

Imputation and Phasing Notes

Jinliang Yang

July 22, 2015

Infer mom's genotype from Jeff

We have obs. mom and obs. (selfed) kids. We want to know $P(G|\theta)$, and $P(G|\theta) \propto P(\theta|G) \times (G)$, where θ is observed data. This consists of observed genotypes (G') of both mom and kids. So: $P(G|\theta) \propto \left(\prod_{i=1}^k P(G'_i|G) \right) \times P(G'_{mom}|G) \times P(G)$ This function is to impute mom's genotype from a progeny array of k kids at a single locus. inferred_mom=1 -> 00, 2->01, 3->11

problems

1. errors should not be used for infer
2. p unknown, could only be estimated from data.

to do:

1. haplotype needs to consider imputation error.?
2. lost info (i.e. sites heter but inferred homo)
3. missing data.