Scopel et al. GENETICS
Supplemental Figure Legends

Figure S1. Increases in aneuploidy prevalence are not an artefact of clonal expansion of aneuploid strains. A. Maximum likelihood tree for the 621 strains used for modeling that shows the full karyotype information for each strain. See Fig 6 for all clade labels. **B.** Maximum likelihood tree for 453 strains (422 diploids and 31 polyploids) after removing closely related strains (genetic distance < 0.000007, see Methods). Logistic regression analysis of the 422 diploids showed differences in the prevalence of Chr 2-16 gain among genetic clades (df=25, P=0.001). It confirmed increases in Chr 2-16 gain for all groups identified in the analysis of 621 strains; however, the increase was not significant for French Dairy and African Beer (df=1, P=0.07) or Mosaic Beer (df=1, P=0.2), likely due to reduced statistical power with fewer strains – these groups are highlighted with dashed boxes for comparison to A. Nonetheless, this model shows that high aneuploidy rates in genetically dissimilar lineages of *S. cerevisiae* are not easily explained by over-representation of a few closely-related aneuploids. The minimal adequate model after removing close relatives showed probabilities of chromosome gain that are similar to those in the larger dataset of 621 strains: an ancestral predicted probability of Chr 2-16 gain of 10% and elevated predicted probabilities for Sake ($P_{Chr2\text{-}16}$= 0.54) and Ale beer and Mixed Origin clades ($P_{Chr2\text{-}16}$ = 0.34).

Figure S2. Ancient admixture cannot fully explain high-aneuploidy clades. A. Log likelihood of TreeMix admixture graph models as a function of the number of possible migration events. The graph shows that the model with 8 migrations has considerably higher likelihood than simpler models, and indicating that invoking more migration events does not greatly change the likelihood of the data. B. Admixture graph showing the relationship between lineages with 8 migration events. Clades are colored as in Fig 6A, and high-aneuploidy groups are indicated with grey boxes. Branch lengths were estimated by maximum likelihood and correspond to genetic drift between populations. Arrows represent gene flow predicted from one lineage to another, where arrow colors indicate admixture weight (an estimate of the fraction of the genome from the source population). C. Same as B, but showing only the 4 migration events confirmed by three-population (f3) tests (REICH *et al.* 2009) of allele frequency differences between populations that imply gene tree discordance caused by admixture (f3 < 0, Z-score < -2), see Methods. Relationships estimated while accounting for ancient admixture are similar to those implied by the topology of the whole-genome tree (Fig. 6A), except that Mosaic Beer, Ale Beer, and Mixed Origin clades can be combined to one larger "beer and bakery" group and African Beer and French Dairy strains are not indicated as sister clades. Multiple migration events from the ancestor of low-aneuploidy Asian Fermentation and high-aneuploidy Sake to the beer and Mixed Origin lineages recapitulate previously known multiple admixture events (Fay *et al.* 2019). We see no strong evidence for admixture to French Dairy or African Beer from other high-aneuploidy groups.

Figure S3. Sake karyotypes in the context of the genome-wide maximum likelihood tree. A close up of the Sake strains in the maximum likelihood tree from Fig 6A suggests that gains of chromosome 3 and 11 have arisen in distinct sake strains (see also Fig S1B). Black squares indicate aneuploid strains. Grey circles indicate nodes with 100% bootstrap support.

Figure S4. Distribution of normalized growth scores pooled across all conditions from Peter *et al.* Strain-specific growth rates measured in each of 35 conditions were normalized to each strain's growth rate in rich medium, and then these growth scores were divided by the median score of the euploid group from that clade measured in that condition. All scores, which represent strain-, clade-, and condition-normalized growth rates were pooled and plotted, as

shown in Fig 4 with euploid strain distributions in grey and strains with Chr 2-16 amplification shown in clade-specific colors. Asterisk, aneuploid growth defect at FDR < 0.05, + at FDR < 0.07. 'Anc' indicates all strains excluding European Wine and those in high-aneuploidy groups as outlined in Fig 4. Trends in growth-rate effects and sporulation efficiency for European Wine versus non-European Wine ancestral frequency (Anc) strains remained consistent when analyzing only diploid strains and excluding six European Wine isolates with genome content (Mb) greater than that of the largest Anc aneuploid.