G3.7

Effectiveness of a new strategy for the computation of transmission probabilities using partial pedigrees. A simulation study

Mayer M., Tuchscherer A. and Reinsch N.
Research Unit Genetics and Biometry,
FBN Dummerstorf, Germany

mmayer@fbn-dummerstorf.de

Requirements for the recursive algorithms

2 Requirements:

1. **Known and ordered pedigree information (individual, sire, dam identification).**
2. **transmission probabilities for the QTL alleles (conditional on marker information) for all non-founder individuals.**

Objectives of this presentation

- **Short presentation of a new strategy for the computation of transmission probabilities, especially when marker information is missing.**
- **Simulation study to show the effectiveness of the proposed strategy using a data structure of practical relevance.**

Our strategy for the computation of transmission probabilities

1. **The complete pedigree is partioned into relatively small partial pedigrees.**
2. **Using these partial pedigrees (IBD-)identity states according Gillois (1964) are computed for the parent-progeny pairs (e.g. MERLIN, SIMWALK2).**
3. **From these identity states the transmission probabilities are computed.**

Simulation design I

The present simulation study is based on the pedigree structure and marker information of a data set, which was used for marker assisted BLUP evaluation for the trait somatic cell score in the German Holstein population.

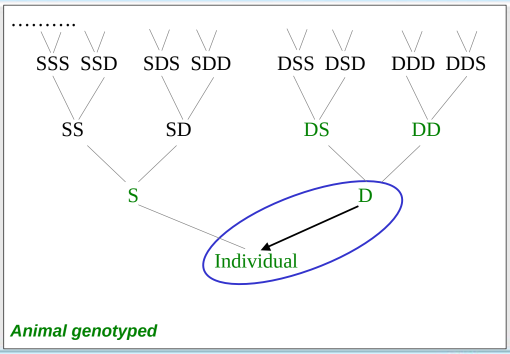
Simulation design II

- 12008 records with the identification of the individual itself and its sire and its dam.
- 6520 individuals genotyped for one particular marker.
- Highly polymorpic marker (# = 15; Allelfr. 33%, 14%, 12% ...)

- The simulation was based on the given pedigree structure. For the base animals a paternal and a maternal marker allele was randomly assigned according to the frequency distribution above. Thereafter with equal probability either the paternal or the maternal marker allele was transmitted to the progeny.

- 100 simulated data sets.

Computation of transmission probabilities Complete marker information I



Animal genotyped

Computation of transmission probabilities Complete marker information II

The transmission probabilities in the past are irrelevant for the actual (parent -> progeny) transmission probabilities.

(The identity states and thus relationship coefficients calculated from these states are not independent from the size of the (partial) pedigree).

Incomplete marker information and family types

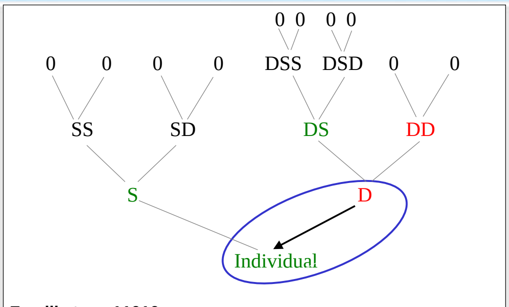
Binary code (5 positions):

1. **position: individual genotyped (1=yes, 0=no)?**
2. **position: sire genotyped?**
3. **position: dam genotyped?**
4. **position: dams sire genotyped?**
5. **position: dams dam genotyped?**

Most common family types

Family type 11010:	**2187**
Family type 11000:	**2135**
Family type 11110:	**686**
Family type 10010:	**635**
Family type 10000:	**371**
Family type 11100:	**246**
Family type 10110:	**93**
Family type 11111:	**79**
…	

Partial pedigrees for the computation of transmission probabilities



Familiy type 11010:
Partial pedigree and marker information (no marker info)

Differences between estimated and true transmission probabilities for family type 11010 (n = 2187), r = 0

difference	**rel. frequency**
0 – 0.05	**33.6 %**
0.05 – 0.15	**11.1 %**
0.15 – 0.25	**23.3 %**
0.25 – 0.35	**3.6 %**
0.35 – 0.45	**9.6 %**
0.45 – 0.55	**3.8 %**
0.55 – 0.65	**6.3 %**
0.65 – 0.75	**1.6 %**
0.75 – 0.85	**6.0 %**
0.85 – 0.95	**1.3 %**
0.95 – 1.00	**0.0 %**

mean: 0.234

Differences between estimated and true transmission probabilities across all family types (n = 6508)

difference	rel. frequency
0 – 0.05	12.3 %
0.05 – 0.15	2.7 %
0.15 – 0.25	5.7 %
0.25 – 0.35	4.0 %
0.35 – 0.45	3.3 %
0.45 – 0.55	66.2 %
0.55 – 0.65	2.3 %
0.65 – 0.75	1.8 %
0.75 – 0.85	1.5 %
0.85 – 0.95	0.3 %
0.95 – 1.00	0.0 %

mean: 0.411

Differences in the estimates of transmission probabilities between standard and 5-generation partial pedigrees

Family type	Mean absolute differences	
	estimates	estimated-true
11010	0.004	0.001
11000	0.005	0.001
11110	0.011	0.006
10010	0.006	0.001
10000	0.009	0.002
11100	0.056	0.030
10110	0.016	0.009
11111	0.000	0.000
all	0.008	0.003

Conclusions

- In situations were marker information is missing the proposed algorithm can be used with high effectiveness for the computation of transmission probabilities and thus gametic or genotypic relationship matrices.
- This strategy can deal with pedigrees of arbitrary size and is very efficient under computational aspects.

References

Tuchscher, Mayer and Reinsch: Calculation of QTL-allele transmission probabilities from detailed IBD-states using partial pedigrees for complete and incomplete marker information, in revision.

Mayer, Tuchscher and Reinsch: Effectiveness of a new strategy for the computation of transmission probabilities using partial pedigrees, submitted.