitness assays were performed using single colony isolates from the evolved populations and a common ancestor. One single colony was isolated from frozen stocks of each well of the evolution experiments (1N(A), 1N(B), 2N(A), 2N(B), 4N(A), 4N(B), 4N(C)) at generation 250. The evolved clones were cultured for 24 h in 500 μ l of Synthetic Complete + 2% raffinose, diluted into fresh medium, and competed with the ancestor expressing the complementary fluorescent protein. Competitions were initially performed using approximately the same number of cells from the ancestor and the evolved clone, but because the evolved clones grew significantly faster than the ancestor strains, the competitions were repeated using approximately five times more ancestor cells than evolved clone cells, with an initial population size of 1×10^5 . Serial dilutions were performed each day and the YFP:CFP ratio was determined by flow cytometry, yielding an estimate of the number of evolved (N_{t1}) cells relative to the ancestors (N_{t0}) as a function of time. The data were analysed in Matlab using a custom script that performed a linear least-squares fit of $\log(N_{t1}/N_{t0})$ over multiple dilution cycles. The fitness relative to the ancestor is defined as $s = d/dt [\log_2(N_{t1}/N_{t0})]$, where t is measured in days¹⁷.

Flow cytometry analysis of DNA content. Cells were prepared for propidium iodide staining as described but with modifications to optimize preparation of samples in 96-well plates. Thirty thousand cells were analysed using a BD LSRII High Throughput Sampler. FlowJo cell cycle analysis used the Dean-Jett-Fox model to estimate the mean G1 and G2 fluorescence peaks of each strain. Control parental 1N, 2N, 3N, and 4N strains were analysed in triplicate with the evolved strains.

Microarray aCGH. Fluorescently labelled DNA was prepared for CGH as described previously⁵¹. Genomic DNA from all experimental strains was compared with the same pool of genomic DNA from the ancestral strain background PY3295 (BY4741, Research Genetics). Agilent yeast DNA 4 \times 44K microarrays (ChIP-on-chip Kit) were used for the hybridization according to the manufacturer's instructions (Agilent Technologies) with several modifications (M. Dunham online protocols, <http://dunham.gs.washington.edu/protocols.shtml>). Briefly, 2.0 μ g of HaeIII-digested (New England Biolabs) genomic DNA was labelled with 2.1 μ l of Cy3 or Cy5 (CyDye-Cy3-dUTP or CyDye-Cy5-dUTP, Amersham GE Healthcare). Three hundred nanograms of Cy3-labelled DNA (experimental strains) was mixed with 300 ng of Cy5-labelled DNA (control DNA) and the volume was brought to 44 μ l with nuclease-free water. Blocking buffer and hybridization buffer 2 \times HiRPM (Agilent Technologies) were added, and 100 μ l was applied to each sub-array; the microarray was hybridized at 65 °C for 17 h and then washed, scanned, and analysed according to the manufacturer's instructions. Agilent Feature Extraction data were converted from \log_{10} ratios to \log_2 ratios and plotted using Treeview⁵² and a custom Matlab script.

A log₂ ratio of zero (baseline) indicates no difference in DNA copy number between reference and experimental samples^{44,51}.

Quantitative PCR. All qPCRs were performed on an Applied Biosystems ViiA-7 real-time PCR machine in 96-well format with Power SYBR Green PCR Master Mix (Applied Biosystems) and three technical replicates. *HXT6/7* gene copy number was determined relative to the ancestor as previously described⁵³. Genomic DNA was isolated and RNase treated from 30 clones with the highest fitness from each haploid, diploid, and tetraploid experiment. *HXT6/7* on chromosome IVR was amplified using forward primer 5'-GATTATTGCTGGTCCGATCC-3' and reverse primer 5'-GAGTAATCGCCAATGGGTCT-3'; the control loci *UBP1* on chromosome IVL was amplified using forward primer 5'-GCGCTCTGTCATTGTTCCAC T-3' and reverse primer 5'-GACTTTCAGCTTCGTCACAAA-3'. Raw *HXT6/7* values were normalized to *UBP1* for each clone and then normalized to the ancestor ($\Delta\Delta Ct$). We found the amplification in 3% of the 1N-, 30% of the 2N-, and 43% of the 4N-evolved clones ($n = 30$). This significant bias for *HXT6/7* amplification in the 2N and 4N populations (t -test, $P = 0.005$ and $P = 1 \times 10^{-4}$, respectively) might have been due to mutations in 1N cells that prevented the acquisition of the *HXT6/7* amplification because of negative epistasis²⁴.

SNF3 gain-of-function mutations were previously shown to increase expression of *HXT4* (ref. 25). Therefore we analysed *HXT4* gene expression levels in diploid-evolved clone 2N_233 (carrying the *SNF3-G439E* mutation), relative to the diploid ancestor (PY6006). Strains were grown up from -80°C stocks overnight in 5 ml Synthetic Complete + 2% raffinose and then diluted into fresh 25 ml Synthetic Complete + 2% raffinose. Cells were cultured at early log phase and RNA was extracted using an RNeasy Mini Kit (Qiagen). Complementary DNA was prepared using a SuperScript III First-Strand synthesis system (Life Technologies). *HXT4* was amplified using primers 5'-TAAGTCAGCGCAGACGATCCA-3' and 5'-TTCACC CCAGGAGGCATTACCA-3'; *ACT1* was amplified using primers 5'-ACGTCGC CTTGGACTTCGAACA-3' and 5'-TGGAACAAAGCTTCTGGGGCTC-3'. Raw qPCR values were normalized to *ACT1* levels and then normalized to the 2N ancestor (PY6006). Relative to the ancestor, the clone bearing *SNF3-G439E* had eightfold higher *HXT4* expression in raffinose medium.

Overview of WGS. We performed WGS and identified *de novo* variants for 74 evolved clones and two ancestors. Initially six evolved clones and one tetraploid ancestor were sequenced on ABI's SOLiD 4 platform. Subsequent sequencing of 68 evolved clones and the haploid ancestor was performed on an Illumina HiSeq 2500. The specifics of the analysis pipeline for each sequencing platform are provided below. Regardless of the underlying platform, the overall analysis strategy was as follows. Briefly, the raw reads underwent quality analysis and barcode/adaptor removal. High-quality reads were mapped to the *S. cerevisiae* reference genome (downloaded June 2010). Reads containing PCR-based artefacts were removed and alignments underwent local realignment around insertions and deletions (indels), resulting in the highest quality alignment. SNPs and indels were called and combined across the evolved strains and within the parental strains to identify a set of variants in the strain background relative to the reference. Each evolved strain (all 74) was individually compared with the parental set to identify the set of potential *de novo* variants. These evolved strain calls were filtered by quality metrics and manually inspected. All variants of moderate or poor quality as well as a few good quality variants were analysed by Sanger sequencing. Chromosomal aneuploidy was inferred from changes in read depth using a windowing approach.

Per cell, the evolved tetraploids had more mutations than haploids or diploids. However, per haploid genome, the evolved tetraploids on average accumulated a similar number of mutations (1.50 average SNPs per haploid genome, on the basis of final evolved ploidy) as the evolved diploids (1.44 average SNPs per haploid genome) and fewer mutations than the evolved haploids (2.05 average SNPs per haploid genome); however, neither comparison was significant. It is likely that the number of mutations in the evolved tetraploids was underestimated because of the high rate of chromosome loss in these strains. Interestingly, there was a higher average for the tetraploid-evolved clones that became $\sim 2N$ in ploidy (2.00 average SNPs per haploid genome) than the diploid-evolved clones (1.44 average SNPs per haploid genome), suggesting that near-diploid cells that underwent a tetraploid intermediate might have acquired more mutations than cells that remained diploid throughout the experiment. However, despite the trend, this effect did not reach statistical significance ($P = 0.22$).

Illumina sequencing (2 × 100). Library preparation. Clones selected for WGS were cultured overnight from -80°C stocks in 2 ml Synthetic Complete + 2% raffinose medium. Genomic DNA was isolated using phenol–chloroform–isoamylalcohol (24:25:1) and bead beating. Libraries were prepared as described⁵⁴. Briefly, DNA was sheared with Diagenode Bioruptor (UCD-200) to a median size of 300–500 bp, end-repair was performed with an NEB Next End repair kit (NEB E6050L), and fragments were A-tailed with Klenow fragment (M0212L). Custom adaptors with in-line barcodes were ligated overnight. Adaptor ligated fragments were size selected on 1% TBE agarose gel stained with Sybr Gold (Invitrogen S-11494) for fragments between

400 and 600 bp and isolated using Qiagen Gel Extraction Kit (28706). Libraries were amplified for 12 cycles with Illumina PE PCR primers 1.0 and 2.0. Libraries were pooled and underwent additional size selection for fragments of 400–600 bp.

Raw data. The genomes were sequenced on an Illumina HiSeq 2500 at the University of Colorado at Denver Next Generation Sequencing Facility. The data, which had an inline barcode, were de-multiplexed by the sequencing facility into individual sample R1/R2 files—one file for each read in the pair. The barcodes were removed before mapping using *Fastx_trimmer* (version 0.0.13.2, http://hannonlab.cshl.edu/fastx_toolkit/). Read trimming from the 5' end of the R2 reads was performed on a sample-specific manner trimming anywhere from 0 to 28 base pairs using in-house script and *Fastx_trimmer*.

Mapping. Reads were mapped to the *S. cerevisiae* reference sequence for the laboratory yeast strain S288c reference genome (*S. cerevisiae* genome obtained 28 July 2010 from the *Saccharomyces* Genome Database FTP site: http://downloads.yeastgenome.org/sequence/S288C_reference/genome_releases/; actual genome: http://downloads.yeastgenome.org/sequence/S288C_reference/genome_releases/S288C_reference_genome_R63-1-1_20100105.tgz). The reads were mapped using the *Bowtie2* (version 2.0.2)⁵⁵ local alignment strategy, allowing for multiple mapping, and setting the following options: `-very-sensitive-local -I 180 -X 1,000 -score-min G,70,8`. The mapped reads then underwent file format conversion into the binary format for downstream analysis using *Samtools* view, sort, and index (version 0.1.18)⁵⁶.

Alignment tailoring. Post-alignment to the genome, duplicate pairs resulting from PCR overamplification were removed using *Samtools* rmdup, eliminating 1–5% of the paired reads. The reads were re-aligned over potential indel sites using the *Genome Analysis Toolkit* *RealignerTargetCreator* and *IndelRealigner* (version 2.4-9)^{57,58}.

Variant calling and refinement. Variant calling was performed on the tailored read mappings using *Genome Analysis Toolkit* *UnifiedGenotyper* (version 2.4-9)^{57,58}. For the haploids and diploids, SNPs were called using default parameters; for the higher ploidy strains, the ploidy option was increased to $5n$, which allowed identification of mutations at allelic frequencies down to 5% alternative allele representation. Variant lists were combined on the basis of ploidy-type using *Genome Analysis Toolkit* *CombineVariants*. SNPs and short indels were compared with the parental set of mutations using in-house scripts to generate a set of non-Parental mutations. These mutations were filtered for alternative allele support and allelic frequency (more than two reads supporting an alternative allele for coverage $10\times$ to $20\times$, and more than four reads supporting an alternative allele for coverage more than $20\times$). The filtered mutations were manually inspected using the *Integrative Genome Viewer* (version 2.1.19)⁵⁹ to refine the set and further remove mapping artefacts such as strand representation bias, regional mapping quality issues from non-unique mapping, and artefacts of homopolymer and simple repeat alignments. We Sanger sequenced variants with low read support (fewer than five reads supporting an alternative allele), as well as a subset of the other medium- and high-confidence variants. The final set of evolved variants discovered was annotated versus the gene file specific to the genome using an in-house script.

Chromosomal CNV identification. Identification of chromosomal CNVs used HTSeq (version 0.6.1)⁶⁰ in conjunction with custom scripts. HTSeq performs coverage estimations on a per-gene basis, and the custom scripts provided normalized log₂ fold change between each sample and the parental haploid strain. Estimates on chromosomal copy number were inferred using the median value for the log₂ fold change on a chromosome-by-chromosome basis. We implemented the Cochran–Armitage test to determine whether chromosome XIII had a trend for higher copy number, relative to the copy number observed for all chromosomes in the tetraploid-evolved clones (Fig. 2d and Supplementary Table 2). This trend analysis is similar to a χ^2 test, but tests whether there is a significant trend or direction to the observed data set (chromosome XIII copy number).

Sequencing quality assessment. Because our sequencing was highly multiplexed, quality assessment on the sequencing data was necessary to eliminate strains without adequate genome coverage. For the haploids and diploids, we determined the adequate depth of coverage to recover mutations in two ways. First, we took the set of 'strain-background' mutations, which were identified by filtering the parental variant calls for a conservative, high-quality (quality > 100), homozygous set of locations. Each strain was then queried for the ability to recapitulate these variants, reporting a percentage overlap between each strain's variant calls and the set of background variants. Any strain with less than 97% of the background mutations was dropped from consideration. Additionally, we examined the impact of subsampling down to various depths to investigate the impact of lower coverage on recovering variants. This was done using *Picard's* *DownsampleSam.jar* (version 1.72, <http://broadinstitute.github.io/picard/>) on two higher-coverage diploid strains to randomly down-sample the coverage to $100\times$, $50\times$, $25\times$, and $10\times$ coverage. We examined the SNP call overlap and found that for the strain-unique SNPs they could be captured even at a level of $10\times$ coverage. Using this information, we set minimum coverage requirements for each strain on a genome-wide scale to eliminate strains

without adequate genomic representation. We analysed depth of coverage on all of the mapped data with BEDTools genomeCoverageBed (version 2.16.2)⁶¹. The per-base coverage was then analysed using an in-house script to produce statistics on minimum coverage per allele, average coverage, etc.

SOLiD sequencing (2×50). Library preparation. We performed a pilot experiment on seven strains using SOLiD paired-end sequencing (Supplementary Table 1). Clones selected for SOLiD sequencing were cultured overnight from −80 °C stocks in 4 ml Synthent Complete + 2% raffinose medium. Genomic DNA was isolated using QIAGEN Genomic-Tip 100 according to the manufacturer's instructions. SOLiD library preparation and sequencing was performed by the Molecular Biology Core Facility at the Dana-Farber Cancer Institute according to the manufacturer's instructions (Applied Biosystems, Life Technologies).

Mapping. The sequencing reads were mapped to the *S. cerevisiae* reference genome (see 'Mapping' in the Illumina section above) using multiple different mapping software including BWA (version 0.5.9)⁶², NovoAlignCS (version 1.01.05)⁶³, Bfast (version 0.6.5a)⁶⁴, and BowtieCS (version 0.12.7)⁶⁵. BowtieCS and BWA were used in the downstream variant calling and copy number changes, while NovoAlignCS and Bfast served as added support in manual inspection of variants.

Alignment tailoring. After mapping, the reads were post-processed for local realignment using SRMA (version 0.1.15)⁶⁶ and Samtools BAQ (version 0.1.18)⁵⁶.

Variant calling and refinement. SNPs, small insertions and deletions (indels) were called from the post-processed reads using Samtools Mpileup (version 0.1.18)⁵⁶, VARIID (version 1.0.7f)⁶⁷, and Freebayes (version 0.8.9, <http://bioinformatics.bc.edu/marthlab/FreeBayes>). Samtools and VARIID variant calls were used to identify the strain background (parental variants relative to the reference). These variants were filtered on the basis of reads supporting the allele in both directions, quality score of the call, and adequate read coverage over the call. Once filtered, all of the variant calls for the evolved strains were merged and compared with the parent. Variations were verified by manual inspection followed by Sanger validation for both a set of randomly sampled loci and regions of disagreement between different combinations of the mapping software and the variant callers (that is, dinucleotide SNPs and multiple indels within a single read). The resulting set was later used for identification of strain-unique variants in the evolved strains.

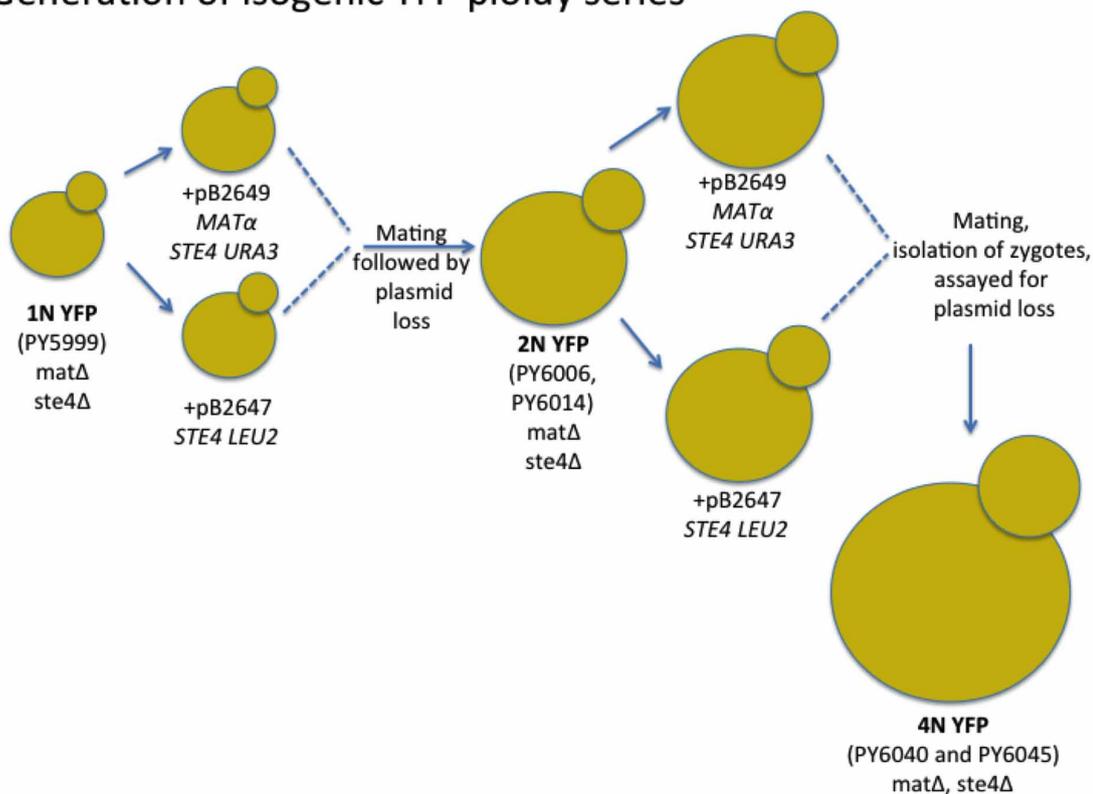
To identify strain-unique variants, Freebayes, a variant caller capable of higher-ploidy (ploidy > 2N), was used. Freebayes has the ability to set the assumed ploidy over a genomic region to adjust the expected distribution for allelic frequency. The assumed ploidy was determined using aCGH as well as the copy number changes. The Freebayes called variants on each evolved progeny were then cross-referenced with the parental variants to produce strain-unique variants (Supplementary Table 1). These variants were then manually examined in Integrative Genomics Viewer⁵⁹ and validated by Sanger sequencing, PCR amplification and Sanger sequencing of approximately 200 bp on either side of the sequence variants used DNA from the evolved clone and the ancestor.

Chromosomal CNVs. Copy number changes, first identified in the aCGH data, were confirmed in the genomic DNA sequencing using BedTools genomeCoverageBed⁶¹ in combination with custom in-house scripts and DESeq (version 1.10.1)⁶⁸. Briefly, the normalized genomic copy number of all annotated genes in each strain was compared back with the parent. These comparisons then were plotted using an in-house script (see Chromosomal CNV identification above).

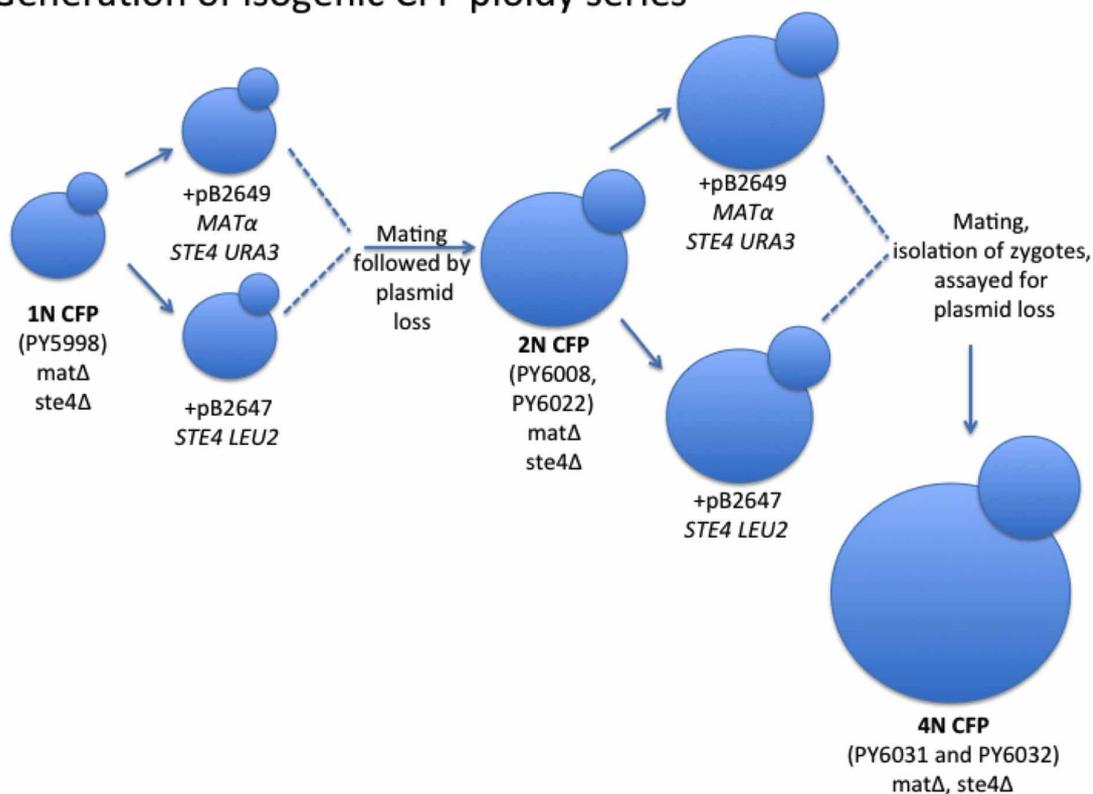
31. Durrett, R., Foo, J., Leder, K., Mayberry, J. & Michor, F. Intratumor heterogeneity in evolutionary models of tumor progression. *Genetics* **188**, 461–477 (2011).
32. Zhu, Y. O., Siegal, M. L., Hall, D. W. & Petrov, D. A. Precise estimates of mutation rate and spectrum in yeast. *Proc. Natl Acad. Sci. USA* **111**, E2310–E2318 (2014).
33. Barbera, M. A. & Petes, T. D. Selection and analysis of spontaneous reciprocal mitotic cross-overs in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **103**, 12819–12824 (2006).
34. Rouzine, I. M., Wakeley, J. & Coffin, J. M. The solitary wave of asexual evolution. *Proc. Natl Acad. Sci. USA* **100**, 587–592 (2003).
35. Desai, M. M., Fisher, D. S. & Murray, A. W. The speed of evolution and maintenance of variation in asexual populations. *Curr. Biol.* **17**, 385–394 (2007).
36. Fogle, C. A., Nagle, J. L. & Desai, M. M. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics* **180**, 2163–2173 (2008).

37. Vetterling, W. T. *Numerical Recipes Example Book* 2nd edn, Ch. 7 (Cambridge Univ. Press, 1992).
38. Wakeley, J. *Coalescent Theory: An Introduction* (Roberts, 2009).
39. Moura de Sousa, J. A., Campos, P. R. & Gordo, I. An ABC method for estimating the rate and distribution of effects of beneficial mutations. *Genome Biol. Evol.* **5**, 794–806 (2013).
40. Goyal, S. *et al.* Dynamic mutation-selection balance as an evolutionary attractor. *Genetics* **191**, 1309–1319 (2012).
41. Ewens, W. J. *Mathematical Population Genetics* 2nd edn (Springer, 2004).
42. Efron, B. & Tibshirani, R. *An Introduction to the Bootstrap* (Chapman & Hall, 1993).
43. Frenkel, E. M., Good, B. H. & Desai, M. M. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. *Genetics* **196**, 1217–1226 (2014).
44. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
45. Mumberg, D., Muller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119–122 (1995).
46. Cross, F. R. 'Marker swap' plasmids: convenient tools for budding yeast molecular genetics. *Yeast* **13**, 647–653 (1997).
47. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
48. Storic, F., Lewis, L. K. & Resnick, M. A. In vivo site-directed mutagenesis using oligonucleotides. *Nature Biotechnol.* **19**, 773–776 (2001).
49. Storic, F. & Resnick, M. A. The *delitto perfetto* approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol.* **409**, 329–345 (2006).
50. Pavelka, N. *et al.* Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**, 321–325 (2010).
51. Selmecki, A., Bergmann, S. & Berman, J. Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. *Mol. Microbiol.* **55**, 1553–1565 (2005).
52. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
53. Wenger, J. W. *et al.* Hunger artists: yeast adapted to carbon limitation show trade-offs under carbon sufficiency. *PLoS Genet.* **7**, e1002202 (2011).
54. Hittinger, C. T. *et al.* Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* **464**, 54–58 (2010).
55. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
56. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43**, 491–498 (2011).
59. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
60. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
62. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
63. Novocraft. Novocraft short read alignment package. <http://www.novocraft.com> (2009).
64. Homer, N., Merriman, B. & Nelson, S. F. BFAST: an alignment tool for large scale genome resequencing. *PLoS ONE* **4**, e7767 (2009).
65. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
66. Homer, N. & Nelson, S. F. Improved variant discovery through local re-alignment of short-read next-generation sequencing data using SRMA. *Genome Biol.* **11**, R99 (2010).
67. Dalca, A. V., Rumble, S. M., Levy, S. & Brudno, M. VARIID: a variation detection framework for color-space and letter-space platforms. *Bioinformatics* **26**, i343–i349 (2010).
68. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

Generation of isogenic YFP ploidy series



Generation of isogenic CFP ploidy series

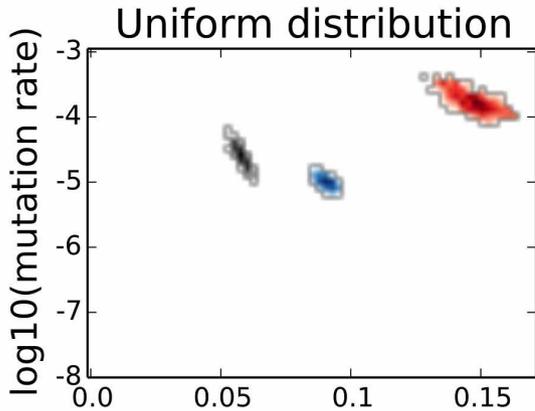


Extended Data Figure 1 | Schematic representation of the construction of isogenic haploid, diploid, and tetraploid strains used in this study. Relevant strain numbers are indicated for the CFP- and YFP-containing ancestors.

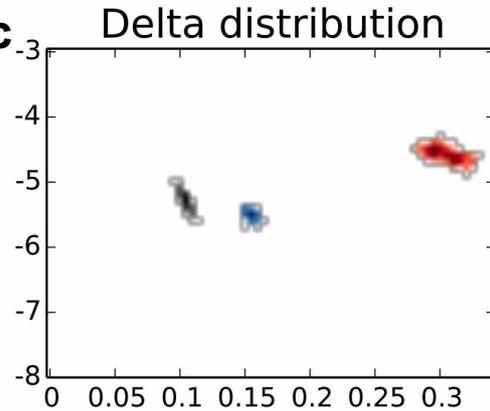
a

Model and Distribution	1N		2N		4N	
	μ	s	μ	s	μ	s
New Exponential	$2.5 \cdot 10^{-3}$ [$3.98 \cdot 10^{-4}$, $3.98 \cdot 10^{-3}$]	0.014 [0.012, 0.018]	$2.5 \cdot 10^{-3}$ [$1 \cdot 10^{-3}$, $3.16 \cdot 10^{-3}$]	0.024 [0.022, 0.026]	$5.0 \cdot 10^{-3}$ [$1.99 \cdot 10^{-3}$, $7.9 \cdot 10^{-3}$]	0.03 [0.028, 0.035]
New Uniform	$2.0 \cdot 10^{-5}$ [$1 \cdot 10^{-5}$, $3.2 \cdot 10^{-5}$]	0.06 [0.115, 0.13]	$7.9 \cdot 10^{-6}$ [$6.3 \cdot 10^{-6}$, $1 \cdot 10^{-5}$]	0.0925 [0.18, 0.195]	$1.2 \cdot 10^{-4}$ [$7.94 \cdot 10^{-5}$, $2.5 \cdot 10^{-4}$]	0.15 [0.275, 0.325]
New Delta	$3.9 \cdot 10^{-6}$ [$2.5 \cdot 10^{-6}$, $6.3 \cdot 10^{-6}$]	0.11 [0.105, 0.115]	$2.5 \cdot 10^{-6}$ [$1.9 \cdot 10^{-6}$, $3.1 \cdot 10^{-6}$]	0.16 [0.155, 0.165]	$1.99 \cdot 10^{-5}$ [$1.58 \cdot 10^{-5}$, $3.1 \cdot 10^{-5}$]	0.31 [0.295, 0.325]
Equivalence Principle model, Delta	$2.5 \cdot 10^{-7}$ [$1.58 \cdot 10^{-7}$, $3.16 \cdot 10^{-7}$]	0.1 [0.95, 0.105]	$2.5 \cdot 10^{-7}$ [$1.99 \cdot 10^{-7}$, $2.81 \cdot 10^{-7}$]	0.13 [0.125, 0.135]	$7.9 \cdot 10^{-7}$ [$6.3 \cdot 10^{-7}$, $1 \cdot 10^{-6}$]	0.2 [0.195, 0.21]

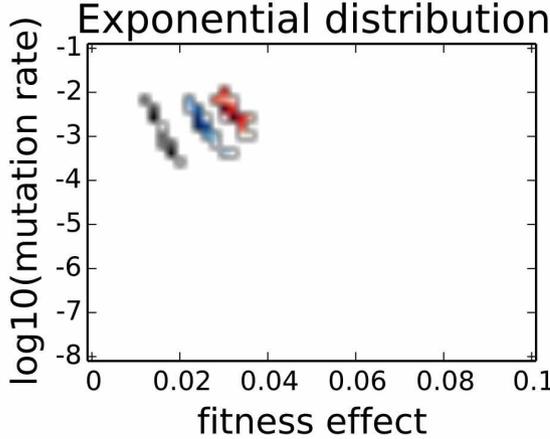
b



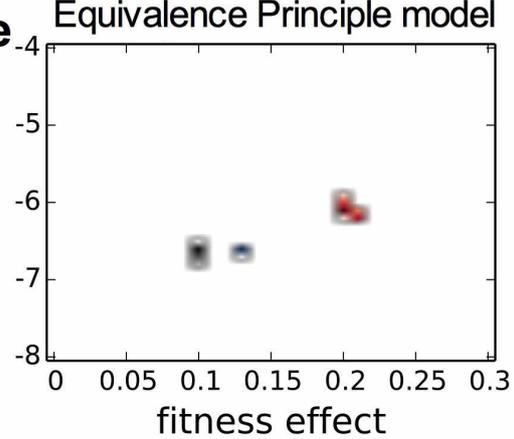
c



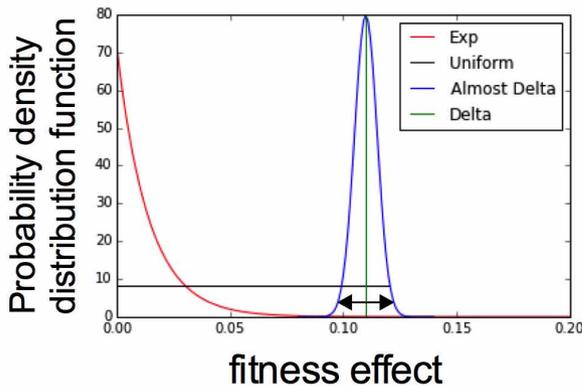
d



e

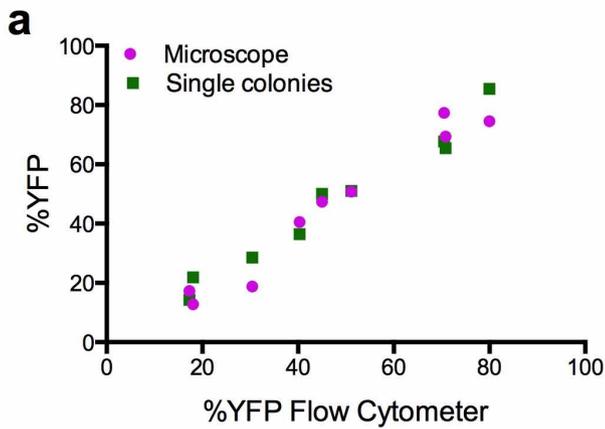


f



Extended Data Figure 2 | Estimates from our mathematical modelling of the best-fit value of the beneficial mutation rate (μ) and the selection coefficient (s) of each ploidy evolution experiment. **a**, Table of μ and s values that had the best-fit between the simulations and the experimental data; brackets, 95% confidence intervals. Values were determined on the basis of different assumptions about the underlying distribution of beneficial mutations, which included **(b)** uniform, **(c)** delta, and **(d)** exponential distributions. Estimates of μ and s were also obtained with **(e)** the equivalence principle model¹⁷, which assumes a delta distribution of beneficial mutations. Each two-dimensional plot includes the error range obtained by parametric bootstrap of 1,000 independent simulated data sets (Methods). The 95% confidence intervals of μ and s from those 1,000 data sets were defined as the error ranges. **f**, Schematic diagram of the three distributions of fitness effects that we used in our mathematical modelling: exponential (red), uniform (black), and delta (green) distributions. For illustration, we also provide a narrow Gaussian distribution (blue) that is close to a delta function. The real distribution of fitness effects probably has a more complex structure than any of

the examples shown. The diagram illustrates the fact that the shape of the assumed distribution mandates differences in mutation rates. For example, if the mutations that mainly drive adaptation fall within the region of the double arrow, only a small proportion of the mutations from the exponential distribution will fall within this range, necessitating a much higher mutation rate to generate mutations in this region. By contrast, the delta distribution lies in the middle of the double arrow range; therefore, all of the mutations that arise from this distribution are strong enough to contribute to adaptation, resulting in a relatively lower mutation rate. The uniform distribution is intermediate between these two extremes. Only a small portion of the mutations of the uniform distribution is within the double arrow region, but the probability of these mutations is orders of magnitude larger than the exponential. Therefore, the mutation rate of the uniform is closer to the delta than to the exponential distribution. The values used to generate this figure are the best-fit values of μ and s of the haploid populations in the different three distributions. See Methods for more details.



b

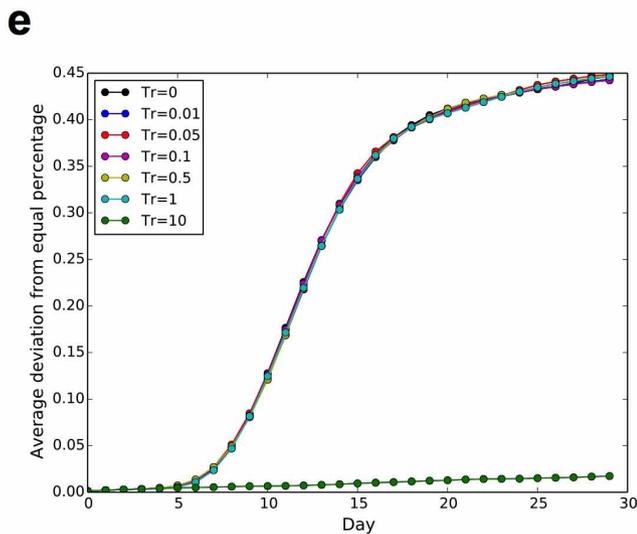
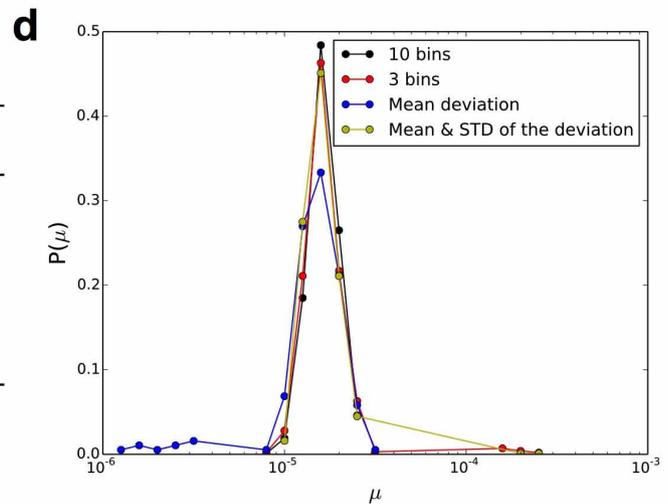
Standard deviation of the percent YFP

	100:0*	85:15	75:25	50:50	25:75	0:100
1N	0.04	0.36	0.52	0.50	0.66	0.09
2N	0.04	0.40	0.48	0.53	0.46	0.06
4N	0.00	0.30	0.51	0.56	0.41	0.04

*Ratio of CFP:YFP expressing cells

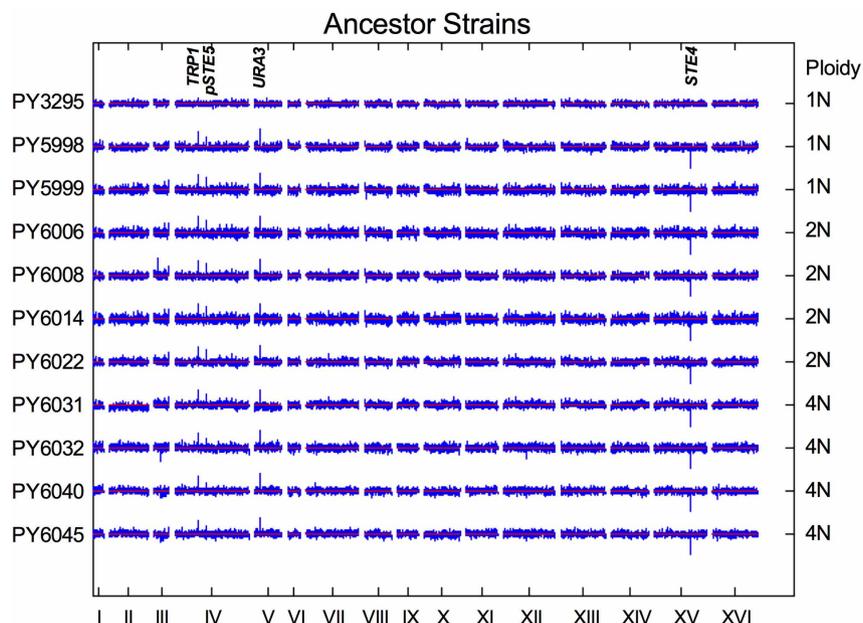
c

Ratio of U_b/U_d	μ	s
1	$2.52e-06 \pm 3.0e-07$	0.1604 ± 0.0027
0.9	$2.54e-06 \pm 2.9e-07$	0.1596 ± 0.0027
0.5	$2.52e-06 \pm 2.9e-07$	0.1594 ± 0.0025
0.1	$2.53e-06 \pm 3.2e-07$	0.1595 ± 0.0029
0.01	$2.54e-06 \pm 2.4e-07$	0.1605 ± 0.0030



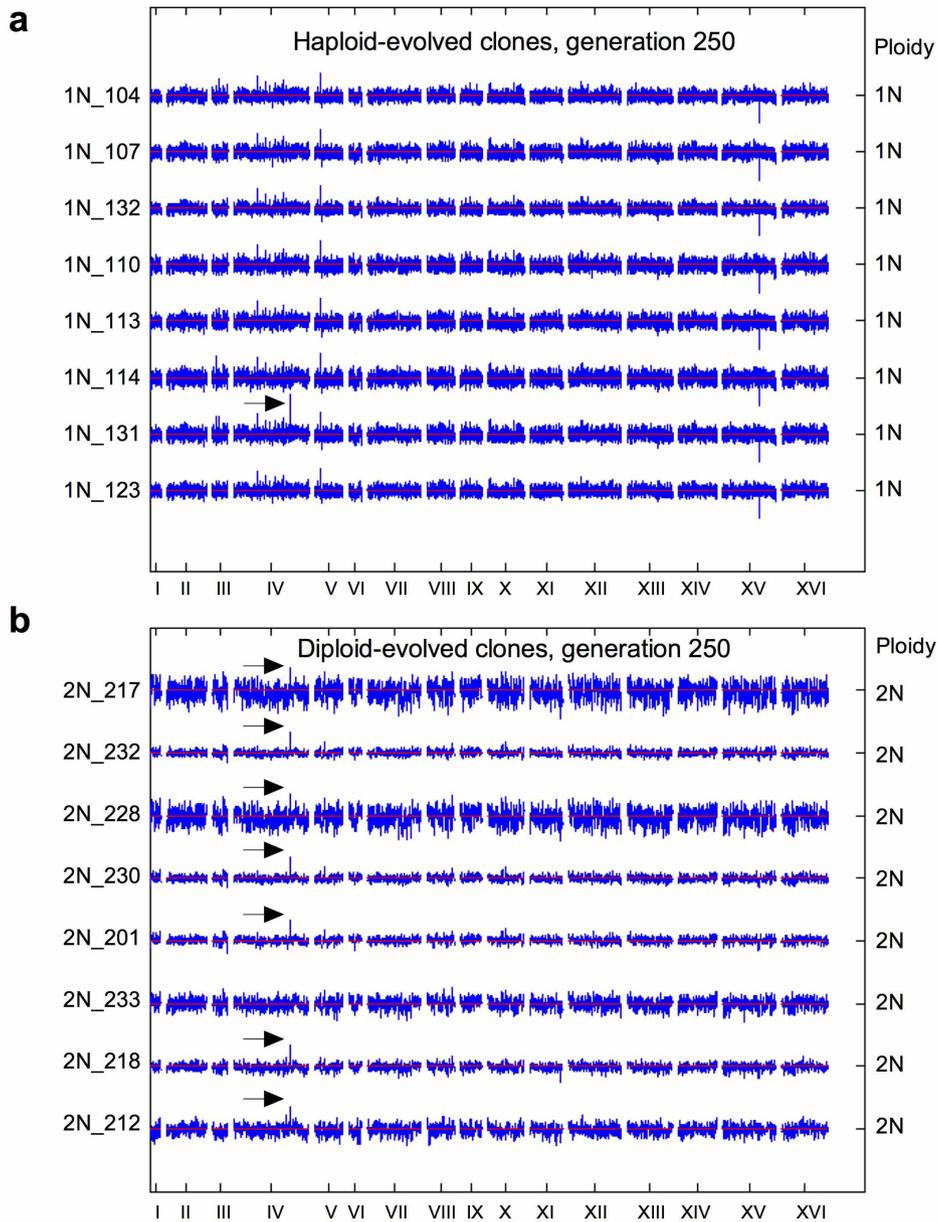
Extended Data Figure 3 | Experimental and computational analyses of the noise in our experimental measurements and of the methods used in our mathematical modelling (see Supplementary Information). **a**, Three different methods were used to determine the percentage of YFP-expressing cells in mixtures of the 1N, 2N, and 4N CFP and YFP ancestor strains. Cells were analysed by flow cytometry (10,000 cells) and fluorescence microscopy (300 cells), and by single colony analysis (96 colonies) of the mixture before galactose induction. The percentage YFP determined by all three methods was highly correlated (Pearson correlation coefficient = 0.98). **b**, Table showing variation in flow cytometry replicate measurements: the standard deviation of the percentage YFP obtained from 48 replicate populations of six different CFP:YFP ratios, for each ploidy type. **c**, Table showing the average and standard deviation of the best-fit values for different ratios between beneficial (U_b) and deleterious (U_d) mutations, obtained from 100 independent data sets. **d**, Evaluation of different summary statistics by calculating the distribution of best-fit values from 1,000 replicate simulations. Four different summary

statistics were used to analyse 1,000 replicates of a parameter pair, s and μ (see Methods). The summary statistic using ten bins has the highest mode and no outliers, and was used to generate our best-fit values. **e**, Criteria for exclusion of near-neutral mutations for implementation of the branching evolutionary model with an exponential distribution of mutations. Shown is the average deviation from equal percentages of YFP and CFP-expressing cells with different thresholds for neutral mutations. The threshold (Tr) represents the fraction of the average fitness effect (s), meaning every mutation whose absolute value was smaller than $Tr \times s$ was excluded. For this scenario (with parameters $\mu = 2 \times 10^{-5}$ and $s = 0.08$), we can exclude every mutation with a fitness effect smaller than s (that is, $Tr = 1$, light blue) without changing the outcome relative to excluding no mutations ($Tr = 0$). However, when excluding all mutations with fitness effects smaller than $10 \times s$ ($Tr = 10$, dark green), the result changes substantially. Thus, for high mutation rates ($N\mu > 1$), we can exclude weak mutations²⁰.



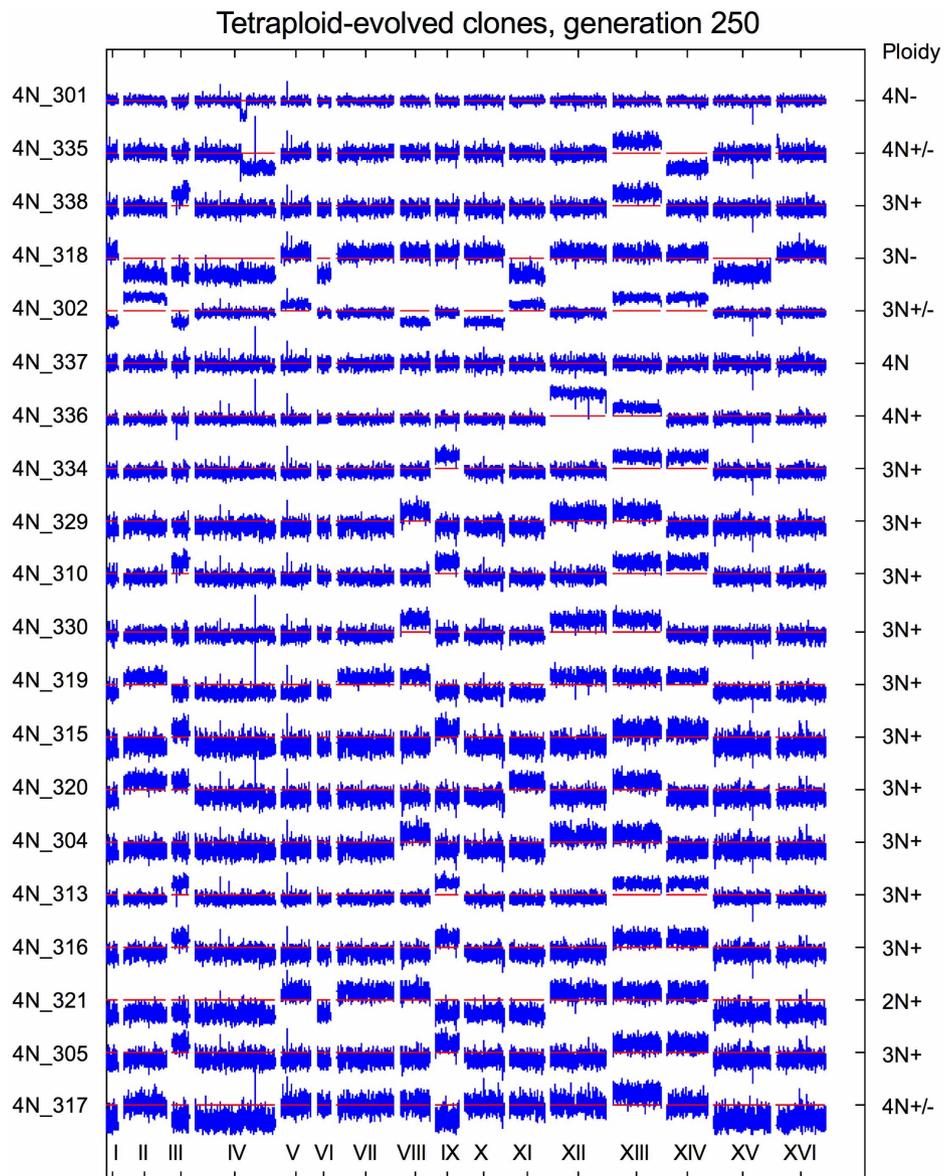
Extended Data Figure 4 | aCGH karyotype of the ancestor strains used in this study. Aneuploidy was not detected in the parental 1N, 2N, or 4N strains. Genomic DNA from each strain was compared with that of an isogenic ancestor PY3295 (BY4741 *MATa ura3 his3 trp1 leu2 LYS2*) and log₂ DNA copy number ratios were plotted using a custom Matlab script. To account for

regions of complete deletion, the data were cropped at log₂ ratios of ± 2.0 and averaged across each chromosome using a sliding window of nine oligonucleotides. A log₂ ratio of zero is indicated by the red line. Loci altered during strain construction are indicated (*TRP1*, *pSTE5*, *URA3*, *STE4*). Strain ploidy, determined by flow cytometry, is indicated on the right.



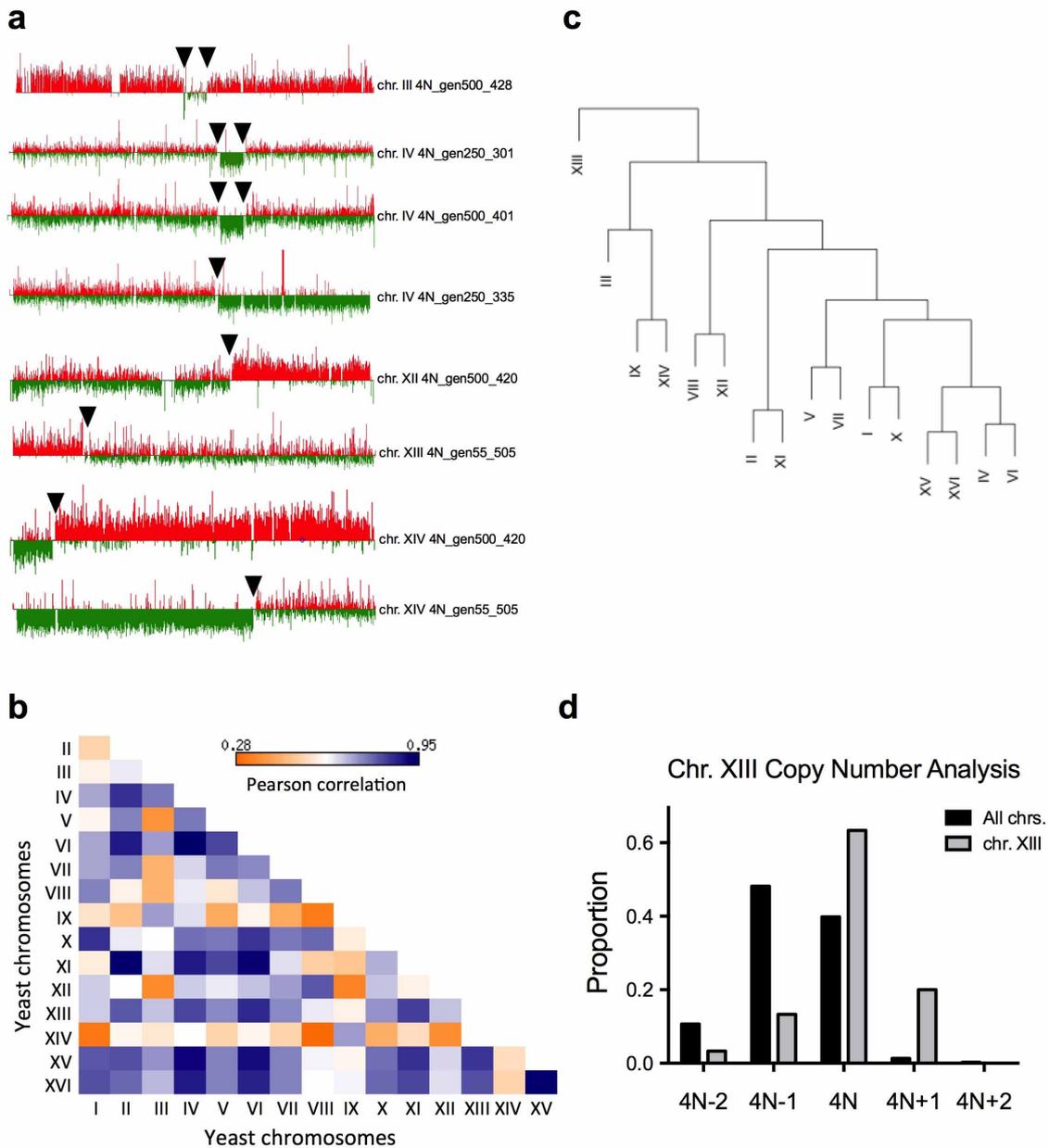
Extended Data Figure 5 | aCGH karyotype of haploid- and diploid-evolved clones at generation 250. **a**, aCGH of eight haploid-evolved clones. Data are displayed as in Extended Data Fig. 4. No aneuploidy was detected. Clone 1N_131 acquired the *HXT6/7* amplification (arrow). **b**, aCGH of eight

diploid-evolved clones. No aneuploidy was detected, but all clones except 2N_233 acquired the *HXT6/7* amplification. Log₂ ratios were averaged across each chromosome using a sliding window of 29 oligonucleotides. The ploidy of the evolved clone, determined by flow cytometry, is indicated on the right.



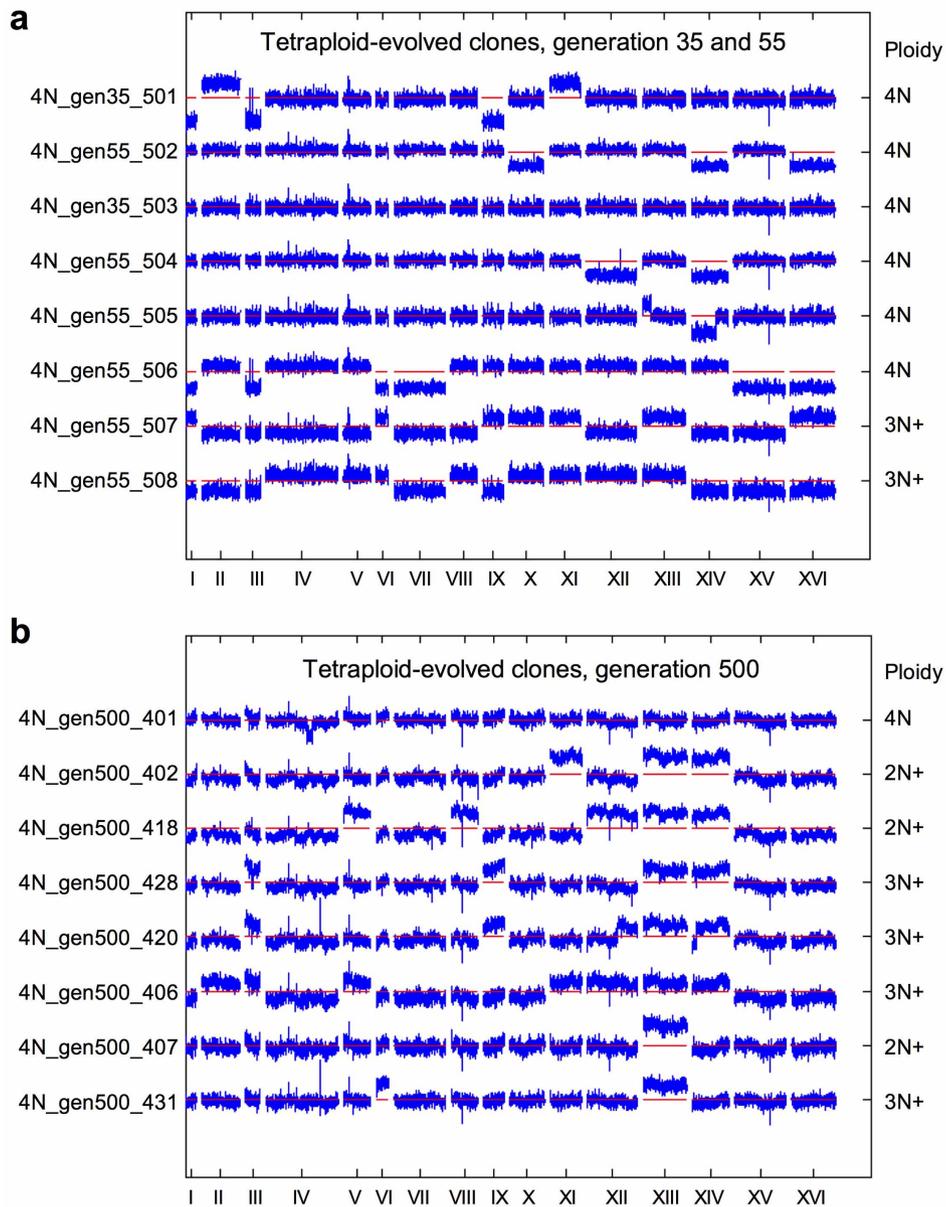
Extended Data Figure 6 | aCGH karyotype for 20 tetraploid-evolved clones at generation 250. aCGH data are displayed as in Extended Data Fig. 4. Note that whole chromosome or large segmental chromosome gain and loss events are observed in all clones except clone 4N_337. Ploidy of the evolved clone,

determined by flow cytometry, is indicated on the right, with +/- indicating chromosome aneuploidy. Some highly aneuploid clones had widely different chromosome copy numbers for different chromosomes (for example, some chromosomes were disomic, others trisomic and tetrasomic).



Extended Data Figure 7 | Analysis of recurrent and concerted chromosome loss events in the tetraploid-evolved clones. **a**, Evolved tetraploids acquired large segmental aneuploidies (regions greater than the ~ 7 kb *HXT6/7* amplification); aCGH data for individual chromosomes with large segmental aneuploidies in 4N-evolved clones (plotted using Treeview⁵²). All breakpoints occurred at or near Ty sequences (arrowheads). **b**, The pairwise patterns (Pearson correlation) of all chromosome copy number alterations in the 4N-evolved clones at generation 250 ($n = 30$, Supplementary Table 2). The copy numbers of some chromosomes were correlated (for example,

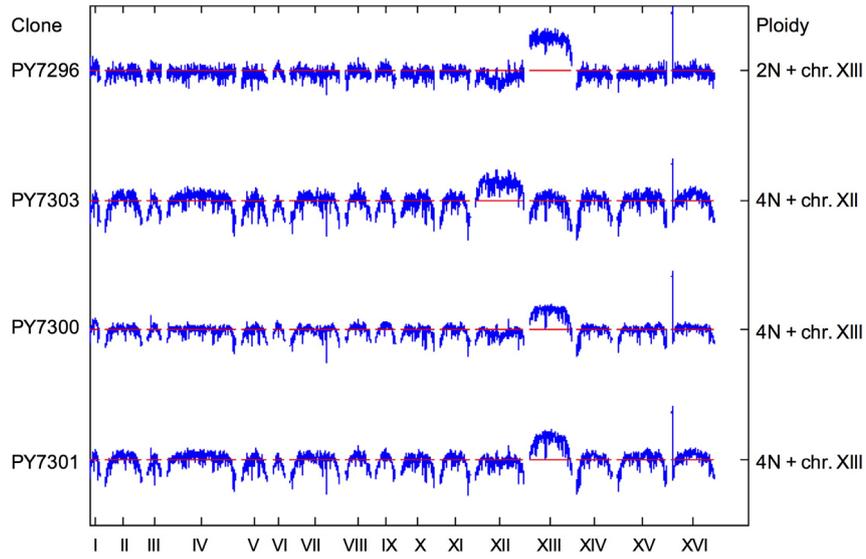
chromosome XV and chromosome XVI), whereas others were anti-correlated (for example, chromosome VIII and chromosome IX), possibly reflecting the need for gene expression balance. **c**, Hierarchical clustering showing the copy number relationship among the chromosomes. **d**, Proportion of all chromosomes in the evolved tetraploid clones with the indicated copy number (black). The copy number of chromosome XIII (grey) in the 4N-evolved clones at generation 250 was significantly different from that of all other aneuploid chromosomes (Cochran–Armitage test, $P < 1 \times 10^{-7}$).



Extended Data Figure 8 | aCGH karyotype for tetraploid-evolved clones at generations 35, 55, and 500. All 4N-evolved clones at (a) generations 35 and 55 and (b) generation 500 were aneuploid for multiple chromosomes or carried large segmental chromosome aneuploidies, except for clone

4N_gen35_503, which remained tetraploid. Data are displayed as in Extended Data Fig. 4. Ploidy of the evolved clone, determined by flow cytometry, is indicated on the right, with +/- indicating chromosome aneuploidy.

Aneuploid Strain Construction



Extended Data Figure 9 | aCGH from isogenic 2N and 4N strains with an extra copy of chromosome XIII or chromosome XII. Data are displayed as in Extended Data Fig. 5b.

Extended Data Table 1 | Yeast strains and plasmids used in this study

Strain (Ploidy) or Plasmid	Parental strain	Relevant genotype	Source
BY3295 (1N)	BY4741	<i>MATa his3Δ leu2Δ met15Δ ura3Δ</i>	Pellman collections
PY5997 (1N)	BY3295	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R</i>	This study
PY5998 (1N)	PY5997	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY5999 (1N)	PY5997	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6006 (2N)	PY5999	(2x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6008 (2N)	PY5998	(2x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6014 (2N)	PY5999	(2x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6022 (2N)	PY5998	(2x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6031 (4N)	PY6008	(4x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6032 (4N)	PY6022	(4x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6040 (4N)	PY6006	(4x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6045 (4N)	PY6014	(4x) <i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY7284 (1N)	PY5999	<i>SNF3-G439E</i>	This study
PY7289-PY7290 (2N)	PY5999	<i>SNF3-G439E/SNF3</i>	This study
PY7285-PY7288 (2N)	PY5999	<i>SNF3-G439E/SNF3-G439E</i>	This study
PY7291-PY7294 (4N)	PY5999	<i>SNF3-G439E/SNF3/SNF3/SNF3</i>	This study
PY7295 (2N)	S288c	<i>RLY4737 MATa/α ura3Δ his3Δ trp1Δ leu2Δ</i>	Ref. 50
PY7296 (2N)	PY7295	<i>RLY4888 MATa/α + chr. XIII trisomy</i>	Ref. 50
PY7297-PY7299 (4N)	PY7295	<i>MATa/a/a/α</i>	This study
PY7300-PY7302 (4N)	PY7295	<i>MATa/a/a/α + chr. XIII pentasomy</i>	This study
PY7303-PY7305 (4N)	PY7295	<i>MATa/a/a/α + chr. XII pentasomy</i>	This study
PB1500		<i>YFP-tADH-SpHIS5, Amp^R</i>	Yeast Resource Center
PB1499		<i>CFP-tADH Kan^R Amp^R</i>	Yeast Resource Center
PB2452		<i>CFP-tADH SpHIS5, Amp^R</i>	Pellman collection
PB2694		<i>pGAL1-ceCFP-tADH-SpHIS5, Amp^R</i>	This study
PB2697		<i>pGAL1-eYFP-tADH-SpHIS5, Amp^R</i>	This study
PB2314		<i>MATa::pSTE5-URA3, Amp^R</i>	Ref. 12
PB1308		<i>ura3::TRP1 Amp^R</i>	Ref. 46
PB2577		<i>MATα URA3 CEN Amp^R</i>	This study
B1819		<i>LEU2 CEN Amp^R</i>	Pellman collection
PB2647		<i>STE4 LEU2 CEN Amp^R</i>	This study
PB2649		<i>STE4 URA3 CEN MATα, Amp^R</i>	This study
PB1640		<i>hphMX4 Amp^R</i>	Ref. 47
PB1942		<i>pGAL-HO HIS3 Amp^R</i>	Gift of the Fink lab
PB1650		<i>pGAL-HO URA3 LEU2</i>	Gift of the Elion lab
pCORE		<i>kanMX4 KIURA3</i>	Ref. 48