

Vergleich von drei Methoden für genomweite Assoziationsstudien in selektierten Populationen

MALENA ERBE¹, FLORENCE YTOURNEL¹, E.C.G. PIMENTEL¹, A.R. SHARIFI¹ und H. SIMIANER¹

Zusammenfassung

Selektion kann die Struktur des Kopplungsungleichgewichts in Nutztierpopulationen beeinflussen. Assoziationen zwischen physisch nicht gekoppelten Loci können dadurch zu einer erhöhten Anzahl an falsch positiven Signalen bei genomweiten Assoziationsstudien führen. In dieser Arbeit wurden drei Methoden verglichen, die für Assoziationsstudien genutzt werden können: Single Marker Regression (SMR), eine zweistufige Methode (GRAMMAR), bei der um Familieneffekte korrigierte Residuen in einer SMR analysiert werden, und eine Methode (MTDT), bei der geschätzte Mendelian sampling Effekte als phänotypische Information in einem quantitativen Transmission Disequilibrium Test verwendet werden. Drei verschiedene Szenarien wurden simuliert: idealisierte Zufallspaarung, Einschränkung der Anzahl an Elterntieren und gerichtete Selektion. Die Anzahl falsch positiver Signale erhöhte sich im Vergleich zu der idealisierten Population bei allen Methoden in den Selektionsszenarien. Da die SMR eine hohe Anzahl falsch positiver Signale in Populationen mit eingeschränkter Populationsgröße produziert, sind Ergebnisse von genomweiten Assoziationsstudien mit dieser Methode in Nutztierpopulationen mit Vorsicht zu betrachten. Mit GRAMMAR und MTDT erhält man weniger falsch positive Signale als mit SMR. MTDT verbindet eine ausreichende Power mit einer akzeptablen Rate an falsch positiven Signalen. Bei GRAMMAR hingegen war die Power sehr niedrig. Die Bonferroni-korrigierte Signifikanzgrenze schien für diese Methode zu konservativ zu sein.

Schlüsselwörter: Genomweite Assoziationsstudien, SNPs, Simulation, Selektion

Summary

Comparison of three whole genome association mapping approaches in selected populations

Selection is known to influence the linkage disequilibrium (LD) pattern in livestock populations. Spurious LD may lead to a higher number of false positive signals in whole genome association mapping experiments. We compared three approaches for whole genome association mapping in a simulation study: single marker regression (SMR), a two-step approach, which analyzes residuals corrected for family effects with an SMR (GRAMMAR), and a combined linkage and LD approach, which applies the quantitative transmission disequilibrium test to the Mendelian sampling term (MTDT). Three different scenarios were simulated: idealized random mating, limited number of parents and directional selection. The number of false positive associations increased when the number of parents was limited. Since SMR produced a high number of false positive signals in small populations, results of whole genome scans in livestock analyzed with SMR

¹ Abteilung Tierzucht und Haustiergenetik, Department für Nutztierwissenschaften, Georg-August-Universität Göttingen, Albrecht-Thaer-Weg 3, 37075 Göttingen, E-Mail: merbe@gwdg.de

should be considered with caution. GRAMMAR was the most accurate, but also the least powerful approach. The Bonferroni corrected significance threshold seemed to be too stringent for this approach. Results obtained with MTDT changed only slightly with selected populations. MTDT combined sufficient power with a manageable number of false positive associations in all scenarios.

Keywords: Whole genome association study, SNPs, simulation, selection

1 Einleitung

Kopplungsungleichgewicht zwischen Genorten ist die Voraussetzung für die Identifizierung von QTL-Positionen in genomweiten Assoziationsstudien. In einer hochdichten Markerkarte besteht die Möglichkeit, dass ein QTL mit mindestens einem der ihm nahe liegenden Marker in hohem Kopplungsungleichgewicht steht und die Position des QTLs durch diese Assoziation entdeckt werden kann. Auch wenn der Begriff „Kopplungs-Ungleichgewicht“ ein Ungleichgewicht zwischen zwei gekoppelten Loci suggeriert, ist dies nicht immer der Fall – es kann auch zwischen physisch nicht gekoppelten Genorten existieren. Assoziationen zwischen nicht gekoppelten Loci beeinflussen die Ergebnisse von genomweiten Assoziationsstudien negativ, da signifikante Signale weit entfernt von einer realen QTL-Position auftreten können. Da normalerweise kein Vorwissen über die möglichen QTL-Positionen besteht, ist es nicht möglich, einfach zu entscheiden, ob ein signifikanter SNP eine korrekt oder eine falsch positive Assoziation widerspiegelt.

Nutztierpopulationen sind häufig einer starken gerichteten Selektion unterworfen. Diese reduziert die effektive Populationsgröße, wodurch die Struktur des Kopplungsungleichgewichts beeinflusst wird (TERWILLIGER et al., 1998). Die beschriebenen falsch positiven Assoziationen sind vor allem in solchen Populationen zu erwarten. Umso wichtiger ist es, dass Methoden, die für genomweite Assoziationsstudien im Nutztierbereich verwendet werden, falsch positive Signale vermeiden können. Für die Kartierung von QTLs in Assoziationsstudien wurden verschiedenste Ansätze publiziert und sowohl in simulierten als auch in realen Datensätzen angewendet. Darunter befinden sich Methoden, die ausschließlich auf Kopplungsungleichgewicht aufbauen (z.B. Regressionsmethoden: GRAPES et al., 2004; ZHAO et al., 2007) wie auch Methoden, die neben Kopplungsungleichgewicht auch Informationen über Kopplung berücksichtigen (z.B. IBD-basierte Ansätze (MEUWISSEN und GODDARD, 2000) oder TDT-basierte Methoden (z.B. RABINOWITZ, 1997; ABECASIS et al., 2000)). Methoden, die für einen ersten Scan des Genoms verwendet werden, sollten eine ausreichende Power mit einer akzeptablen Anzahl an falsch positiven Signalen kombinieren. Weiterhin sollten sie mit möglichst wenig Rechenaufwand auch in großen Datensätzen leicht anwendbar sein.

In dieser Arbeit sollen drei Methoden, die für Assoziationsstudien verwendet werden können, hinsichtlich ihrer Power und der Anzahl an falsch positiven Signalen in drei verschiedenen Simulationsszenarien evaluiert werden.

2 Material und Methoden

Simulationen

Ein Überblick über die Simulationen ist in Abb. 1 zu finden.

Über 1000 Generationen wurde eine zufallspaarende Population ($N = 1000$) simuliert. Kopplungsungleichgewicht konnte dabei durch genetische Drift und Mutation entstehen.

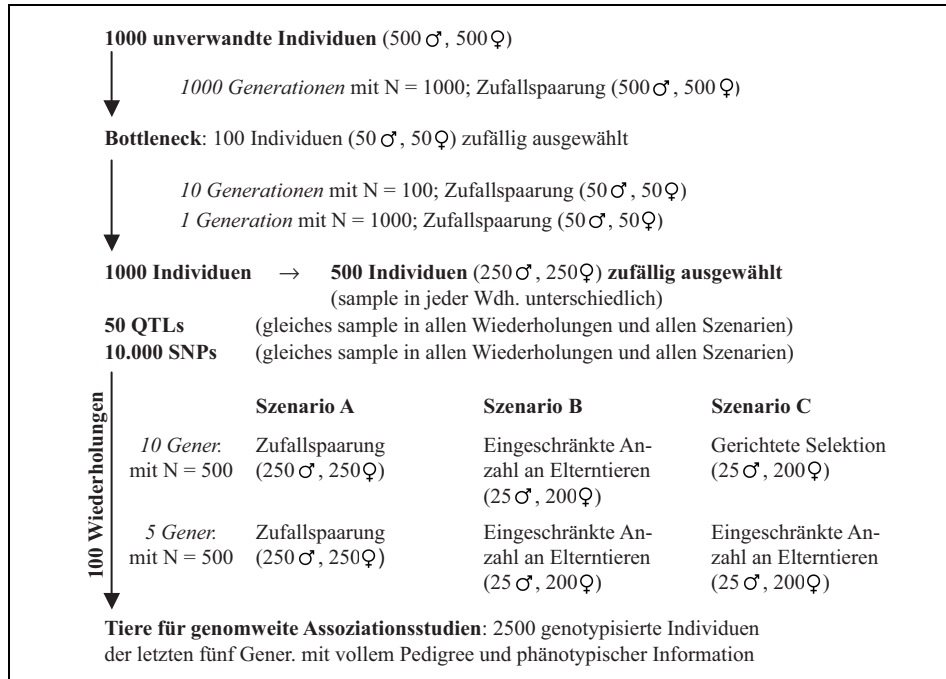


Abb. 1. Überblick über die Simulationszenarien.
Overview over the simulation scenarios.

Nach 1000 Generationen führte ein genetischer Flaschenhals zu einer Verkleinerung der Populationsgröße (N = 100). Für Generation 1011 wurden schließlich 1000 Individuen durch Zufallspaarung von 50 Vätern und 50 Müttern produziert. Aus dieser „founder generation“ wurden für die folgenden Simulationsschritte in jeder Wiederholung (n = 100) 250 männliche und 250 weibliche Tiere zufällig ausgewählt. Ab diesem Punkt wurden für weitere 15 Generationen in jeder Wiederholung je drei Szenarien simuliert.

Szenario A stellte eine idealisierte Population mit vollständiger Zufallspaarung dar, während in Szenario B die Anzahl der potentiellen Elterntiere für die nächste Generation begrenzt war. Da die Elterntiere in Szenario B zufällig ausgewählt wurden, könnte dieses Szenario eine Population darstellen, in der Selektion vorherrscht, jedoch nicht auf das beobachtete Merkmal. In Szenario C wurden die Elterntiere, deren Anzahl auch hier beschränkt war, auf Grund eines geschätzten Zuchtwertes für das beobachtete Merkmal ausgewählt. Um die Heritabilität in den letzten Generationen, die für die genomweiten Assoziationsstudien verwendet wurden, weitgehend konstant zu halten, wurde auch in Szenario C in diesen Generationen nicht mehr gerichtet selektiert, sondern es wurde nur noch die Anzahl an Elterntieren eingeschränkt. Alle Tiere, die genotypisiert waren, die eine phänotypische Beobachtung vorweisen konnten und deren Eltern genotypisiert waren und eine phänotypische Beobachtung aufwiesen, wurden für die genomweiten Assoziationsstudien verwendet. Dies waren in unserer Simulation alle Tiere der letzten fünf Generationen, d.h. 2500 Tiere. Die Assoziationsanalysen wurden in jedem Szenario in jeder Wiederholung durchgeführt.

Ab Beginn der Simulation wurden 60.000 gleichmäßig verteilte biallelische SNPs auf zehn Chromosomen, die jeweils 1 Morgan lang waren, simuliert. In Generation 1011

wurden 50 SNPs zufällig ausgewählt, um ab dieser Generation als QTL zu wirken. Auf Chromosom 10 wurde kein QTL simuliert. Die QTLs hatten nur additive Effekte; Dominanz und Epistasie wurden nicht simuliert. Die Effekte wurden zufällig aus einer Gammaverteilung gezogen, wobei jeweils ein QTL-Allel die Hälfte des aus der Gamma-Verteilung erhaltenen Wertes als additiven Effekt und das zweite QTL-Allel den gleichen Effekt, aber mit anderem Vorzeichen erhielt. Aus den 59.950 verbleibenden SNPs wurden auch in Generation 1011 je Chromosom 1000 SNPs zufällig ausgewählt, nachdem alle SNPs mit einer Frequenz des selteneren Allels von unter 5% entfernt worden waren. Diese 10.000 SNPs wurden dann für die Assoziationsanalysen verwendet.

Die additiv-genetische Varianz wurde als Summe der additiv-genetischen Varianzen definiert, die durch die QTLs erklärt wurden. Die additiv-genetische Varianz, die durch einen QTL verursacht wurde, wurde als $\sigma_a^2 = 2pqa^2$ definiert, wobei p und q die Allelfrequenzen von Allel 1 bzw. 2 darstellen und a den Allelsubstitutionseffekt (FALCONER und MACKAY, 1996). Der wahre Zuchtwert eines Tieres entsprach der Summe seiner Alleleffekte. Für die Selektion in Szenario C wurde ein geschätzter Zuchtwert aus wahren Zuchtwert und einem zufälligen Fehler $x \sim N(0, \sigma_x^2)$ berechnet. Die Varianz des Fehlers wurde so bestimmt, dass die Genauigkeit der Zuchtwertschätzung in Generation 1011 0,8 für die männlichen Tiere und 0,6 für die weiblichen Tiere betrug. Um die Phänotypen, die für die genomweiten Assoziationsanalysen verwendet wurden, zu simulieren, wurde zum wahren Zuchtwert ein zufälliger umweltbedingter Fehler $e \sim N(0, \sigma_e^2)$ addiert; σ_e^2 wurde so festgesetzt, dass die Heritabilität in Generation 1021 0,3 betrug.

Methoden

Single Marker Regression (SMR):

Es wurde folgendes Modell verwendet:

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \boldsymbol{\varepsilon}$$

wobei \mathbf{y} ein Vektor mit den beobachteten Phänotypen ist, \mathbf{W} die Inzidenzmatrix für den Intercept und den SNP-Effekt – codiert als 0 für Genotyp 11, 1 für Genotyp 12 und 2 für Genotyp 22 –, \mathbf{g} ein Vektor, der den Intercept und den Regressionskoeffizienten enthält, und $\boldsymbol{\varepsilon}$ ein Vektor, der die zufälligen Restfehler enthält. Die SMR wurde für jeden der 10.000 SNPs durchgeführt. In diesem Modell wurden Populations- und Familienstrukturen nicht berücksichtigt. SMR in dieser Form wurde z.B. von GRAPES et al. (2004) oder ZHAO et al. (2007) in Simulationsstudien verwendet.

Genomewide Rapid Association Using Mixed Model and Regression (GRAMMAR) (AULCHENKO et al., 2007):

GRAMMAR ist eine zweistufige Methode, die Pedigreeinformationen berücksichtigt. Zuerst wird folgendes gemischte Modell gelöst:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

wobei \mathbf{y} ein Vektor der beobachteten Phänotypen ist, \mathbf{X} die Inzidenzmatrix für die fixen Effekte, \mathbf{b} ein Vektor der Koeffizienten für die fixen Effekte, \mathbf{Z} die Inzidenzmatrix für den zufälligen additiven polygenen Effekt, \mathbf{u} ein Vektor, der den additiven polygenen Effekt jedes Tieres enthält, also dessen Zuchtwert, und \mathbf{e} ein Vektor mit den zufälligen Resteffekten. In unserer Studie wurde nur die Generation als fixer Effekt ins Modell genommen. Die Varianz-Kovarianz-Struktur der additiv-polygenen Effekte entsprach $A\sigma_u^2$, wobei A die pedigreebasierte Verwandtschaftsmatrix ist.

Aus den Ergebnissen des gemischten linearen Modells werden als nächstes geschätzte Residuen abgeleitet:

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}} - \mathbf{Z}\hat{\mathbf{u}}$$

Im zweiten Schritt wird eine lineare Regression durchgeführt. Das Modell ist ähnlich dem in der SMR, jedoch wird statt des Phänotyps der geschätzte Residualterm als abhängige Variable verwendet. Das Modell ist

$$\hat{\mathbf{e}} = \mathbf{W}\mathbf{g} + \boldsymbol{\varepsilon}$$

wobei $\hat{\mathbf{e}}$ ein Vektor mit den zuvor geschätzten Residuen ist, \mathbf{W} die Inzidenzmatrix für den Intercept und den SNP-Effekt, \mathbf{g} ein Vektor, der den Intercept und den Regressionskoeffizienten enthält, und $\boldsymbol{\varepsilon}$ ein Vektor der zufälligen Restfehler.

Applying a Quantitative Transmission Disequilibrium Test to the estimated Mendelian Sampling Term (MTDT) (SIMIANER und PIMENTEL, 2009):

Auch diese Methode ist zweistufig. Zuerst werden BLUP-Zuchtwerte geschätzt, damit aus ihnen die Mendelian sampling terms für jedes Tier geschätzt werden können. Das gemischte Modell entspricht demjenigen im ersten Schritt bei GRAMMAR. Der Mendelian sampling term \hat{m}_i eines Tieres i ist definiert als

$$\hat{m}_i = \hat{u}_i - 0,5 (\hat{u}_j + \hat{u}_k)$$

wobei \hat{u}_i der geschätzte Zuchtwert von Tier i ist und \hat{u}_j, \hat{u}_k die geschätzten Zuchtwerte seiner Eltern sind.

Die Idee von Transmission Disequilibrium Test (TDT)-basierten Methoden ist es, die Übertragung von Allelen von Eltern auf ihre Nachkommen zu beobachten und nachzuvollziehen und die Gruppe der Nachkommen, die Allel 1 und nicht Allel 2 erhalten haben, mit der Gruppe der Nachkommen zu vergleichen, die Allel 2 und nicht Allel 1 erhalten haben. Im Fall des MTDT werden die Differenzen zwischen den mittleren geschätzten Mendelian sampling terms der beiden Gruppen analysiert. Die Differenz sollte sich nur dann signifikant von Null unterscheiden, wenn der Marker und der QTL sowohl physisch eng gekoppelt als auch im Kopplungsungleichgewicht sind. Für den MTDT werden ein Nachkomme und seine Eltern als Trio genutzt – bei vollständig genotypisierten Triplets ist die Übertragung der Allele bis auf den Fall, dass sowohl der Nachkomme als auch beide Eltern heterozygot sind, eindeutig rekonstruierbar. Für weitere Details siehe SIMIANER und PIMENTEL (2009).

Bestimmung der Power

Jedes Chromosom wurde in 100 Intervalle geteilt. Jedes Intervall war 1 centiMorgan lang. Da die exakten QTL-Positionen bekannt waren, waren es auch die Intervalle, die einen wahren QTL enthielten. Der QTL wurde als entdeckt betrachtet, wenn mindestens ein SNP im QTL-Intervall oder in einem der direkt angrenzenden Intervalle signifikant war. Als Signifikanzgrenze wurde eine Bonferroni-korrigierte Signifikanzgrenze bei einem genomweiten α von 1% verwendet.

Empirische Falsch-Positiv-Rate

Jedes Intervall (x) wurde betrachtet, das keinen QTL enthielt und nicht direkt neben einem QTL-Intervall lag und das mindestens einen signifikanten SNP umfasste. Gab es

zwischen dem Intervall x und dem nächstliegenden QTL-Intervall (auf der einen und der anderen Seite) mindestens ein Intervall, in dem keine signifikanten SNPs vorkamen, so wurden alle signifikanten SNPs im Intervall x als falsch positiv bezeichnet. Zur Berechnung der empirischen Falsch-Positiv-Rate wurde die Anzahl der signifikanten SNPs, die als falsch positiv gewertet wurden, durch die Anzahl aller signifikanten SNPs geteilt.

3 Ergebnisse

Überblick

Abb. 2 zeigt Assoziationsprofile für eine spezifische Wiederholung in Szenario B für alle Methoden. Sie geben einen grundlegenden Überblick über die Eigenschaften der Methoden. Insgesamt sind die $-\log(p)$ -Werte bei GRAMMAR im Durchschnitt im Vergleich zu den beiden anderen Methoden deutlich niedriger, was zu einer geringen Power, aber auch kaum zu falsch positiven Assoziationen führt. MTDT zeigt deutlich höhere Power, eine akzeptable Varianz um die wahren QTL-Positionen und keine falsch positiven Assoziationen auf Chromosom 10. SMR zeigt die höchste Power, aber auch die größten Variationen um die QTL-Positionen und falsch positive Assoziationen auch in QTL-freien Regionen, z.B. auch auf Chromosom 10.

Power – Anzahl entdeckter QTLs

Ergebnisse sind in Abb. 3 zu finden. GRAMMAR zeigte in allen Szenarien die niedrigste Power. Die Anzahl entdeckter QTLs schien jedoch unbeeinflusst von der Populationsstruktur. Für MTDT und SMR war die Power in allen Szenarien signifikant höher als für GRAMMAR. Mit SMR konnten in Szenario A im Durchschnitt drei QTLs mehr entdeckt werden als in Szenario B, zwischen Szenario B und C war jedoch kaum ein Unterschied in der Power zu sehen. Auch die Power für MTDT war in Szenario B höher als in Szenario A, jedoch war der Anstieg nicht so deutlich wie mit SMR. Sowohl mit SMR als auch mit MTDT sank die Anzahl entdeckter QTLs in Szenario C im Vergleich zu Szenario B wieder leicht, wobei die Abnahme bei MTDT stärker war.

Empirische Falsch-Positiv-Rate

Die empirische Falsch-Positiv-Rate (Abb. 4) in Szenario A ist bei GRAMMAR sehr niedrig. SMR und MTDT zeigen eine ähnliche Falsch-Positiv-Rate, die signifikant höher ist als bei GRAMMAR. In Szenario B stieg die empirische Falsch-Positiv-Rate für alle Methoden deutlich an. SMR zeigte die höchste Rate mit über 50%. Für GRAMMAR stieg die empirische Falsch-Positiv-Rate in Szenario C nochmals an, während sie für SMR und MTDT wieder leicht abnahm. In Szenario C zeigte sich zwischen GRAMMAR und MTDT kein signifikanter Unterschied, während SMR signifikant von den beiden anderen Methoden abwich.

4 Diskussion

Ziel dieser Arbeit war es, drei Methoden für genomweite Assoziationsstudien zu evaluieren, indem in einer Simulationsstudie ihre Power und eine empirische Falsch-Positiv-Rate in verschiedenen Szenarien betrachtet wurden.

Mit SMR erhielt man vor allem in den selektierten Populationen bei der Anzahl an falsch positiven Signalen deutlich schlechtere Ergebnisse als mit den beiden anderen

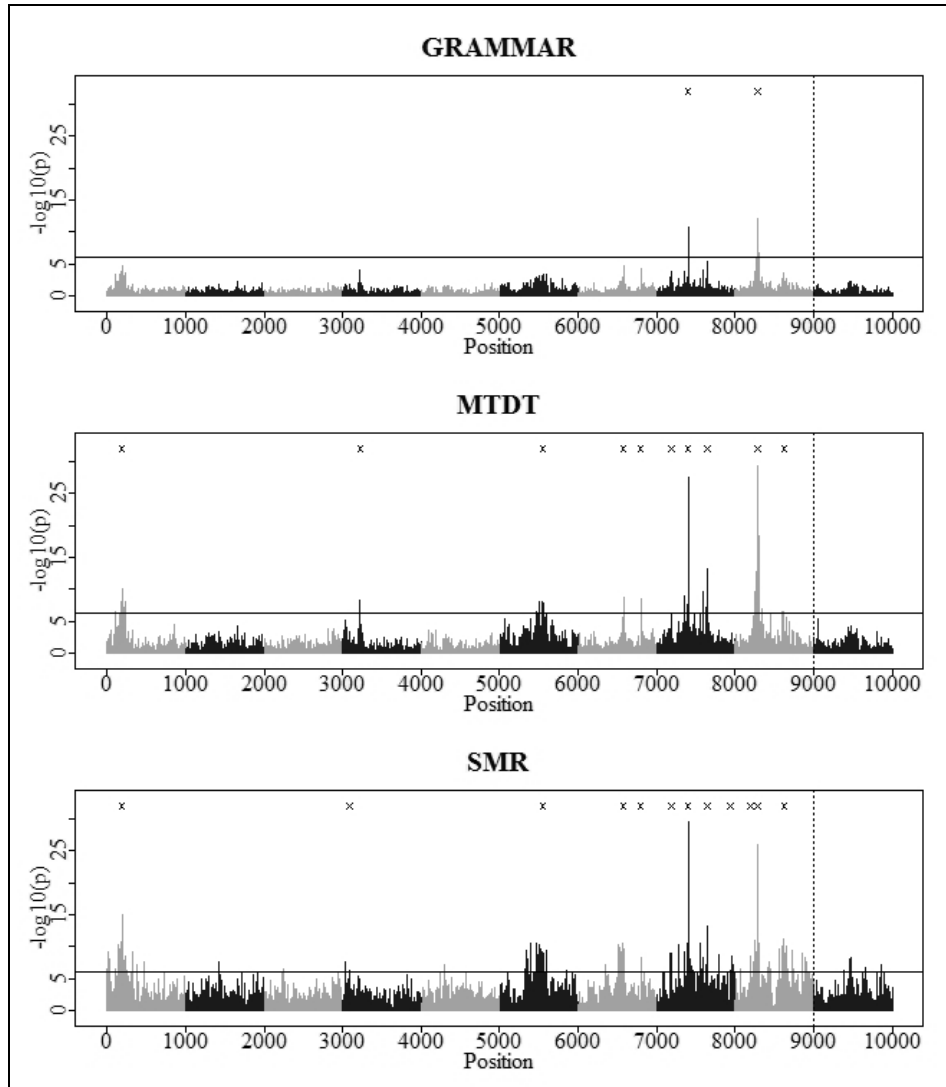


Abb. 2. Assoziationsprofil der $-\log(p)$ -Werte für alle Methoden in einer spezifischen Wiederholung in Szenario B. Die Chromosomen sind farblich getrennt. Die gestrichelte Linie trennt Chromosom 10 ab, auf dem keine QTLs simuliert wurden. Die horizontale Linie zeigt die Bonferroni-korrigierte Signifikanzgrenze bei einem genomweiten α von 1%. Die Kreuze zeigen wahre QTLs, die durch eine Assoziation entdeckt werden konnten.

Profile of $-\log(p)$ -values for the three approaches for a particular replicate in scenario B. Chromosomes are separated by different colours. The dotted line separates chromosome 10 which was free of QTL. The horizontal line shows the Bonferroni corrected threshold on a 1% genome-wide error level. Crosses represent detected true QTL positions.

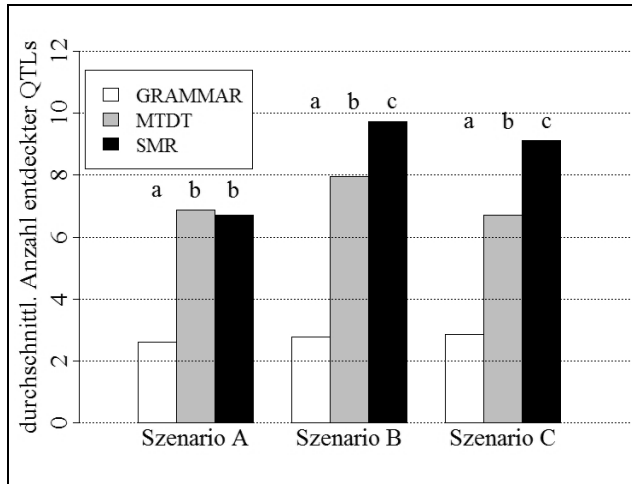


Abb. 3. Durchschnittliche Anzahl entdeckter QTLs. Szenario A: vollständige Zufallspaarung. Szenario B: eingeschränkte Anzahl an Elterntieren. Szenario C: gerichtete Selektion. Unterschiedliche Buchstaben (a, b, c) symbolisieren im jeweiligen Szenario signifikant unterschiedliche Werte ($\alpha = 0,01$).

Number of detected QTL.
Scenario A: idealized random mating.
Scenario B: limited number of parents.
Scenario C: directional selection.
 Different letters (a, b, c) symbolize significantly different means on a 1% error level.

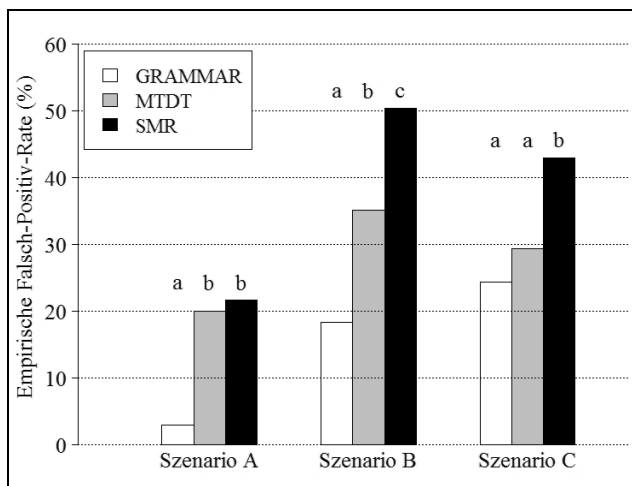


Abb. 4. Empirische Falsch-Positiv-Rate. Szenario A: vollständige Zufallspaarung. Szenario B: eingeschränkte Anzahl an Elterntieren. Szenario C: gerichtete Selektion. Unterschiedliche Buchstaben (a, b, c) symbolisieren im jeweiligen Szenario signifikant unterschiedliche Werte ($\alpha = 0,01$).

Empirical false positive rate.
Scenario A: idealized random mating.
Scenario B: limited number of parents.
Scenario C: directional selection.
 Different letters (a, b, c) symbolize significantly different means on a 1% error level.

Methoden. Während SMR und MTD in der idealisierten Population noch weitgehend gleichwertig waren, zeigte MTD in den selektierten Populationen deutliche Vorteile. Das belegt, dass mit MTD in Situationen mit kleinen effektiven Populationsgrößen und großen Familienstrukturen bessere Ergebnisse zu erwarten sind.

Methoden wie SMR, die Populationsstrukturen nicht beachten, sind äußerst empfindlich gegenüber einem hohen Maß an Verwandtschaft zwischen den Tieren in der analysierten Stichprobe. Deshalb sollten Ergebnisse, bei denen Daten von Nutztierpopulationen mit SMR analysiert wurden, mit Vorsicht betrachtet werden. LEDUR et al. (2009) beschrieben ähnliche Probleme mit hohen Falsch-Positiv-Raten bei mit Single Marker Association (SMA) analysierten Rohdaten. Auch sie nutzten ein simuliertes Datenset, bei dem die Tiere zwar durch Zufallspaarung erzeugt wurden, aber jeweils nur eine begrenzte

Anzahl an Elterntieren zur Verfügung stand, was Szenario B in unserer Arbeit ähnelt. Das Assoziationsprofil ihrer SMA (Abb. 1 in LEDUR et al., 2009) ist dem in dieser Arbeit (Abb. 2) sehr ähnlich. Mit einem gemischten Modell, in das neben dem SNP-Effekt auch ein zufälliger polygener Effekt und andere fixe Effekte mit eingingen, konnten LEDUR et al. (2009) die Anzahl an signifikanten SNPs deutlich verringern, und das Assoziationsprofil wurde deutlich klarer.

Die Anzahl entdeckter QTLs erhöhte sich mit allen Methoden, wenn die Populationsgröße beschränkt war. Dies beruhte vor allem auf dem höheren Ausmaß an Kopplungsungleichgewicht in diesem Szenario. Mit SMR war der Anstieg der Anzahl entdeckter QTLs jedoch außerordentlich hoch. Die Power ist möglicherweise überschätzt, da festzustellen war, dass der starke Anstieg auch auf die Entdeckung kleinster QTLs zurückzuführen war. Wenn sich kleinste QTLs im Umfeld eines größeren QTLs befanden, mit dem viele SNPs assoziiert waren und dessen Variation sehr groß war, war es möglich, dass der kleine QTL definitionsgemäß als entdeckt gewertet wurde, obwohl die Assoziation nicht auf ihn zurückzuführen ist.

Für TDT-basierte Methoden wird oftmals eine verringerte Power im Vergleich mit anderen Methoden für Assoziationsmapping beschrieben (z.B. LONG und LANGLEY, 1999). Das konnte für den auch auf einem TDT basierenden MTDT in dieser Arbeit nicht bestätigt werden. Dieser zeigte ausreichende Power, auch wenn die Population irgendeiner Art von Selektion unterworfen war. Auch als MTDT zur Analyse des QTLMAS XII-Datensets (CROOKS et al., 2009) verwendet wurde, war die Power ausreichend (SIMIANER und PIMENTEL, 2009) und zum Teil sogar besser als mit anderen Methoden, die auch an diesem Datenset getestet wurden.

In einer genomweiten Assoziationsstudie werden sehr viele Tests auf der Grundlage des gleichen Materials durchgeführt. Deshalb muss in irgendeiner Form für dieses multiple Testen korrigiert werden. Eine einfache Korrekturmöglichkeit ist die Bonferroni-Korrektur, bei der das gewünschte genomweite Level für den Fehler 1. Art durch die Anzahl an Tests dividiert wird, um die Signifikanzgrenze für jeden einzelnen Test festzulegen. Jedoch wird für diese Art der Korrektur vorausgesetzt, dass alle Tests unabhängig sind, was jedoch bei einer hochdichten Markerkarte nicht der Fall sein wird. Die Bonferroni-Korrektur ist also in solch einer Studie eher überkonservativ. Trotzdem entschieden wir uns für die Bonferroni-Korrektur, da es so möglich war, zwischen den Wiederholungen, Szenarien und Methoden die Ergebnisse mit einer konstanten Signifikanzgrenze zu vergleichen. Die Korrektur war jedoch für GRAMMAR deutlich zu konservativ. Wie in Abb. 2 zu sehen ist, zeigte GRAMMAR oftmals Peaks an den gleichen Stellen wie MTDT, die jedoch die Bonferroni-Grenze nicht erreichten. LAM et al. (2009) analysierten die Daten des QTLMAS XII workshops (CROOKS et al., 2009) mit GRAMMAR und benutzten eine Bonferroni-korrigierte Signifikanzgrenze bei einem genomweiten α von 5%. GRAMMAR produzierte bei 15 entdeckbaren QTLs neun QTL-Peaks, die die Signifikanzgrenze überschritten. Beim gleichen Datenset identifizierte MTDT 12 der 15 möglichen QTLs mit einer Bonferroni-korrigierten Signifikanzgrenze bei einem genomweiten α von 1% (SIMIANER und PIMENTEL, 2009), was die höhere Power von MTDT im Vergleich zu GRAMMAR bestätigt.

Um die Struktur der p-Werte in einem Kartierungs-Experiment zu berücksichtigen, sind zwei Wege üblich, um die Signifikanzgrenze anzupassen. Es gibt einerseits die Methoden, die versuchen, die False Discovery Rate (FDR) zu kontrollieren, was erstmals von BENJAMINI und HOCHBERG (1995) publiziert wurde. Andererseits kann die Signifikanzgrenze empirisch bestimmt werden, indem ein Permutationstest durchgeführt wird, wie er von DOERGE und CHURCHILL (1995) vorgeschlagen wurde.

Alle Methoden für genomweite Assoziationsstudien, die in dieser Arbeit betrachtet wurden, haben den Vorteil, sehr schnell und wenig rechenaufwändig zu sein. Somit sind

sie geeignet, ganze Genome in einer angemessenen Zeit zu analysieren, auch wenn das Set von Markern groß ist und eine große Anzahl an Tieren und eine Vielzahl an Merkmalen analysiert werden müssen. Dieser Vorteil geht jedoch verloren, wenn rechenaufwändige Prozeduren wie zum Beispiel ein Permutationstest zur Bestimmung der Signifikanzgrenze angewendet werden. Deshalb ist zu empfehlen, für einen ersten Überblick über das ganze Genom eine schnelle und wenig aufwändige Methode (z.B. GRAMMAR oder MTDT) zu verwenden und dann in dabei identifizierten Kandidatenregionen sowohl für die Bildung der Analysemodelle als auch für die Bestimmung der Signifikanzgrenze eine differenzierte Methodik zu verwenden.

Danksagung

Diese Arbeit war Teil der FUGATO-plus Projekte GenoTrack und Brain und wurde vom Bundesministerium für Bildung und Forschung (BMBF), dem Förderverein Biotechnologieforschung e.V. (FBF), Bonn, und den Industriepartnern Lohmann Tierzucht GmbH, Cuxhaven, Verband hannoverscher Warmblutzüchter e.V., Verden, Verband der Züchter des Oldenburger Pferdes e.V., Vechta, Verband der Züchter des Holsteiner Pferdes e.V., Elmshorn, Westfälisches Pferdestammbuch e.V., Münster, und Trakehner Verband, Neumünster, finanziell unterstützt.

Literatur

- ABECASIS, G.R., L.R. CARDON and W.O.C. COOKSON, (2000): A General Test of Association for Quantitative Traits in Nuclear Families. *Am. J. Hum. Genet.* **66**, 279-292.
- AULCHENKO, Y.S., D.-J. DE KONING and C. HALEY, (2007): Genomewide Rapid Association Using Mixed Model and Regression: A Fast and Simple Method For Genomewide Pedigree-Based Quantitative Trait Loci Association Analysis. *Genetics* **177**, 577-585.
- BENJAMINI, Y. and Y. HOCHBERG, (1995): Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc. B* **57**, 289-300.
- CROOKS, L., G. SAHANA, D.-J. DE KONING, M.S. LUND and Ö. CARLBORG, (2009): Comparison of analyses of the QTLMAS XII common dataset. II: genome-wide association and fine mapping. *BMC Proc.* **3** (Suppl 1), S2.
- DOERGE, R.W. and G.A. CHURCHILL, (1995): Permutation Tests for Multiple Loci Affecting a Quantitative Character. *Genetics* **142**, 285-294.
- FALCONER, D.S. and T.F.C. MACKAY, (1996): *Introduction to Quantitative Genetics*. 4th ed. Harlow, Essex, UK: Longman Group.
- GRAPES, L., J.C.M. DEKKERS, M.F. ROTHSCHILD and R.L. FERNANDO, (2004): Comparing Linkage Disequilibrium-Based Methods for Fine Mapping Quantitative Trait Loci. *Genetics* **166**, 1561-1570.
- LAM, A.C., J. POWELL, W.-H. WEI, D.-J. DE KONING and C.S. HALEY, (2009): A combined strategy for quantitative trait loci detection by genome-wide association. *BMC Proc.* **3** (Suppl 1), S6.
- LEDUR, M.C., N. NAVARRO and M. PÉREZ-ENCISO, (2009): Data modeling as a main source of discrepancies in single and multiple marker association methods. *BMC Proc.* **3** (Suppl 1), S9.
- MEUWISSEN, T.H.E. and M.E. GODDARD, (2000): Fine Mapping of Quantitative Trait Loci Using Linkage Disequilibria With Closely Linked Marker Loci. *Genetics* **155**, 421-430.
- RABINOWITZ, D., (1997): A Transmission Disequilibrium Test for Quantitative Trait Loci. *Hum. Hered.* **47**, 342-350.

- SIMIANER, H. and E.C.G. PIMENTEL, (2009): Robust QTL fine mapping by applying a quantitative transmission disequilibrium test to the Mendelian sampling term. *J. Anim. Breed. Genet.* **126**, 432-442.
- TERWILLIGER, J.D., S. ZÖLLNER, M. LAAN and S. PÄÄBO, (1998): Mapping Genes through the Use of Linkage Disequilibrium Generated by Genetic Drift: 'Drift Mapping' in Small Populations with No Demographic Expansion. *Hum. Hered.* **48**, 138-154.
- ZHAO, H.H., R.L. FERNANDO and J.C.M. DEKKERS, (2007): Power and Precision of Alternate Methods for Linkage Disequilibrium Mapping of Quantitative Trait Loci. *Genetics* **175**, 1975-1986.