Research Institute for the Biology of Farm Animals (FBN), Dummerstorf, Germany

CHRISTINE BAES and NORBERT REINSCH

TIGER: A software system for fine-mapping quantitative trait loci

Abstract
The localisation of quantitative trait loci which contribute significantly to phenotype variation of economically important traits in domestic species has become an important goal in animal genomics. Several such loci have been roughly identified using linkage analyses; however the focus has now shifted towards fine mapping and pinpointing causal mutations. In the context of a cooperative national research project, the software system TIGER was developed. TIGER is a UNIX script linking several individual Fortran programmes and is used for comprehensive variance component analysis of fine mapping data. Starting with raw genotype data, pedigree and marker map information and ending with a residual maximum likelihood-based test for each putative quantitative trait locus position, the software provides the user with an 'all in one' package capable of linkage analysis, linkage disequilibrium analysis and combined linkage/linkage disequilibrium analysis. The software system has been employed in 4 fine mapping projects on 4 distinct cattle chromosomes.

Keywords: fine mapping, quantitative trait loci, variance component

Zusammenfassung
Titel der Arbeit: **TIGER: Ein Programmsystem zur Feinkartierung von Quantitative Trait Loci**
Strategien, die Informationen über Loci für quantitative Merkmale (QTL) mittels genetischer Marker einbeziehen, eröffnen neue Perspektiven. Einige solcher Loci sind schon grob mittels Kopplungsanalysen identifiziert worden, jedoch hat sich der Schwerpunkt der Untersuchungen in Richtung Feinkartierung und der genauen Erkennung von kausalen Mutationen verschoben. Im Rahmen eines kooperativen nationalen Forschungsprojektes wurde das Software-System TIGER entwickelt. TIGER ist ein UNIX Skript, das mehrere Fortran-Programme verbindet und wird für Varianzkomponenenten-Analysen von Feinkartierungsdaten eingesetzt. Anhand von Genotypen, einer passenden Markerkarte und Abstammungsinformationen kann eine Testgröße, basierend auf der Residual Maximum Likelihood, für jede putative QTL-Position ermittelt werden. Das Software-System bietet sich als umfassendes Paket an und kann sowohl für Kopplungsanalysen und Kopplungsungleichgewichtanalysen als auch für kombinierte Kopplungsgleichgewicht/Kopplungsungleich-gewichtanalysen eingesetzt werden. Das System wurde schon in vier Feinkartierungsprojekten angewandt.

Schlüsselwörter: Feinkartierung, Quantitative Trait Loci, Varianzkomponenten

Introduction
The vast majority of economically important traits in plants and animals are quantitative in nature and can be influenced by many individual chromosomal regions (quantitative trait loci, QTL). The localisation of QTL which contribute significantly to phenotypic variation is an important goal of both the scientific and the agricultural community, although mapping these QTL is only the first step in the search for the underlying genes (COPPIETERS et al., 1999). Several QTL for economically important traits in domestic species have been roughly positioned using linkage analysis (LA). However, the confidence intervals achieved in such studies contain far too many genes to allow the identification of causal mutations and the results are family-specific. The focus has now shifted to fine mapping these QTL as the next step toward identifying genes underlying complex traits. Methods that utilize historical

recombination information on the basis of linkage disequilibrium (LD) allow implementation of fine mapping analyses on existing outbred populations. These methods have become popular for QTL fine mapping in livestock, where the creation of mapping populations such as advanced intercross lines (DARVASI and SOLLER, 1995) or others (DARVASI, 1998) is not practicable.

The pedigree structure of typical dairy cattle populations is ideally suited for QTL mapping. Through the wide-scale implementation of artificial insemination, the paternal ancestry of many animals in the population can be traced back to a handful of bulls. The effective population size is small, although, as for example in the Holstein Friesian breed, many millions of animals exist. Furthermore, the phenotypic data of female animals for most relevant traits has been systematically recorded for use in routine breeding value estimation. The performance of the sires' progeny can thereby be included in the calculation of the sires' breeding value. This 'granddaughter design' (Figure 1) lends itself to use in mapping QTL and requires less genotyped animals than a 'daughter design', in which only large sets of paternal half sisters are used (GEORGES et al., 1995, WELLER et al., 1990).
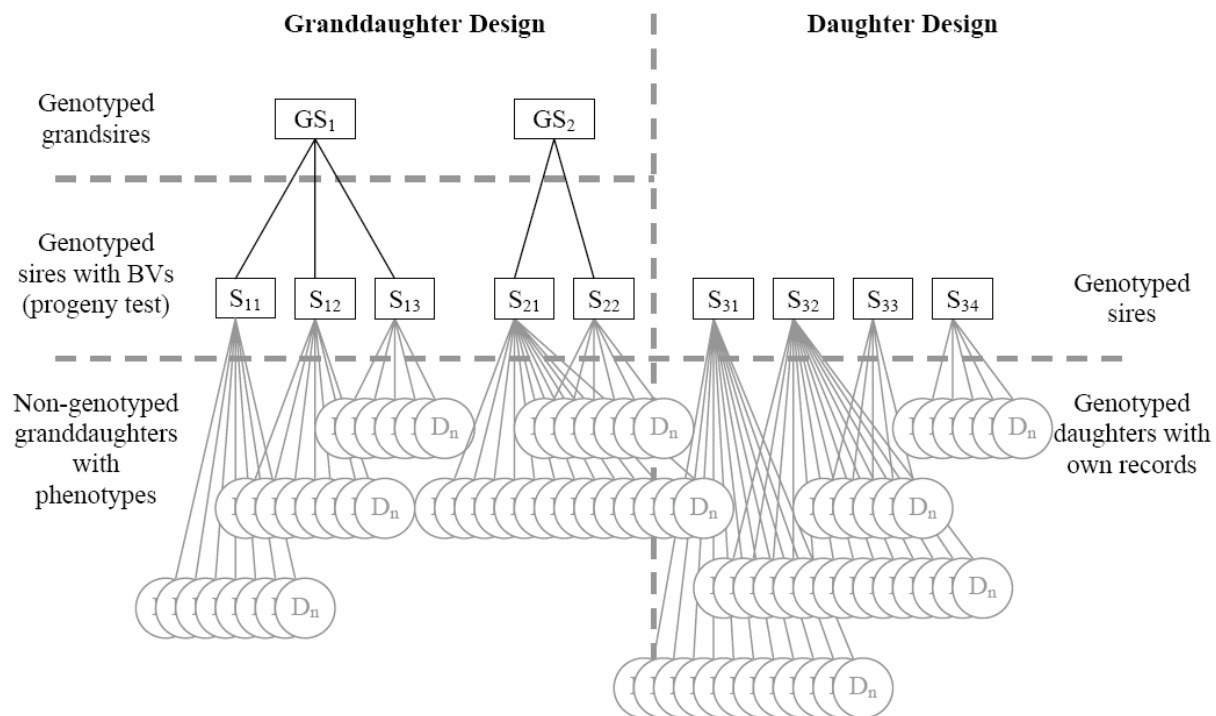

Fig. 1: The granddaughter design (left) is based on a series of paternal half brother pedigrees. Genotype information is available for grandsires and sires but not for granddaughters. Sire breeding values are estimated through progeny testing of at least 50 granddaughters. The daughter design (right) consists of large sets of paternal half sisters. (Das „Granddaughter Design" [links] basiert auf väterlichen Halbbrüdern. Genotypinformationen sind für Großväter und Väter, aber nicht für Enkelinnen vorhanden. Väterliche Zuchtwerte werden durch die Nachkommenprüfung von mindestens 50 Enkelinnen geschätzt. Das „Daughter Design" [rechts] besteht aus einer großen Anzahl von väterlichen Halbschwestern.)

Various statistical models can be utilized to refine QTL positioning in mapping experiments (Freyer et al. 2003), whereby both regression-based models (RM) and variance component models (VC) are usually preferred. Single-marker or multi-marker association studies have been documented (XU, 2003; GRAPES et al., 2004; Zhao et al., 2007). Although the RM method is robust, disequilibrium between specific haplotypes and the putative QTL may be larger than between single markers and QTL in some cases.

In VC analyses, a random additive allelic effect is estimated for every unique haplotype and a random polygenic effect for each animal is included to correct for correlations between polygenic and allelic coancestry. The covariance between the effects of any two haplotypes in a pedigree is assumed equivalent to the probability that the haplotypes have identical origin (identical by descent, IBD) and are included in the gametic relationship matrix **G** (FERNANDO and GROSSMAN, 1989). Founder haplotypes in LA (i.e. maternal haplotypes of sires and maternal and paternal haplotypes of grandsires in a granddaughter design) are considered unrelated and their pair-wise IBD probabilities (covariances) are set to zero. For fine mapping experiments, MEUWISSEN and GODDARD (2000, 2001) suggested incorporating historical recombination as a means of exploiting the non-random allelic association between QTL and closely linked markers (i.e. population-wide LD). The authors showed how to compute IBD sub-matrices containing IBD covariances between founder haplotypes conditional on marker data (see Figure 2). This method has consequently been extended to include linkage information (MEUWISSEN et al., 2002; BLOTT et al., 2003; OLSEN et al., 2004).
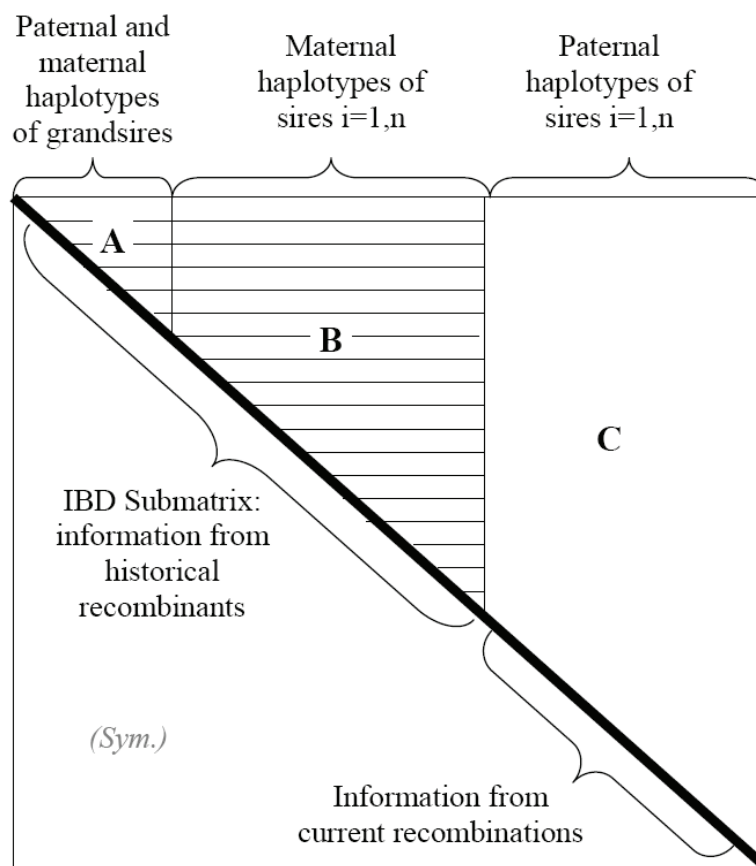

Fig. 2: Schematic representation of the gametic relationship matrix. Haplotypes are sorted according to generations with founder haplotypes to the left (blocks A and B) and haplotypes with known ancestors to the right (block C). The blocks marked with horizontal lines (A and B, IBD submatrix) are empty in a linkage analysis (i.e. elements = 0) and are dense for linkage disequilibrium and combined linkage/linkage disequilibrium experiments (i.e. elements ≠ 0). Block C is always dense. (Vereinfachte Darstellung der gametischen Verwand-schaftsmatrix. Haplotypen werden nach Generationen sortiert, wobei Gründerhaplotypen den Blöcken A und B und Haplotypen mit bekannten Vorfahren dem Block C zugeordnet sind.)

Variance components and individual haplotype effects can be estimated by incorporating the IBD sub-matrix into the gametic relationship matrix **G** and applying restricted maximum likelihood methods (REML). The disadvantage of this method is

that, as opposed to RM methods, certain assumptions regarding population structure must be made (i.e. number of generations since the mutation and effective population size of founder population, selection, etc.). More importantly, an accurate marker map is mandatory, especially with regard to the correct order of markers. Although the first few drafts of bovine linkage maps appeared as early as 1994 (i.e. BARDENSE et al., 1994; BISHOP et al., 1994; GEORGES et al., 1995, etc.), with later versions using improved techniques (i.e. KAPPES et al., 1997; IHARA et al., 2004; EVERTS-VAN DER WIND et al., 2005), gaps and areas of uncertainty in publicly available maps still persist.

A whole genome scan for QTL affecting functional traits including udder health was conducted in 2001 in cooperation with German AI and breeding organizations, scientific institutes for animal breeding and animal computing centers (THOMSEN et al., 2001). This project, instigated by the German cattle breeders' federation (ADR), enabled the identification of several QTL affecting functional traits in the German Holstein population, including those on Bos taurus autosomes 2, 18 and 27 affecting somatic cell score in milk (KÜHN et al., 2003; XU et al., 2006; KÜHN et al., 2008). Although extensive information was gained from these experiments, no candidate genes could be identified due to the relatively large confidence intervals typical of QTL mapping experiments. In 2005, a successor project (FUGATO M.A.S.-Net) was initiated with the goal of identifying the positional and functional candidate genes associated with SCS in the regions previously identified. In the context of this second project, a comprehensive software package was developed to process the genotypic data resulting from these efforts using VC methods. The aim of this paper is to present and describe the software system TIGER, which was designed for conducting LA, LD and combined LA/LD analyses for fine mapping experiments. The package is composed of both new and modified pre-existing Fortran programmes within a UNIX shell script. The system has been used for fine mapping on 4 distinct bos Taurus chromosomes.

Program sequence

The system operates in 6 main steps:

1 Allele frequencies and transmitting probabilities for each putative QTL position are determined.
2 The IBD sub-matrices for each putative QTL position are calculated and are stored.
3 The IBD sub-matrices are tested for positive definiteness; matrices with negative determinants are bent until all eigenvalues are positive. All IBD sub-matrices are then inverted and stored.

Finally, 3 loops are conducted over all putative QTL positions. The first loop conducts LA, the second LD analysis and the third and final loop provides the used with combined LA / LD analysis information, however the same programmes are used for each loop:

4 In the first loop, LA is conducted by setting up and inverting the condensed gametic relationship matrix (BAES et al., 2007; TUCHSCHERER et al., 2004) using transmitting probabilities calculated in step 1, merging the gametic index information with the phenotypes and running ASReml.

5   In the second loop, the IBD sub-matrices calculated in step 2 are incorperated into the gametic relationship matrix. Transmitting probabilities of non-founder gametes (i.e. values calculated in step 1) are set to 0.5 (LD). The gametic index information is then merged with the phenotypes and ASReml is run. In the third and final loop, the IBD sub-matrices calculated in step 2 are again incorporated into the condensed gametic relationship matrices and the transmitting probabilities of non-founder gametes are calculated as in step 1. As in steps 4 and 5, the gametic index information is then merged with the phenotypes and ASReml is run.

A schematic flowchart of the package is displayed in Figure 3. The program sequence also describes the acronym 'TIGER', which stands for **T**ransmitting probabilities, **I**BD sub-matrices, $\mathbf{G}^*$ matrix and **E**stimation of variance components using **R**EML. The following sections offer a more detailed description of each step.

Transmitting probabilities and allele frequencies

Although QTL genotypes are unobservable, markers linked to the putative QTL can be genotyped and used to infer the unknown QTL IBD status. The problem is that marker information is often incomplete. Unknown linkage phases, non-informative markers and missing marker genotypes complicate the calculation of transmitting probabilities, which refers to the probability that the sire passed his first (second) allele at a given locus to his offspring. These probabilities are required for calculation of $\mathbf{G}$, which is in turn required for solving the mixed model equations in VC analysis for LA, LD and LD/LA studies.

In typical fine mapping experiments, markers are dense. Furthermore, due to the size of the half sib families, reconstruction of haplotypes is fairly precise for both parents and progeny. Computation of probabilities for multi-marker haplotypes is implemented as described in REINSCH (2002); the program implemented was developed especially for use in designs in which 2 generations are genotyped (i.e. half sib families, backcross, etc.). Input files include genotype information, a short pedigree (only sires and grandsires) and a file containing marker and putative QTL positions. Once the input files are read in, genotypes of offspring are checked for compatibility with those of the parents using Mendelian laws of inheritance. The pedigree is tested for parent-offspring conflicts and the origin of progeny alleles is determined for each marker locus. The recombination frequencies for all intervals are calculated using Kosambi's map function. The most probable paternal and maternal haplotypes for parents are then reconstructed. Probabilities for linkage phase 1 are computed from genotypes and pedigree information. After all possible haplotypes of the sires are recorded, the most probable haplotype (linkage phase) produced by the candidates parents is calculated assuming haplotypes of the parents are known and correct. Apart from their use in the recursive calculation of $\mathbf{G}$, the transmitting probabilities also provide a measure of marker informativity. The program was further modified to estimate population-wide allele frequencies using all possible maternal haplotypes of offspring:

$$\text{frequency}(k) = \frac{1}{n} \sum_{ij} \delta_{ij}^k \cdot h_{ij}$$

where: k = a given marker allele, n = the number of final progeny, δ = an indicator variable equal to 1 if allele k is present in a given haplotype and zero otherwise and h = the probability of haplotype j occurring for animal i. The most probable maternal haplotypes are also saved to file for use in calculation of the IBD sub-matrices.
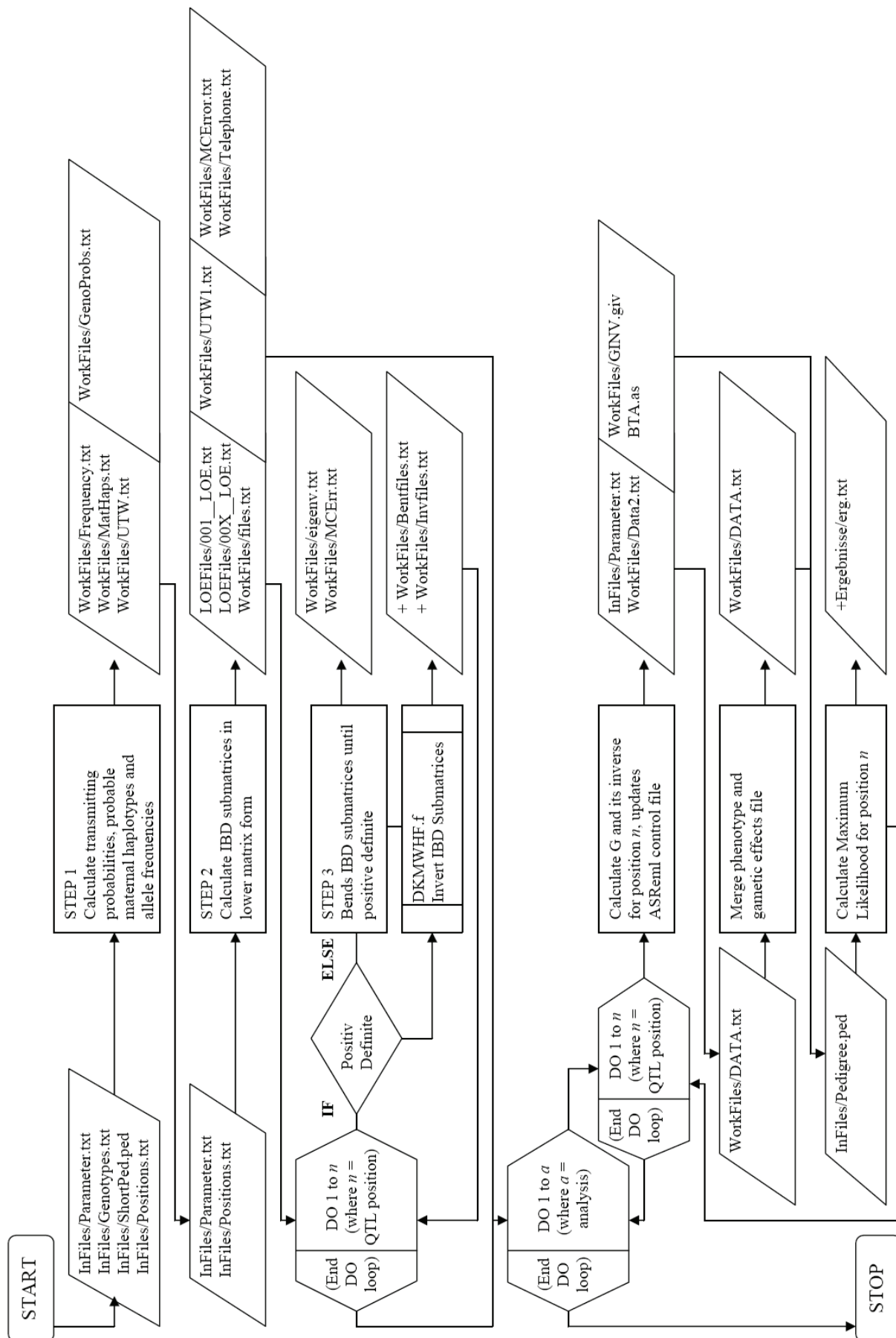
Fig. 3: Schematic representation of the TIGER system (Vereinfachte Darstellung des TIGER-Systems)

Calculation of the IBD sub-matrices
MEUWISSEN and GODDARD (2000) described a method to calculate the IBD sub-matrix based on deterministic predictions which takes the number of markers flanking the putative QTL position, the extent of LD in the population based on the expectation under finite population size, and the time point of mutation occurrence into account. A Fortran (90/95) program based on the gene dropping procedure described by MEUWISSEN and GODDARD (2000, 2001) was developed and incorporated into the TIGER system. The method we employed involves setting up IBD sub-matrices for the maternal haplotypes of the sires and both maternal and paternal haplotypes of the grandsires in the real data set being examined. Parameter options allow the user to define the effective population size, the number of generations of gene-dropping, the number of repetitions and the mapping function to be used. The number of markers left and right of the putative QTL can also be user-defined. GRAPES et al. (2006) found that fitting a 4-6 marker haplotype as a sliding window across the region resulted in the greatest accuracy compared to that of the other haplotype sizes tested. However for practical applications it may be advantageous to define the size of the 'window' surrounding the putative QTL in cM instead of the number of markers. This allows for more flexibility when the marker distribution over the chromosome is not constant. Haplotypes were sorted into groups according to the number of identical markers to the left and to the right of the putative QTL.

The method and basic principles of gene dropping are relatively simple. Since each parent possesses two alleles at a given locus and since only one of these alleles will be passed on to an offspring, each allele has a one in two chance of being transmitted to each descendant, this probability being independent of the heritage of the other parent. In other words, this is quite simply a simulation of the Mendelian transmission of a gene at a given locus. The method can be extended to the case of several loci, thus making it possible to measure the change in linkage disequilibrium. Application of this method, however, rests on certain assumptions. Selection at the simulated locus is not taken into account, since we are seeking only to replicate the effect of change due to the Mendelian transmission of genes. In reality, besides random variations, allele frequencies are subject to evolutionary forces such as migration, mutation, and natural selection, which produce specific and non-random variation in allele frequencies. The effect of selection can be considered negligible; however, it is important to realize that it is not accounted for here.

One IBD sub-matrix is calculated for every putative QTL locus given by the user. These matrices are stored individually and are dense. The average inbreeding coefficient within simulated generations and the Monte Carlo errors of each haplotype constellation are saved to file for control purposes.

Positive definiteness of IBD coefficient matrices
The numeric instability of the IBD sub-matrices was mentioned but not discussed in depth by MEUWISSEN and GODDARD (2000). The IBD coefficient sub-matrices in LD and LA/LD analyses are very dense because all founder alleles (paternal and maternal alleles of grandsires and maternal alleles of sires) are assigned covariances. Furthermore, these matrices are not necessarily positive definite, which impedes the inversion step required for their use as a part of the gametic relationship matrix. However, IBD coefficient matrices with determinants less than one can be 'bent' until

the lowest eigenvalue exceeds a preset limit (i.e. zero to achieve positive definiteness) (HAYES and HILL, 1981; ESSL 1991).

We adopted the BENDPDF program of ESSL (1991) and added the matrix inversion subroutine DKMXHF of MEYER (2008, personal communication) to ensure positive definiteness of the IBD sub-matrices, proper calculation of matrix determinants and correct inversion. The numbers of negative eigenvalues in each sub-matrix as well as the matrix determinants are saved to file for control purposes. In our experience, approximately 20 % of all IBD sub-matrices within must be bent; the number of negative eigenvalues in our submatrices normally ranged from from 0 to 4.5 % per matrix.

The G matrix

Because allelic (gametic) effects are random, their covariance must be included for estimation in a variance component model. This information is included in the gametic relationship matrix **G**. TUCHSCHERER et al. (2004) showed how to eliminate possible linear dependencies in this matrix by considering haplotypes with extremely high or extremely low transmitting probabilities identical to those of the respective parent. If transmission probabilities very close to one or zero occur, a predefined threshold (e.g. $\varepsilon = 0.03$) can be given and all transmission probabilities below 0.03 are 'condensed' to 0 and all those above 0.97 to 1. Haplotypes corresponding to such transmission probabilities are considered 'non-unique'. The resulting condensed gametic relationship matrix, $\mathbf{G}^*$, can have between n and 2n rows and columns, depending on how many unique gametes occur, whereby n is the number of genotyped animals in the pedigree. This method prevents numerical difficulties due to quasi-linear dependencies which may occur due to the inclusion of almost identical rows and columns in the gametic relationship matrix.

The condensed covariance matrix therefore contains one row and one column for every unique gamete (haplotype) of genotyped animals and is sorted according to gametic generation (i.e. founder gametes followed by gametes with known ancestors). If only grandsires and sires are genotyped, all maternal gametes of sires are considered founders and therefore unrelated; the pairwise IBD probabilities of these alleles are zero. COBRA (BAES and REINSCH, 2007) was designed to calculate the condensed matrix $\mathbf{G}^*$ based on the algorithm of TUCHSCHERER et al. (2004). The COBRA program was slightly modified to include IBD sub-matrices described above for LD and combined LA/LD analyses (see Figure 2 for the arrangement of haplotype effects in $\mathbf{G}^*$). An ASReml parameter file corresponding to the data on the given marker interval (i.e. number of QTL effects in the $\mathbf{G}^*$ matrix, etc.) is created for each position.

Likelihood calculation

Variance component analyses for continuous traits are normally based on the mixed model, and maximum likelihood or related methods of inference are employed. In particular, the so-called residual or restricted maximum likelihood (REML) is widely used for analyses of continuous traits. The popularity of mixed model analyses by REML has been furthered by the availability of appropriate software. Although implemented in general statistical packages, specific programs for VC estimation in an animal breeding context have been developed (GILMOUR et al., 2006; KOVACS and GROENEVELD, 2003; MADSEN and JENSEN, 2006). These programs use a mixture of Fisher's scoring and Newton-Raphson to maximize the restricted likelihood,

called the average information restricted maximum-likelihood (AI-REML) algorithm. Due to the somewhat delicate procedure required for setting up and inverting $\mathbf{G}^*$, a program allowing user-defined covariance structures and their inverses was required. This prerequisite was filled by ASReml (GILMOUR et al., 2006). The following mixed model equations were applied:

$$\begin{bmatrix} \mathbf{X'D^{-1}X} & \mathbf{X'D^{-1}Z} & \mathbf{X'D^{-1}W} \\ \mathbf{Z'D^{-1}X} & \mathbf{Z'D^{-1}Z} + \mathbf{A}_u^{-1} & \mathbf{Z'D^{-1}W} \\ \mathbf{W'D^{-1}X} & \mathbf{W'D^{-1}Z} & \mathbf{W'D^{-1}W} + \mathbf{G}_v^{*-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'D^{-1}y} \\ \mathbf{Z'D^{-1}y} \\ \mathbf{W'D^{-1}y} \end{bmatrix}$$

Where $y_{(m\times 1)}$ = a vector of m DYDs for n animals, μ = a fixed effect common to all observations, $u_{(n\times 1)}$ = a vector of random polygenic effects and $v_{(gam\times 1)}$ = a vector of random gametic effects of a marked QTL that is linked to a single polymorphic marker locus with gam equal to the number of unique gametes. Subscripts in parenthesis of the vectors and matrices denote their dimensions. $X_{(m\times 1)}$, $Z_{(m\times n)}$ and $W_{(m\times gam)}$ = known incidence matrices, $\mathbf{A}_{u(n\times n)}^{-1}$ = the inverse of the relationship matrix and $\mathbf{G}_{v(gam\times gam)}^{*-1}$ = the inverse of the condensed gametic QTL relationship matrix. D^{-1} = the inverse of a diagonal matrix containing weights for each observation. Expectations of u, v and e and covariances between them are assumed to be zero. The model depicted above contains only the fixed effect μ, however LD can easily be included by regression on one or more markers instead of through the IBD sub-matrix. The number of regression coefficients is then the number of different marker alleles of the given marker(s).

Conclusion

A comprehensive software package was developed to process and analyse the genotypic data within the context of the FUGATO M.A.S.-Net project using VC methods. This paper presented the software system TIGER, which was designed for conducting linkage, linkage disequilibrium and combined linkage/linkage disequilibrium analyses for fine mapping experiments. The TIGER system employs both new and modified pre-existing Fortran programmes within a UNIX shell script and is an 'all in one' solution for analysing genotypic data.

Acknowledgements

References

ABECASIS, G.R.; CHERNY, S.S.; COOKSON, W.O.; CARDON, L.R.:
Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nat. Genet. **30** (2002), 97-101
ANDERSON, S.; DE BRUIJN, M.H.; COULSON, A.R.; EPERON, I.; SANGER, F.; YOUNG, I.G.:
Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J. Mol. Biol. **156** (1982) 4, 683-717
BAES, C.; REINSCH, N.:
Computing the condensed conditional gametic QTL relationship matrix and its inverse. Arch. Tierz. **50** (2007) 3, 294-308
BARENDSE, W.; ARMITAGE, S.M.; KOSSAREK, L.M.; SHALOM, A.; KIRKPATRICK, B.W.; RYAN, A.M.; CLAYTON, D.; LI, L.; NEIBERGS, H.L.; ZHANG, N.; GROSSE, W.M.; WEISS, J.; CREIGHTON, P.; MCCARTHY, F.; RON, M.; TEALE, A.J.; FRIES, R.; MCGRAW, R.A.; MOORSE, S.S.; GEORGES, M.; SOLLER, M.; WOMACK, J.E.; HETZEL, D.J.S.:
A genetic linkage map of the bovine genome. Nature Genet. **6** (1994), 227-235

BISHOP, M.D.; KAPPES, S.M.; KEELE, J.W.; STONE, R.T.; SUNDEN S.L.F.; HAWKINS, G.A.; TOLDO, S.S.; FRIES, R.; GROSZ, M.D.; YOO, J.; BEATTIE, C.W.:
> A genetic linkage map for cattle. Genetics **136** (1994), 619-639

BLOTT, S.; KIM, J.J.; MOISIO, S.; SCHMIDT-KÜNTZEL, A.; CORNET, A.; BERZI, P.; CAMBIASO, N.; FORD, C.; GRISART, B.; JOHNSON, D.; KARIM, L.; SIMON, P.; SNELL, R.; SPELMAN, R.; WONG, J.; VILKKI, J.; GEORGES, M.; FARNIR, F.; COPPIETERS, W.:
> Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics **163** (2003), 253-266

COPPIETERS, W.; BLOTT, S.; FARNIR, F.; GRISART, B.; RIQUET, J.; GEORGES, M.:
> From phenotype to genotype: towards positional cloning of QTL in livestock? Arch. Tierz. **42** (1999) Special Issue, 86-92

DARVASI, A.; SOLLER, M.:
> Advanced Intercross Lines, an Experimental Population for Fine Genetic Mapping. Genetics **141** (1995) 3, 1199-1207

DARVASI, A.:
> Experimental strategies for the genetic dissection of complex traits in animal models. Nat. Genet. **18** (1998) 1, 19-24

ESSL, A.:
> Choice of an appropriate bending factor using prior knowledge of the parameters. J. Anim. Breed. Gen. **108** (1991), 89-101

EVERTS-VAN DER WIND, A.; LARKIN, D.M.; GREEN, C.A.; ELLIOTT, J.S.; OLMSTEAD, C.A.; CHIU, R.; SCHEIN, J.E.; MARRA, M.A.; WOMACK, J.E.; LEWIN, H.A.:
> A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. Proc. Natl. Acad. Sci. USA, **102** (2005) 51, 18526-18531

FERNANDO, R.L.; GROSSMAN, M.:
> Marker-assisted selection using best linear unbiased prediction. Genet. Sel. Evol. **21** (1989), 467-477

FREYER, G.; KÜHN, C.; WEIKHARD, R.:
> Comparison of different statistical-genetic approaches of QTL detection by evaluating results from a real dairy cattle data set. Arch. Tierz. **46** (2003), 413-423

GEORGES, M.; NIELSEN, D.; MACKINNON, M.; MISHRA, A.; OKIMOTO, R.; PASQUINO, A.T.; SARGEANT L.S.; SORENSEN, A.; STEELE, M.R.; ZHAO, X.; WOMACK, J.E.; HOESCHELE, I.:
> Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics **139** (1995), 907-920

GEORGES, M.:
> Mapping, Fine Mapping, and Molecular Dissection of Quantitative Trait Loci in Domestic Animals. Annu. Rev. Genomics Hum. Genet. **8** (2007), 131-162

GILMOUR, A.; GOGEL, B.; CULLIS, M.; THOMPSON, R.:
> ASReml User Guide Release 2.0. (2006) VSN International Ltd., Hemel Hempstead, HP1 1ES, UK

GRAPES, L.; DEKKERS, J.C.M.; ROTHSCHILD, M.F.; FERNANDO, R.L.:
> Comparing linkage disequilibrium-based methods for fine mapping quantitative trait loci. Genetics **166** (2004), 1561-1570

GRAPES, L.; FIRAT, M.Z.; DEKKERS, J.C.M.; ROTHSCHILD, M.F.; FERNANDO, R.L.:
> Optimal haplotype structure for linkage disequilibrium-based fine mapping of quantitative trait loci using identity by descent. Genetics **172** (2006), 1955-1965

HAYES, J.F.; HILL, W.G.:
> Modification of Estimates of Parameters in the construction of Genetic Selection Indices ('Bending'). Biometrics **37** (1981), 483-493

IHARA, N.; TAKASUGA, A.; MIZOSHITA, K; TAKEDA, H.; SUGIMOTO, M.; MIZOGUCHI, Y.; HIRANO, T.; ITOH ,T.; WATANABE, T.; REED, K.M.; SNELLING, W.M.; KAPPES, S.M.; BEATTIE, C.W.; BENNETT, G.L.; SUGIMOTO, Y.:
> A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res. **14** (2004) 10A, 1987-1998

KAPPES, S.M.; KEELE, J.W.; STONE, R.T.; MCGRAW, R.A.; SONSTEGARD, T.S.; SMITH, T.P.L.; LOPEZ-CORRALES, N.L.; BEATTIE, C.W.:
> A second-generation linkage map of the bovine genome. Genome Res. **7** (1997) 3, 235-249

KOVACS, M.A.; GROENEVELD, E.:
> VCE-5 User's Guide and Reference Manual, Version 5.1. (2003) Mariensee, Germany: Institute of Animal Husbandry and Animal Behaviour, Federal Research Center of Agriculture

KÜHN, C.; BENNEWITZ, J.; REINSCH, N.; XU, N.; THOMSEN, H.; LOOFT, C.; BROCMANN, G.A.;
SCHWERIN, M.; WEIMANN, C.; HIENDLEDER, S.; ERHARDT, G.; MEDJUGORAC, I.; FORSTER, M.;
BRENIG, B.; REINHARDT, F.; REENTS, R.; RUSS, I.; AVERDUNK, G.; BLUMEL, J.; KALM, E.:
    Quantitative trait loci mapping of functional traits in the German Holstein cattle population. J. Dairy
    Sci. **86** (2003), 360-368
KÜHN, C.; REINHARDT, F.; SCHWERIN, M.:
    Marker assisted selection of heifers improved milk somatic cell count compared to selection on
    conventional pedigree breeding values. Arch. Tierz. **51** (2008), 23-32
MEUWISSEN, T.; GODDARD, M.:
    Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci.
    Genetics **155** (2000), 421-430
MEUWISSEN, T.; GODDARD, M.:
    Prediction of identity by descent probabilities from marker haplotypes. Genet. Sel. Evol. **33** (2001), 605-634
MEUWISSEN, T.; KARLSEN, A.; LIEN, S.; OLSAKER, I.; GODDARD, M.:
    Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage
    disequilibrium mapping. Genetics **161** (2002), 373-379
OLSEN, H.G.; LIEN, S.; SVENDSEN, M.; NILSEN, H.; ROSETH, A.; AASLAND OPSAL, M.;
MEUWISSEN, T.H.E.:
    Fine Mapping of Milk Production QTL on BTA6 by Combined Linkage and Linkage Disequilibrium
    Analysis. J. Dairy Sci. **87** (2004), 690-698
REINSCH, N.:
    A multiple-species, multiple-project database for genotypes at codominant loci. J. Anim. Breed. Genet.
    116 (1999), 425-435
REINSCH, N.:
    A general likelihood approach to trait-based multipoint linkage analysis in large groups of half sibs and
    super sisters. Genetics **162** (2002), 413-424
SOBEL, E.; LANGE, K.:
    Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing
    statistics. Am. J. Hum. Genet. **58** (1996), 1323-1337
THOMSEN, H.; REINSCH, N.; XU, N.; BENNEWITZ, J.; LOOFT, C.; GRUPE, S.; KUHN, C.; BROCKMANN,
G.A.; SCHWERIN, M.; LEYHE-HORN, B.; HIENDLEDER, S.; ERHARDT, G.; MEDJUGORAC, I.; RUSS, I.;
FORSTER, M.; BRENIG, B.; REINHARDT, F.; REENTS, R.; BLUMEL, J.; AVERDUNK, G.; KALM, E.:
    A whole genome scan for differences in recombination rates among three Bos taurus breeds. Mamm
    Genome. **12** (2001), 724-728
TUCHSCHERER, A.; MAYER, M.; REINSCH, N.:
    Identification of gametes and treatment of linear dependencies in the gametic QTL-relationship matrix
    and its   inverse. Genet. Sel. Evol. **36** (2004), 621-642
WELLER, J.I.; KASHI, Y.; SOLLER, M.:
    Power of daughter and granddaughter designs for determining linkage between marker loci and
    quantitative trait loci in dairy cattle. J. Dairy Sci. **73** (1990), 2525-2537
XU, S.:
    Estimating polygenic effects using markers of the entire genome. Genetics **163** (2003), 789-801
XU N.; PAUL, S.; BENNEWITZ, J.; REINSCH, N.; THALLER, G.; REINHARDT, F.; KÜHN, CH.;
SCHWERIN, M.; ERHARDT, G.; WEIMANN, C.; THOMSEN, H.; MISCHRA, S. AND KALM, E.:
    Confirmation of quantitative trait loci for somatic cell score on bovine chromosome18 in the German
    Holstein. Arch.Tierz. **49** (2006), 111-119
ZHAO, H.H.; FERNANDO, R.L.; DEKKERS, J.C.M.:
    Power and Precision of Alternate Methods for Linkage Disequilibrium Mapping of Quantitative Trait
    Loci. Genetics **175** (2007), 1975-1986

Authors:
CHRISTINE BAES
Prof. Dr. NORBERT REINSCH*
Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere (FBN)
Forschungsbereich Genetik und Biometrie
Wilhelm-Stahl-Allee 2
18196 Dummerstorf
Germany

*Corresponding author
email: reinsch@fbn-dummerstorf.de