Genetic Parentage Analysis Software

Cervus 3.0

Table 1 The implications of various types of samples from natural populations for the reconstruction of parentage

How many of the unknown parents can I sample?	Can I sample large groups of full or half-sibs?	How many parents of each offspring are known a priori?	Preferred technique	What can I do if I don't powerful have sufficiently markers for the preferred technique?	Possible computer programs (see Table 2)
All or some	Yes or no	Both	Genotype some individuals to verify parent–offspring relationships	Nothing	No program necessary
All	Yes	One	Exclusion corroborated by reconstruction of parental genotypes from progeny arrays	Categorical allocation, fractional allocation, estimate number of parents	NEWPAT, CERVUS, FAMOZ, PARENTE, PATRI or KINSHIP for exclusion or allocation. GERUD for reconstruction
All	Yes	Neither	Complete exclusion. If progeny arrays contain half-sibs, reconstruction of parental genotypes	Categorical allocation, fractional allocation, estimate number of parents	PROBMAX, PAPA, FAMOZ or PARENTE for exclusion or allocation. GERUD for reconstruction
All	No	One	Complete exclusion	Categorical allocation, fractional allocation, kinship techniques	NEWPAT, CERVUS, FAMOZ, PARENTE, PATRI or KINSHI for exclusion or allocation
All	No	Neither	Complete exclusion	Categorical allocation, fractional allocation, kinship techniques	PROBMAX, PAPA, FAMOZ or PARENTE for exclusion or allocation.
Some	Yes	One	Reconstruction of parental genotypes. Complete exclusion or categorical/fractional allocation	Categorical allocation, fractional allocation, estimate number of parents	NEWPAT, CERVUS, FAMOZ, PARENTE, PATRI or KINSHI for exclusion or allocation. GERUD for reconstructio
Some	Yes	Neither	Complete exclusion or categorical/fractional allocation Reconstruction of parental genotypes if progeny arrays contain half-sibs	Categorical allocation, fractional allocation, estimate number of parents	NEWPAT, CERVUS, FAMOZ, PARENTE, KINSHIP, PROBMAX[1] or PAPA[1] for exclusion or allocation. GERUD for reconstruction
Some	No	One or neither	Complete exclusion or categorical/fractional allocation.	Kinship techniques	NEWPAT, CERVUS, FAMOZ, PARENTE, PATRI (if one parent is known), KINSHIP, PROBMAX[1] or PAPA[1] for exclusion or allocation
None	Yes	One	Reconstruction of parental genotypes	Estimate number of parents	GERUD for reconstruction
None	Yes	Neither	Reconstruction of parental genotypes if progeny arrays contain half-sibs	Estimate number of parents	GERUD for reconstruction
None	No	One or neither	Use kinship or relatedness techniques	Nothing	KINSHIP or RELATEDNESS[2]

[1]Parent-pair allocation techniques are particularly sensitive to incomplete sampling of candidate parents and should be used carefully here.
[2]RELATEDNESS is available from http://gsoft.smu.edu/

Jones et al. 2003. Mol Ecol.12:2511

Table 2 Computer programs for reconstructing parentage in natural populations

Method	Program	Available functions						Required info.	Type of genetic markers	Ability to accommodate 'error'			Comments
		PM[a]	PP[b]	PR[c]	IC[d]	EC[e]	EP[f]			Null alleles	Genotyping error	Mutation	
Exclusion	PROBMAX[g]		X					Genotypes of offspring and sexed parents	Diploid codominant Diploid dominant	Good	Moderate	Moderate	Can specify parental mating combinations
	NEWPAT[h]	X			X	X		Genotypes of offspring and sexed parents	Diploid codominant Sex-linked loci	Moderate	Moderate	Moderate	Can calculate confidence intervals for null allele frequency
	KINSHIP[i]	X			X			Genotypes of offspring and parents	Diploid codominant	None	None	None	Can handle haploids as well as diploids
Categorical allocation	CERVUS[j]	X			X	X	X	Genotypes of offspring and sexed parents	Diploid codominant	Moderate	Good	Good	Excellent manual and user interface Calculates expected null allele frequency
	PAPA[k]		X			X		Genotypes of offspring and parents	Diploid codominant	Poor	Good	Good	Easy to use and excellent interface
	FAMOZ[l]	X	X		X	X	X	Genotypes of offspring and parents	Diploid dominant Diploid codominant Cytoplasmic	Poor	Good	Good	Can estimate cryptic gene flow Difficult file format
	PARENTE[m]	X	X		X			Genotypes of offspring and parents	Diploid codominant	Poor	Good	Good	Can take into account dates of birth and death
Fractional allocation	PATRI[n]	X			X			Genotypes of parent-offspring pairs and sexed parents	Diploid codominant	None	None	None	Can test relative reproductive success of different groups Can also be used for categorical allocation
Parental reconstruction	GERUD[o]			X		X	X	Genotypes of known parent with a large group of its progeny	Diploid codominant	None	None	None	Uses multilocus data to determine the minimum number of sires for a family

[a]Paternity/maternity (see Information box).
[b]Parental pair allocation (see Information box).
[c]Parental reconstruction (see Information box).
[d]Ability to assign statistical confidence for particular parent–offspring pairs.
[e]Ability to assess the expected confidence in assignments on an experiment-wide basis.
[f]Ability to calculate exclusion probabilities.
[g]Danzmann (1997); available at http://www.uoguelph.ca/~rdanzman/software/PROBMAX/
[h]Worthington Wilmer *et al.* (1999); available at http://www.zoo.cam.ac.uk/zoostaff/amos/newpat.html
[i]Goodnight & Queller (1999); available at http://gsoft.smu.edu/
[j]Marshall *et al.* (1998); available at http://helios.bto.ed.ac.uk/evolgen/cervus/cervus.html
[k]Duchesne *et al.* (2002); available at http://www.bio.ulaval.ca/contenu-fra/professeurs/Prof-l-bernatchez.html
[l]Gerber *et al.* (2000); available at http://www.pierroton.inra.fr/genetics/labo/Software/Famoz/
[m]Cercueil *et al.* (2002); available at http://www2.ujf-grenoble.fr/leca/membres/manel.html
[n]Signorovitch & Nielsen (2002); available at http://www.biom.cornell.edu/Homepages/Rasmus_Nielsen/files.html
[o]Jones (2001); available at http://www.biology.gatech.edu/professors/labsites/jones/parentage.html

Jones et al. 2003. Mol Ecol.12:2511

Cervus 3.0

- Likelihood based parentage assignment
  - Maternity/Paternity
  - Parent Pair Assignment
  - Have to have at least some parents' genotypes

- Likelihood equations from:
  - Marshall et al. 1998
  - Revised by Kalinowski et al. 2007
    - New version accounts for problems in Marshall's method of calculating genotyping error (it inflated the rate in which errors were expected to be observed – it did not exclude parents when it should have)

Assumptions/Limitations

- Genotyping errors are independent and constant across loci
- No linkage
- Hardy-Weinberg Equilibrium
  - Can handle minor deviations at a few loci
- Codominant, diploid data

Analyses Steps

1. Allele Frequency Analysis
2. Simulation
3. Parentage Analysis

Allele Freq Analysis Output

Allele Frequency Analysis: On Screen Output

Locus	k	N	HObs	HExp	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI	HW	F(Null)
A-	8	135	0.63	0.59	0.55	0.79	0.62	0.42	0.19	0.50	NS	-0.050
B-	17	135	0.89	0.87	0.86	0.40	0.24	0.09	0.02	0.32	NS	-0.012
C-	9	134	0.79	0.84	0.82	0.48	0.31	0.14	0.04	0.34	NS	+0.027

- NE = Non-exculsion probability:
  probability of not excluding a single unrelated candidate parent (probability a parent that is NOT the real parent is included as a potential parent)
  - NE-1P: NE prob for a single parent
  - NE-2P: NE prob for a single parent when one parent is known
  - NE-PP: NE prob for a parent pair
- NE prob is not used for likelihood calculations but may be useful for comparisons to other data or choosing loci

Simulations

- Purpose:
  - Estimate the resolving power of loci given their allele frequencies.
  - Estimate critical values of LOD or Delta, allowing to statistically analyze the confidence of assignment.
- Generate pair of parent genotypes and several unrelated parent genotypes using allele frequencies
- Creates series of progeny from the parent pair
- Calculates likelihood of each true parent and the unrelated parents for each offspring

Simulations (cont.)

- Scores from the simluated data are used to determine the critical LOD or Delta score for the real parentage analysis
- Simluation options:
  - Self-fertilization
  - Inbreeding/Related parents

Parentage Analysis

- Types of analyses:
  - One parent is known (paternity/maternity)
  - Neither parent is known (but have potential parent genotypes)
    - Sexes unknown
    - Sexes known

Parentage Output

Offspring ID	Loci typed	Candidate mother ID	Loci typed	Pair loci compared	Pair loci mismatching	Pair LOD score	Pair Delta	Pair confidence
Deer5	29	Deer25	29	29	9	-2.89E+01	0.00E+00	
Deer34	29	Deer119	29	29	0	2.30E+01	2.30E+01	*
Deer35	29	Deer21	29	29	1	9.47E+00	9.47E+00	*

Offspring ID	Candidate father ID	Loci typed	Pair loci compared	Pair loci mismatching	Pair LOD score	Pair Delta	Pair confidence
Deer5	Deer65	29	29	4	-6.35E+00	0.00E+00	
Deer34	Deer31	29	29	6	-9.91E+00	0.00E+00	
Deer35	Deer87	29	29	3	5.71E+00	5.71E+00	*

Offspring ID	Trio loci compared	Trio loci mismatching	Trio LOD score	Trio Delta	Trio confidence
Deer5	29	11	-2.37E+01	0.00E+00	
Deer34	29	11	-4.56E+00	0.00E+00	
Deer35	29	4	2.05E+01	9.85E+00	*

Things to Consider

- “Guessing” at the proportion of parents sampled
- Linkage?
- Dealing with Null Alleles

- Run it several times with various options