File S1

Details of files submitted as Supporting Information

We provide two collections of data and processed output files from our simulation study. Many other files can be made available by request.

File S2 (**data_files.tar.gz**; see next page 3SI) consists of the marker map and frequencies, the true ibd states among our 500 pairs of individuals sampled for the study, and the generated haplotypes of the 100 individuals who provide these 500 pairs. The files included in the S2 tar archive are

samp500_9_trueibd.txt	The true genotypic ibd state, by marker, for the sample of 500 pairs.
samp500_15_trueibd.txt	The true haplotypic ibd state, by marker, for the sample of 500 pairs
posandfreq.txt	position in cM and allele frequencies of the markers simulated
haps_LD_[gamma].txt	The simulated haplotypes for the 100 individuals generated at LD level gamma,
where gamma = 00, 05, 10,100	where gamma = 0 (very high LD), 0.05 (moderately high), 0.1 (moderately low), and 1.00 (no LD)

File S3 (**sweave_ldresults.tar.gz**; see following page 4SI) contains the output files of inferred states from running ibd_haplo in MORGAN 3.0.3 (Fall 2011 release), to produce "qibd" files and then calling the states using the R-package IBDhaploRtools (Fall 2011 release) as described in the paper. The qibd files provide the probabilities of each ibd state, at each marker, for each of the 500 pairs of individuals. We have not provided the original qibd files, since each of the 22 files is approx 400 Mb (and 30Mb even when compressed).
Instead we have provided the called state results from each of these runs in the form of an Rdata file, together with a Sweave document that can be used together with the the R-package IBDhaploRtools (Fall 2011 release) and used to regenerate the tables and figures of our study. We have provided only the 22 files of inferred ibd states that provide the results tabulated in the paper. The files included in the S3 tar archive are

data.files.needed.txt	list of the files needed to run the Sweave code
marker.pos.Rdata	marker data (as in File S2) in .Rdata format
samp500_15_trueIBD.txt	true hapotypic ibd states (see details for file S2 above)
samp500_9_trueIBD.txt	true genotypic ibd states (see details for file S2 above)
Rplots.pdf,Sweave.sty, create_figures.*	Sweave files (input and output) required to recreate the figures of the paper relating to the performance of ibd_haplo to infer segments of ibd.
create_figures-fig*.pdf create_figures-fig*.eps	PDF and EPS versions of figures 1, 4, and 5 of the paper, as recreated using the Sweave document and IBDhaploRtools R package.
inf_states_h_gamma*_ch.Rdata	Inferred states, using haplotypic data, on each of the four data sets at varying LD levels gamma, for four values of gamma =”*”= 0.0, 0.05, 0.1 and 1.0. Run at model parameters alpha=0.05, beta=0.05, and delta=0.1, as described in the paper.
inf_states_g_gamma*_ch.Rdata	Inferred states, as above, for the same data analyzed as unphased genotypes.
inf_states_h_fkin*_ch.Rdata	Inferred states, using haplotypic data, for the LD (gamma) value 0.1, for two values of beta =”*”=0.02 and 0.05. With alpha=0.05, delta=0.1 as above.
inf_states_g_fkin*_ch.Rdata	Inferred states, as above, for the same two beta values, with the data analyzed as unphased genotypes.
inf_states_h_ffrate*_ch.Rdata	Inferred states, using haplotypic data, for the LD (gamma) value 0.1, for five values of rate parameter alpha = “*” = 0.01, 0.05, 0.1, 0.5, 2.0. With beta=0.05, delta =-0.1 as above.
inf_states_g_ffrate*_ch.Rdata	Inferred states, as above, for the same five alpha values, with the data analyzed as unphased genotypes.

File S2

data_files.tar.gz

File S3

sweave_ldresults.tar.gz

Files S3 and S3 are available for download as a compressed (gzipped) tar archive at http://www.genetics.org/content/suppl/2012/01/31/genetics.111.137570.DC1.

M. D. Brown et al. 3 SI