

**Supplementary Material: Detectability of runs of homozygosity is influenced by analysis parameters as well as population-specific demographic history**

Avril M. Harder<sup>1\*</sup>, Kenneth B. Kirksey<sup>2</sup>, Samarth Mathur<sup>3</sup>, Janna R. Willoughby<sup>1</sup>

1. College of Forestry, Wildlife and Environment, Auburn University, Auburn, Alabama, USA

2. Walker College of Business, Appalachian State University, Boone, North Carolina, USA

3. Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, Ohio, USA

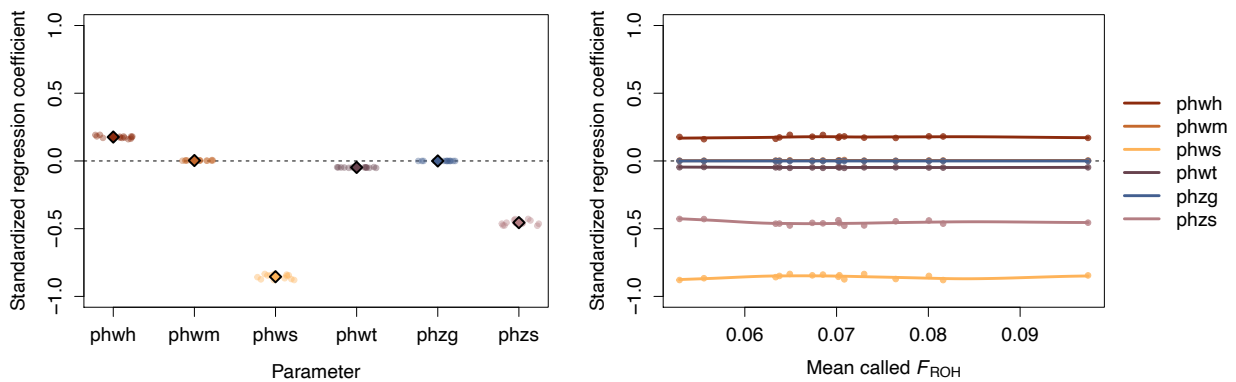
\*Author for correspondence: Avril M. Harder, College of Forestry, Wildlife and Environment, Auburn University, Auburn, Alabama, USA, (773) 688-8564, [avrilharder@gmail.com](mailto:avrilharder@gmail.com)

## Empirical data sensitivity analysis

We applied to same process described in Box 1 for the simulated data to the empirical data, beginning with the same 486 PLINK parameter setting combinations. As with the simulated data, we present results for ROHs called at 15X coverage, but results were highly concordant across all coverage levels. We tested multiple sets of parameter settings, with all of these combinations presented in Table S2 and three of these iterations presented below.

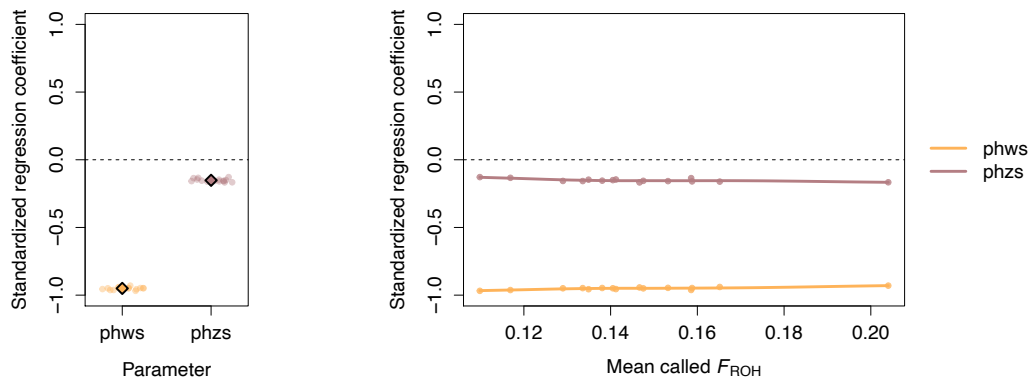
### Iteration 1

For the initial set of 486 parameter combinations, we observed very little variation in the effects of parameter settings across individuals ( $n = 15$ ). However, for the parameters *phws* and *phzs*, there was some evidence for individual variation in response to changes in these settings (see vertical spread of individual points in below panels) and varying the values of these parameters also led to strong effects on  $F_{ROH}$ . Specifically, increases in the scanning window length in SNPs (*phws*) and in the number of SNPs required to support a ROH 100 kb in length led to decreases in  $F_{ROH}$  estimates (see negative values in lower left panel). For these two parameters, we next tested (i) a set of values for each parameter towards the lower end of the range tested in Iteration 1 (Iteration 2) and (ii) sets of values a bit higher than those tested in Iteration 2 (Iteration 3).



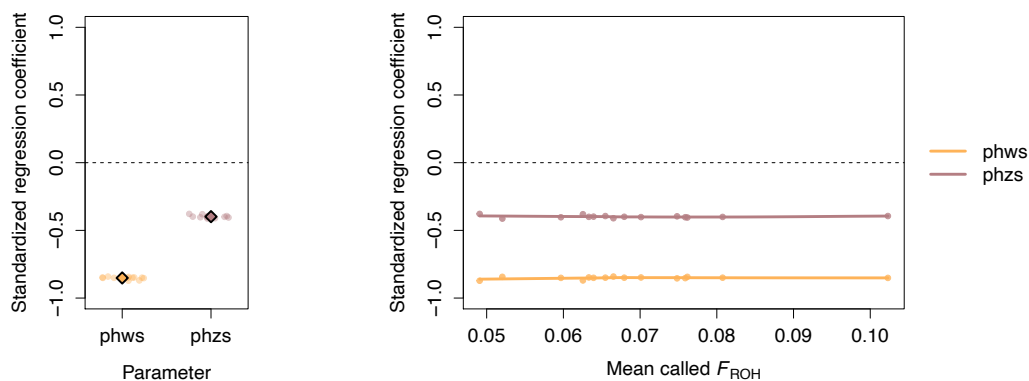
### Iteration 2

In Iteration 2, we tested a new set of values for both *phws* and *phzs* at the lower end of the ranges tested in Iteration 1. Although *phws* still has a strong effect on  $F_{ROH}$  estimation, there is very little variation across individuals for both of these parameters. The effect of *phzs* on  $F_{ROH}$  estimation has also decreased relative to Iteration 1 (*i.e.*, SRC for *phzs* is closer to zero in Iteration 2).



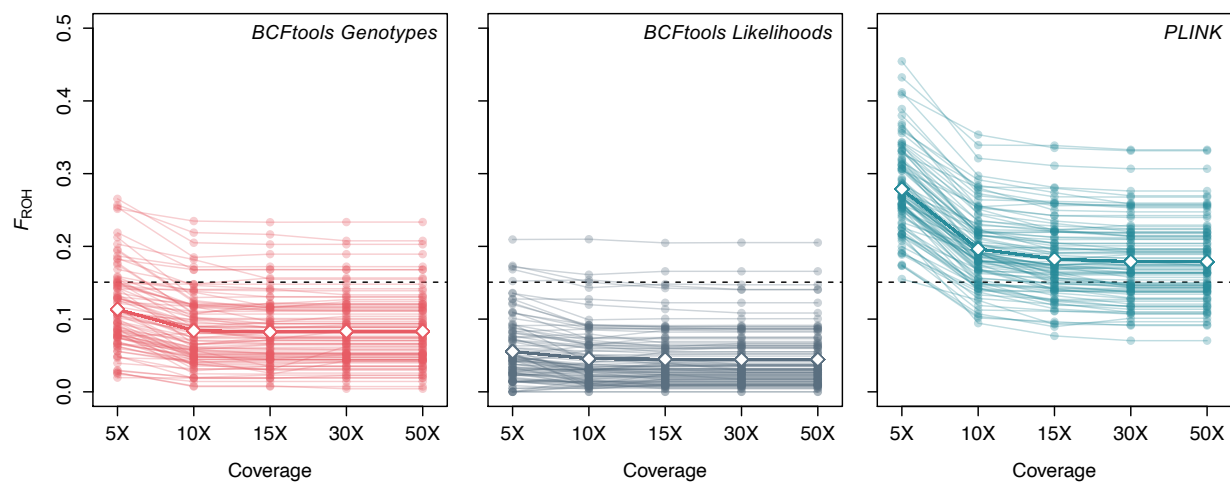
### Iteration 3

In Iteration 3, we tested a new set of values for both  $phws$  and  $phzs$  that were slightly higher than those tested in Iteration 2. As in Iterations 1 and 2,  $phws$  has a strong effect on  $F_{ROH}$  estimation. Relative to Iteration 2, the effect of  $phzs$  on  $F_{ROH}$  estimation has increased (*i.e.*, SRC for  $phzs$  is farther from zero than in Iteration 2). The effect of  $phzs$  is also negative, indicating that increases in the value of this parameter corresponds to decreases in  $F_{ROH}$ . In other words, increasing the number of SNPs required to support a ROH 100 kb in length leads to more ROHs being discarded or ignored. This is consistent with the lower  $F_{ROH}$  estimates produced by Iteration 3 combinations relative to Iteration 2. Because the effects of  $phzs$  were lower in Iteration 2 and to avoid discarding legitimate ROHs, we selected parameter values from the ranges tested in Iteration 2 (*i.e.*,  $phws = 50$  and  $phzs = 100$ ).

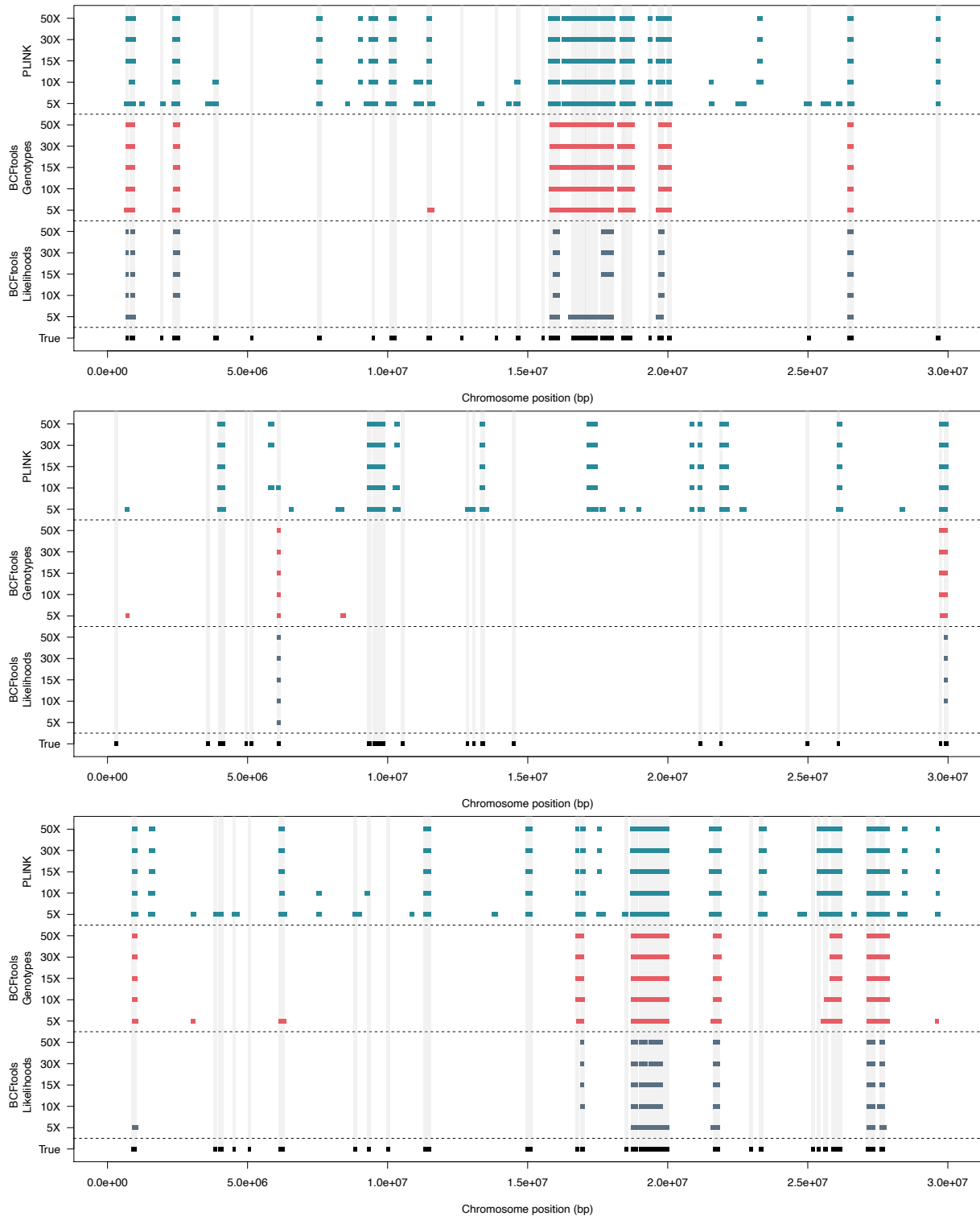


### Additional iterations

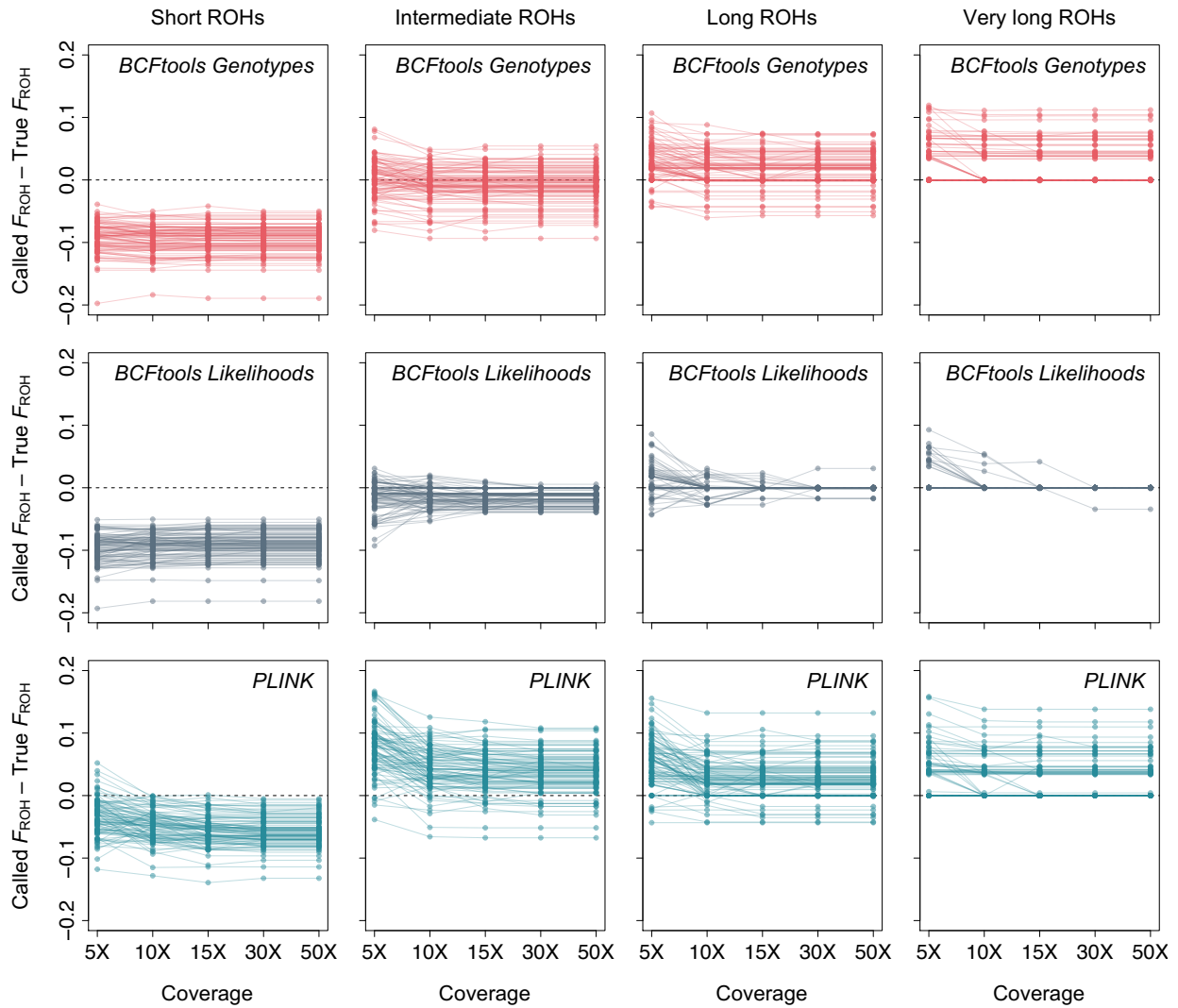
As indicated in Table S2, we tested several other combinations of settings for these two parameters, including some extreme values nearly an order of magnitude larger than what was tested in Iterations 2 and 3. Many of these combinations produced obviously incorrect  $F_{ROH}$  estimates (*e.g.*, increasing the ranges for  $phws$  and  $phzs$  to 900-1,000 led to  $F_{ROH}$  estimates  $< 0.006$  across all samples), but testing these exaggerated values allowed for visualization of the effects of these parameters on  $F_{ROH}$  estimation. It can sometimes be difficult to imagine how changing parameter settings will affect  $F_{ROH}$ , particularly in combination with other varying parameter settings, but our approach allows for quick and efficient comparisons that facilitate familiarity with PLINK parameters and their effects on a specific data set.



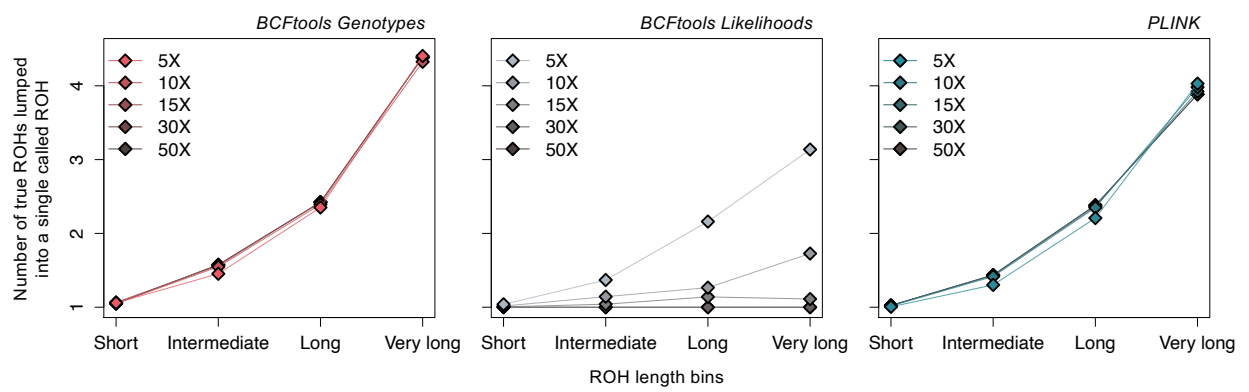
**Figure S1.** Called  $F_{ROH}$  values for all three methods across coverage levels. True mean  $F_{ROH}$  is indicated by horizontal dashed line and white diamonds indicate mean called values.



**Figure S2.** Positions of true and called ROHs for three exemplar simulated individuals across methods and coverage levels. Shaded grey polygons correspond to true ROHs present in each individual. Each plot demonstrates false negative and false positive calls as well as the ROH-lumping issue noted for PLINK and BCFTools Genotypes.



**Figure S3.** Called  $F_{\text{ROH}}$  – true  $F_{\text{ROH}}$  across methods, ROH length bins, and coverage levels for simulated data. Dashed horizontal line is at  $y = 0$  and values above this line indicate overestimation of  $F_{\text{ROH}}$  whereas values below this line indicate underestimation. Length bins were defined as: (i)  $100 \text{ kb} \leq \text{short ROHs} < 250 \text{ kb}$ ; (ii)  $250 \text{ kb} \leq \text{intermediate ROHs} < 500 \text{ kb}$ ; (iii)  $500 \text{ kb} \geq \text{long ROHs} < 1 \text{ Mb}$ ; (iv)  $1 \text{ Mb} \geq \text{very long ROHs}$ .



**Figure S4.** Mean number of true ROHs lumped into a single called ROH across methods, coverage levels, and ROH length bins.