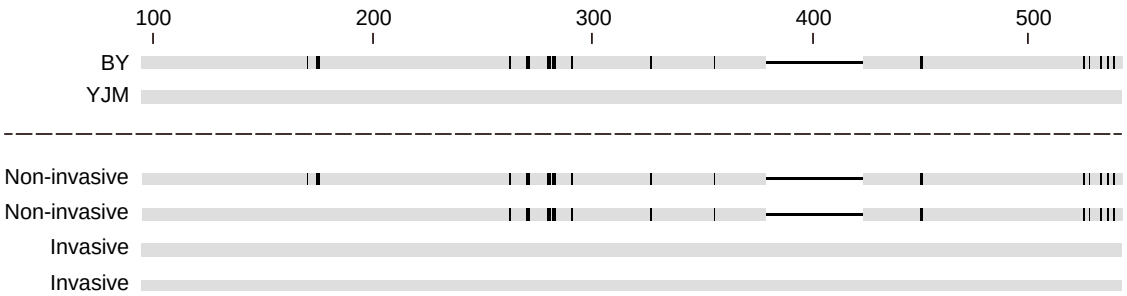

Figure S4. **Replacement of the *FLO11* coding region in segregant 2 with the BY allele causes loss of invasion.** To verify that *FLO11*^{BY} was correctly integrated and replaced using our one-step allele replacement, we PCR amplified the 5' end of the gene, and Sanger sequenced multiple invasive and non-invasive transformants. Only the transformants carrying the BY SNPs (marked in black) toward the 5' end showed loss of invasion, implying that only individuals with most of the *FLO11* gene replaced exhibited loss of invasion. Flo11 protein is comprised of three domains, which are reflected in the sequence of the *FLO11* gene. The N-terminal portion of the protein encodes a hydrophobic signal sequence, is exposed at the cell surface, and binds to ligands. The middle domain largely contains variable length tandem repeats that are enriched for serines and threonines, and is the part of the protein where heavy glycosylation occurs. The C-terminal portion of the protein is a GPI anchor that localizes Flo11 to the cell wall. The highly repetitive nature of the middle portion of *FLO11* makes it difficult to accurately determine the length and sequence of the gene using short Illumina reads. In the regions that we were able to confidently align, we identified 69 SNPs between the BY and the YJM allele, of which 31 were non-synonymous. In addition, we identified that the YJM allele of *FLO11* has a 45bp insertion in the N-terminal region between amino acid position 123 and 124. We also found that no sequencing reads from the YJM mapped to 635 base positions in comparison to BY, which is most likely due to deletions given that the YJM allele of *FLO11* was ~700 bases smaller in comparison to the BY allele (Figure S4). In particular, large stretches of the middle domains were missing from amino acid positions 207 to 315, 359 to 372, 409 to 449, 795 to 808, 824 to 845, and 881 to 899 in the YJM allele. We have not yet determined how these changes alter the functionality of Flo11. We note that this portion of the gene is known to be highly variable across yeast strains, affecting many *FLO11*-dependent traits, such as biofilm formation, flocculation, and invasion.