

FIGURES

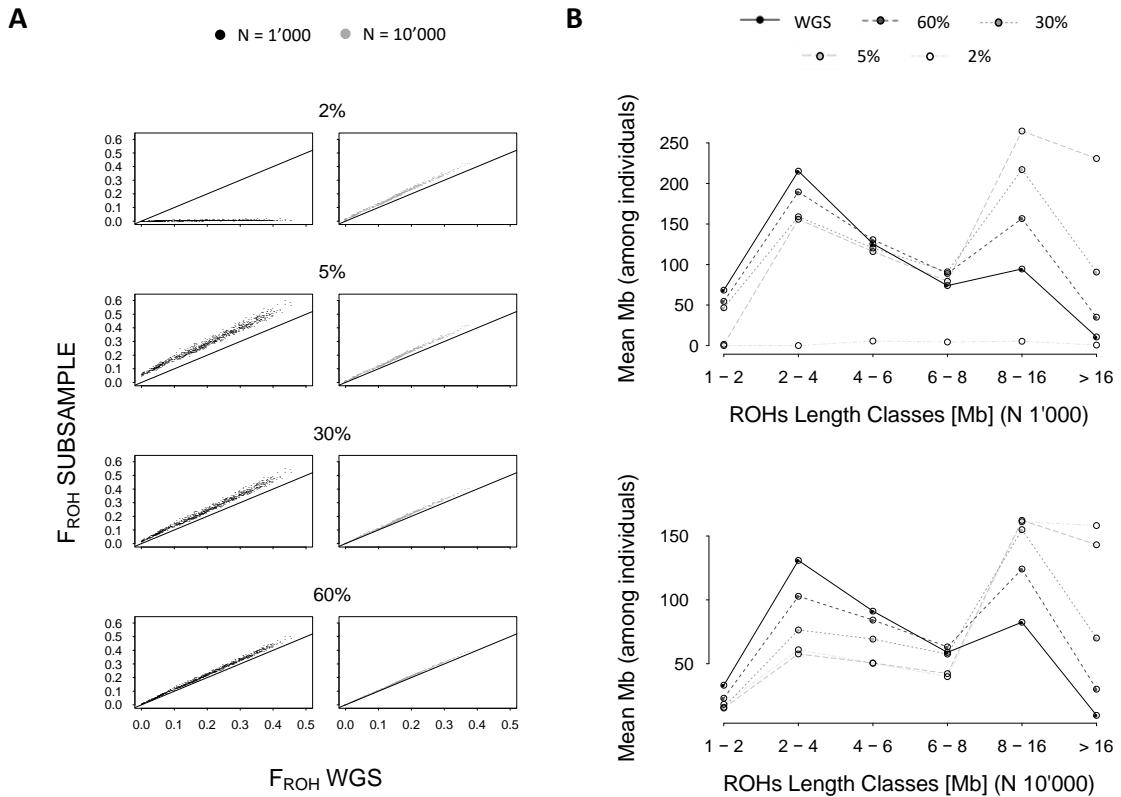


Figure 1: A: The relationship between F_{ROH} estimated with a subsample of the SNPs (indicated in percentages above plots) in the y axis and F_{ROH} calculated with all the SNPs in the x axis for both populations (small is in back on the left and large is in grey on the right). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. F_{ROH} subsampled is represented as the mean among subsampling replicates. r^2 , slopes and intercepts are reported in Table 2. **B:** ROHs distributions. The mean individual sum of lengths of ROHs on the y axis per ROHs length classes on the x axis. Small population is above and large population is below. Individual means are among subsampling and simulations replicates. Points for the same subsampling are linked with specific lines. Legend for both plots on top of the first plot.

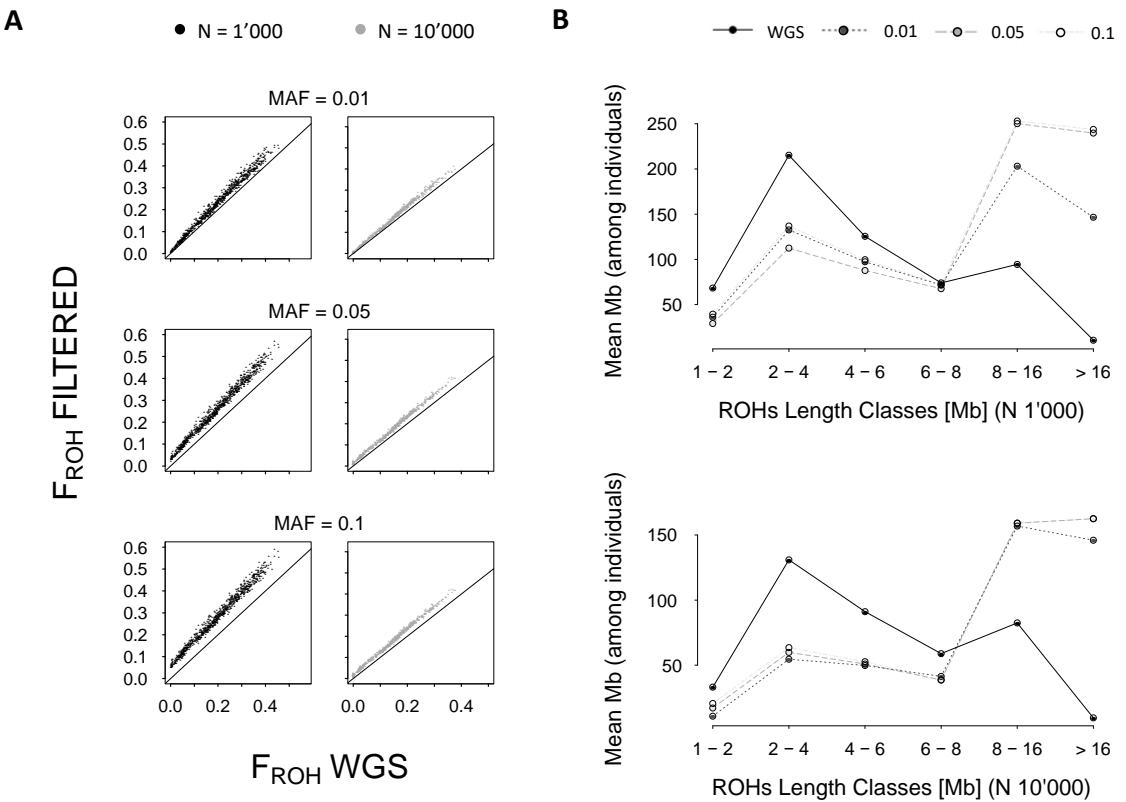


Figure 2: A: The relationship between F_{ROH} estimated with filtering performed on MAF (MAF above plots) on the y axis and F_{ROH} calculated with no filtering on the x axis for both populations (small is in back on the left and large is in grey on the right). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. r^2 , slopes and intercepts are reported in Table 3. **B:** ROHs distributions. The mean individual sum of lengths of ROHs on the y axis per ROHs length classes on the x axis. Small population is above and large population is below. Individual means are among simulations replicates. Points for the same filtering are linked with specific lines. Legend for both plots on top of the first plot.

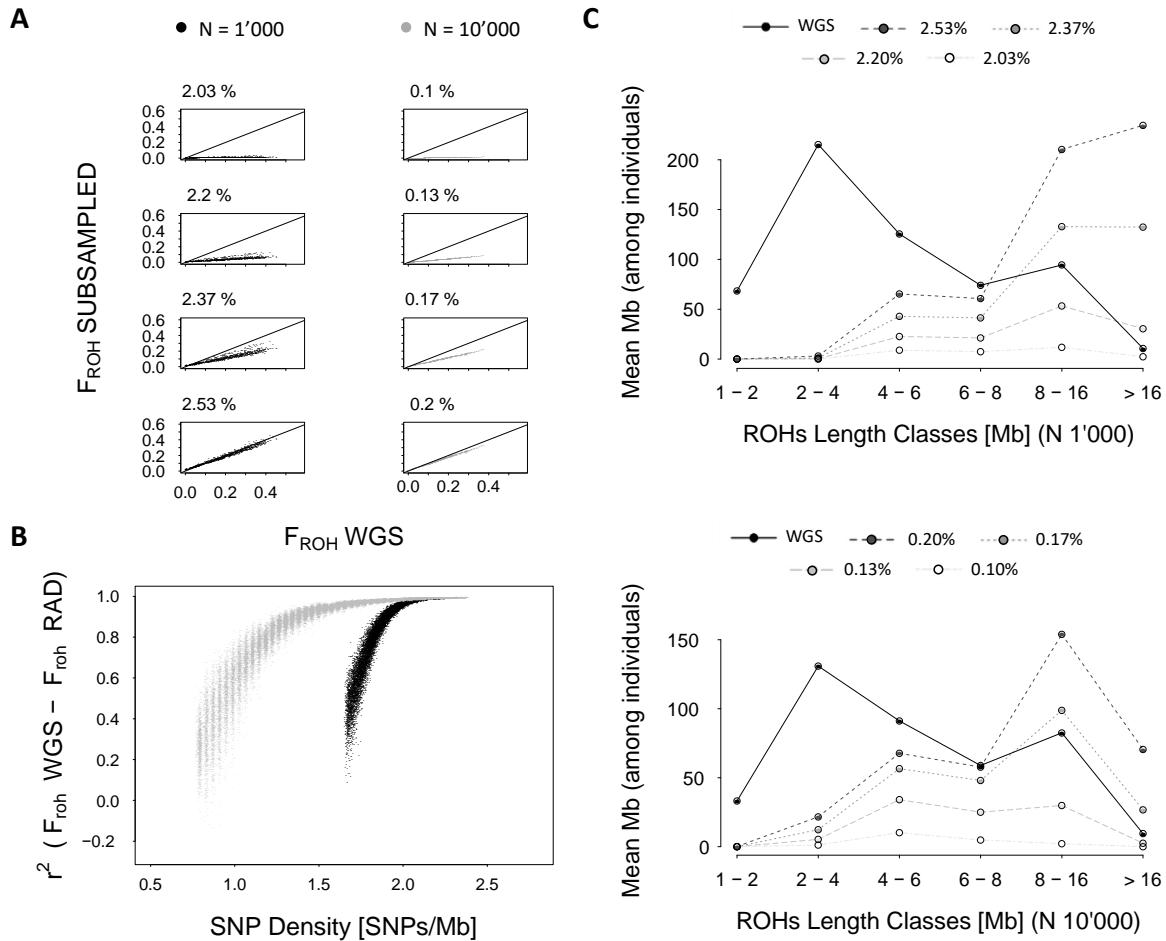


Figure 3: A: The relationship between F_{ROH} estimated with a subsample of the genome (indicated in percentages above plots) mimicking RAD-sequencing on the y axis and F_{ROH} calculated with all the data (WGS-like) on the x axis for both populations (small is in back on the left and large is in grey on the right). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. F_{ROH} RAD-sequencing is represented as the mean among subsampling replicates. r^2 , slopes and intercepts are reported in Table 4. **B:** Correlation between F_{ROH} WGS (on the x axis) and F_{ROH} with RAD-sequencing-like data (on the y axis) dependent on mean SNP density for the subsampling replicate [SNPs/Mb]. All simulations are represented in this graph and each point represents the correlation for one subsampling replicate. Same colours as panel A, small population in black and large population in grey. **C:** ROHs distributions. The mean individual sum of lengths of ROHs on the y axis per ROHs length classes on the x axis. Small population is above and large population is below. Individual means are among subsampling and simulations replicates. Points for the same subsampling are linked with specific lines. Legends above each plot.

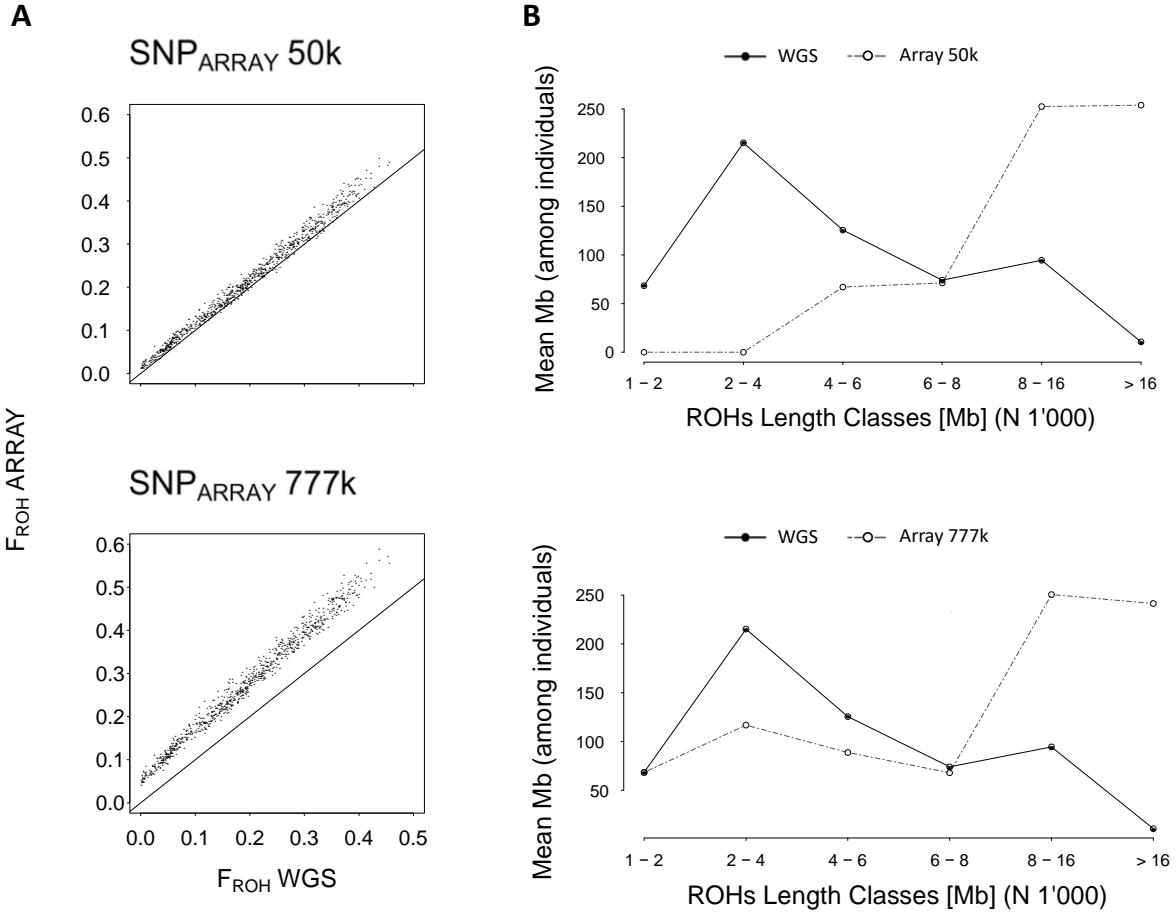
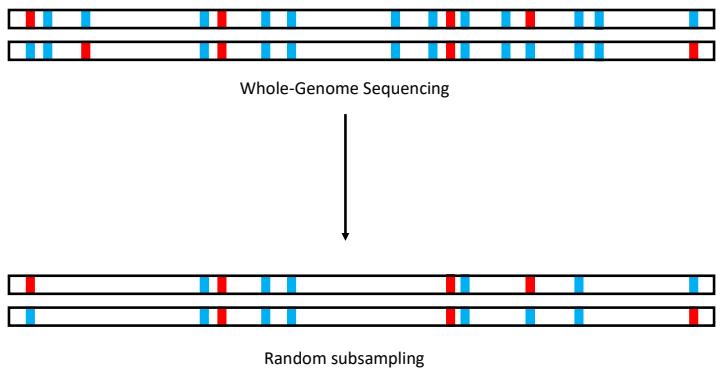


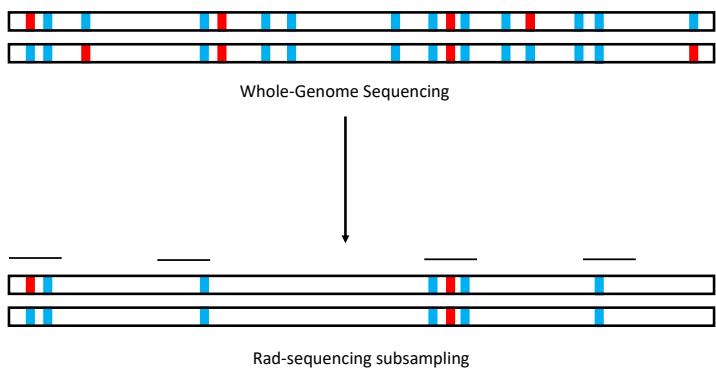
Figure 4: The relationship between F_{ROH} estimated with SNP Arrays like subsampling on the y axis and F_{ROH} calculated with all data (WGS) on the x axis for the small population. The graph above represents SNP Array 50k subsampling and the second graph below SNP Array 700k subsampling. The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. Mean r^2 , slopes and intercepts are reported in Table 5. r^2 , slopes and intercepts per replicate are reported in Table S5. B: ROHs distributions. The mean individual sum of lengths of ROHs on the y axis per ROHs length classes on the x axis. Small SNP Array (50k) is above and large SNP Array (700k) is below. Individual means are among simulations replicates. Legends on top of the plots.

SUPPLEMENTARY FIGURES

A



B



C

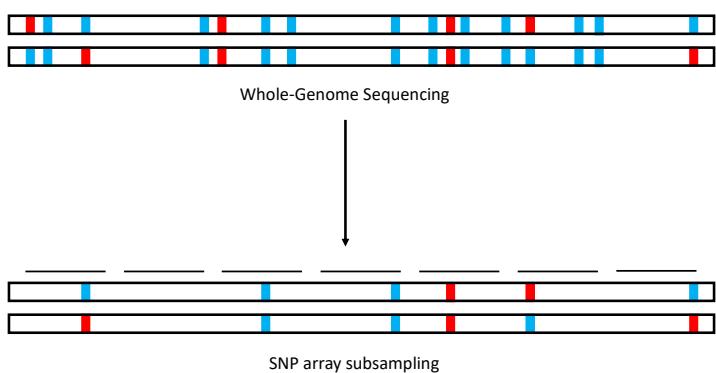


Figure S1: Scheme of three of the four subsampling techniques for a small genomic segment. Each square is a SNP, blue SNPs are the reference allele and red SNPs are the derived allele. **A:** Random subsampling, SNPs are completely randomly selected. **B:** Rad-sequencing, 500bp windows (represented by the black lines above the genome) are randomly assigned in the genome and each SNP within these windows gets subsampled. **C:** SNP array, windows (represented by the black lines above the genome) are equally spaced in the genome. If there is SNPs in the window, one is randomly selected.

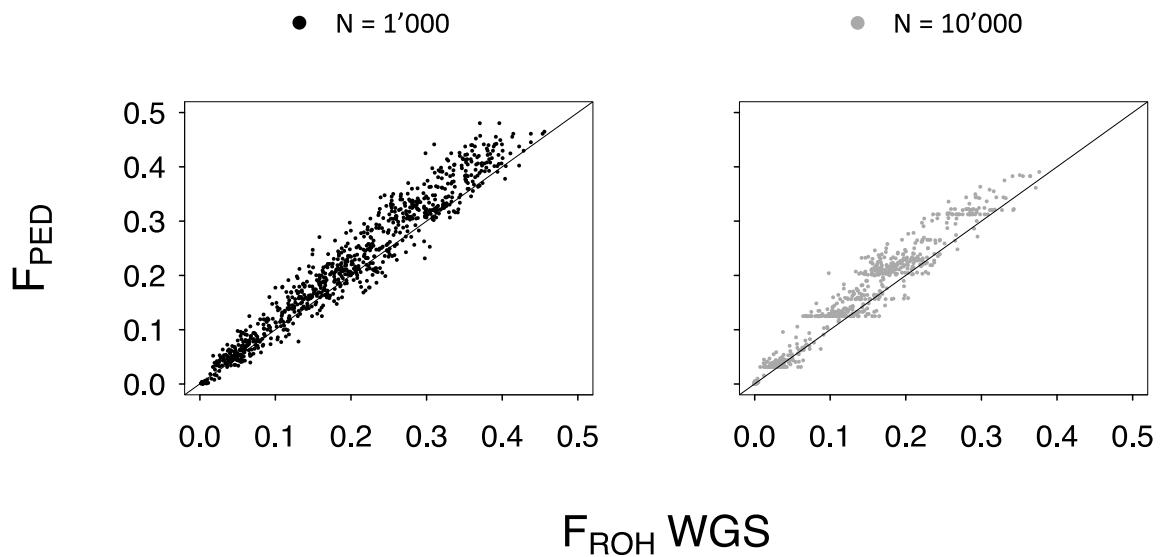


Figure S2: The relationship between F_{PED} on the y axis and F_{ROH} calculated with all the data (WGS-like) on the x axis for both populations (small is in back on the left and large is in grey on the right). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. $r^2: 0.975$, slope: 1.077 and intercept: 0.007 for the small population. $r^2: 0.975$, slope: 1.079 and intercept: 0.009 for the large population.

F_{ROH} SUBSAMPLE

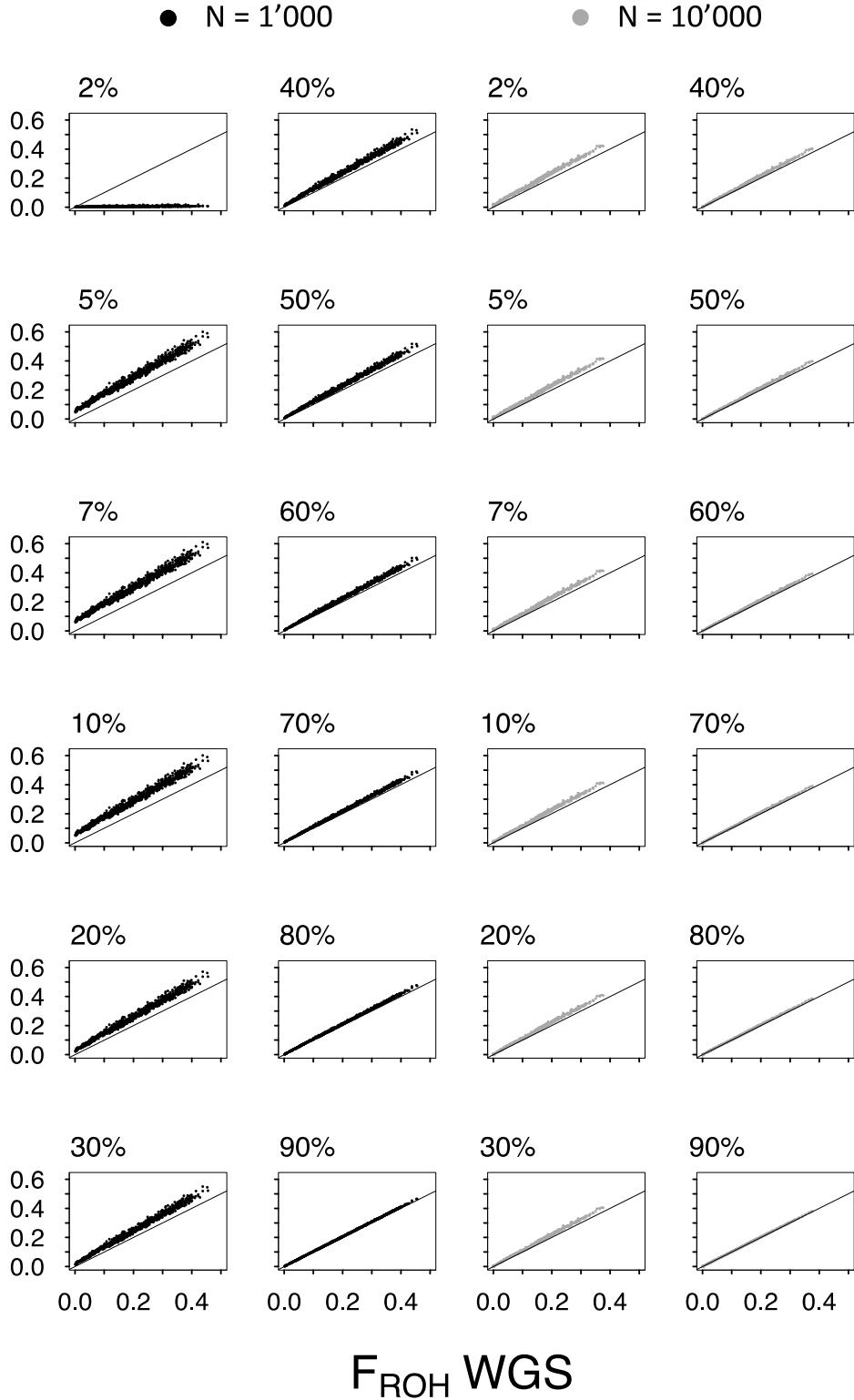


Figure S3: The relationship between F_{ROH} calculated with a subsample of the SNPs (indicated in percentages above plots) in the y axis and F_{ROH} calculated with all the SNPs in the x axis for both populations (small is in back on the two left columns and large is in grey on the two right columns). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. F_{ROH} subsampled is represented as the mean among subsampling replicates. r^2 , slopes and intercepts are reported in Table S2.

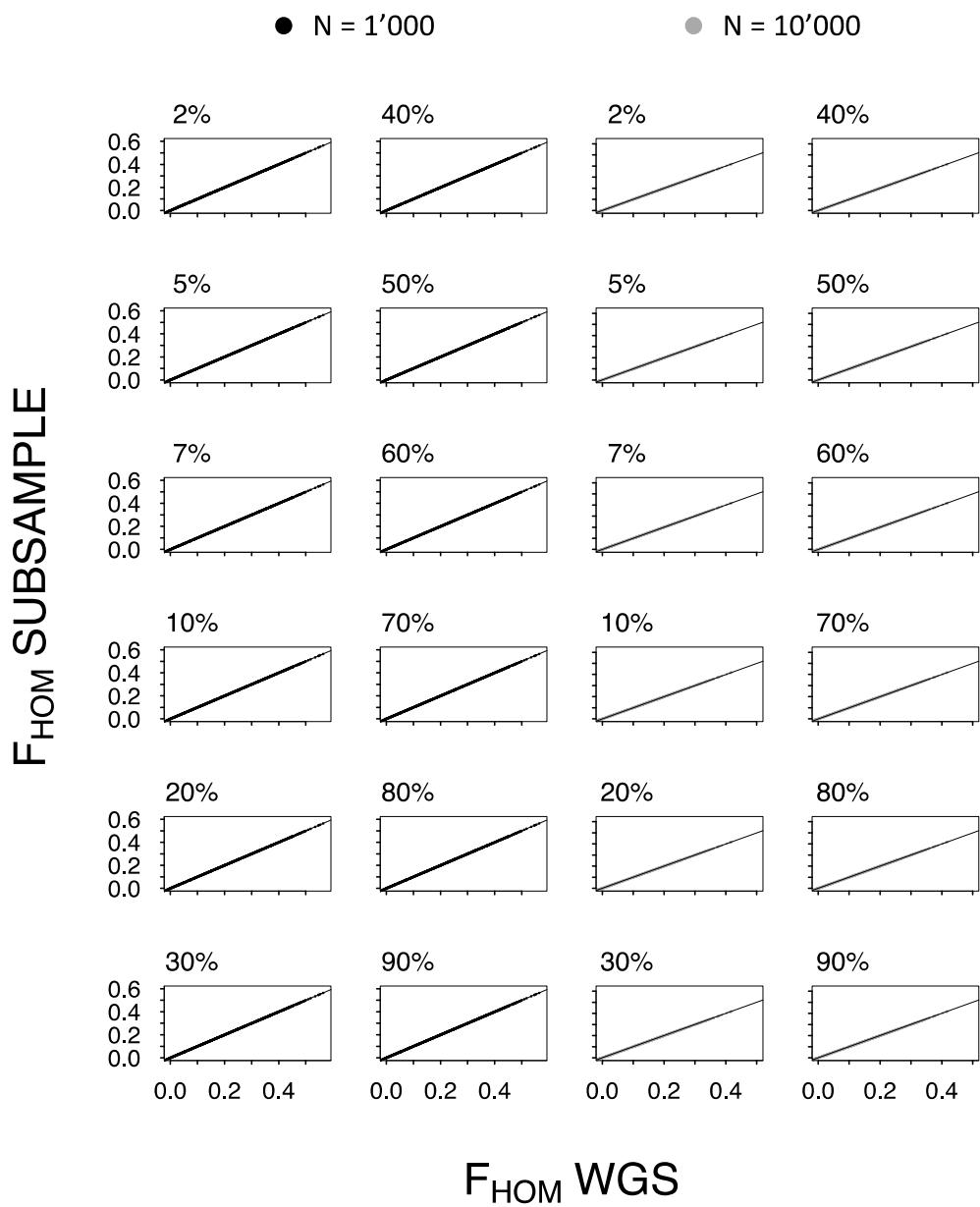


Figure S4: The relationship between F_{HOM} calculated with a subsample of the SNPs (indicated in percentages above plots) in the y axis and F_{HOM} calculated with all the SNPs in the x axis for both populations (small is in back on the two left columns and large is in grey on the two right columns). The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. F_{HOM} subsampled is represented as the mean among subsampling replicates.

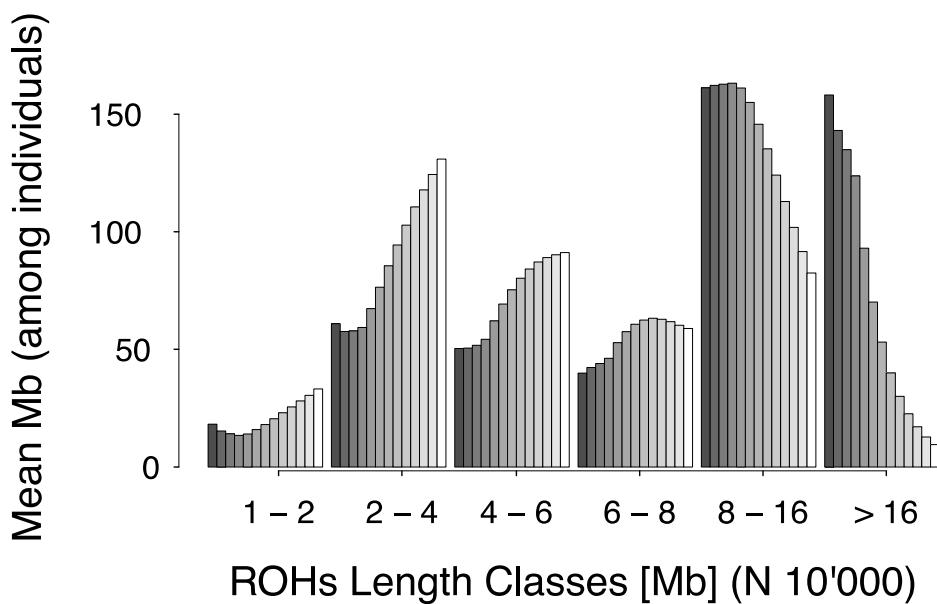
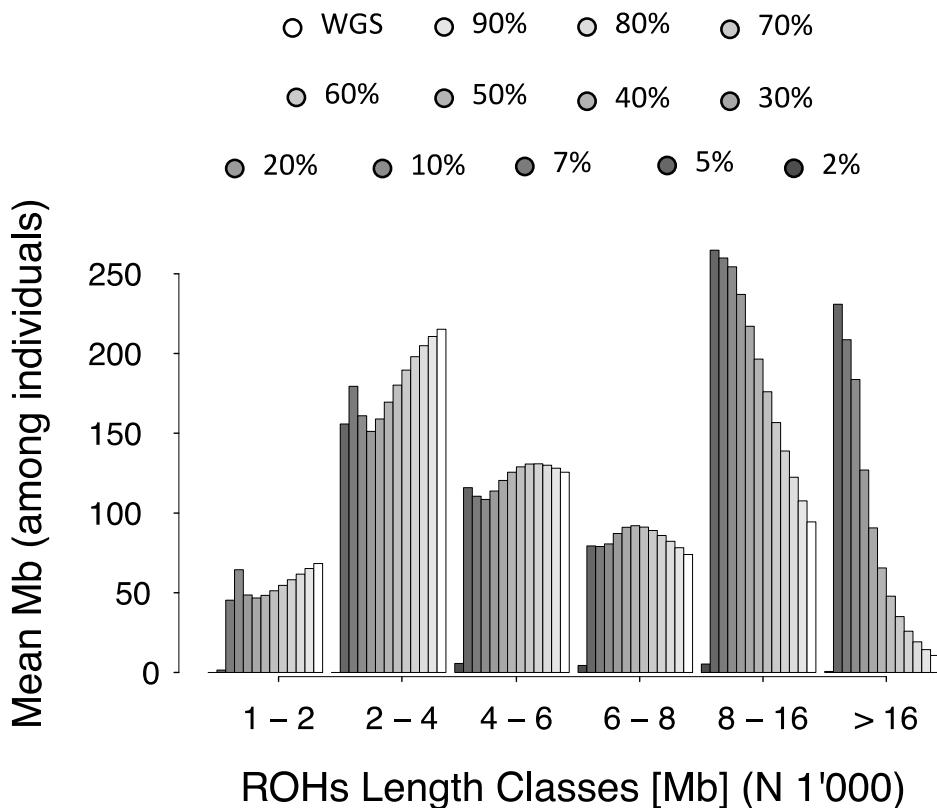


Figure S5: ROHs distributions. The mean individual sum of lengths of ROHs on the y axis per ROHs length classes on the x axis. Small population is above and large population is below. Individual means are among subsampling and simulations replicates. Legend for both plots on top of the first plot.

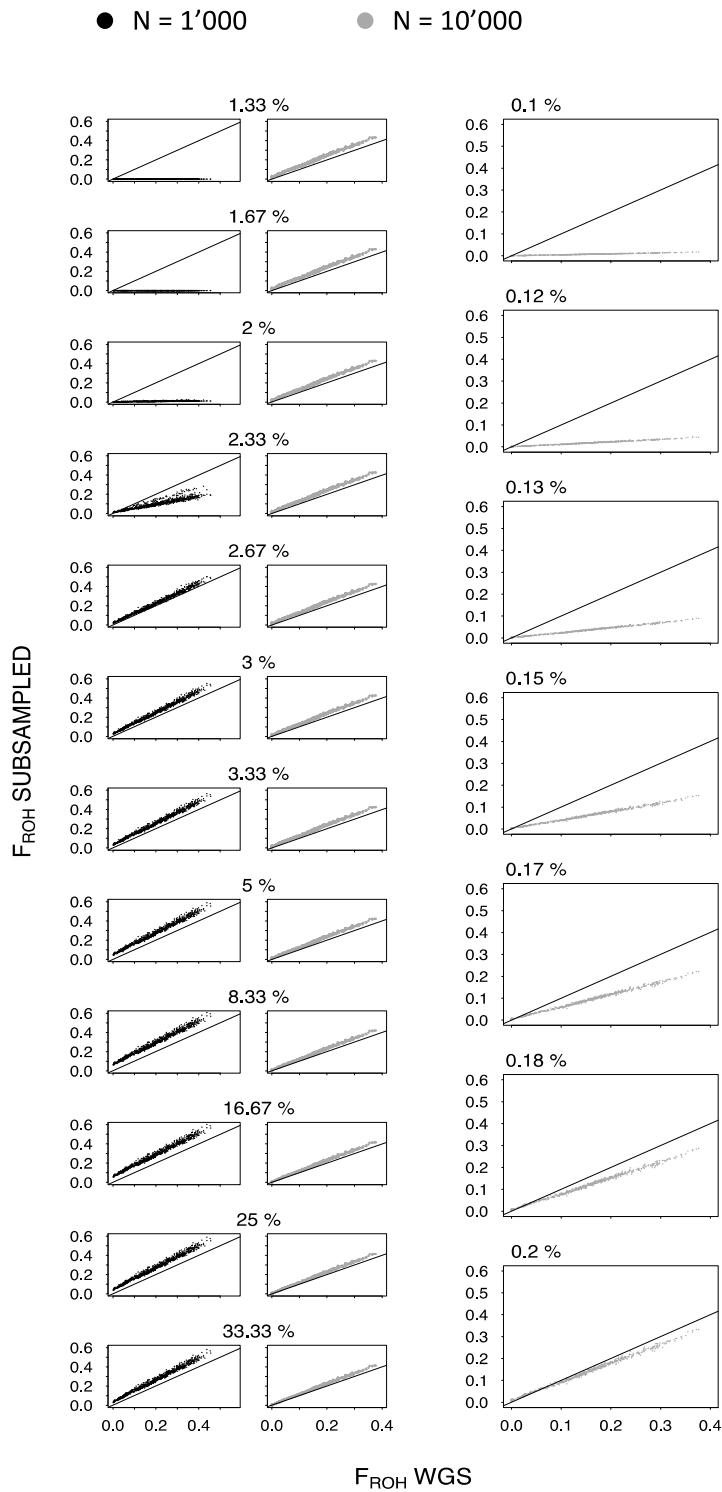


Figure S6: The relationship between F_{ROH} estimated with a subsample of the genome (indicated in percentages above plots) mimicking RAD-sequencing on the y axis and F_{ROH} calculated with all the data (WGS-like) on the x axis for both populations (small is back on the left and large is in grey on the right). The percentages represent the proportion of genome sequenced. The black lines represent the perfect correlation between both. Each point represents one individual. All simulation replicates are represented in these graphs. F_{ROH} RAD-sequencing is represented as the mean among subsampling replicates. r^2 , slopes and intercepts are reported in Table S4.

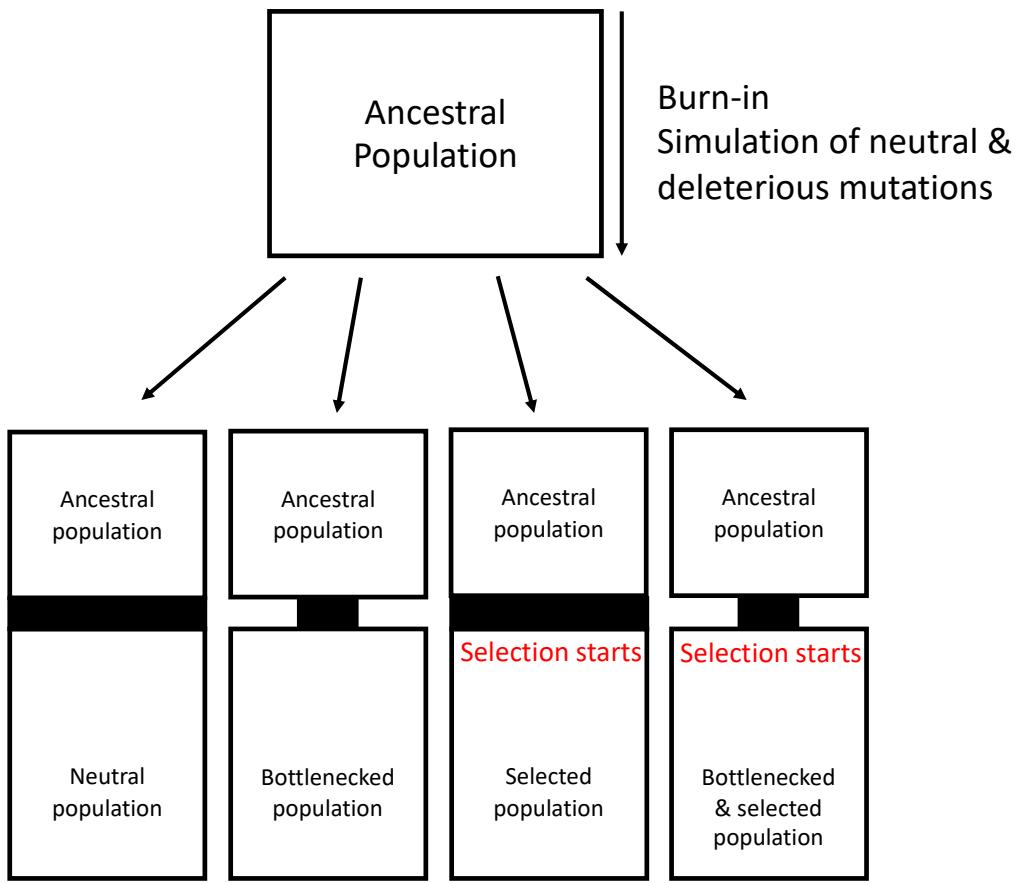


Figure S7: Scheme of Simulations. One burn in will be performed and then be used to construct the four populations. No change for the first population, a bottleneck for the second population, apparition of selection for the third population and a bottleneck and the apparition of selection for the fourth population.