

Genomic Mismatch Scanning

Semesterarbeit von Michael Breitenstein
Betreuer: Bernd Fischer

Zeitraum der Arbeit: Oktober 2004 bis Februar 2005

Breeding programs usually rely on the phenotypic selection of individuals to reduce the probabilities of a certain diseases. Since several years people are interested in a genotypic selection of breeding animals. Linkage disequilibrium is the most common way to locate genes on the genome that are responsible for a certain trait. At certain marker positions the genotype of individuals is measured. Linkage analysis is based on a likelihood ratio test on the occurrences of different allele at the markers. Usually the number of markers is very low, but the number of test individuals is very large.

Cheung et al. [1] investigated the genomic mismatch scanning (GMS) technique, that enriches fragments of DNA that are identical for two individuals. Under the assumption that a gene is inherited identical by descent from a common founder, two related and affected individuals should share a region of DNA around the gene locus of interest. The length of this region depends on the relationship between the individuals. The closer the two individuals are related, the larger the common DNA around the gene locus. Grant et al. [3] developed a method for direct IBD mapping without considering the relation between individuals.

In this Semester thesis the pedigree information between individuals should be used to validate the mismatch scans and to select individuals from a large pedigree for genomic mismatch scanning. Donnelly [2] developed a hidden Markov model to compute the probability that two individuals share *some* part of DNA identical by descent. Based on this model the probability that two individuals share a part of genome of length l in common should be computed by simulation. For a number of given mismatch scans and a given pedigree, the probability of each gene to be the disease gene should be computed. Further for a given pedigree individuals should be selected such that success of the experiments is maximized.

The formal requirements include a short final presentation of about 20 minutes or more to summarize the results. The results will be written in a short report of approximately 5 pages. Well documented, running source code has to be delivered to the advisor of this project.

References

- [1] V. Cheung, J.P. Gregg, K.J. Gogolin-Ewens, J. Bandong, C. A. Stanley, L. Baker, M. J. Higgins, N. J. Nowak, T. B. Shows, W. J. Ewens, S. F. Nelson, and R. S. Spielman. Linkage-disequilibrium mapping without genotyping. *Nature Genetics*, 18(3):225–230, 1998.
- [2] P. Donnelly. The probability that some individuals share some section of genome identical by descent. *Theoretical Population biology*, 23:34–64, 1983.
- [3] G. Grant, E. Manduchi, V. Cheung, and W. Eweans. Significance testing by direct identity-by-descent mapping. *Ann. Hum. Genet.*, 63:441–454, 1999.