



Supplemental Figure S3. Relationship between D17Z1 array size and variation

A. Scatterplot of D17Z1 size versus variation in our HSA17 dataset. There is a weak correlation between small D17Z1 array size and increased variation (proportion of variant HORs) within the array. Blue triangles denote HSA17s with a stable D17Z1 centromere (i.e. D17Z1 is the site of centromere assembly); green squares denote HSA17s with an unstable D17Z1 centromere, and red circles represent HSA17s with high D17Z1 variation, an inactive D17Z1 array, and a stable D17Z1-B centromere. The data appear to be skewed by the extremely large and variant D17Z1 array in Z1_4.0 that is mitotically unstable (Figure 4). If this array is excluded from the analysis due to its unusual behavior, the correlation between large D17Z1 size and homogeneity (less variation) increases ($R^2=0.57$; Pearson correlation = 0.7557, Spearman correlation 0.69337).