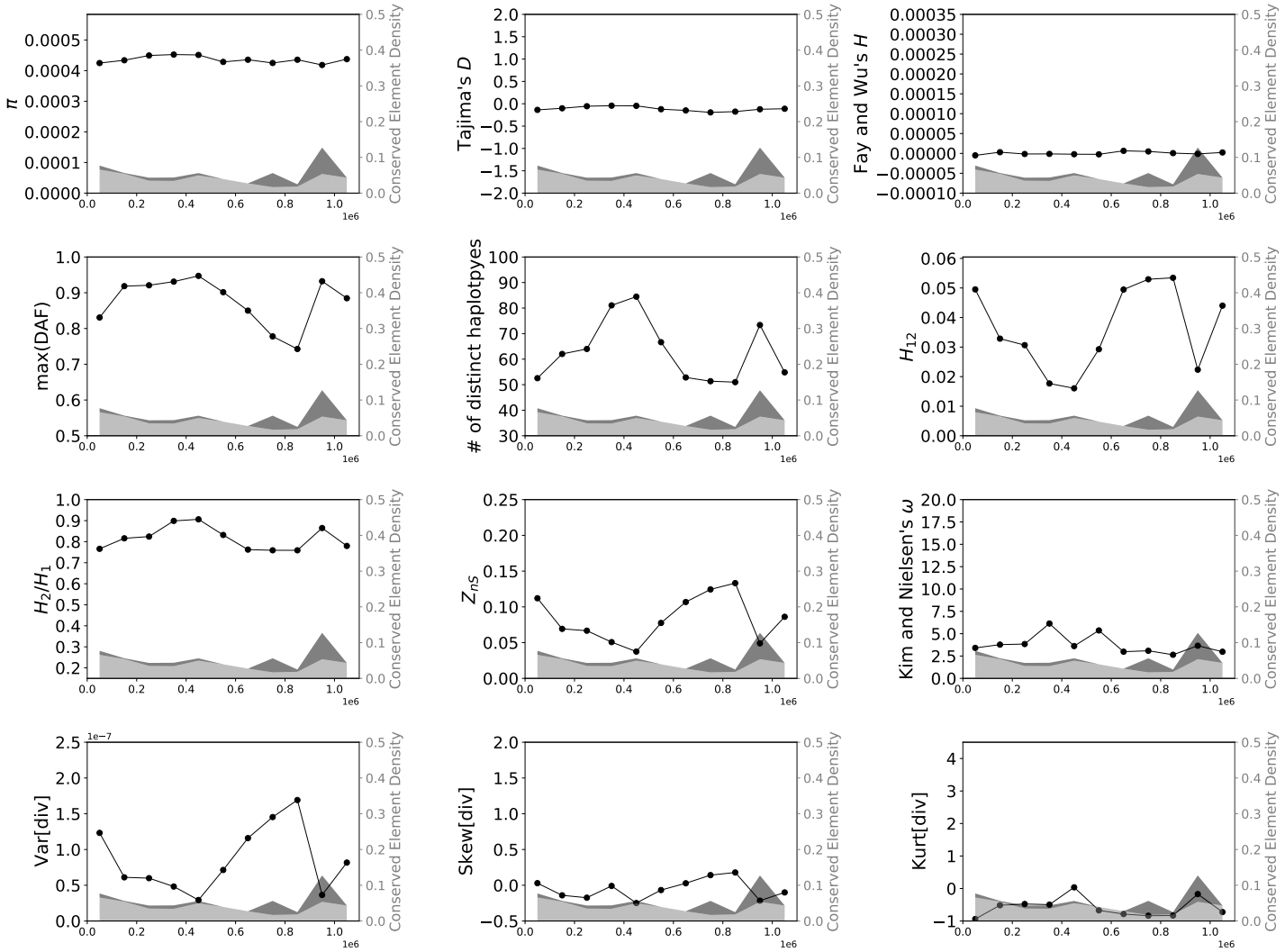
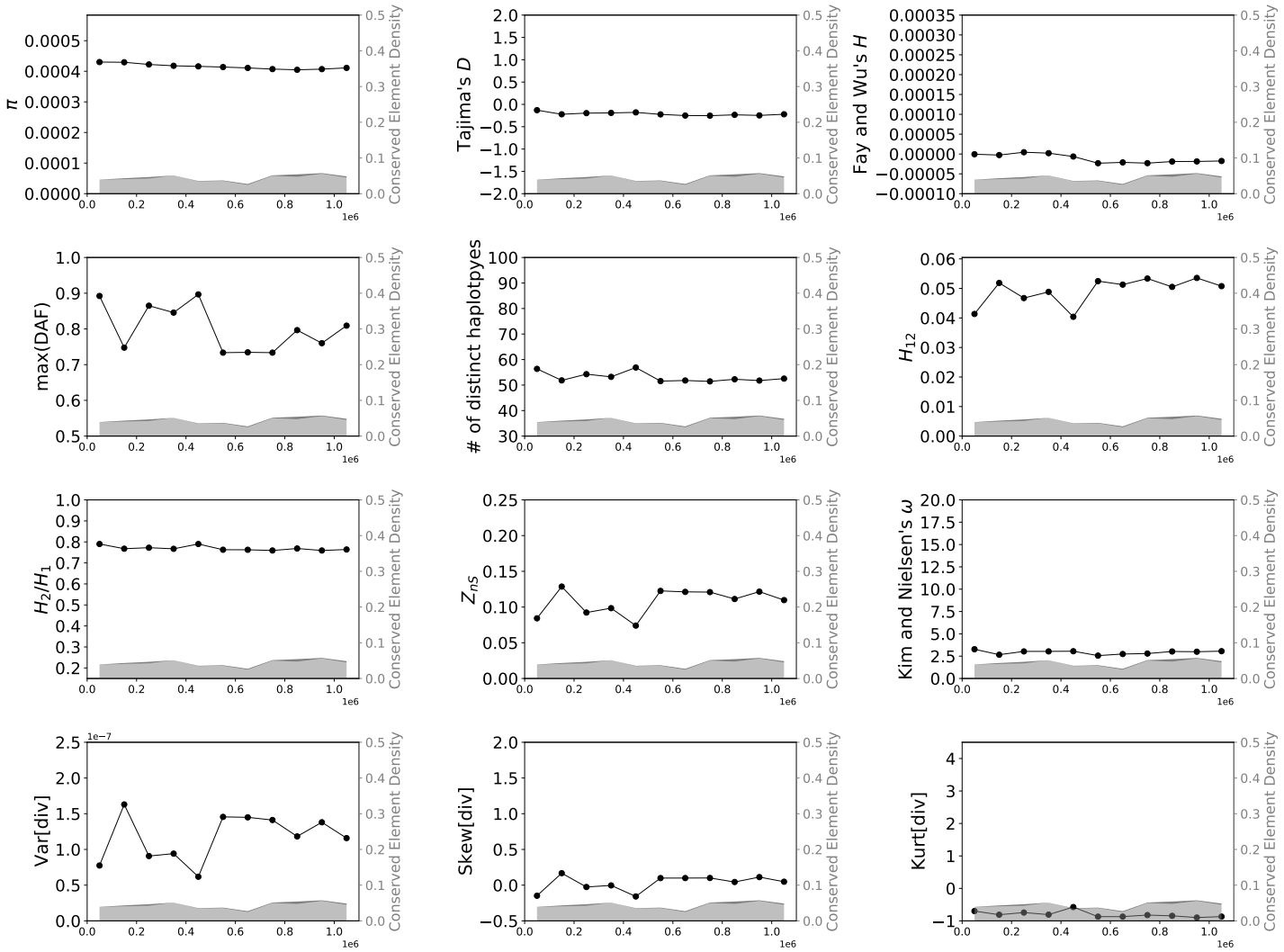
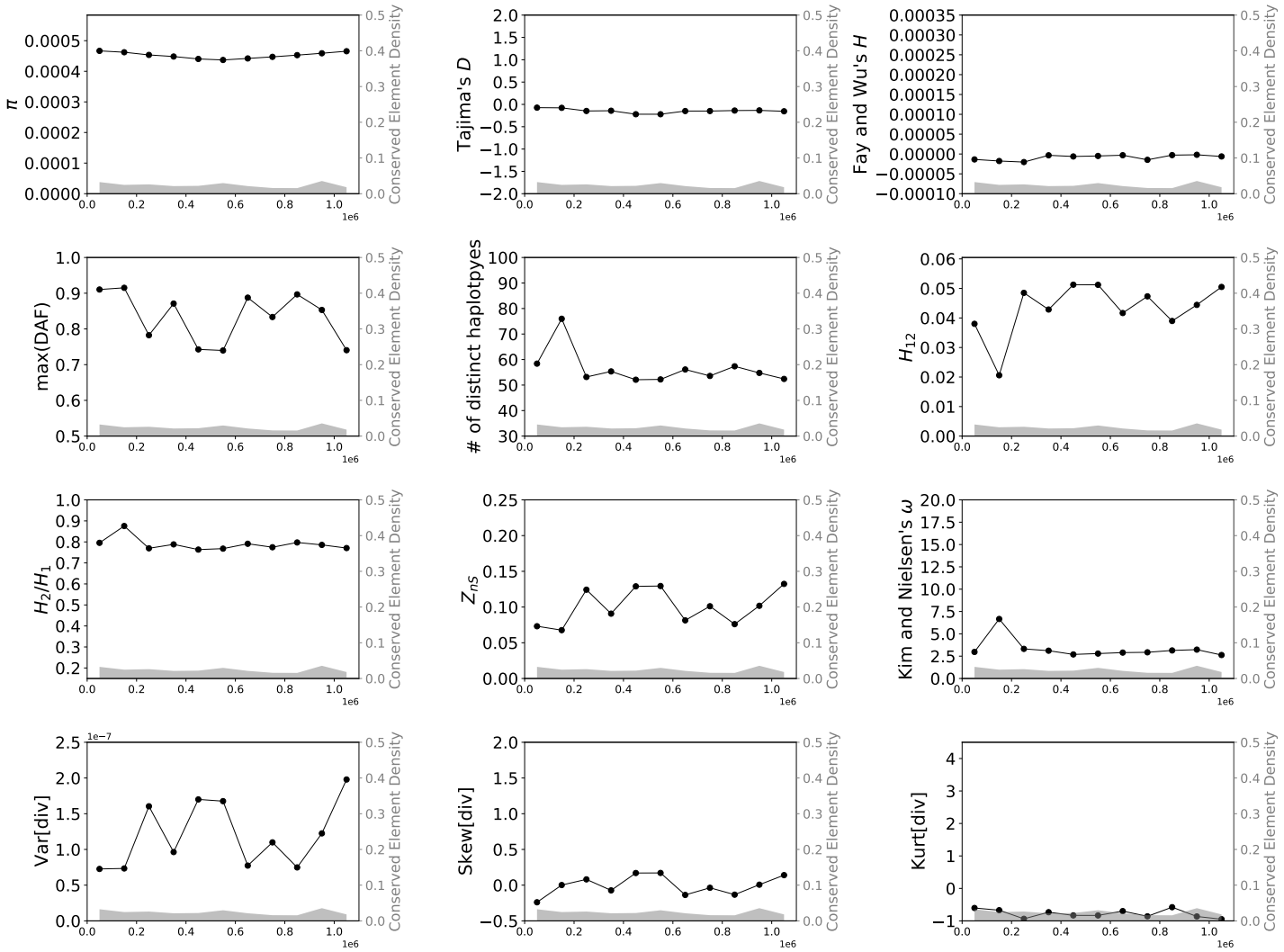


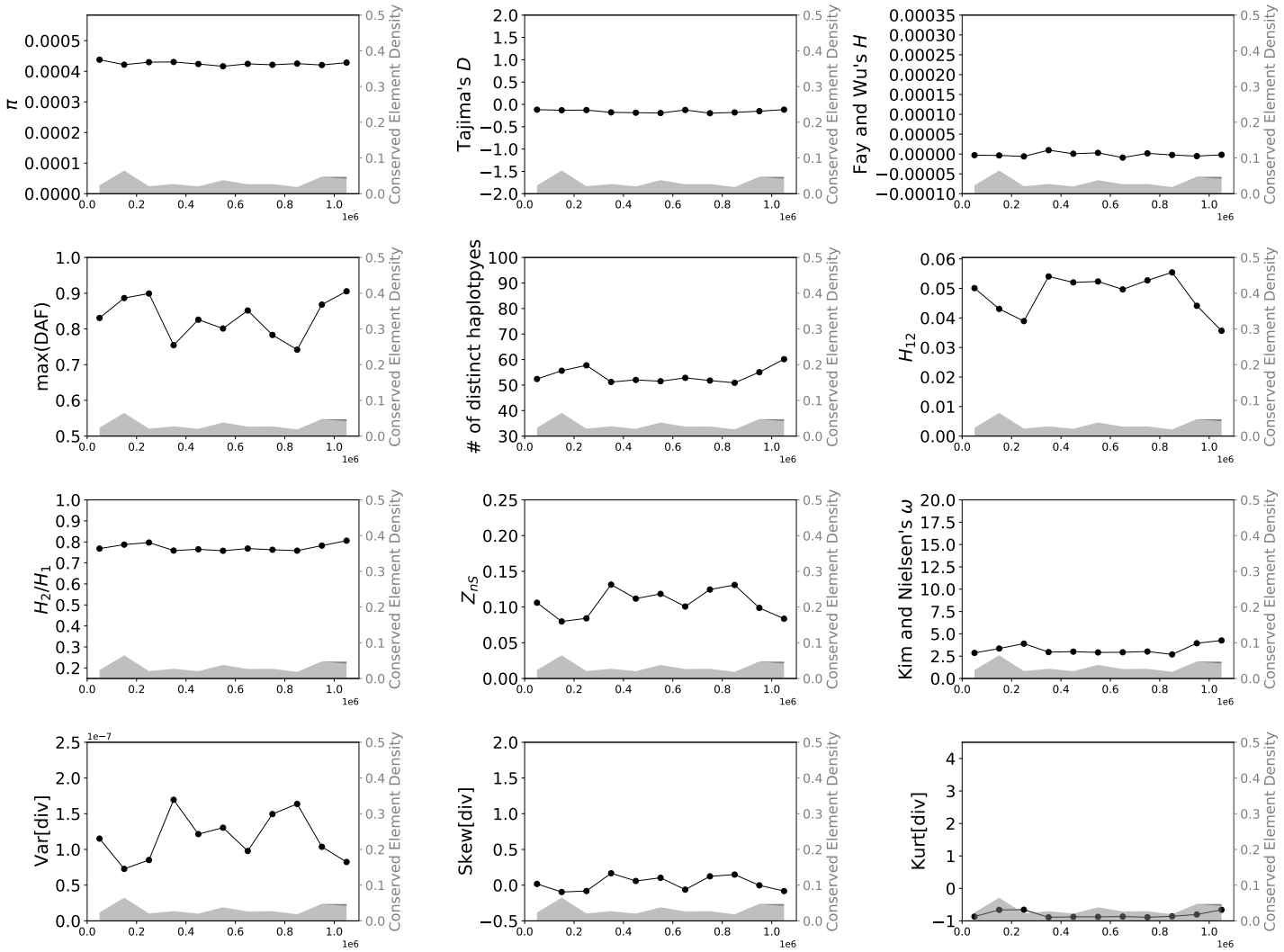
Supplementary Figure 1: Values of 15 statistics calculated from a simulated data set of 1000 replicates designed to match the locations of conserved sites from chr7:37000001-38100000 in the human genome (hg19 assembly version), simulated under equilibrium demography. Each replicate used the same overall recombination rate and map (chosen to match this region), however the mutation rate varied from replicate to replicate in the same manner as described in the Methods. Position labels on the x-axes are relative to the beginning of the window. The density of exonic and conserved non-coding DNA is shown in the shaded dark gray and light gray regions, respectively.



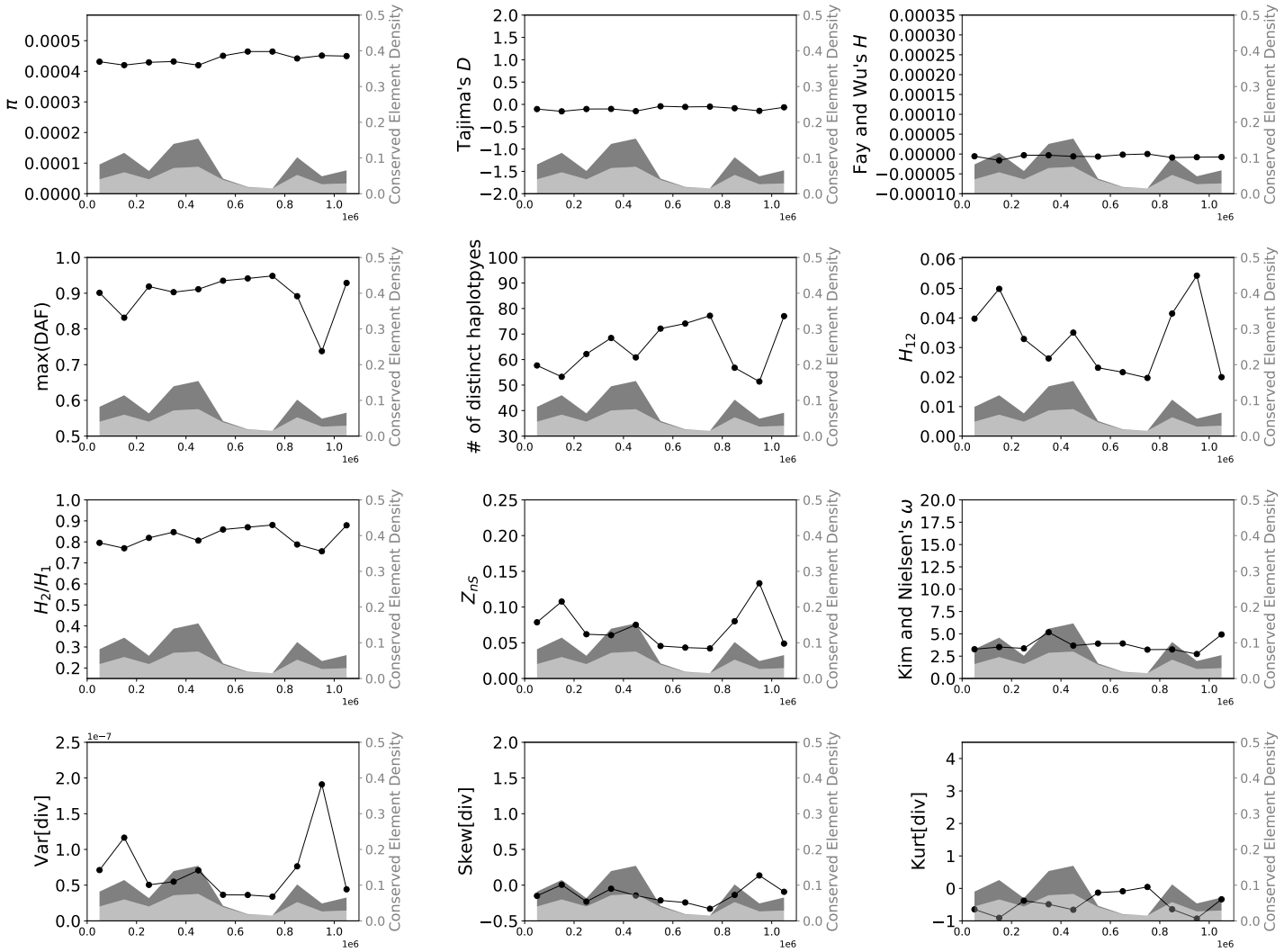
Supplementary Figure 2: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr4:98200001–99300000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



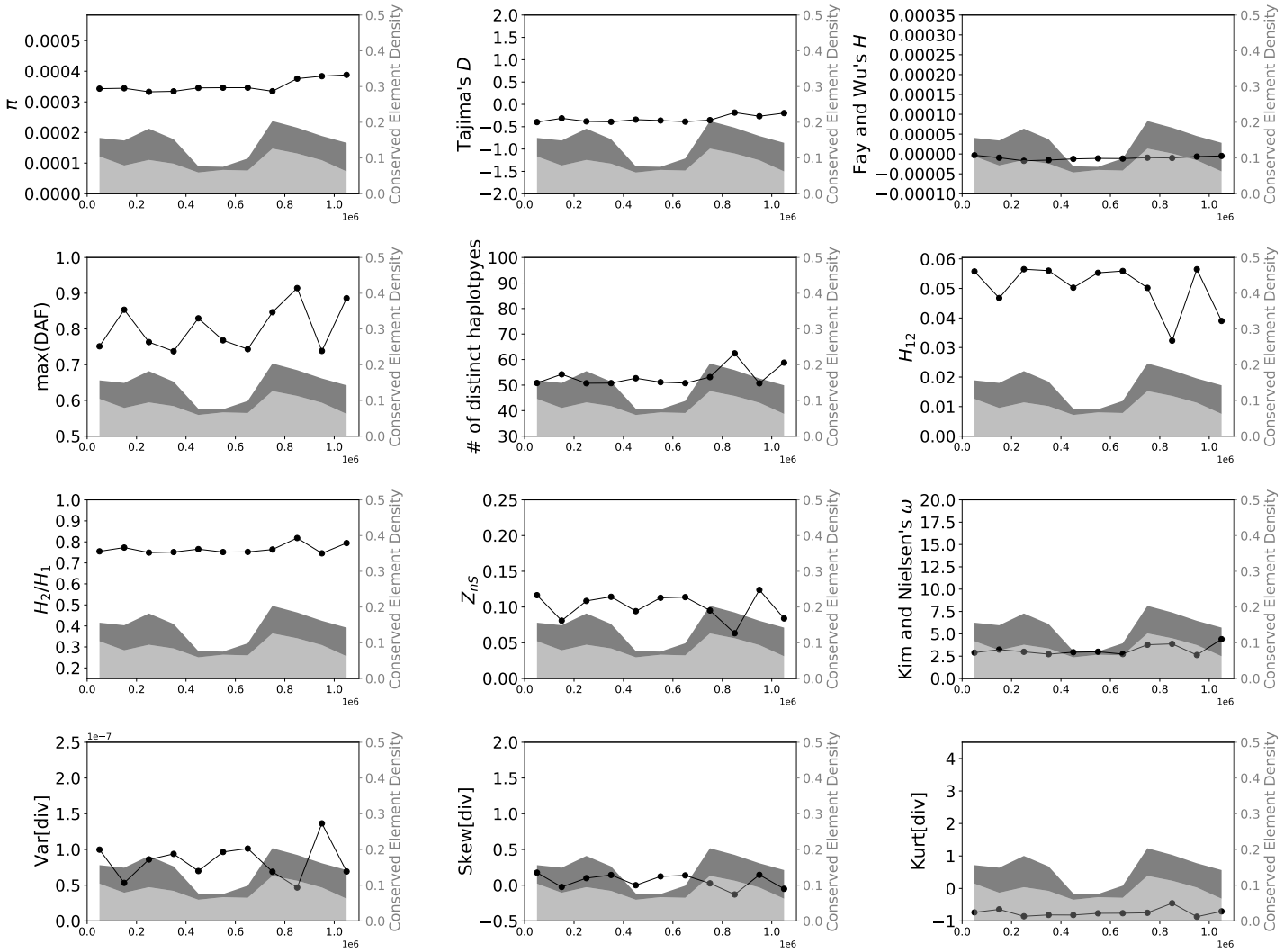
Supplementary Figure 3: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr4:131500001–132600000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



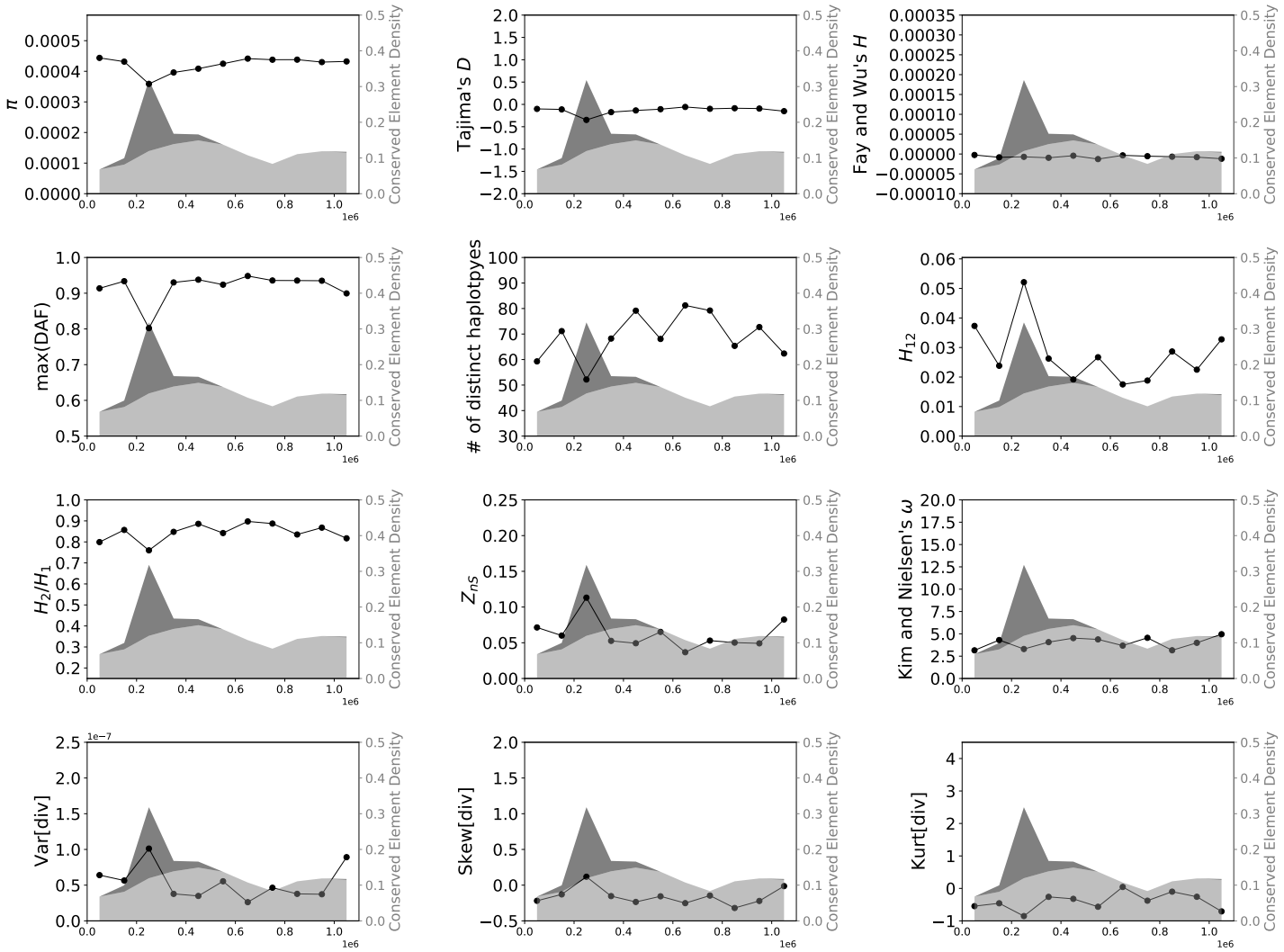
Supplementary Figure 4: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3:80100001–81200000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



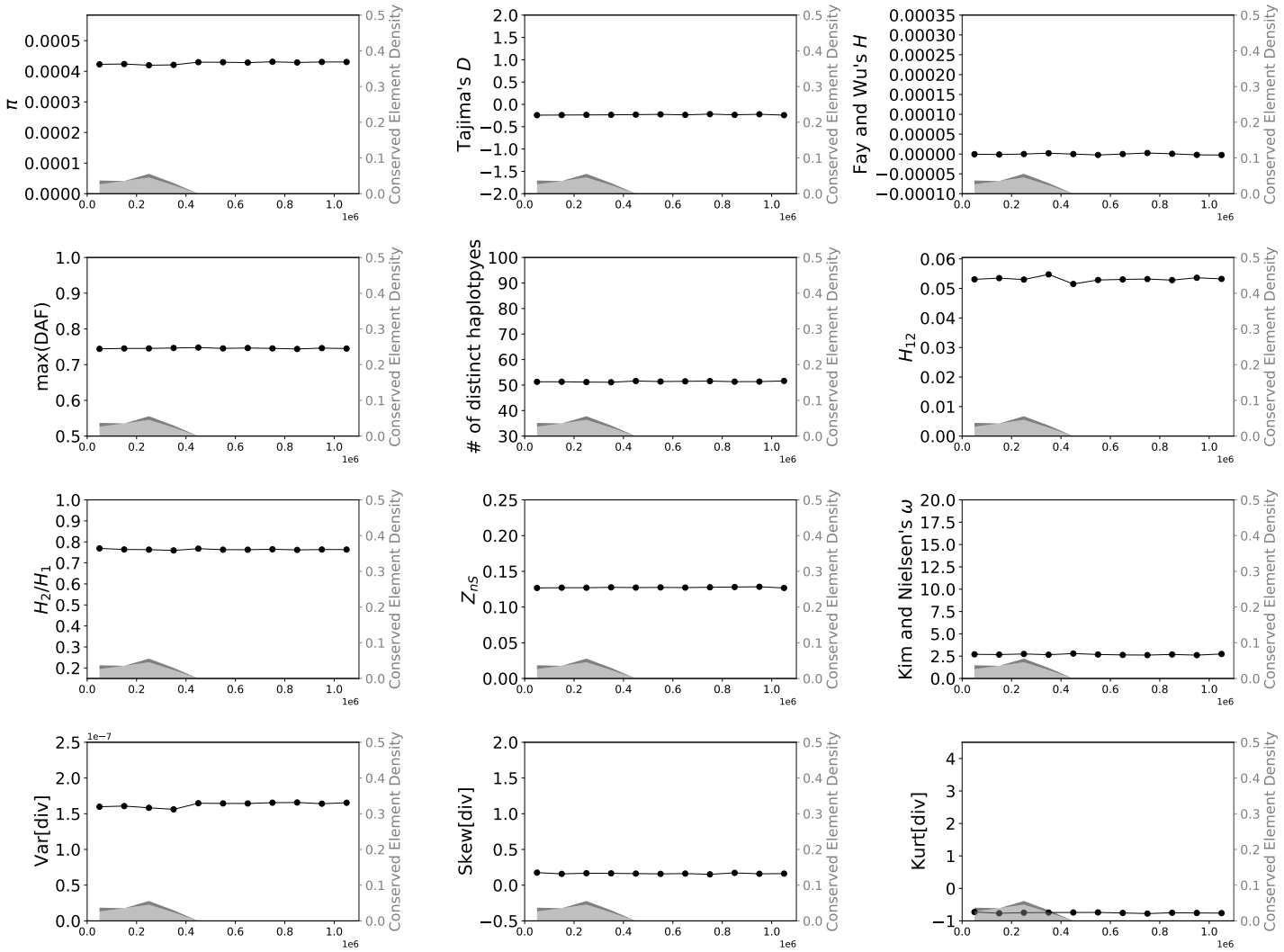
Supplementary Figure 5: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3:12500001–13600000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



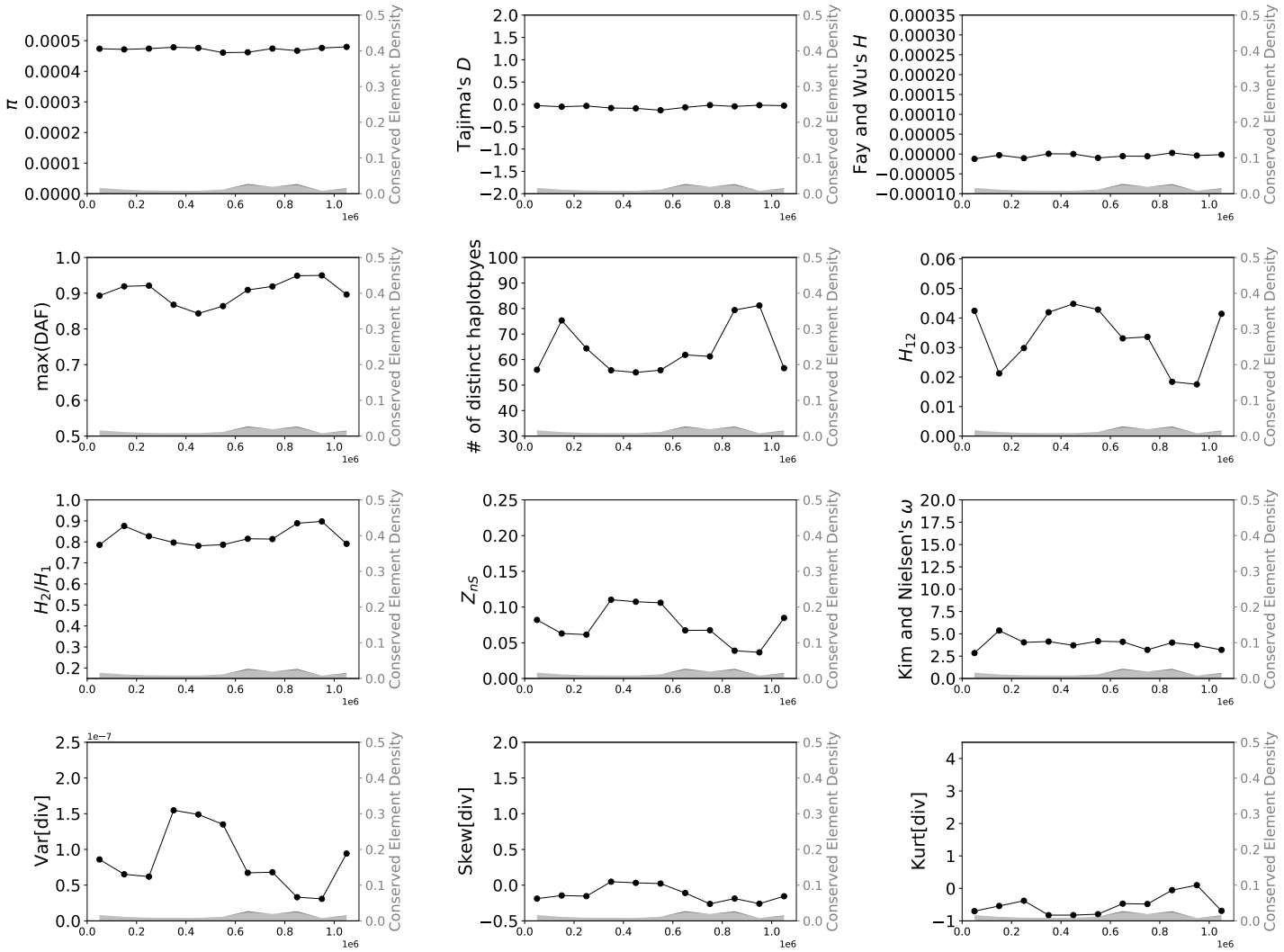
Supplementary Figure 6: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr1:45800001–46900000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



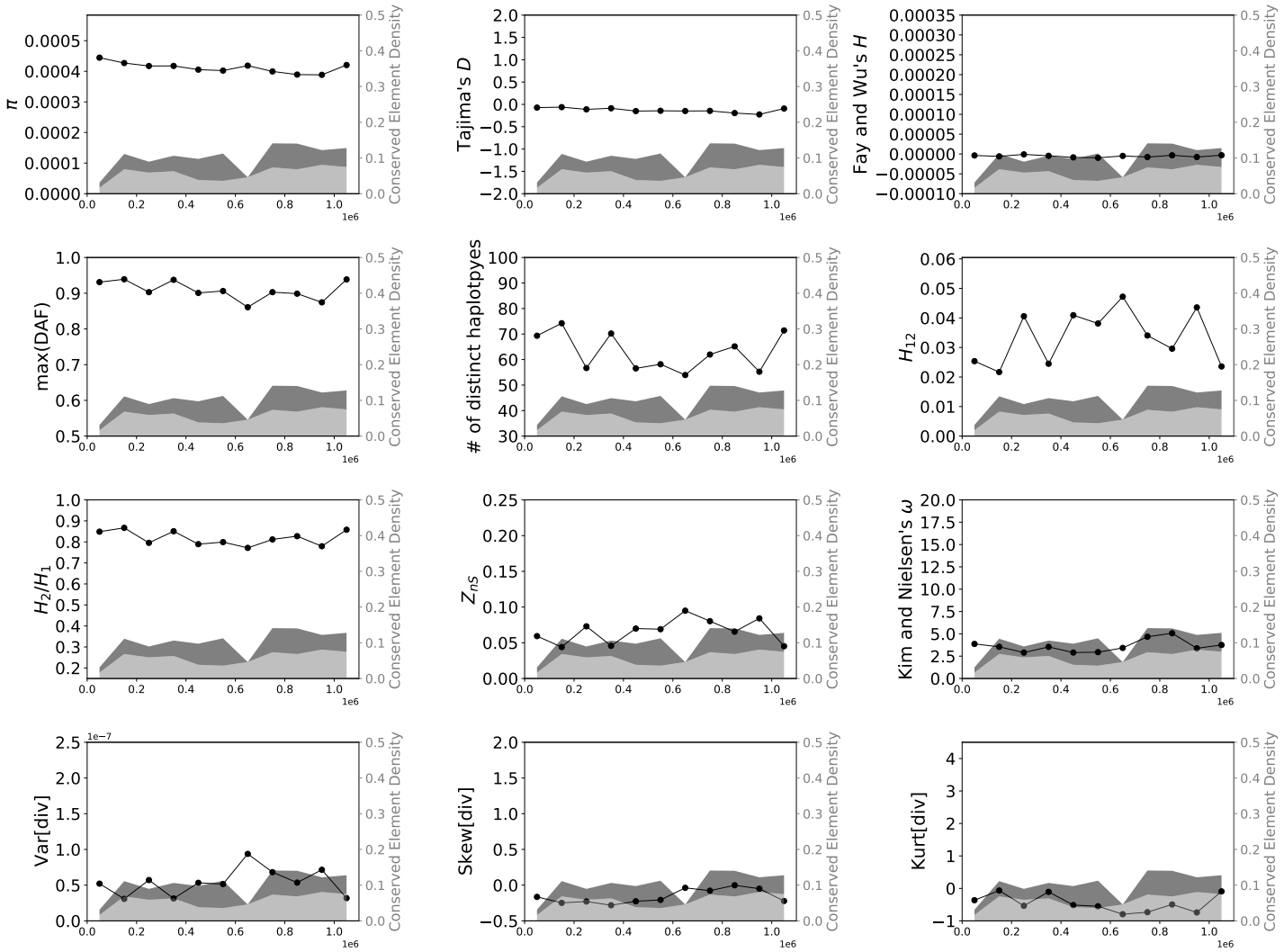
Supplementary Figure 7: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr15:60500001–61600000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



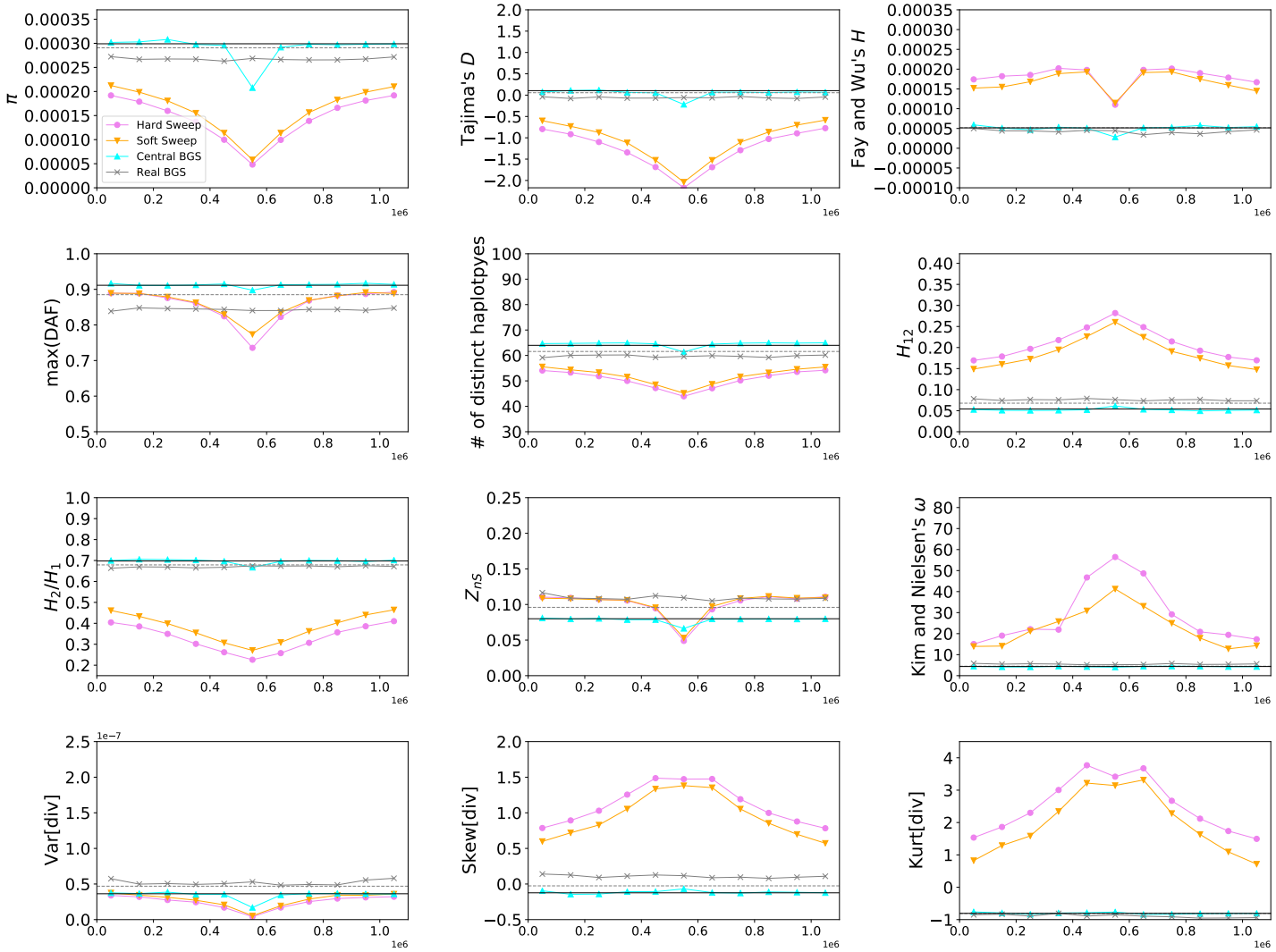
Supplementary Figure 8: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr11:50000001–51100000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



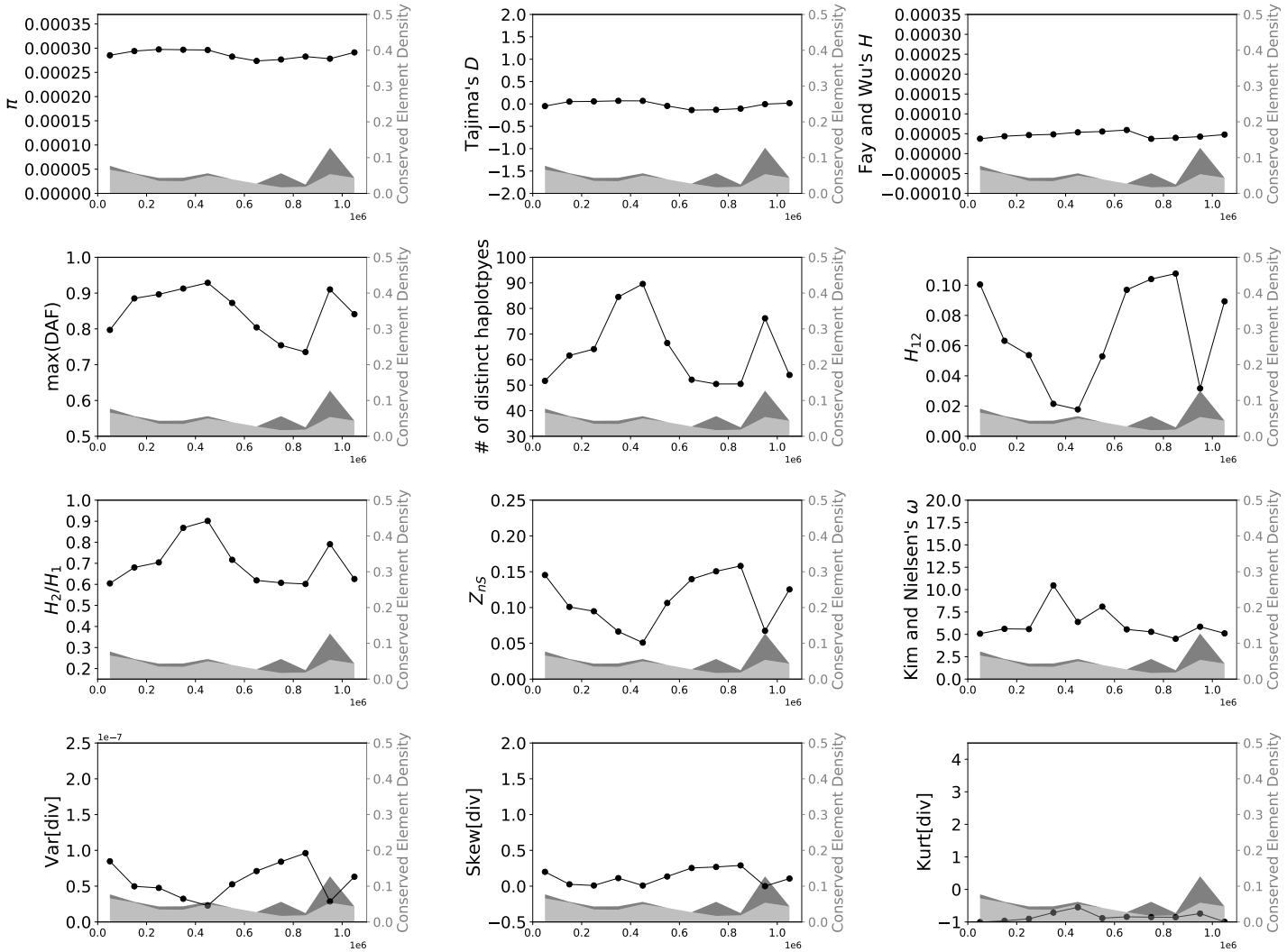
Supplementary Figure 9: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr11:23900001–25000000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



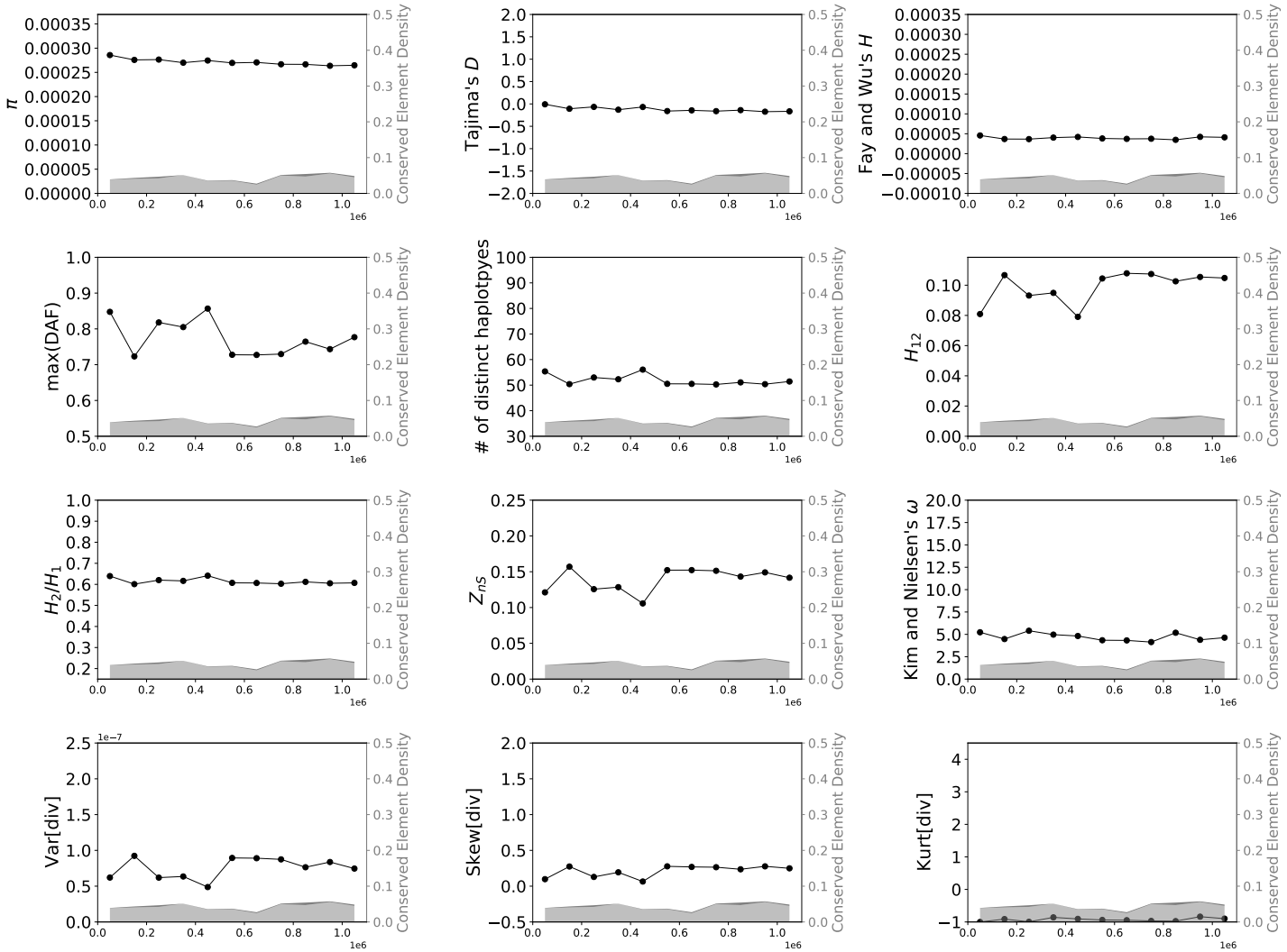
Supplementary Figure 10: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr10:49800001–50900000 in the human genome (hg19 assembly version), simulated under equilibrium demography. See legend of Supplementary Figure 1 for more detail.



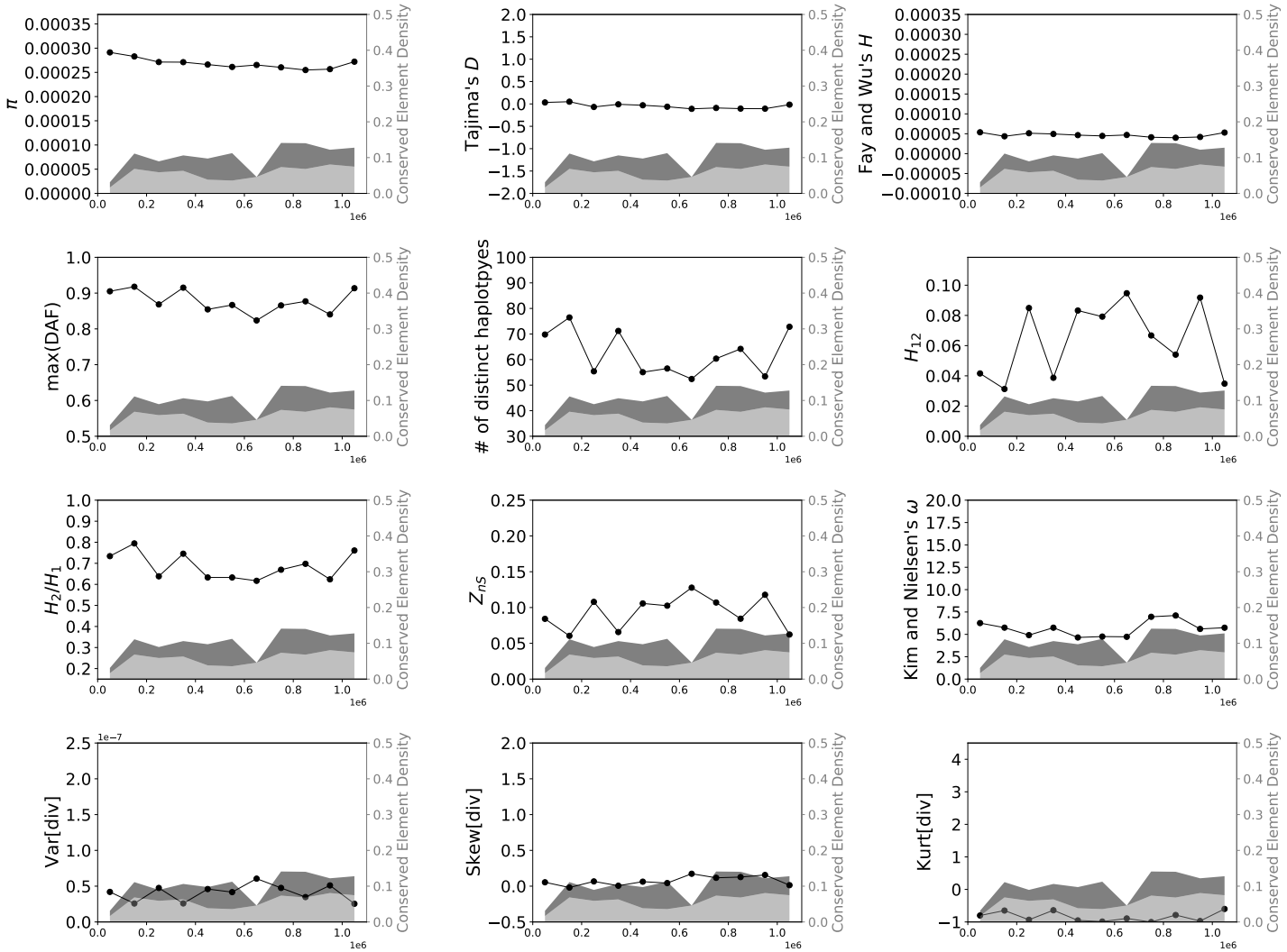
Supplementary Figure 11: Values of 15 statistics calculated in each of the four data sets simulated under Tennessen et al.'s model of European demography (Tennessen *et al.* 2012). Here, the mean value of each statistic across each simulated replicate is shown in the 11 adjacent 100 kb windows within the 1.1 Mb chromosomal region. In each panel the neutral expectation of the statistic obtained from forward simulations are shown as a horizontal black line, while the neutral expectation from coalescent simulations is shown as a dashed gray line—although for most statistics these two expectations were essentially identical, for some there was a slight discrepancy between the two, perhaps reflecting the subtle differences in the models used by the two simulators.



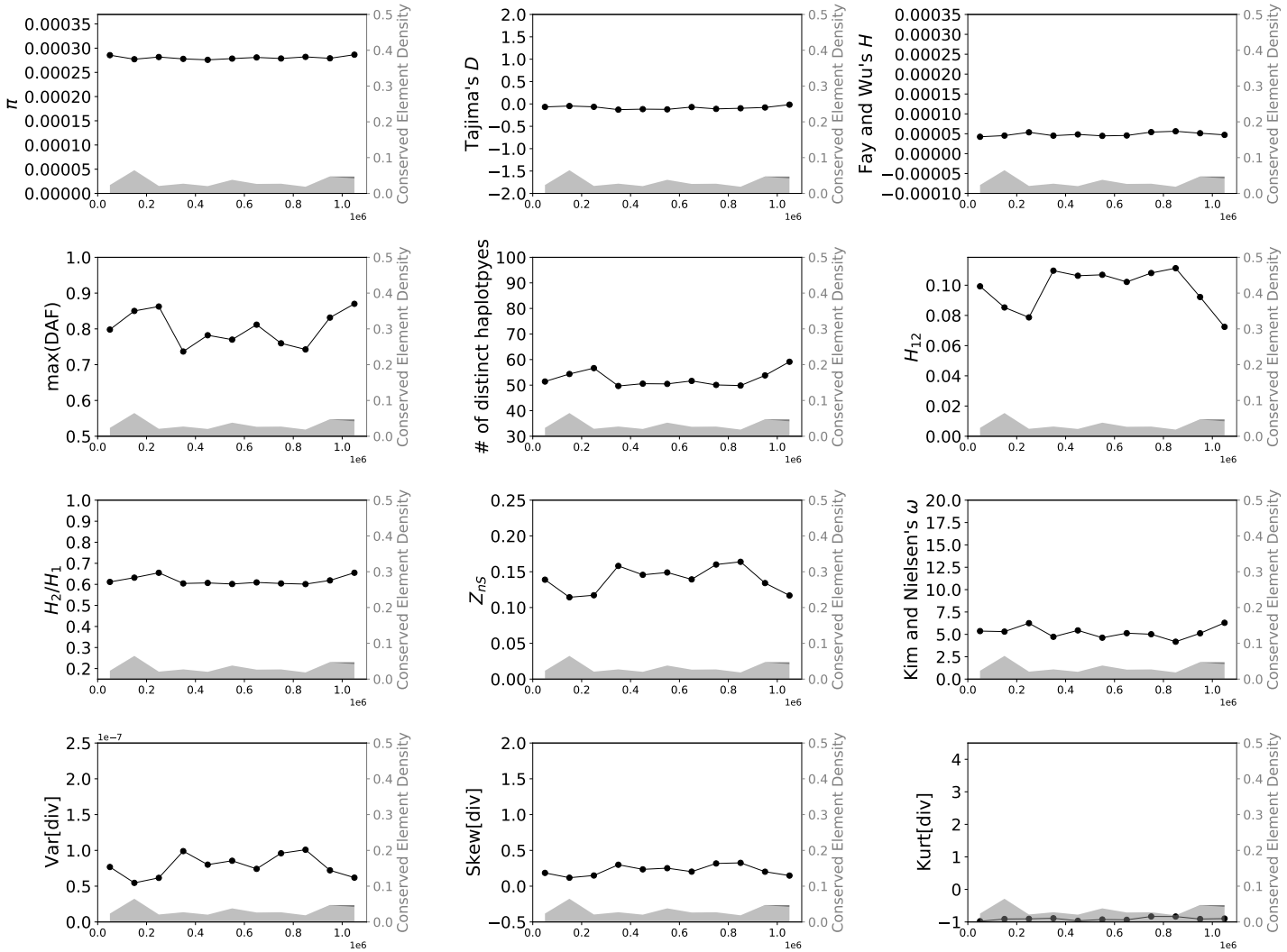
Supplementary Figure 12: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr7:37000001–38100000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



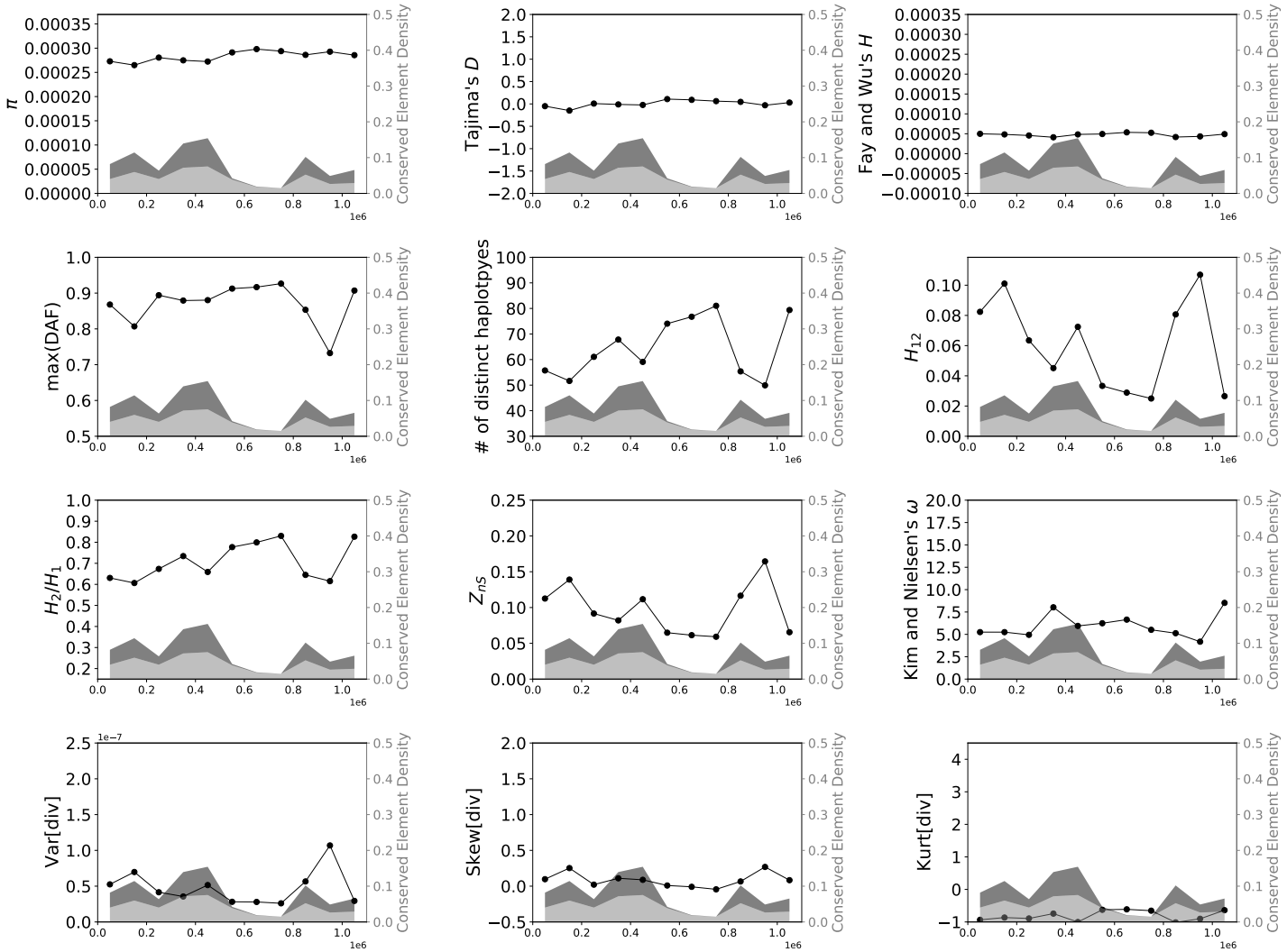
Supplementary Figure 13: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr4:98200001–99300000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



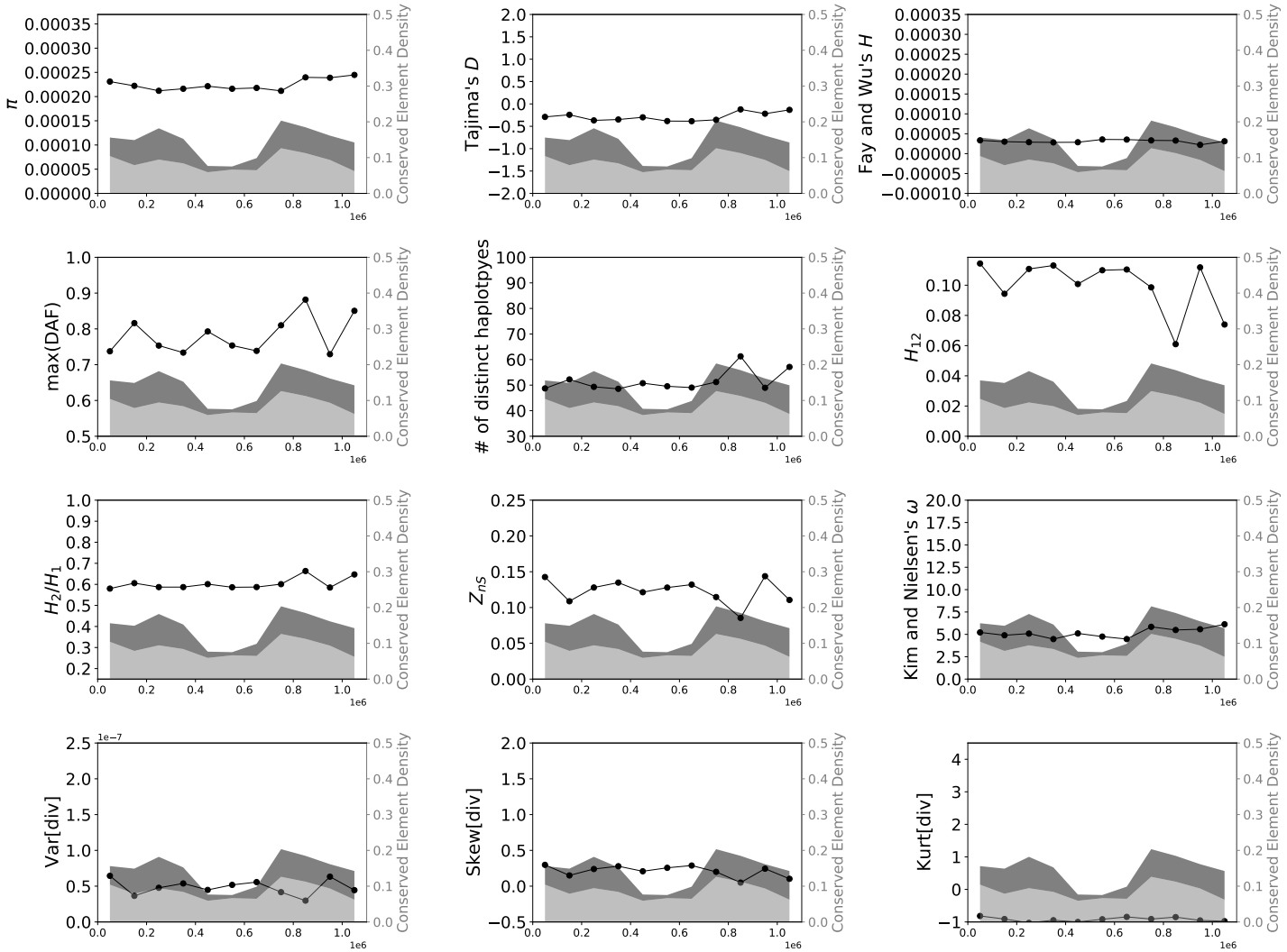
Supplementary Figure 14: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr10:49800001–50900000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



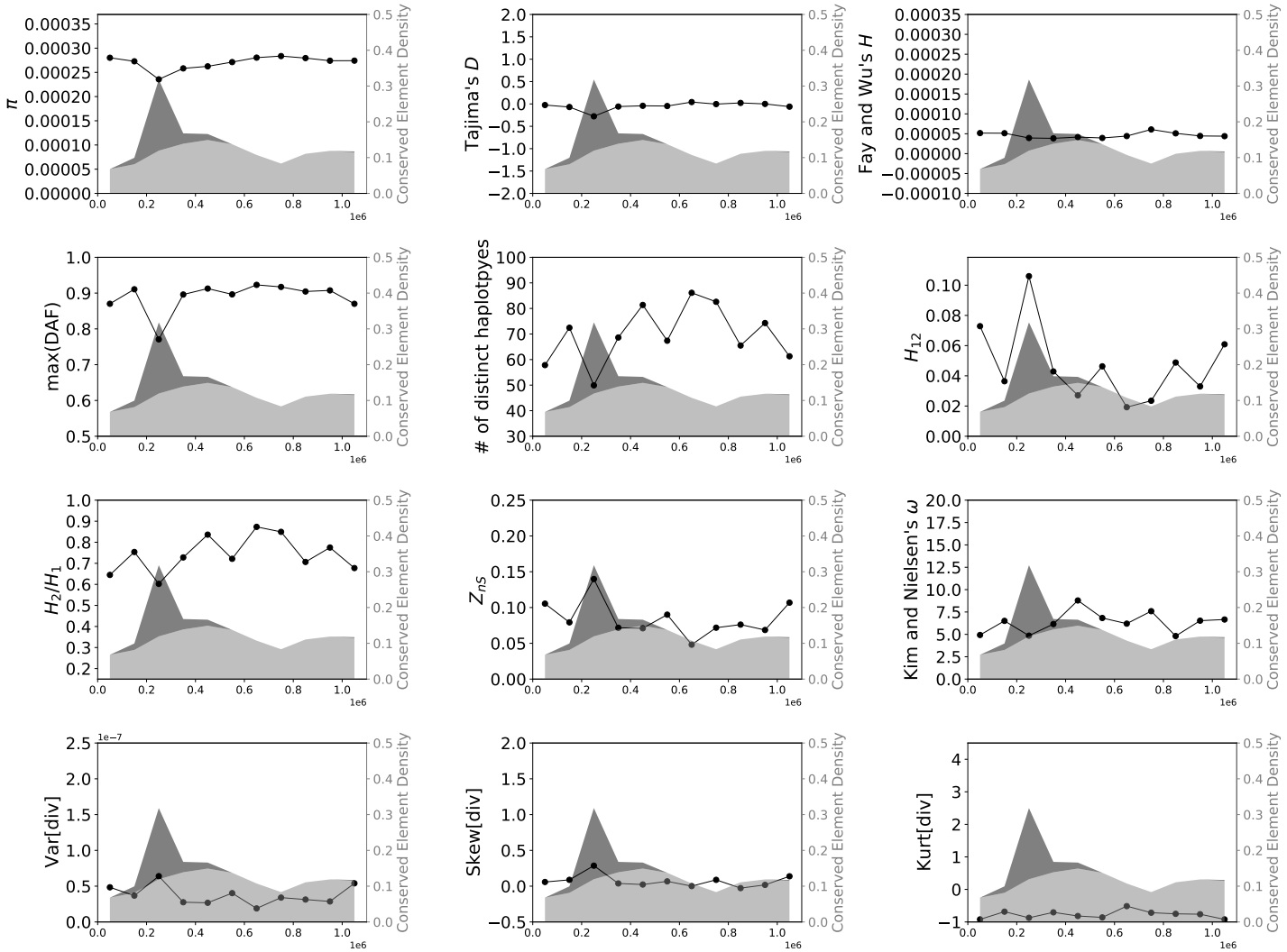
Supplementary Figure 15: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3:80100001–81200000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



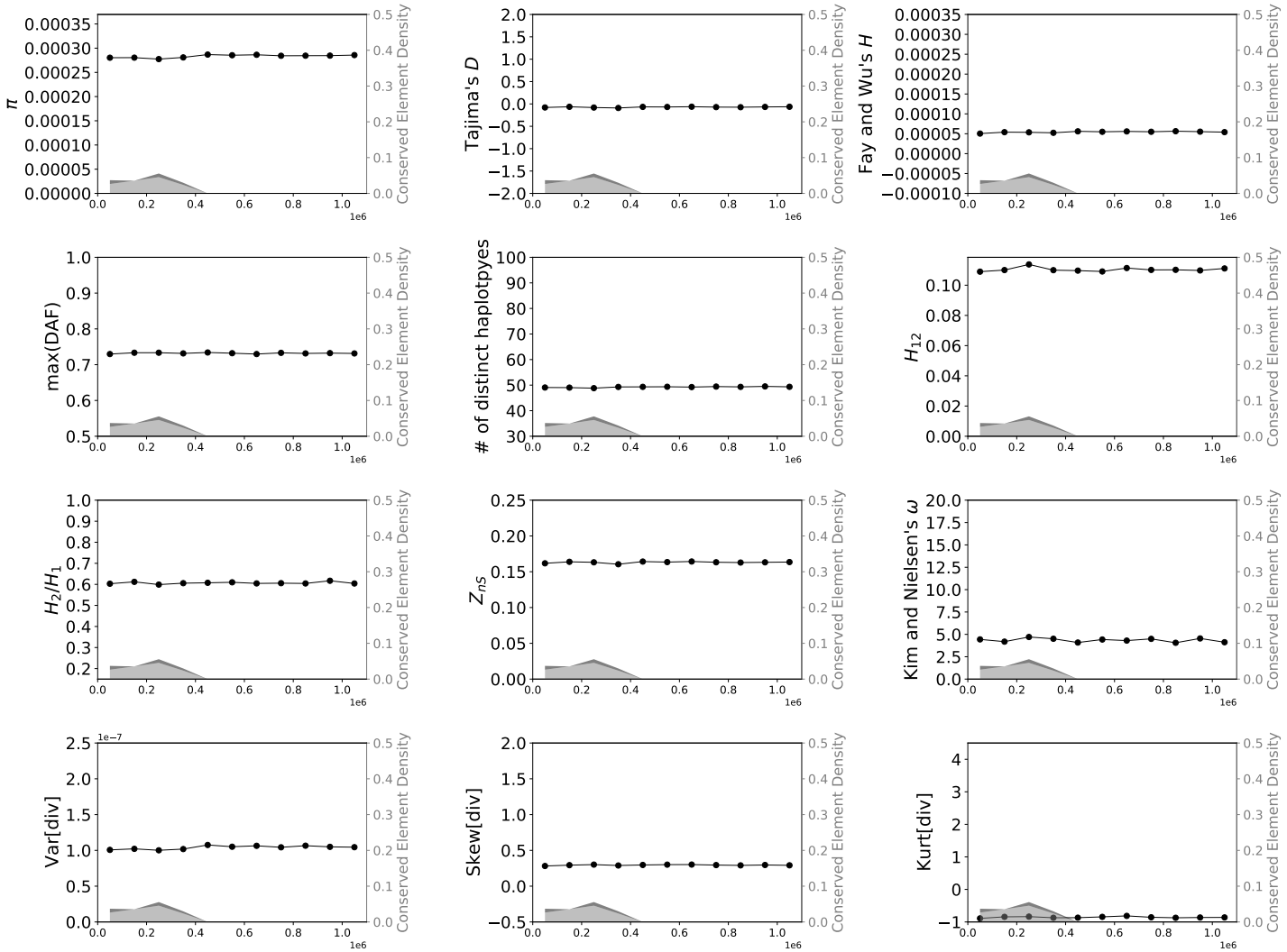
Supplementary Figure 16: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3:12500001–13600000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



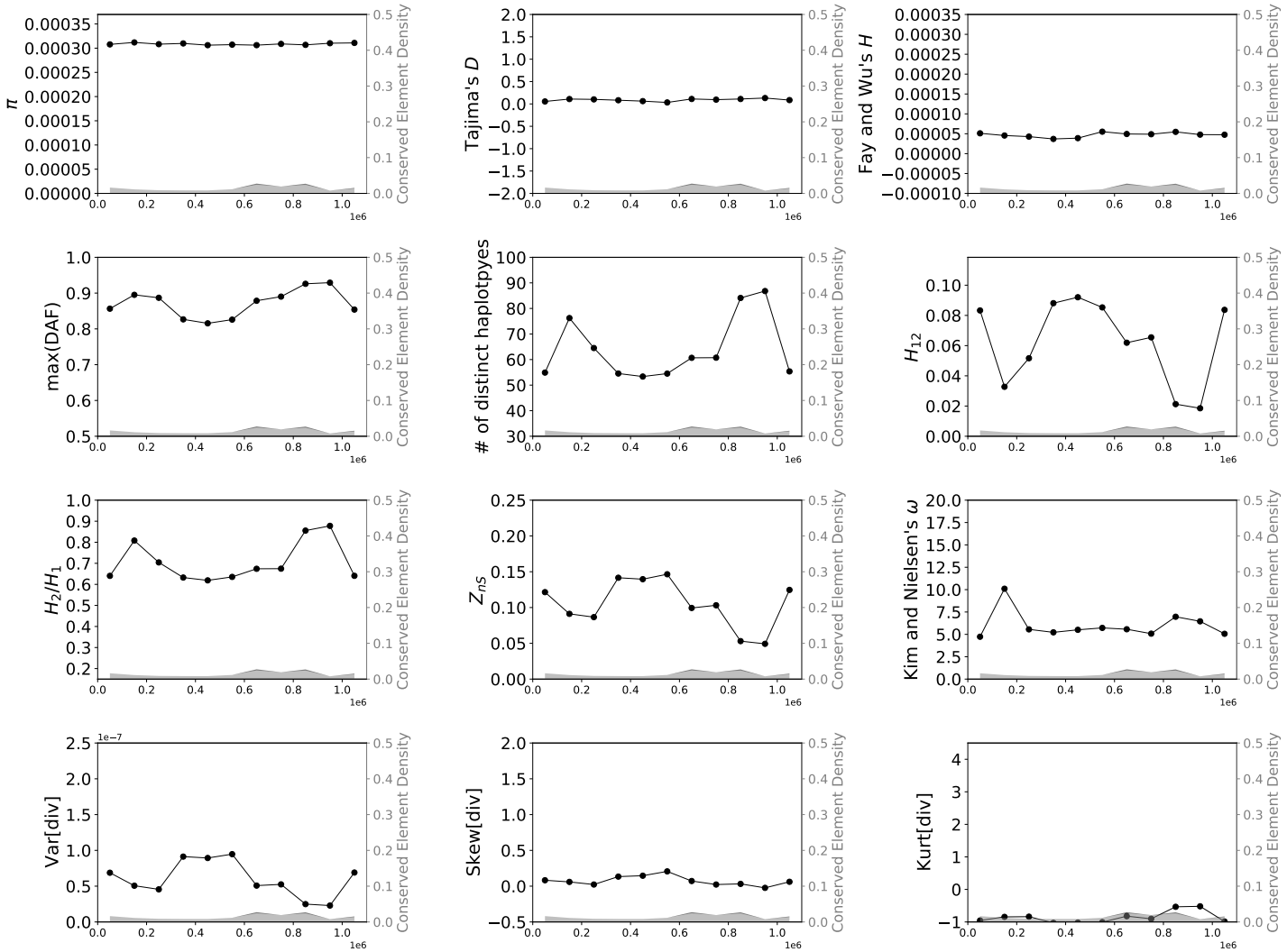
Supplementary Figure 17: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr1:45800001–46900000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



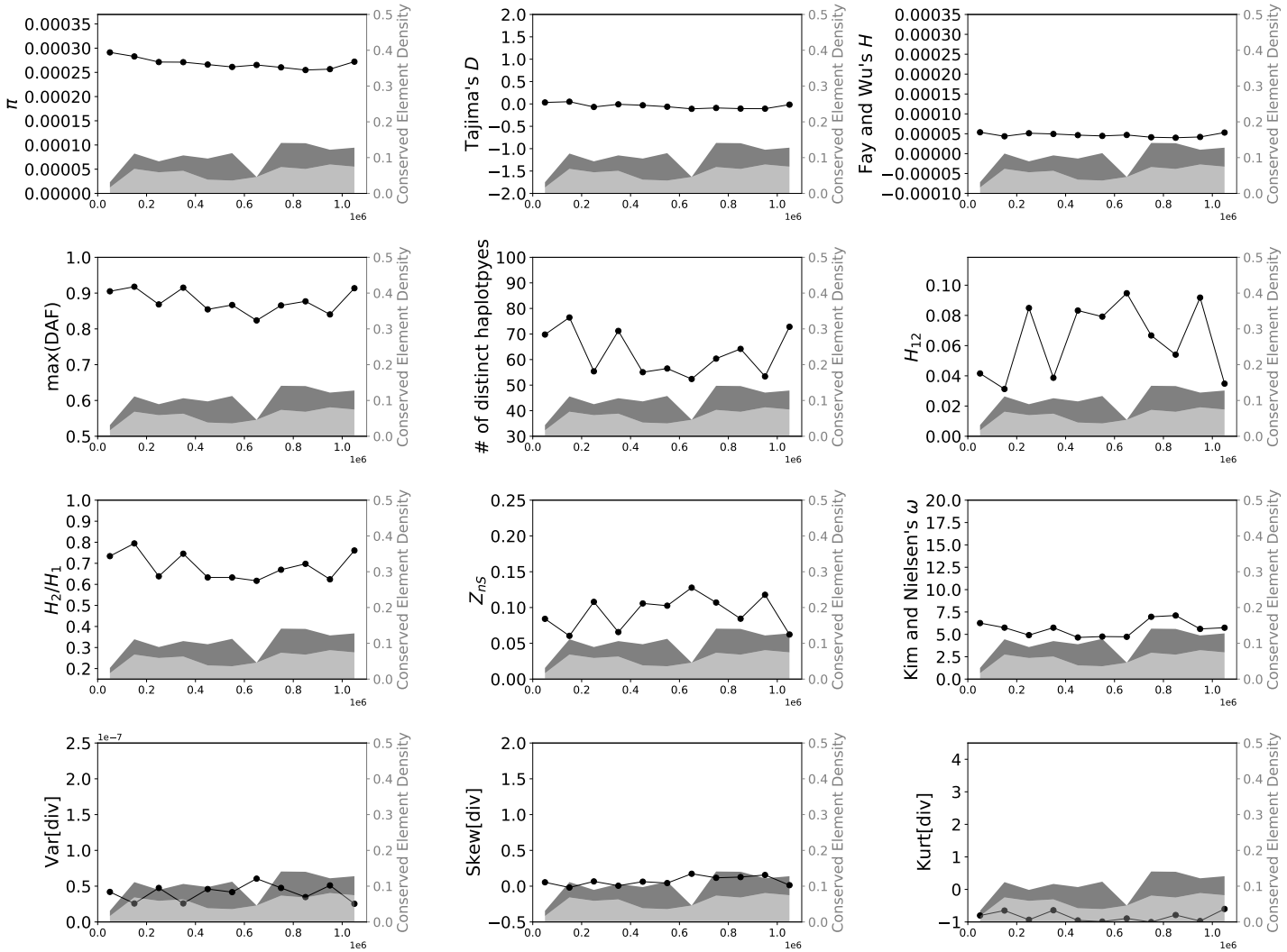
Supplementary Figure 18: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr15:60500001–61600000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



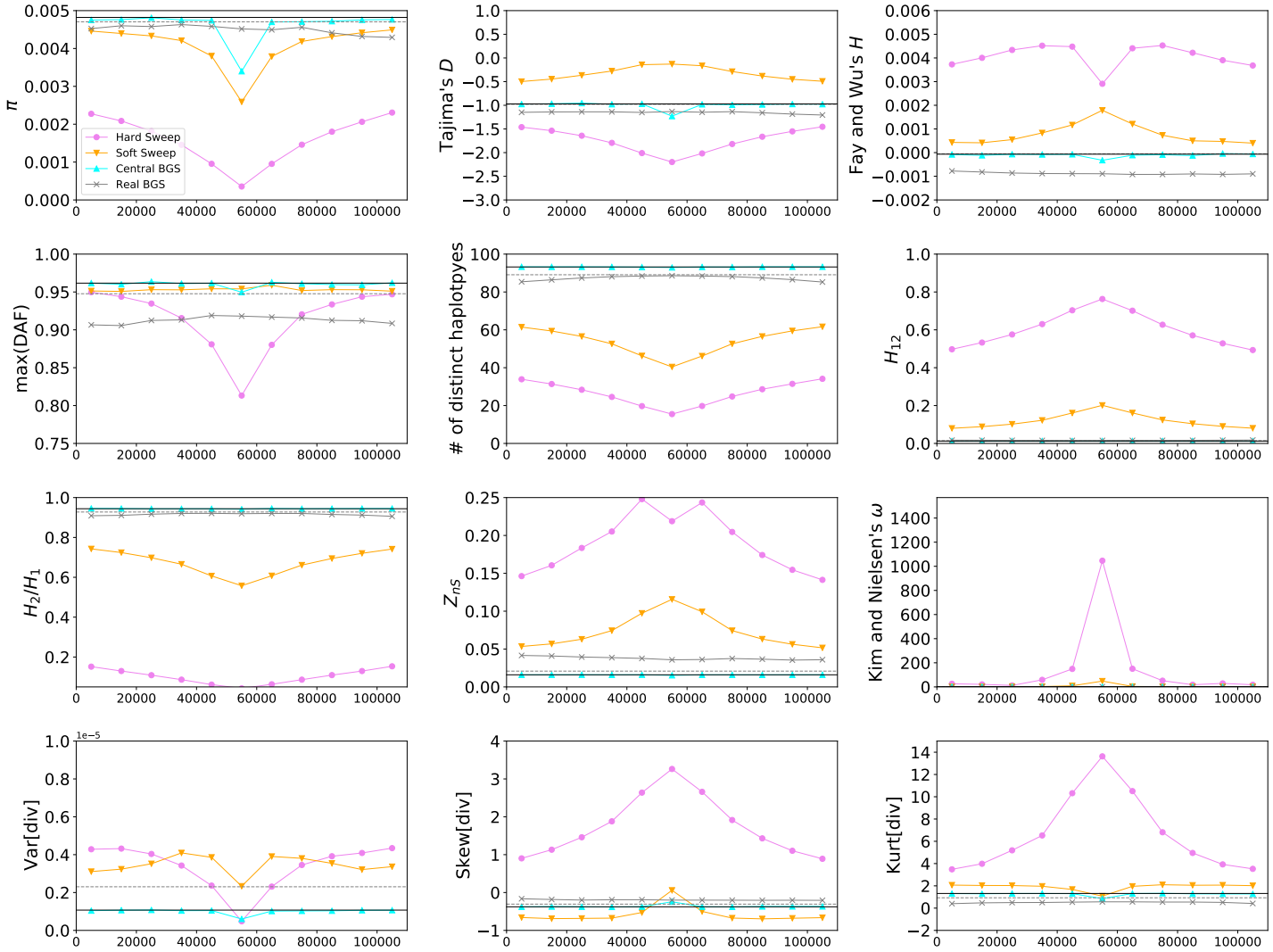
Supplementary Figure 19: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr11:50000001–51100000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



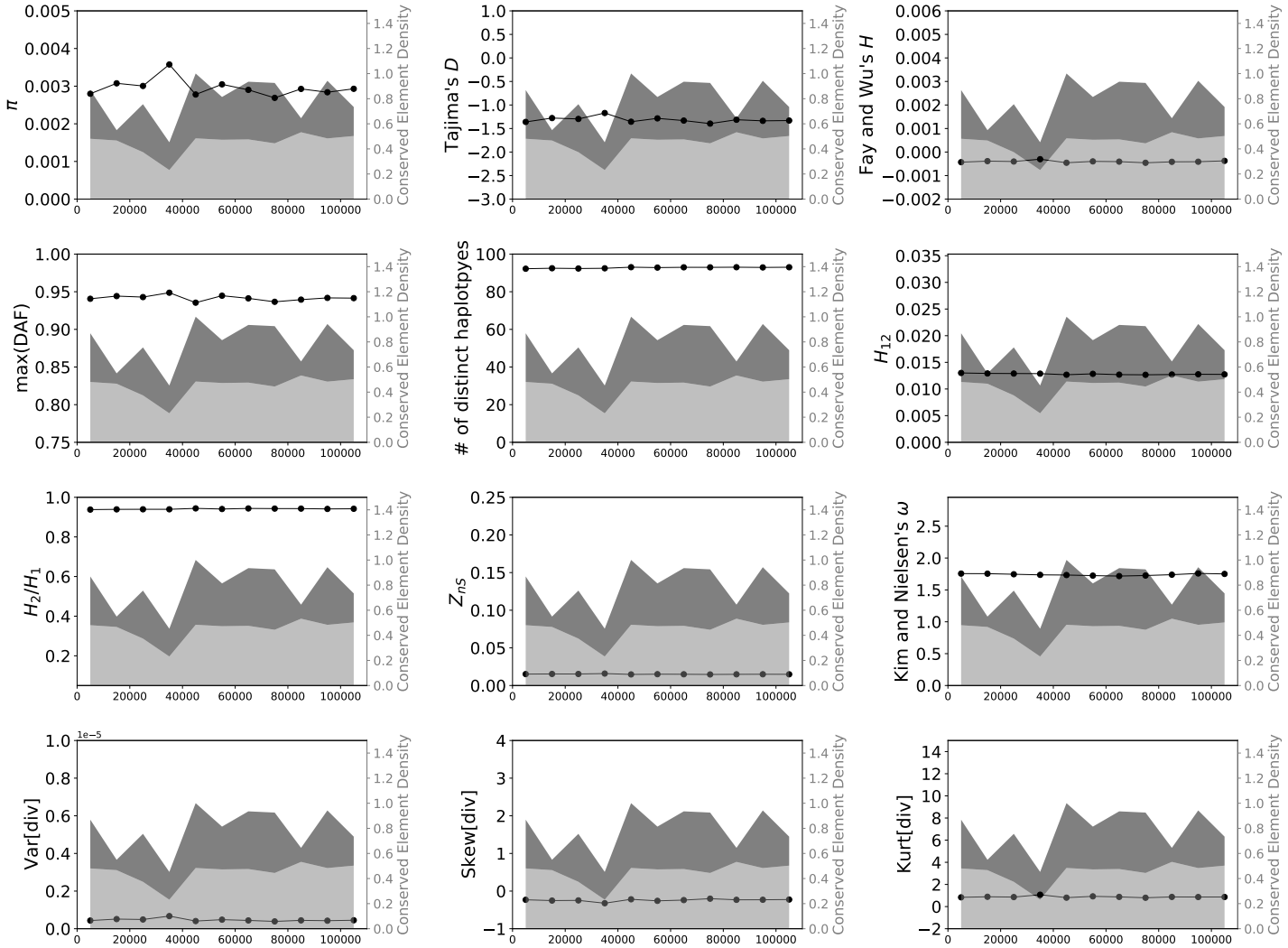
Supplementary Figure 20: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr11:23900001–25000000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



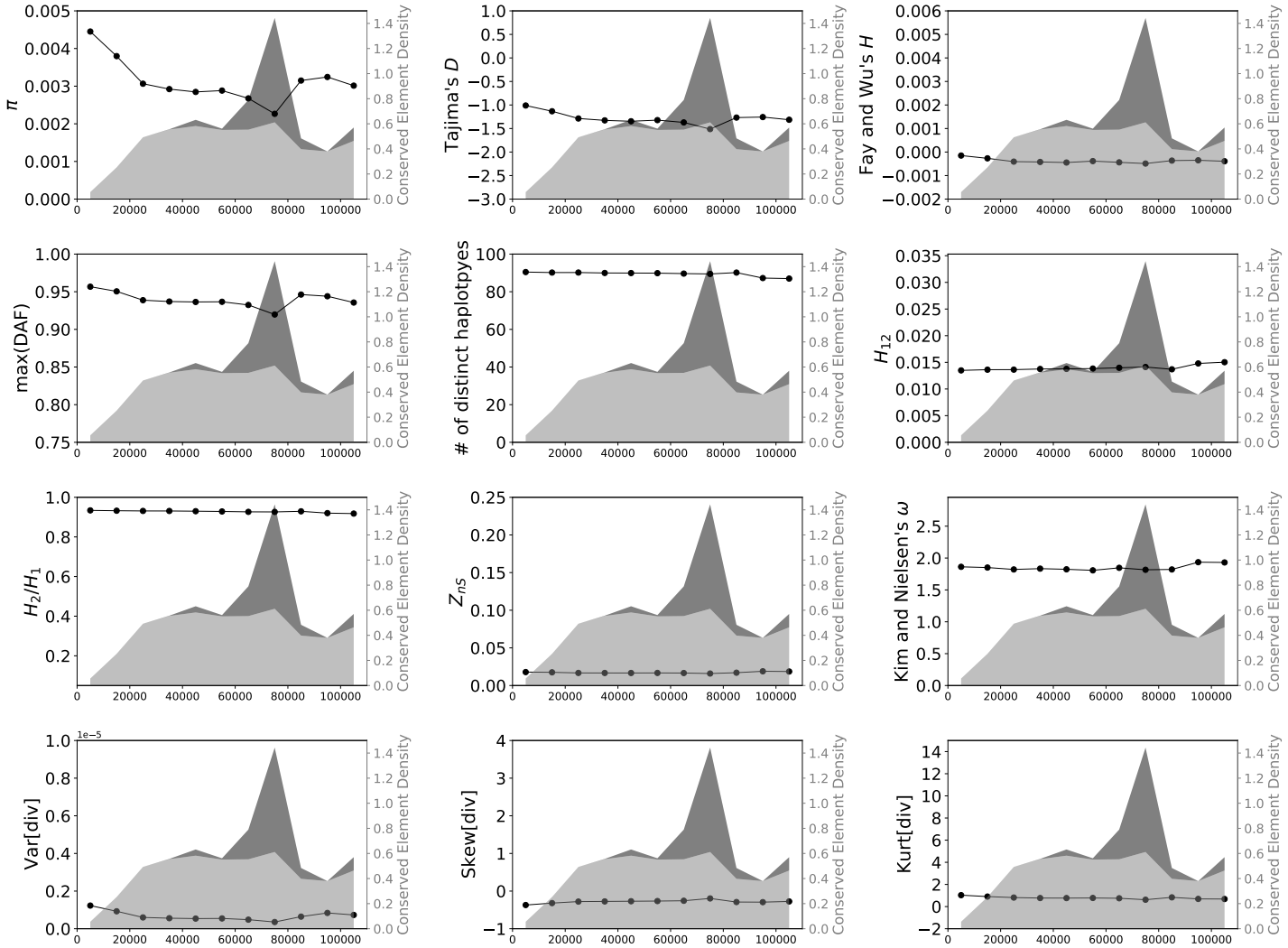
Supplementary Figure 21: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr10:49800001–50900000 in the human genome (hg19 assembly version), simulated under European population size history (Tennessen *et al.* 2012). See legend of Supplementary Figure 1 for more detail.



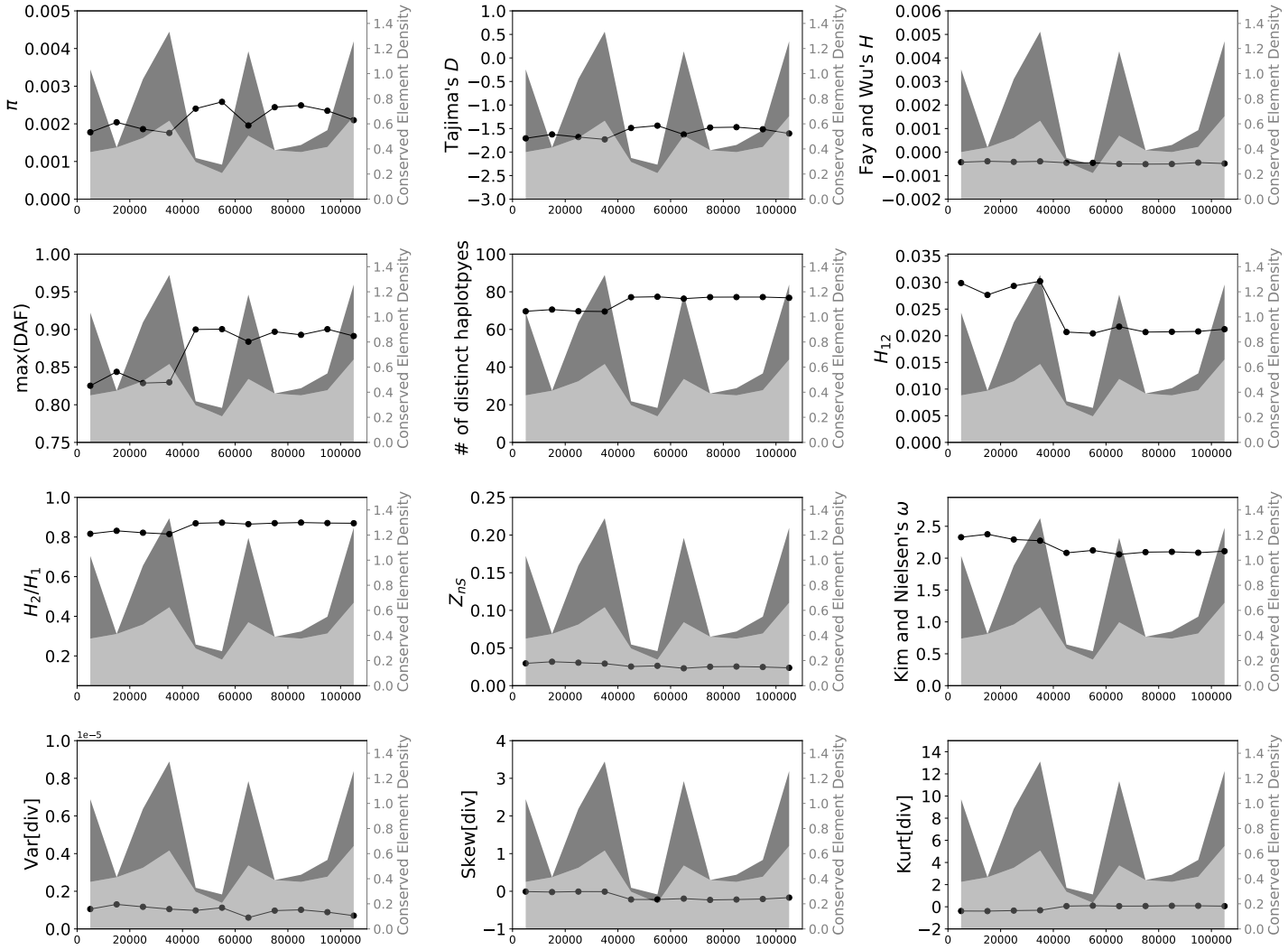
Supplementary Figure 22: Values of 15 statistics calculated in each of the four data sets simulated under Sheehan and Song's bottleneck model for African *Drosophila melanogaster* (Sheehan and Song 2016). Here, the mean value of each statistic across each simulated replicate is shown in the 11 adjacent 10 kb windows within the 110 kb chromosomal region. In each panel the neutral expectation of the statistic obtained from forward simulations are shown as a horizontal black line, while the neutral expectation from coalescent simulations is shown as a dashed gray line. The density of exonic and conserved non-coding DNA is shown in the shaded dark gray and light gray regions, respectively.



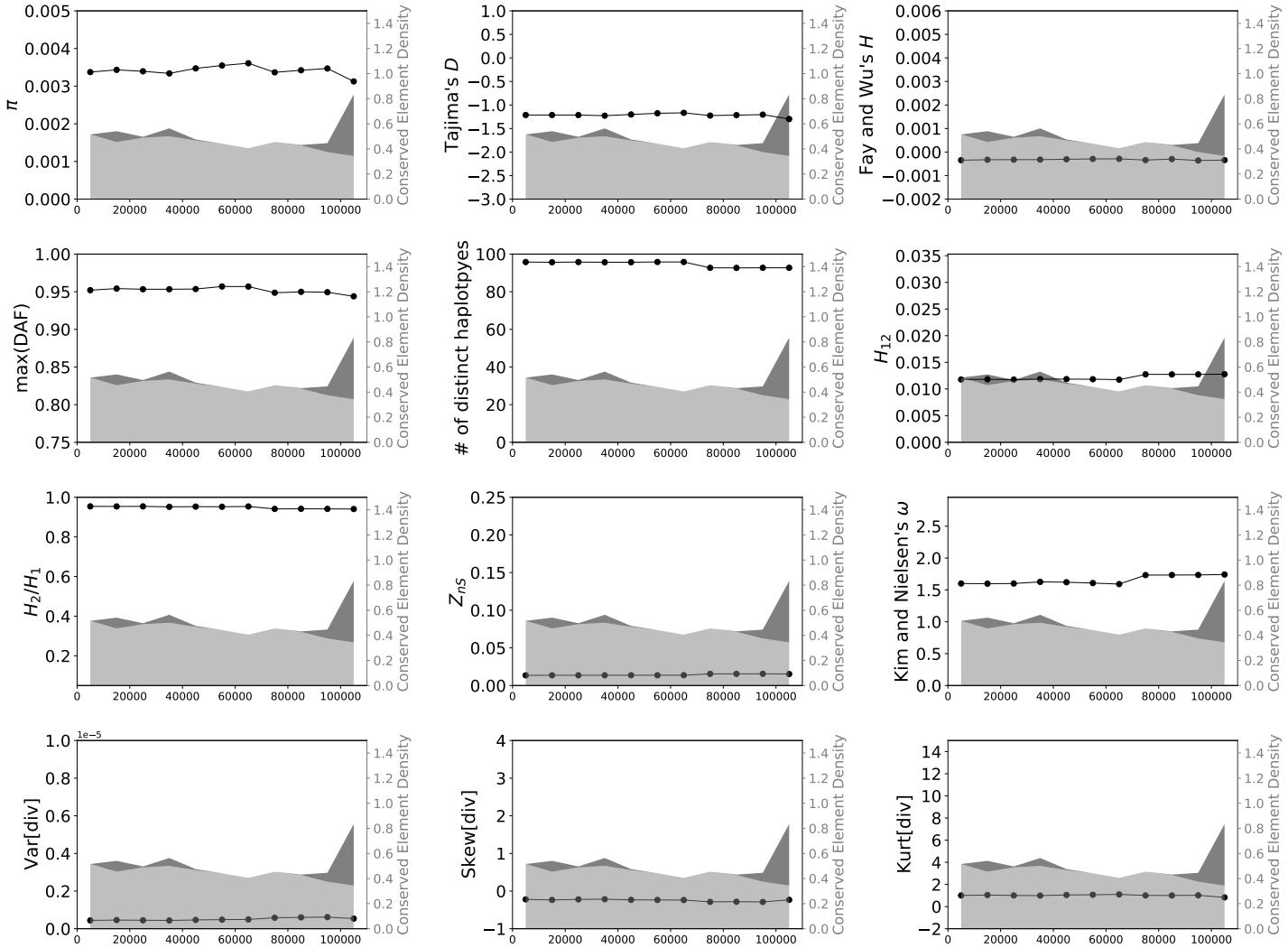
Supplementary Figure 23: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:7460001–7570000 in the *Drosophila melanogaster* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



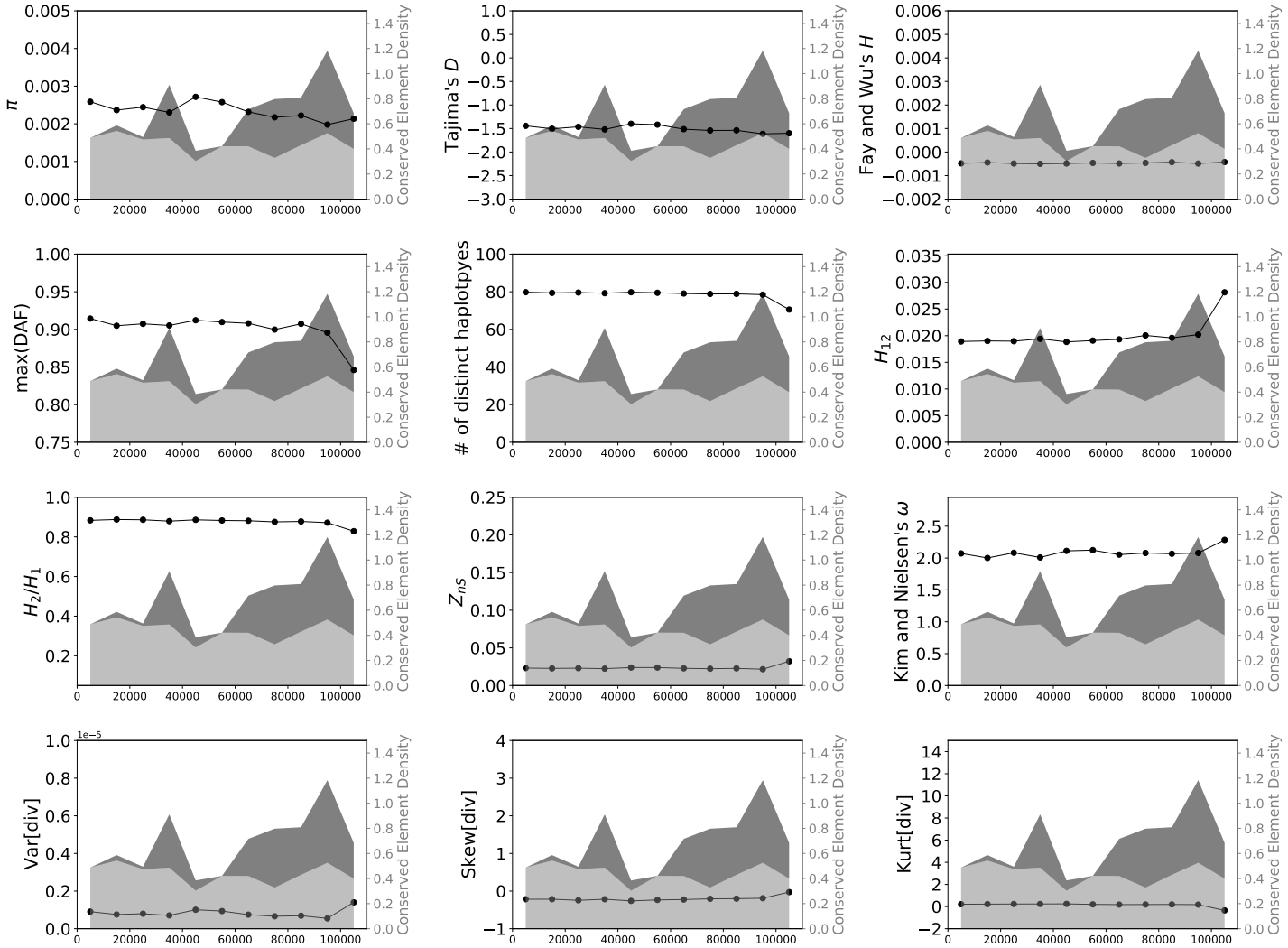
Supplementary Figure 24: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:6410001–6520000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



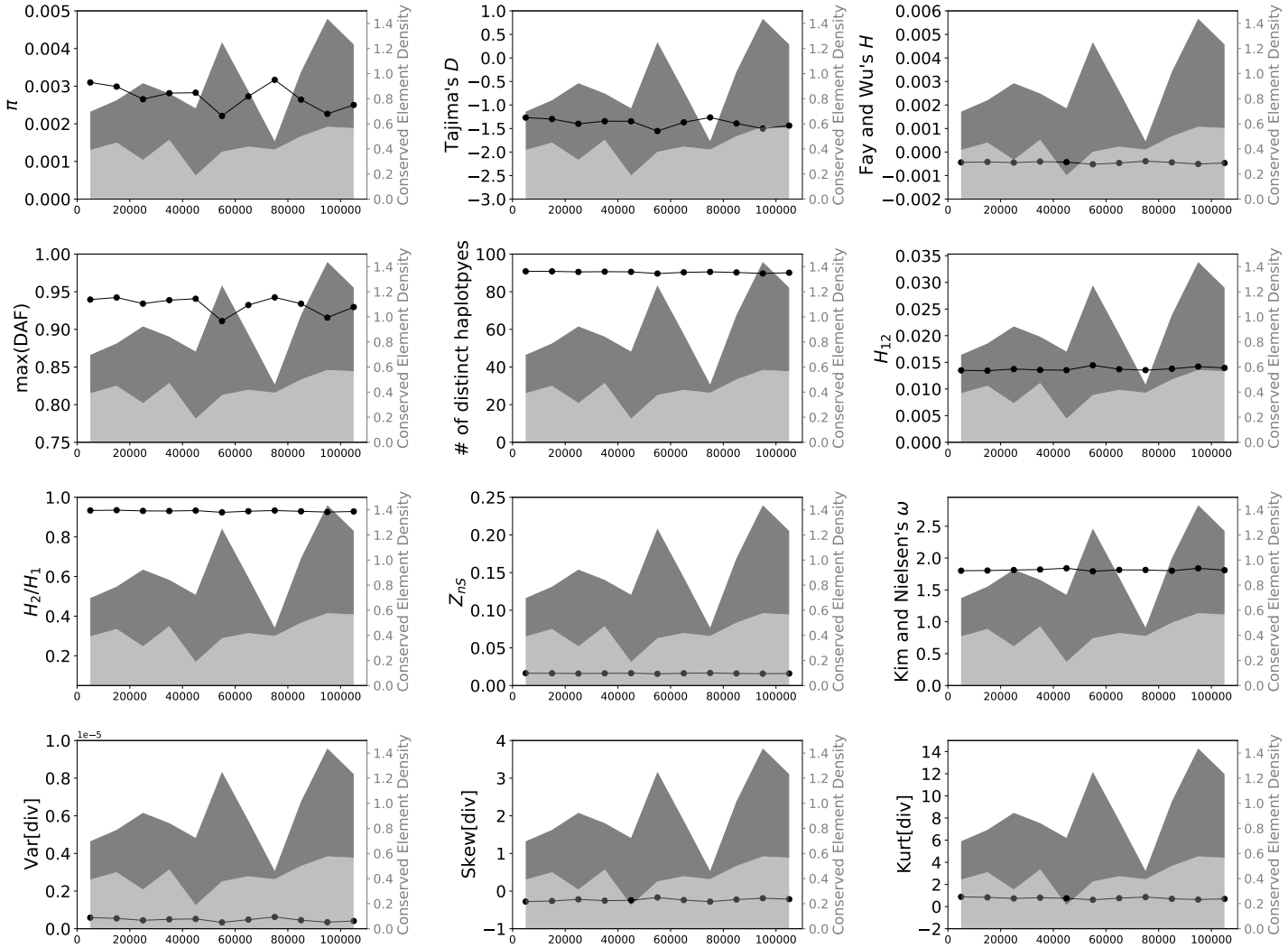
Supplementary Figure 25: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:20960001–21070000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



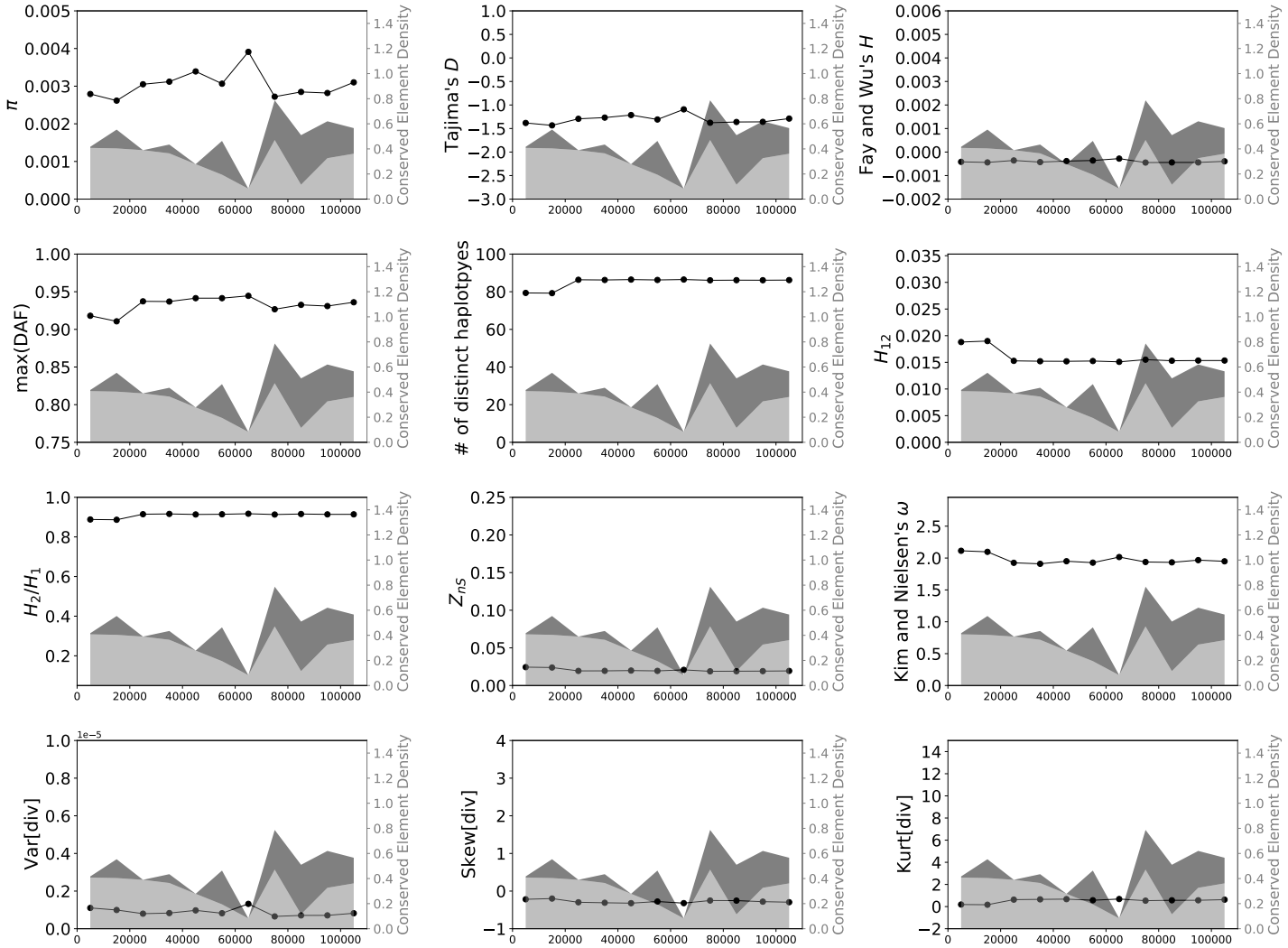
Supplementary Figure 26: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:17130001–17240000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



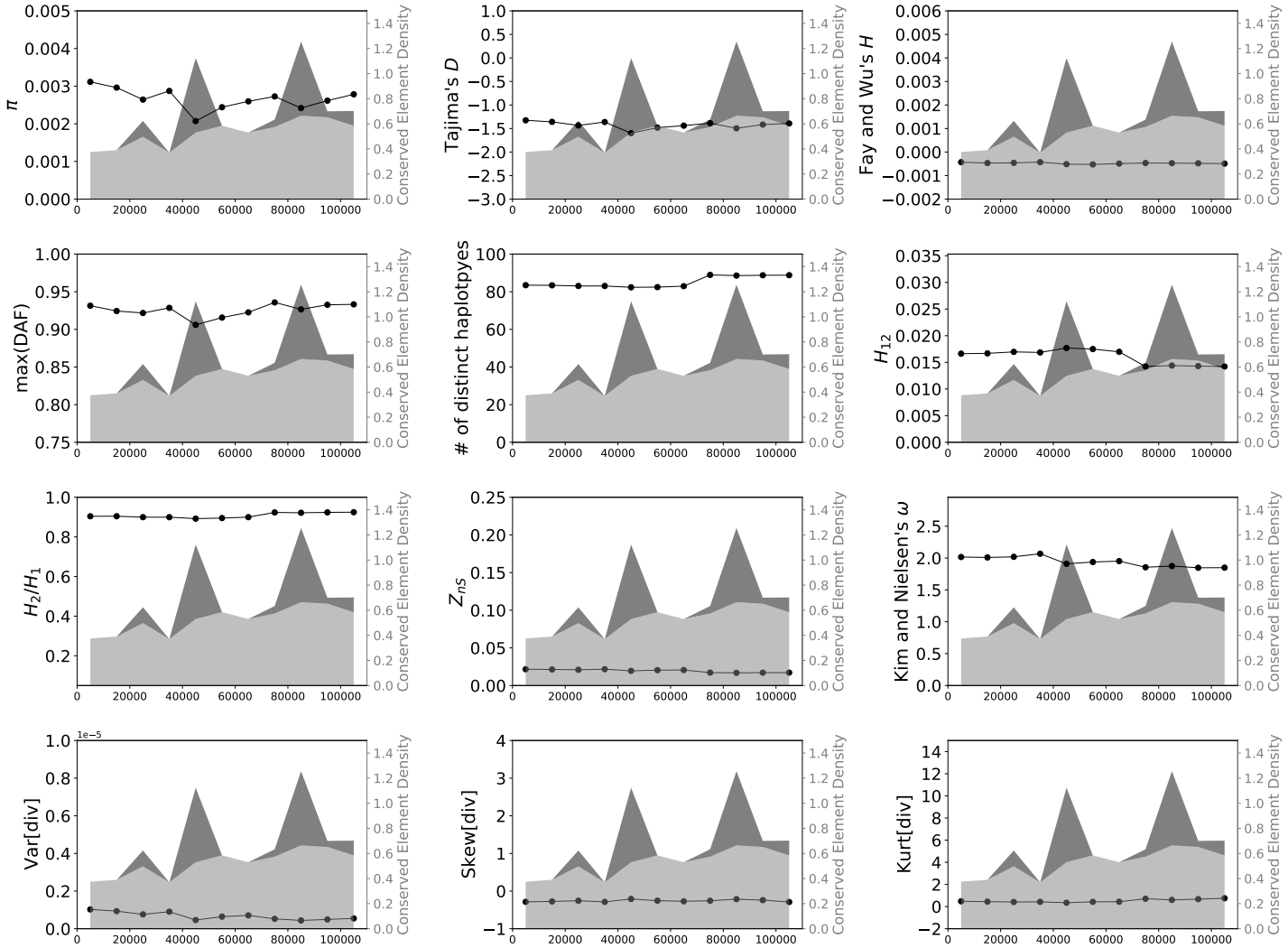
Supplementary Figure 27: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:14900001–15010000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



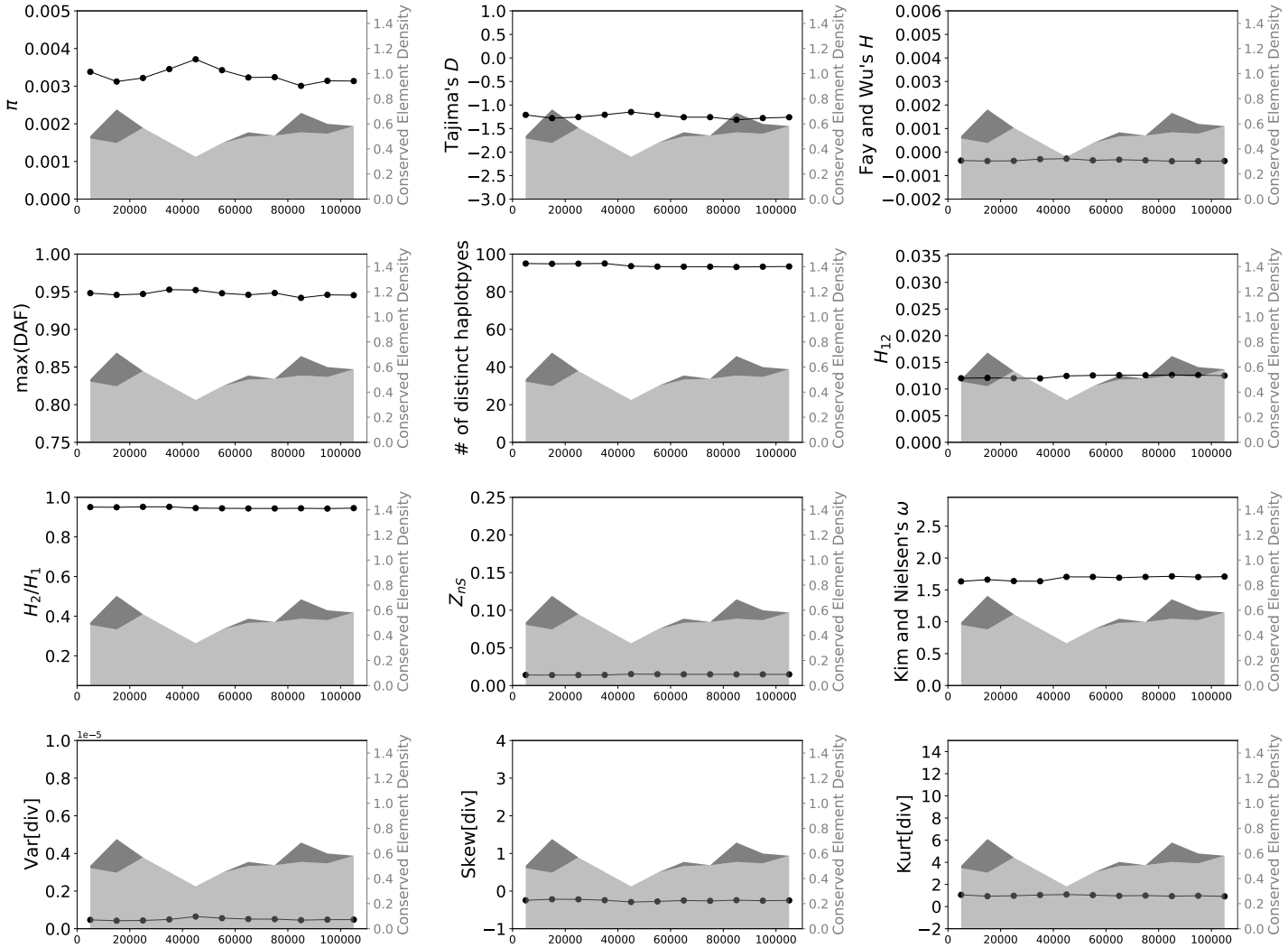
Supplementary Figure 28: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr2R:20360001–20470000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



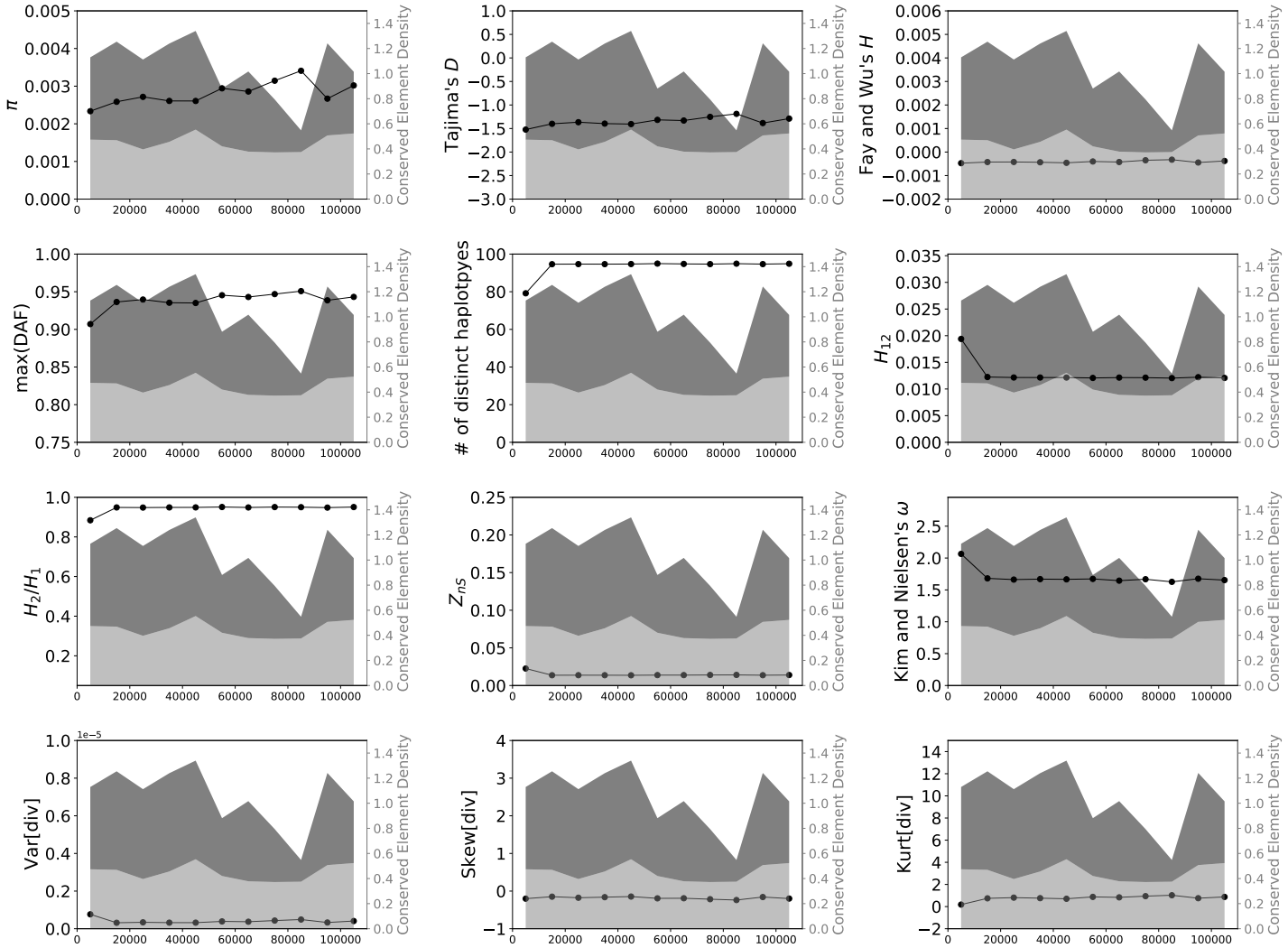
Supplementary Figure 29: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3L:17980001–18090000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



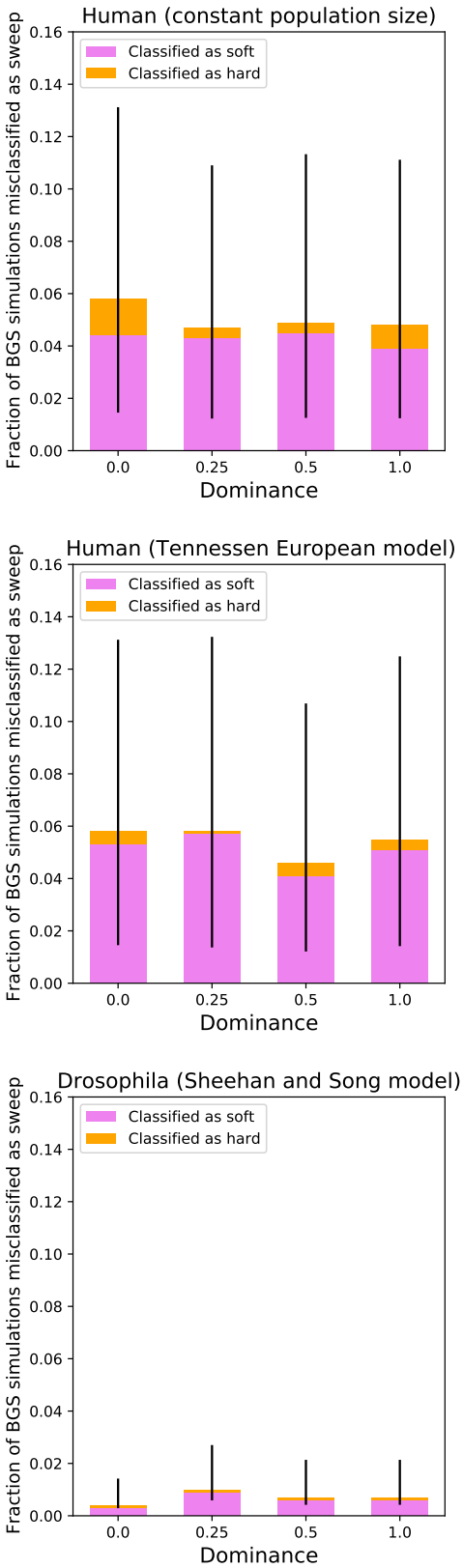
Supplementary Figure 30: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr2R:15630001–15740000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



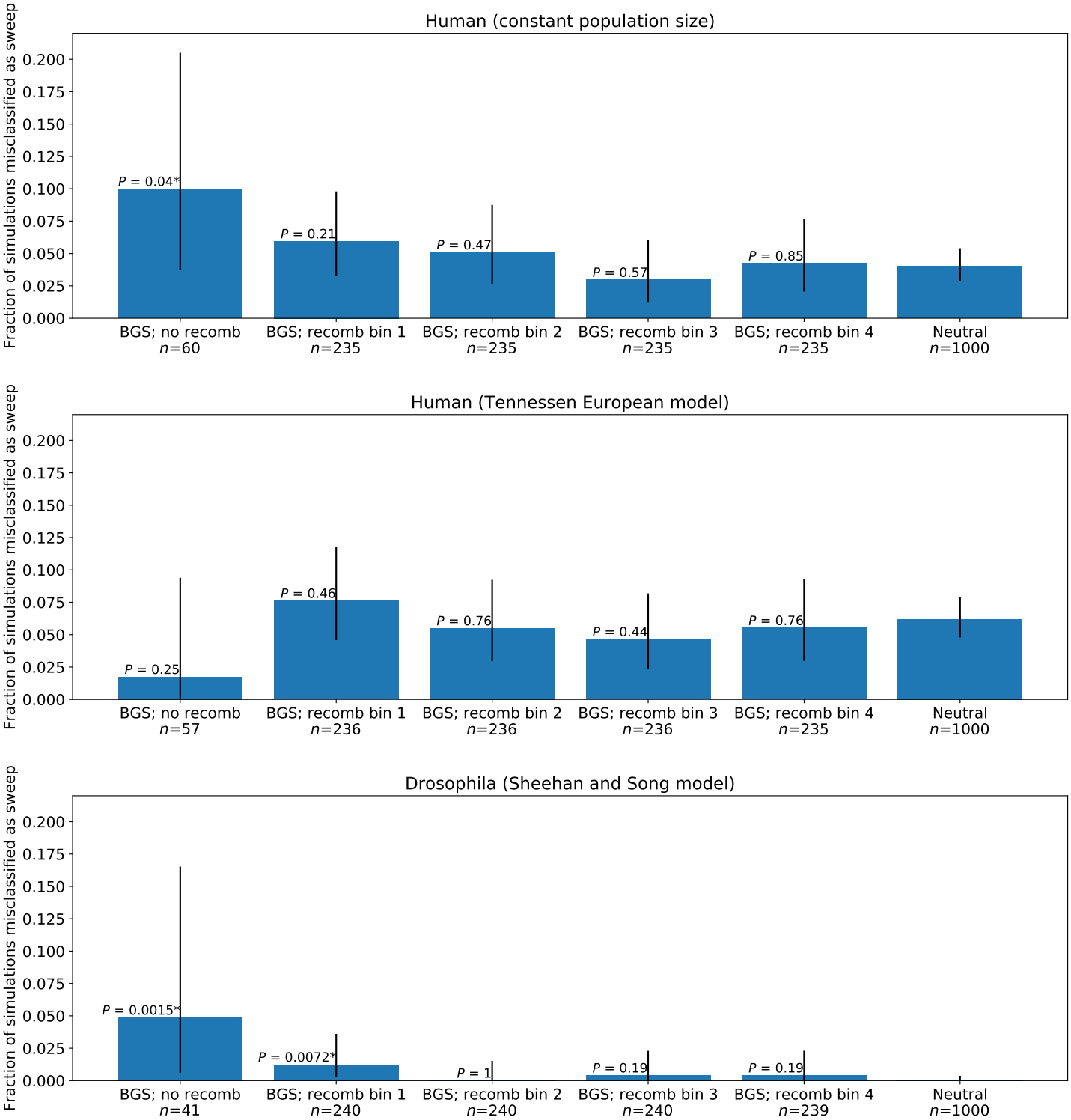
Supplementary Figure 31: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr3R:18660001–18770000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



Supplementary Figure 32: Values of 15 statistics calculated from a simulated data set designed to match the locations of conserved sites from chr2L:10390001–10500000 in the *Drosophila* genome (assembly release 5), simulated under African population history (Sheehan and Song 2016). See legend of Supplementary Figure 1 for more detail.



Supplementary Figure 33: The fractions of simulations of our Real BGS model misclassified as hard and soft sweeps, shown for simulated data sets with different fixed dominance coefficients for deleterious mutations. The results for a dominance coefficient of 0.25 summarizes results from the same set of simulations reported elsewhere in the text, while the other simulations were generated by repeating the same procedure described in the Methods but altering the dominance coefficient accordingly. The error bars show the 95% binomial confidence intervals for our estimates of the total fraction of simulations misclassified as a sweep of either type.



Supplementary Figure 34: The fractions of simulations of our Real BGS–weak CNE model misclassified as a sweep of either type for each demographic model, shown after binning our data. The rightmost bar shows the fraction of misclassified neutral simulations for comparison, and all P-values show significance of the comparison with neutrality. The error bars show the 95% binomial confidence intervals.