

Parentage Analysis



Mariah Meek and Jessica
Petersen

March 9, 2007

Parentage Analysis to Answer Ecological Questions

- Variance in Reproductive Success
- Estimate Number of Breeders
- Prevalence of Multiple Paternity
- Frequency of Extra-pair Matings
- Determination of Breeding Patterns



One Parent Known





Neither parent known



Methods

- Exclusion
- Categorical likelihood
- Fractional likelihood
- Genotypic reconstruction



Exclusion

- Mendelian rules of inheritance in diploid data
- Uses incompatibilities to reject parent-offspring hypotheses
- Most successful when few candidate parents and high polymorphism
- More potential parents = More loci needed
- Most desirable method, if possible
- Genotyping errors, null alleles, and mutations contribute to false exclusions
 - More Data makes higher likelihood of errors
- Some programs allow specification of number of mismatches necessary for exclusion

Categorical Likelihood

- Calculates log likelihood ratio (LOD score): $(\text{likelihood of being the parent}) / (\text{likelihood unrelated})$
- Offspring assigned to parent with highest LOD score
- Assigns entire offspring to a particular male
- Allows for some genotyping errors or mutation
- Male's likelihood of parentage is positively correlated with homozygosity

Fractional allocation



- Assigns some fraction of each offspring to all non-excluded candidate parents
- Portion assigned is proportional to its likelihood of parentage compared to all other non-excluded candidates
- Not biologically correct, but can allow greater statistical properties for evaluation of some hypotheses

Genotypic Reconstruction

- Uses multilocus genotypes to reconstruct genotypes of an unknown parent contributing to full or half sib progeny arrays.
- Offspring must be collected in family groups
- Algorithm reconstructs minimum number of parental genotypes necessary to explain the data set
- Computationally intensive, prohibitively time consuming if 6+ fathers
- Susceptible to scoring errors, null alleles, and mutations, but often only problem for one or few offspring at a locus

Data issues

- Proportion of candidate adults sampled
 - CERVUS: you estimate
 - PATRI: calculates estimate of population size and uncertainty
 - PAPA/PROBMAX: assign parent pairs, if one parent unsampled, correct assignment is impossible
- Null alleles:
 - Detect with departures from HWE
 - Once detected, recode homozygotes as heterozygotes with the null allele
 - Or remove from analyses
- Linked loci or Linkage Disequilibrium
 - Linkage disequilibrium decreases expected probability of exclusion and accuracy
 - Linked loci can increase accuracy of assignment if linkage phase and recombination rate are known
 - BUT, computer programs not well equipped to deal with LD, so exclude from data set

Data Issues

- Mutation and Genotyping Error

- Causes rejection of true parent
- Different programs deal with differently

- Extended family structure

- Most studies assume adults are not related to one another and no other relatives of offspring present
- Most problematic when half or full siblings included in pool of potential parents
- Strict exclusion methods still work, but require high number of loci
- Some programs can deal with using estimate of relatedness