

Statuserhebung

Genetische Diversität Schwein

Schriftenreihe, Heft 9/2012



Statuserhebung zur genetischen Diversität in sächsischen Nutztierpopulationen am Beispiel Schwein

Gudrun A. Brockmann, Uwe Müller, Armin O. Schmidt
Ottmar Distl
Uwe Bergfeld, Ulf Müller

1	Einleitung	6
2	Material und Methoden	7
2.1	Population untersuchter Schweine.....	7
2.2	Genotypisierungen mittels SNP-Chip-Technologie	7
2.3	Phänotypische Daten.....	8
2.4	Haplotypenanalyse	10
2.5	Assoziationsanalysen	11
3	Ergebnisse	12
3.1	Phänotypische Beziehungen zwischen Merkmalen	12
3.2	Genomstruktur der untersuchten Schweinepopulation	13
3.2.1	Allelfrequenzen von SNPs	13
3.2.2	Haplotypenblockstruktur	14
3.3	Populationssubstrukturen	15
3.4	Genomweite Assoziationsanalysen für verschiedene Merkmale aus der Leistungsprüfung	15
3.4.1	Prüftagszunahme.....	15
3.4.2	Lebenstagszunahme.....	16
3.4.3	Futtermverehr.....	18
3.4.4	Ultraschall-Fettdicke	19
3.4.5	Fettfläche	20
3.4.6	Ultraschall-Muskeldicke	22
3.4.7	Intramuskulärer Fettgehalt	23
3.4.8	Dripverlust.....	24
3.4.9	pH-Wert 1. Kontrolle	25
4	Diskussion.....	27
5	Quellenverzeichnis	30

Abbildung 1:	Verteilung der auf dem porcinen SNP-Chip vorhandenen SNPs auf 18 Autosomen, die Geschlechtschromosomen und die genomisch nicht zu lokalisierenden SNPs der Gruppe 0	8
Abbildung 2:	Verteilung der Minderheitenallelfrequenzen (MAF) (a) und Heterozygotiegrad aller Einzeltiere (b)	13
Abbildung 3:	Beispiel für einen Abschnitt auf Chromosom 2 mit niedriger Minderheitenallelfrequenz (MAF) und mittlerer MAF im Nachbarabschnitt. Die Länge des Abschnitts mit geringer MAF ist etwa 400.000 Basenpaare lang.	14
Abbildung 4:	Histogramm für die Häufigkeitsverteilung von Haplotypenblöcken entsprechend ihrer Länge	14
Abbildung 5:	Darstellung von Populationssubstrukturen.	15
Abbildung 6:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Prüftagszunahme und genotypische Unterschiede in der Prüftagszunahme	16
Abbildung 7:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Lebenstagszunahme und genotypische Unterschiede in der Lebenstagszunahme	17
Abbildung 8:	Darstellung der Haplotypeneffekte im Merkmal Lebenstagszunahme für den Haplotypenblock, der den signifikantesten SNP enthält.....	17
Abbildung 9:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Futterverzehr und genotypische Unterschiede im Futterverzehr	19
Abbildung 10:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Ultraschall-Fettdicke und genotypische Unterschiede in der Ultraschall-Fettdicke	20
Abbildung 11:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Fettfläche und genotypische Unterschiede in der Fettfläche.....	21
Abbildung 12:	Darstellung der Haplotypeneffekte im Merkmal Fettfläche für den Haplotypenblock, der den signifikantesten SNP enthält	21
Abbildung 13:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Ultraschall-Muskeldicke und genotypische Unterschiede in der Ultraschall-Muskeldicke	23
Abbildung 14:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal intramuskulärer Fettgehalt und genotypische Unterschiede im intramuskulären Fettgehalt	24
Abbildung 15:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Dripverlust und genotypische Unterschiede im Dripverlust.....	25
Abbildung 16:	Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal pH-Wert und genotypische Unterschiede im pH-Wert	26
Abbildung 17:	Darstellung der Haplotypeneffekte im Merkmal pH-Wert für den Haplotypenblock auf Chromosom 15, der den signifikantesten SNP enthält.....	26

Tabellenverzeichnis

Tabelle 1:	Einflussfaktoren auf die erhobenen Prüfmerkmale.....	9
Tabelle 2:	Deskriptive Statistik der auf Umwelteffekte korrigierten Phänotypen in der Leistungsprüfung.....	10
Tabelle 3:	Korrelationskoeffizienten zwischen vorkorrigierten Merkmalen.....	13
Tabelle 4:	p-Werte und Rangfolge von SNPs in Modellen mit und ohne Berücksichtigung der Populationsstruktur (Stratifikation)	22

1 Einleitung

Das im Jahr 2006 verabschiedete Tierzuchtgesetz verpflichtet die zuständigen Behörden der Länder zu einem laufenden Monitoring der genetischen Diversität. Dies betrifft sowohl die unter Leistungszucht stehenden Populationen der Tierarten Rind, Pferd, Schwein, Schaf und Ziege als auch die im Rahmen eines Erhaltungszuchtprogramms bearbeiteten gefährdeten Populationen. Derartige Monitoringprogramme basieren sowohl auf Leistungs- und Pedigreeinformationen aus der Leistungsprüfung als auch auf Ergebnissen molekulargenetischer Untersuchungen.

Die genetische Diversität in einer Population ist entscheidend für die Anpassungsfähigkeit an schwankende Umweltbedingungen und bildet die wichtigste genetische Ressource für die Züchtung. Um vorhandene Population dauerhaft anpassungsfähig zu erhalten und vorhandene Nutztierpopulationen züchterisch effektiver zu nutzen, sind eine Charakterisierung der genetischen Diversität und eine differenzielle Beschreibung der Populationen notwendig. Darüber hinaus bietet die Datenerhebung zur genetischen Diversität mit sehr vielen Markern die Möglichkeit, die genetischen Beziehungen zwischen einzelnen polymorphen Loci und Leistungsparametern genomweit aufzudecken.

Im vorliegenden Projekt wurden erstmals der enorme Fortschritt in der Entwicklung von genetischen Markern beim Schwein und die Verfügbarkeit effizienter Hochdurchsatzplattformen für Chip-basierte Genotypisierungen genutzt, um alle grundsätzlichen Aspekte des genomischen Ansatzes der Tiercharakterisierung an realen Schweinepopulationen im Detail zu untersuchen. Der entscheidende technische Durchbruch wurde mit der Identifizierung und Nutzung von Einzelbasenpaarpolymorphismen, sogenannten Single Nucleotide Polymorphisms (SNPs), als genetische Marker erzielt. Die für das Schwein entwickelten und mit Beginn des Jahres 2009 kommerziell verfügbaren SNP-Chips mit etwa 54.000 SNPs (Illumina, San Diego, CA, USA, www.illumina.com) stellen eine neue Dimension für die Untersuchung der genetischen Diversität des Schweinegenoms dar. Die SNPs, die über das ganze Genom verteilt liegen, erlauben nicht nur eine punktuelle, sondern eine genomweite Charakterisierung der genetischen Struktur und Vielfalt des Genoms. Damit werden auch beim Schwein Untersuchungen möglich, wie sie für das Rind und nahezu zeitgleich für das Huhn 2008 begonnen wurden und beim Mensch mit inzwischen mehr als 1 Million SNP erfolgreich eingesetzt werden.

In der vorliegenden Studie wurde die genetischen Diversität exemplarisch für die sächsische Schweinezuchtpopulation an einer Stichprobe von 288 Prüfebern der Deutschen Landrasse untersucht. Die aus der gleichzeitigen Analyse von etwa 54.000 SNPs erhaltenen Genotypen bilden die Grundlage zur populationsgenetischen Beschreibung des gegenwärtigen Status der Nutztierpopulation und zur genomweiten Analyse der Assoziation zwischen der genetischen Information und wichtigen Nutzeigenschaften. Mit der sehr dichten SNP-Genotypisierung ist es erstmals auch möglich, das Genom nicht nur punktuell an einzelnen SNP-Orten zu betrachten, sondern das gesamte Genom anhand jeweils benachbarter Markergruppen in Segmente, sogenannte Haplotypenblöcke, zu unterteilen, die in Form von Haplotypen charakterisiert werden können.

Auf der Grundlage der mit den SNP-Chips erhaltenen Genominformation können folgende Fragen gestellt werden:

- Welche Allelfrequenzen liegen an den einzelnen SNPs im Genom vor?
- Wie sieht das Haplotypeninventar in den ausgewählten Populationen aus?
- Welche genetischen Effekte werden einzelnen SNPs und Haplotypen zugeordnet?
- Welche SNPs sind optimal, um in einer größeren Tierpopulation Assoziationen zur Merkmalsausprägung zu testen?

Als Ergebnis der Selektion in Nutztierpopulationen wird erwartet, dass für additiv wirkende Genvarianten ein Verlust der Heterozygotie auftritt und für Loci, die an Fitnessmerkmalen beteiligt sind, möglicherweise ein erhöhter Heterozygotiegrad zu beobachten ist. Es wird ferner erwartet, dass die wiederholte Selektion auf additive Geneffekte in Nutztierpopulationen sogenannte Selektionssignaturen hinterlässt, die durch Erhöhung der Homozygotie gekennzeichnet sind. Lange Haplotypenblöcke sind darüber hinaus ein Maß für hohe verwandtschaftliche Beziehungen in der Population.

Mit der Sicht auf ca. 54.000 Loci im gesamten Genom eines Tieres besteht neben der umfangreichen Beurteilung der genomischen Diversität innerhalb einer Population auch die Möglichkeit, die genetische Veranlagung eines einzelnen Tieres im Vergleich zu anderen Tieren zu charakterisieren. Dabei wird die Summe der Einzeleffekte von SNP-Allelen im gesamten Genom zur Schätzung des sogenannten genomischen Zuchtwerts eines einzelnen Tieres verwendet. Diese Methode wird inzwischen beim Rind erfolgreich eingesetzt, um das Nutzungspotenzial junger Tiere mit hoher Sicherheit vorherzusagen. Für die Schätzung genomischer Zuchtwerte ist es jedoch notwendig, eine Summenformel zu entwickeln, die die Alleleffekte der einzelnen, über das gesamte Genom verteilten SNPs innerhalb der betrachteten Population möglichst genau berücksichtigt.

Die durchgeführte Studie sollte auch dazu dienen, durch die Beurteilung von SNP-Effekten zur Vorbereitung der genomischen Selektion beim Schwein beizutragen. Hierfür wurden SNP-Effekte für ausgewählte Merkmale der Mast- und Schlachtleistung geschätzt, um einen ersten Eindruck darüber zu erhalten, inwieweit diese Merkmale mit genomischen Informationen assoziiert sind. Außerdem sollten Haplotypenstrukturen untersucht und entsprechende Effekte abgeleitet werden, um zu prüfen, ob Haplotypen möglicherweise in der genomischen Selektion genutzt werden könnten.

2 Material und Methoden

2.1 Population untersuchter Schweine

Die Untersuchungen wurden mit Prüfebern der Deutschen Landrasse durchgeführt, deren Leistungsprüfung auf zwei Stationen des Mitteldeutschen Schweinezuchtverbandes erfolgte. Die Tiere wurden zwischen September 2009 und Januar 2010 geprüft. In diesem Zeitraum wurden Daten von 584 Ebern erhoben. Von 491 Tieren standen Abstammungsdaten zur Verfügung. Die 491 Prüfeber stammten von 34 Zuchtebern ab.

Unter den 491 Prüfebern wurden 288 Tiere ausgewählt, die mittels SNP-Chips genotypisiert wurden. Repräsentation der gesamten Population und möglichst geringe Verwandtschaft waren Kriterien für die Tierausswahl. Für die Auswahl der zu genotypisierenden Tiere wurden die Verwandtschaftsinformationen über drei Generationen zurückverfolgt. Daraus wurden die F-Like-Werte als Maß für die Verwandtschaft zwischen den Tieren ermittelt. Für SNP-Chip-Genotypisierungen wurden dann solche Tiere ausgewählt, die kleine F-Like-Werte hatten und damit geringe verwandtschaftliche Beziehungen untereinander aufwiesen.

Der mittlere Verwandtschaftskoeffizient unter den ausgewählten 288 Tieren betrug 1,21, unter den nicht ausgewählten Tieren 1,29. Die ausgewählten 288 Tiere stammten von 31 Vätern ab. Die Anzahl der Söhne je Vater schwankte zwischen 1 und 31.

2.2 Genotypisierungen mittels SNP-Chip-Technologie

Die Genotypisierung der 288 ausgewählten Tiere erfolgte mit dem 60K PorcineSNP60 BeadChip der Firma Illumina (Illumina, San Diego, CA, USA). Die Genotypisierung erfolgte nacheinander in zwei Stufen mit jeweils 144 Tieren. Zunächst wurden 144 Tiere mit einem Chip genotypisiert, der ein Vorgänger zu dem jetzt erhältlichen kommerziellen Chip war. Auf diesem Chip waren 64.232 SNPs enthalten. Weitere 144 Tiere wurden mit einem weiterentwickelten kommerziell vertriebenen Chip genotypisiert, der 62.163 SNPs enthielt. Der weiterentwickelte Chip enthielt 2.069 SNPs weniger als der Vorgängerchip, weil die Firma Illumina SNPs mit ungenügenden Qualitätsparametern eliminiert hat. Für die Analyse der Genotypen und nachfolgende Assoziationsanalysen wurden nur solche SNPs verwendet, die auf beiden Chips vorhanden waren.

Zunächst wurden die auf dem Chip vorhandenen SNPs auf der Grundlage der aktuellen, öffentlich zugänglichen NCBI-Variante für die genomische Sequenz beim Schwein (Built Version Sscrufa9) neu kartiert (SCHMITT et. al 2010). In diesem Schritt wird jedem SNP seine genomische Position auf einem Chromosom entsprechend der genomischen Sequenz des Schweins zugeordnet. Dieser Schritt ist von besonderer Bedeutung für die spätere Interpretation der Ergebnisse der Assoziationsanalyse

zur Ermittlung von Genen, die mit einem signifikanten SNP gekoppelt vererbt werden können und für die richtige Ableitung von Haplotypen und Haplotypenblock-Strukturen. 8.203 SNPs konnten keinem Chromosom zugeordnet werden. Sie wurden einer Gruppe ohne Position mit der Bezeichnung „Chromosom 0“ zugeordnet.

Die mit beiden Chips erzeugten Genotypendaten wurden einer gründlichen Qualitätsprüfung unterzogen, bevor statistische Analysen durchgeführt wurden. Die Qualitätsprüfung erfolgte in mehreren Schritten. SNPs, deren Frequenz des Minderheitenallels kleiner als 1 % war und deren Genotypisierung bei mehr als 10 % aller Individuen fehlschlug, wurden eliminiert. Ebenso wurden Tiere von den weiteren Analysen ausgeschlossen, bei denen weniger als 90 % aller SNPs genotypisiert wurden. Im Ergebnis der Qualitätsprüfung standen Genotypen von 283 Ebern mit 53.536 SNPs für nachfolgende Untersuchungen zur Verfügung. Die meisten SNPs, die allen Qualitätskriterien entsprachen, befanden sich auf Chromosom 1, gefolgt von den Chromosomen 14, 4 und 13 (Abbildung 1).

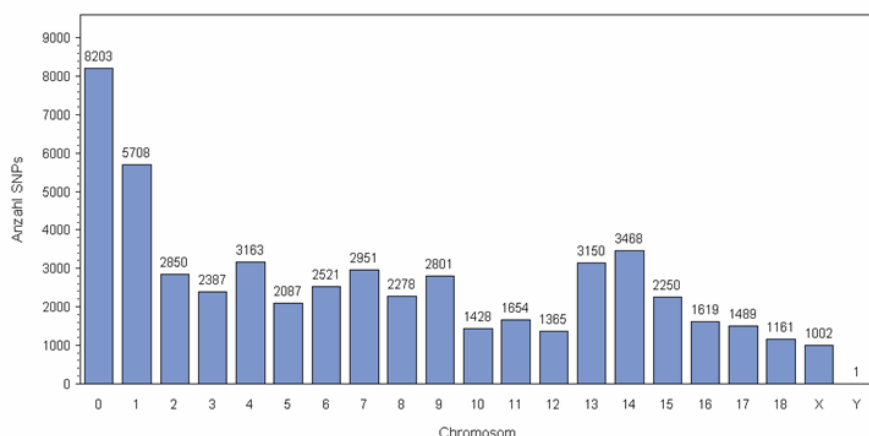


Abbildung 1: Verteilung der auf dem porcinen SNP-Chip vorhandenen SNPs auf 18 Autosomen, die Geschlechtschromosomen und die genomisch nicht zu lokalisierenden SNPs der Gruppe 0

2.3 Phänotypische Daten

Die Originalleistungsdaten stammten aus zwei Prüfstationen im Zeitraum zwischen September 2009 und Januar 2010. Die Prüfung begann bei einem Gewicht von 30 kg und endete bei einem Gewicht von 105 kg. Anschließend wurden die Tiere geschlachtet. Die Prüfung erfolgte nach den „Richtlinien für die Stationsprüfung auf Mastleistung, Schlachtkörperqualität und Fleischbeschaffenheit beim Schwein“ (2004). Aus der Leistungsprüfung wurden folgende phänotypische Daten verwendet:

Lebenstagszunahme tägliche Körpermassezunahme im Zeitraum zwischen Geburt und Schlachtung bei einem Gewicht von 105 kg

Merkmale der Mastleistung:

Prüftagszunahme tägliche Körpermassezunahme im Prüfzeitraum der Körpermasseentwicklung von 30 bis 105 kg
Futterverzehr gesamte Futtermittelaufnahme während der Wachstumsphase zwischen 30 und 105 kg

Merkmale des Schlachtkörperwertes:

Ultraschall-Fettdicke die mittels Ultraschall am lebenden Tier gemessene Fettdicke an der 13. Rippe (Mittelwert aus drei Messwerten)
Fettfläche die im Schlachtkörper gemessene Fettfläche im Kotelett im Abschnitt der 13. bis 14. Rippe
Ultraschall-Muskeldicke die mittels Ultraschall am lebenden Tier gemessene Muskeldicke des Koteletts im Bereich der 13. Rippe
Intramuskulärer Fettgehalt prozentualer Fettgehalt im Kotelett an der 13. Rippe

Merkmale der Fleischbeschaffenheit:

Dripverlust	Verlust von Gewebeflüssigkeit aus dem Kotelett nach Lagerung bei 4 °C über 24 Stunden
pH-Wert 1	pH-Wert im Kotelett 45 Minuten nach Schlachtung

Für Assoziationsanalysen wurden die vorkorrigierten BLUE-Werte genutzt. Diese wurden unter Verwendung des etablierten Zuchtwertschätzmodells berechnet, wobei die Phänotypen aller 584 etwa zeitgleich geprüften Eber aus der Vorauswahl zur Korrektur beitrugen. Tabelle 1 enthält die bei der Vorkorrektur berücksichtigten Faktoren. Im Gegensatz zum Zuchtwertschätzmodell musste der Einfluss des Geschlechtes nicht berücksichtigt werden, weil in den vorliegenden Untersuchungen nur Daten von Ebern ausgewertet wurden. Auswertungen auf der Basis vorkorrigierter Werte hatten gegenüber einer Assoziationsanalyse unter gleichzeitiger Korrektur auf Störgrößen den Vorteil, dass hierbei wie in der Zuchtwertschätzung alle Effekte in gleicher Weise berücksichtigt werden konnten.

Alle vorkorrigierten Merkmale wurden hinsichtlich Normalverteilung überprüft. Weil der intramuskuläre Fettgehalt von der Normalverteilung deutlich abwich, wurde das Merkmal logarithmiert. Die Darstellung von Zusammenhängen zwischen den Merkmalen erfolgte mit Hilfe von Korrelationsanalysen.

Von 16 der 283 genotypisierten Eber waren keine Phänotypen vorhanden. Diese Tiere waren ohne Merkmalserhebung unmittelbar nach der Genotypisierung verkauft worden. Tabelle 2 enthält neben der Anzahl verfügbarer Phänotypen weitere wesentliche Kennzahlen der deskriptiven Statistik. Von den insgesamt neun in die Analysen einbezogenen Merkmalen schwankte die Verfügbarkeit der Daten zwischen 94 % (N=267 von 283) und 45 % (N=128 von 283).

Tabelle 1: Einflussfaktoren auf die erhobenen Prüfmerkmale

Merkmal	Faktoren*
Prüftagszunahme	Lebendmasse am Prüfende, Abteil, Rassengruppe, Tier
Lebenstagszunahme	Betriebs-Jahres-Quartalseffekt, Rassengruppe, Tier
Futtermittelverzehr	Lebendmasse am Prüfende, Abteil, Rassengruppe, Tier
Ultraschall-Fettdicke	Lebendmasse (Feld), Betriebs-Jahres-Quartalseffekt, Rassengruppe, Ultraschall-Seitenspeckdicke, Tier
Fettfläche	Schlachtkörpermasse (warm), Abteil, Rassengruppe, Tier
Ultraschall-Muskeldicke	Lebendmasse (Feld), Rassengruppe, Betriebs-Jahres-Quartals-Effekt, Ultraschall-Seitenspeckdicke, Tier
intramuskulärer Fettgehalt	Schlachtkörpermasse (warm), Abteil, Rassengruppe, Tier
Dripverlust	Schlachtkörpermasse (warm), Abteil, Rassengruppe, Tier
pH-Wert	Schlachthof, Rassengruppe, Tier

* Tier wurde als zufälliger Faktor berücksichtigt, alle anderen Faktoren als fixe Effekte.

Tabelle 2: Deskriptive Statistik der auf Umwelteffekte korrigierten Phänotypen in der Leistungsprüfung

Merkmal	Symbol	Einheit	Anzahl Tiere	Mittelwert	Standard-abweichung	Minimum	Maximum
Prüftagszunahme	PTZ	g/d	267	959	88	678	1247
Lebenstagszunahme	LTZ	g/d	266	689	49	558	841
Futtermverzehr	FUVZ	kg	265	201	43	69	330
Ultraschall-Fettdicke	USF	mm	266	13	3	6	23
Fettfläche	FFL	cm ²	236	18	5	5	35
Ultraschall-Muskeldicke	USMD	mm	148	52	2	43	59
Intramuskulärer Fettgehalt*	IMF	%	236	18	5	5	35
Dripverlust	DV	%	209	4,0	1,7	0,6	9,1
pH-Wert der ersten Kontrolle	PH1K	dimensionslos	128	6,4	0,18	6,0	6,9

* Um Normalverteilung herzustellen, wurden die Daten für die Korrelations- und Assoziationsanalysen log-transformiert.

2.4 Haplotypenanalyse

Wirft man einen detaillierten Blick auf eine Reihe eng zusammen liegender SNPs an einem beliebigen genomischen Locus, so wird man feststellen, dass nicht alle theoretisch möglichen Abfolgen der Allele (Varianten) auf den Chromosomen vorkommen, sondern nur ein Bruchteil davon. Kennt man also die Allele nur einiger benachbarter SNPs auf einem Chromosom, so kann man (bei Kenntnis der tatsächlich vorkommenden Kombinationen) mit einiger Sicherheit die Allele der benachbarten SNPs vorhersagen. Entfernt man sich jedoch zu weit von den Ausgangs-SNPs, so wird dies ab einer bestimmten Entfernung nicht mehr möglich sein. Die Region, in der die Varianten stark korreliert sind, in der also Vorhersage möglich ist, nennt man Haplotypenblock; die tatsächlich vorkommenden Allel-Kombinationen heißen Haplotyp und die Reichweite der Korrelation markiert die Grenze zwischen einem Haplotypenblock und dem nächsten. Die Haplotypenblockstruktur ist ein spezifisches Merkmal einer Population und erlaubt, insbesondere durch den Vergleich mit anderen Populationen, Aussagen über die züchterische Vergangenheit. Viele kurze Haplotypenblöcke lassen auf eine alte Populationsgeschichte schließen, wenige lange Haplotypenblöcke deuten auf eine junge Geschichte hin oder dass die Population vor kurzem durch einen genetischen Flaschenhals gegangen ist.

Die Assoziationsanalyse mit Haplotypen als Ergänzung zur Analyse mit einzelnen SNP-Allelen und Genotypen bietet eine Reihe von Vorteilen, die im Folgenden kurz umrissen werden sollen:

Aus statistischer Sicht relevant ist die Verringerung der Anzahl von Tests, wodurch das Problem des multiplen Testens entschärft wird. Ein biologisches Argument für die Betrachtung von Haplotypen ist, dass in manchen Fällen ein molekularbiologischer Effekt einer Variante nicht durch einen einzigen SNP erklärt werden kann, sondern nur durch das Zusammenwirken dieses SNPs mit einem oder mehreren benachbarten SNPs. Eine signifikante Assoziation würde also in einem solchen Fall unter Umständen bei einer reinen SNP-basierten Analyse nicht beobachtet werden.

Die mathematische Analyse von Assoziationen mit Merkmalen geschieht über das Aufstellen von Verallgemeinerten Linearen Modellen, wobei die statistische Signifikanz der Haplotypen durch einen Wald-Test ihrer Effekte beurteilt wird. Verkompliziert wird eine Assoziationsanalyse auf Haplotypenbasis dadurch, dass beim Genotypisieren nicht festgestellt werden kann, welches Allel von welchem Chromosom (also vom väterlichen oder mütterlichen) stammt. Genau diese Information ist jedoch wichtig für die Ermittlung der Haplotypen. Durch Wahrscheinlichkeitsüberlegungen können die Allele jedoch mit hoher Sicherheit mittels

der Software fastPHASE (SCHEET & STEPHENS 2006) einem Chromosom zugeordnet werden. Dieser Vorgang wird in der Fachsprache „Phasing“ genannt.

Ein weiterer notwendiger Zwischenschritt für die Assoziationsanalyse auf Haplotypenbasis ist das Ableiten der Haplotypenblöcke. Darunter versteht man das Ermitteln der Grenzen zwischen jeweils zwei Regionen, deren SNPs untereinander hoch korreliert sind, z. B. mit der Software haploview (BARRETT et al. 2005). Die SNPs aus unterschiedlichen Haplotypenblöcken sind jedoch nicht miteinander korreliert. Bei diesem Schritt müssen einige Parameter festgelegt werden wie z. B. der Algorithmus oder die Stärke der Korrelation, die zwischen den SNPs aus einem Block mindestens herrschen muss. Diese Entscheidungen können nicht willkürlich getroffen werden, haben aber einen Einfluss auf die erhaltene Blockstruktur einer Population. Dies erschwert demzufolge den Vergleich der Ergebnisse, die etwa für zwei Spezies ermittelt werden.

2.5 Assoziationsanalysen

Ziel einer Assoziationsanalyse ist es festzustellen, ob Tiere, die sich genetisch an einem bestimmten Locus unterscheiden, auch unterschiedlich bezüglich eines interessierenden Merkmals sind oder ob kein Unterschied besteht.

Man spricht von einer genomweiten Assoziationsanalyse, wenn variable Loci wie z. B. SNPs, die dicht und gleichmäßig über das gesamte Genom liegen, hinsichtlich ihrer Beziehung zur Merkmalsausprägung überprüft werden. Prinzipiell können Allele, Genotypen, Haplotypen oder Diplotypen (Kombinationen zweier Haplotypen) auf Assoziation mit einem Merkmal überprüft werden. In der vorliegenden Studie beschränkten sich die Autoren auf das Testen auf Allel- und Haplotypenebene (siehe Abschnitt 3.4). Dafür wurden die Softwareprogramme GenAble (AULCHENKO et al. 2008) und PLINK (PURCELL et al. 2007) verwendet. Ergebnisse einer Assoziationsanalyse sind die geschätzten Effekte, die ein Allel, ein Genotyp oder ein Haplotyp auf ein Merkmal hat sowie der dazugehörige p-Wert, der die statistische Signifikanz angibt.

In der Praxis wird eine Analyse durch verschiedene Umstände erschwert, die im Folgenden kurz skizziert werden sollen.

Idealerweise werden für eine Assoziationsanalyse Phänotypen und Genotypen aus unverwandten Individuen gewonnen. Dies ist bei selektierten Nutztierpopulationen aufgrund der Züchtungsgeschichte nur eingeschränkt möglich. Deshalb muss die Verwandtschaftsstruktur bei den Berechnungen berücksichtigt werden. Das Vorhandensein von Populationssubstrukturen wird auch als Populationsstratifikation bezeichnet. In einer Population vorhandene Substrukturen können durch Suche nach systematischen Unterschieden im Auftreten von Allelen und Allelfrequenzen zwischen den Individuengruppen und der Gesamtpopulation aufgedeckt werden.

In der vorliegenden Studie wurde das in GenAble implementierte egsscore-Verfahren genutzt, welches auf einer von PRICE et al. (2006) beschriebenen Methode für genomweite Assoziationsanalysen mit Berücksichtigung der Populationsstratifikation beruht. Sie basiert 1. auf einer Hauptkomponentenanalyse, in welcher Eigenvektoren für die Populationsstratifikation bestimmt werden, 2. auf der Korrektur von Phäno- und Genotypen und 3. auf einem sich anschließenden Assoziationstest zwischen korrigierten Phäno- und Genotypen. Zusätzlich wurde eine Lambda-Korrektur durchgeführt. Lambda-Werte spiegeln die Abweichung der Allelfrequenz an Markern wider, die nicht mit dem Merkmal assoziiert sind im Vergleich zu allen Markern in der Population. Ein Lambda-Wert um 1 signalisiert, dass keine Verwandtschaft zwischen den Individuen vorliegt. Die Lambda-Korrektur führt in der Regel zu weniger signifikanten p-Werten in Assoziationsanalysen.

Es wurden sowohl Analysen unter Berücksichtigung der Populationsstratifikation als auch unter deren Vernachlässigung durchgeführt. Sofern nicht anders vermerkt, werden stets die p-Werte mit Berücksichtigung der Stratifikation angegeben.

Ein weiteres Problem besteht darin, dass bei einer genomweiten Assoziationsstudie ebenso viele statistische Tests wie die Anzahl der genomischen Loci durchgeführt werden. Dies führt zu dem Phänomen des starken Anwachsens der Anzahl falsch positiver Assoziationen. Mehrere Korrekturverfahren, die sich in ihrer Schärfe unterscheiden, wurden zu diesem Zweck entwickelt. In dieser Studie entschieden sich die Autoren für das Permutationsverfahren. Es besteht im Kern darin, dass die Zuordnung von Genotypen zu Phänotypen 10.000mal zufällig festgelegt wird und jedes Mal die Stärke der Assoziation

berechnet wird. Die daraus erhaltene Verteilung erlaubt die statistische Beurteilung der Assoziation, die für die ursprüngliche Zuordnung berechnet wurde.

Aufgrund des relativ kleinen Stichprobenumfangs von weniger als 300 Tieren war zu erwarten, dass unter Berücksichtigung des multiplen Testens die p-Werte über den allgemein anerkannten Schwellwert von 0,05 angehoben werden. Weil die vorliegende Studie den Charakter einer Machbarkeitsstudie trägt, zeigen die Autoren sowohl die unkorrigierten als auch die auf multiples Testen korrigierten p-Werte. In genomweiten Assoziationsanalysen beim Menschen konnte gezeigt werden, dass unkorrigierte p-Werte kleiner 10^{-4} biologisch relevant sein können und deshalb Berücksichtigung finden sollten (SPELIOTES et al. 2010). In der vorliegenden Studie wird dieser Schwellenwert zur Interpretation der Ergebnisse genutzt.

3 Ergebnisse

3.1 Phänotypische Beziehungen zwischen Merkmalen

Korrelationen zwischen Merkmalen, die auf ihre Umwelteffekte korrigiert wurden, geben erste Hinweise auf einen möglichen genetischen Zusammenhang zwischen diesen Merkmalen (Tabelle 3). Wie zu erwarten, war die Lebenstagszunahme mit der Prüftagszunahme sehr eng korreliert ($r=0,80$). Eine höhere Zunahme ist nur zum Teil auf eine höhere Futteraufnahme zurückzuführen. Der Zusammenhang zwischen Futterverzehr und Prüftagszunahme ist höher ($r=0,35$) als für Lebenstagszunahme ($r=0,22$). Höherer Futterverzehr und höhere Zunahme führen zu höherem Fettansatz, gemessen als Ultraschall-Fettdicke und Fettfläche. Die Korrelationskoeffizienten zwischen diesen Parametern liegen zwischen 0,29 und 0,67. Hoher Futterverzehr ist insbesondere mit großer Fettfläche korreliert ($r=0,67$). Die Korrelation zwischen Ultraschall-Fettdicke und Fettfläche ist ebenfalls hoch ($r=0,58$), was darauf hinweist, dass die Ultraschallmessungen der Fettdicke am lebenden Tier zuverlässige und genaue phänotypische Daten zum Fettansatz liefern. Die Ultraschall-Muskeldicke korreliert mit den Daten der Körpermasseentwicklung, aber nicht mit Futteraufnahme und Fettansatz. Für die beiden Parameter der Muskel-Qualität besteht in der untersuchten Tierpopulation eine starke negative Korrelation ($r=-0,51$). Geringe pH-Werte im Kotelett 45 Minuten nach der Schlachtung sind mit hohen Dripverlusten verbunden. Nur eine marginal signifikante und sehr geringe Korrelation konnte zwischen Ultraschall-Muskeldicke und Dripverlust ($r=0,17$) sowie pH-Wert 45 Minuten nach Schlachtung ($r=-0,15$) gefunden werden.

Tabelle 3: Korrelationskoeffizienten zwischen vorkorrigierten Merkmalen

	PTZ	LTZ	FUVZ	USF	FFL	USMD	IMF	DV	PH1K
Prüftagszunahme (PTZ)	1.00000 267								
Lebenstagszunahme (LTZ)	0.80283 <.0001 265	1.00000 266							
Futtermverzehr (FUVZ)	0.34897 <.0001 265	0.21763 0.0004 263	1.00000 265						
Ultraschall-Fettdicke (USF)	0.33805 <.0001 265	0.40536 <.0001 266	0.46100 <.0001 263	1.00000 266					
Fettfläche (FFL)	0.31225 <.0001 236	0.29439 <.0001 235	0.66900 <.0001 234	0.58193 <.0001 2235	1.00000 236				
Ultraschall-Muskeldicke (USMD)	0.17334 0.0358 147	0.21565 0.0085 148	0.06405 0.4441 145	0.02948 0.7220 148	-0.02142 0.8148 122	1.00000 148			
Intramuskulärer Fettgehalt (IMF)	-0.00652 0.9352 158	0.01719 0.8308 157	0.09971 0.2128 158	0.01120 0.8893 157	0.08616 0.2833 157	-0.05208 0.6953 59	1.00000 158		
Dripverlust (DV)	-0.08004 0.2493 209	-0.09517 0.1715 208	-0.06255 0.3706 207	-0.12882 0.0637 208	-0.07743 0.2663 208	0.16890 0.0750 112	-0.05160 0.5292 151	1.00000 209	
pH-Wert (PH1K)	-0.05048 0.5715 128	0.04780 0.5921 128	0.09961 0.2671 126	-0.00248 0.9779 128	0.00393 0.9649 128	-0.15469 0.0889 122	-0.00854 0.9488 59	-0.51771 <.0001 112	1.00000 128

Aufgeführt sind: Pearsonscher Korrelationskoeffizient, p-Wert und Anzahl berücksichtigter Tiere

3.2 Genomstruktur der untersuchten Schweinepopulation

3.2.1 Allelfrequenzen von SNPs

Jeder SNP-Lokus kann homozygot mit einem Allel oder heterozygot mit zwei alternativen Allelen auftreten. Das Allel mit der geringeren Häufigkeit in der Population wird als Minderheitenallel bezeichnet. Die Frequenzen der Minderheitsallele (MAF), können für die Beurteilung der Selektionswürdigkeit eines SNPs von Bedeutung sein. Genomische Regionen, in denen Homozygotie auftritt oder ein Allel in geringer Häufigkeit vorkommt, können unter Selektionsdruck stehen und Selektionssignaturen markieren. Sie können aber auch durch Inzucht oder genetische Drift verursacht sein.

In der Population der erfolgreich genotypisierten Tiere tritt das Minderheitenallel mit Häufigkeiten der gesamten Breite von 0 (Homozygotie) bis 0,5 (gleiche Frequenz beider alternativer Allele) auf (Abbildung 2a). Die mittlere Frequenz der Minderheitenallele über alle SNPs betrug $0,27 \pm 0,14$.

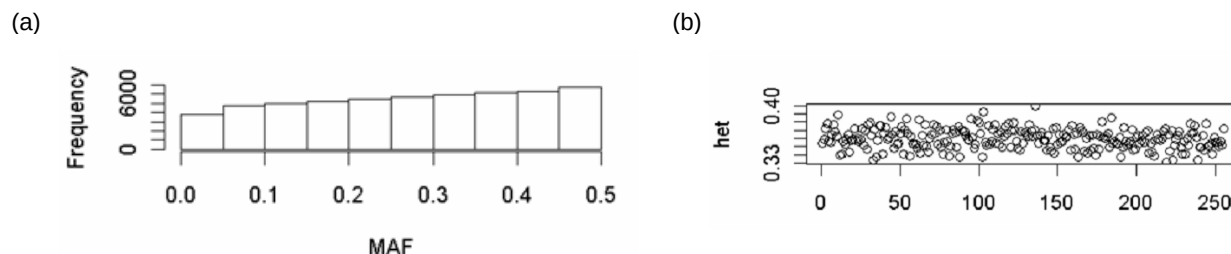


Abbildung 2: Verteilung der Minderheitenallelfrequenzen (MAF) (a) und Heterozygotiegrad aller Einzeltiere (b)

Aus den Genotypen eines Individuums über alle SNP-Loci lässt sich sowohl der Heterozygotiegrad eines Tieres als auch der Inzuchtgrad ermitteln. Der Heterozygotiegrad schwankte zwischen 0,33 und 0,40 und betrug im Mittel über alle genotypisierten Tiere $0,360 \pm 0,012$ (Abbildung 2b). Der daraus resultierende mittlere Homozygotiegrad über alle SNPs eines Tieres betrug entsprechend 0,64. Das spricht für eine geringe Verwandtschaft unter den genotypisierten Tieren.

Das spiegelt sich auch im mittleren Inzuchtkoeffizienten der genotypisierten Prüfeber wider. Der mittlere Inzuchtkoeffizient, der aus dem verfügbaren Pedigree der untersuchten Tiere berechnet wurde, betrug im Durchschnitt 1,2 %; er variierte zwischen 0,015 und 6,79 %.

Bei der Untersuchung der genomweiten Verteilung der Allelfrequenzen fiel auf, dass auf einigen Chromosomen ganze Abschnitte existieren, die durch eine sehr niedrige Minderheitenallelfrequenz ($MAF < 0,10$) charakterisiert sind (Abbildung 3). Auffallend viele solche Regionen befanden sich auf den Chromosomen 1, 7 und 17.

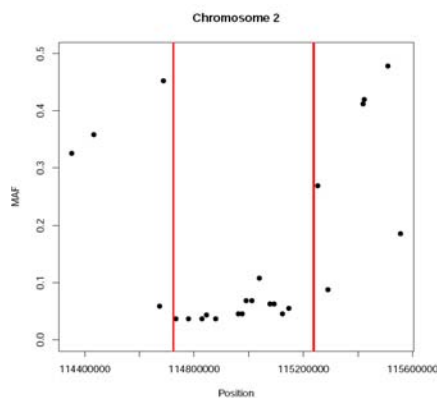


Abbildung 3: Beispiel für einen Abschnitt auf Chromosom 2 mit niedriger Minderheitenallelfrequenz (MAF) und mittlerer MAF im Nachbarabschnitt. Die Länge des Abschnitts mit geringer MAF ist etwa 400.000 Basenpaare lang.

3.2.2 Haplotypenblockstruktur

Haplotypenblockstrukturen geben einen Einblick in die Abstammungsgeschichte einer Population. Wurde ein Kopplungsfaktor D' von 0,8 für die Identifizierung eines Haplotypenblocks unterstellt, so ergaben sich in der vorliegenden Untersuchung 8.602 Haplotypenblöcke. Diese erstreckten sich über 57 % des Gesamtgenoms. Die meisten hatten eine Länge von nur wenigen Basenpaaren, am seltensten traten Blöcke mit einer Länge von 350-400 kb auf (Abbildung 4). Die Häufigkeit von Blöcken über 400 kb Länge stieg wieder etwas an. Insgesamt befanden sich 77 % aller untersuchten SNPs in Haplotypenblöcken.

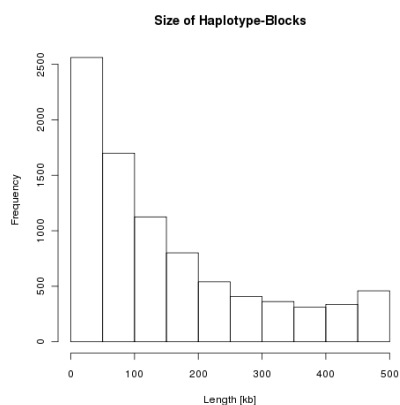


Abbildung 4: Histogramm für die Häufigkeitsverteilung von Haplotypenblöcken entsprechend ihrer Länge

3.3 Populationssubstrukturen

Populationssubstrukturen kennzeichnen Gruppen von Individuen, die untereinander enger verwandt sind als mit dem Rest der Population. Populationssubstrukturen können in allen Populationen vorkommen. Mit Hilfe der individuellen Genotypeninformationen kann die Verwandtschaft zwischen allen Individuen getestet werden. Unter Berücksichtigung der auftretenden Allelvarianten in jedem einzelnen Tier können Gruppen von Tieren (Cluster) gebildet werden, die die Struktur der Population besser widerspiegeln als die unstrukturierte Population. Diese Individuen ähneln sich hinsichtlich ihrer genetischen Ausstattung untereinander stärker als im Vergleich zum Rest der Population. Mathematisch wird dies in der genomischen Verwandtschaftsmatrix (genomic kinship) hinterlegt. Zur Darstellung genetischer Substrukturen wurde daraus eine Distanzmatrix abgeleitet und in weiteren Schritten zwei- oder mehrdimensional dargestellt (Abbildung 5). Die beiden Dimensionen D1 und D2 in Abbildung 5 korrespondieren mit den ersten beiden Achsen der genetischen Variation der Clusteranalyse. Solche Cluster haben ihren Ursprung in gemeinsamen Verwandten. Sie können auf gemeinsame historische Vorfahren für bestimmte Individuen oder Zukauf von Tieren aus anderen Zuchtgebieten zurückzuführen sein.

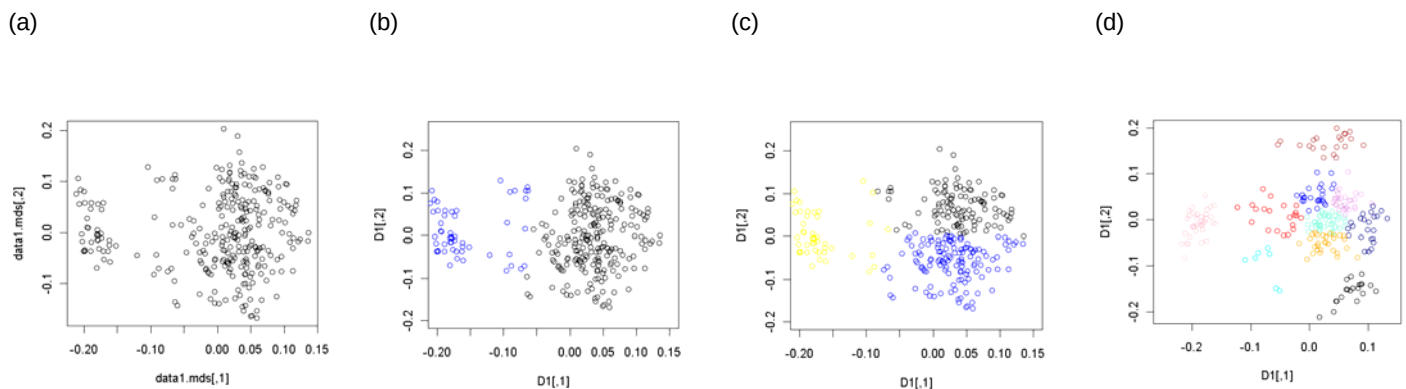


Abbildung 5: Darstellung von Populationssubstrukturen

Jeder Punkt in der Abbildung repräsentiert ein Tier. Je weiter die Punkte auseinander liegen, umso weniger sind die Eber genetisch miteinander verwandt. Unterstellt man zwei Cluster (Abbildung 5b), so gehören die blau gekennzeichneten Eber zu einer und die schwarz gekennzeichneten Eber zu einer anderen genetischen Gruppe. Weitere Cluster sind zu vermuten (Abbildung 5c-5d).

3.4 Genomweite Assoziationsanalysen für verschiedene Merkmale aus der Leistungsprüfung

3.4.1 Prüftagszunahme

Die mittlere Prüftagszunahme eines Tieres betrug 959 g pro Testtag mit einer Standardabweichung von 88 g. Das Minimum wurde mit 678 und das Maximum mit 1.247 g pro Testtag ermittelt. Für 267 Tiere waren Phänotypen vorhanden.

Ohne Berücksichtigung des multiplen Testens (Abbildung 6a) sind auf mehreren Chromosomen SNPs zu finden, deren phänotypische Unterschiede zwischen den Genotypenklassen signifikant verschieden sind und die in der Assoziationsanalyse p-Werte unter 10^{-4} aufweisen. Unter Berücksichtigung des multiplen Testens gibt es keine p-Werte mehr, die unter dem kritischen Grenzwert für genomweit signifikante Ergebnisse (Signifikanzgrenzwert nach Korrektur $p < 0,10$) liegen (Abbildung 6b). Verschiedene humane Studien haben jedoch gezeigt, dass Signifikanzwerte unter 10^{-4} in genomweiten Assoziationsanalysen ohne Korrektur auf multiples Testen durchaus biologisch relevante Zusammenhänge zeigen, die in unabhängigen Stichproben häufig bestätigt werden (SPELIOTES et al. 2010). Unter Berücksichtigung dieses Grenzwertes sind die in Abbildung 6b deutlich herausragenden SNPs durchaus als interessante SNPs, die mit dem Merkmal Prüftagszunahme gekoppelt sind, einzustufen. Das betrifft SNPs auf den Chromosomen 3, 4, 5, 9, 14 und 17.

Die höchste Signifikanz zeigte ein SNP auf Chromosom 3 (Abbildung 6b), der auch nach dem Permutationstest deutlich vor anderen Loci rangierte. An diesem SNP-Lokus sind homozygote Anlagenträger des Allels A den homozygoten Trägern des Allels G um 67 g täglicher Zunahme im Prüfzeitraum überlegen (Abbildung 6c). Der Alleleffekt ist überwiegend additiv. Die Frequenz des Minderheitenallels G beträgt 3,2 %. Dadurch tritt auch der Genotyp GG mit der geringeren Prüftagszunahme nur mit einer Frequenz von 10 % auf. Die geringe Frequenz führt zu einem höheren Schätzfehler für die Phänotypenleistung der Tiere dieser Genotypenklasse.

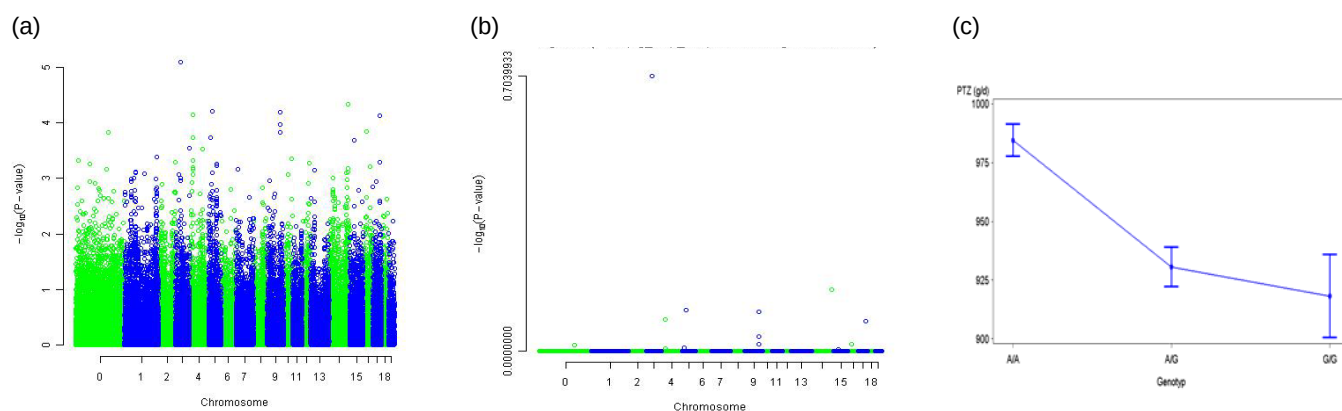


Abbildung 6: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Prüftagszunahme und genotypische Unterschiede in der Prüftagszunahme

- Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- Darstellung der Genotypeneffekte bei 267 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=147, AG=98, GG=22; Allelfrequenzen: A=0,73 G=0,27)

3.4.2 Lebensstagszunahme

Die mittlere Lebensstagszunahme eines Prüfebers betrug 689 g pro Tag mit einer Standardabweichung von 49 g, einem Minimum von 558 g und einem Maximum von 841 g pro Tag und Tier in der Population. Von 266 Tieren waren Daten verfügbar. Die höchsten Signifikanzen wiesen in der genomweiten Assoziationsstudie zwei SNPs auf Chromosom 13 auf (Abbildung 7). Beide SNPs liegen nur 59.195 Basenpaare voneinander getrennt, sodass sie eng gekoppelt vererbt werden. Von beiden Markern geht die gleiche Information aus. Ohne Korrektur auf multiples Testen weisen beide SNPs exakt die gleichen p-Werte $<10^{-4}$ auf. Die Allelfrequenzen beider SNPs sind ebenfalls gleich. Die Frequenz des Minderheitenallels beträgt 2,5 %. Ein weiterer SNP auf Chromosom 14 und ein SNP, dessen Position im Genom unbekannt ist, sind ebenfalls mit p-Werten $<10^{-4}$ mit der Lebensstagszunahme assoziiert. Führt man jedoch die Korrektur der p-Werte auf multiples Testen durch, bleibt keiner der SNPs signifikant (Abbildung 7b).

Der Alleleffekt der beiden signifikantesten SNP auf Chromosom 13 ist überwiegend additiv. Homozygote Träger des Minderheitenallels C haben gegenüber homozygoten Trägern des Allels A eine um 28 g höhere Lebensstagszunahme (Abbildung 7c). Die Sicherheit der Schätzwerte für den wertvolleren CC-Genotyp ist jedoch gering, weil nur 11 Genotypen vorhanden waren. Sollte der CC-Genotyp in der gesamten Population des Deutschen Landschweins ebenfalls nur selten vertreten sein und die Effekte in ähnlicher Größenordnung vorliegen, dann wären an diesem Locus besonders große Selektionsfortschritte im Merkmal Lebensstagszunahme zu erwarten, wenn auf das Allel C selektiert wird.

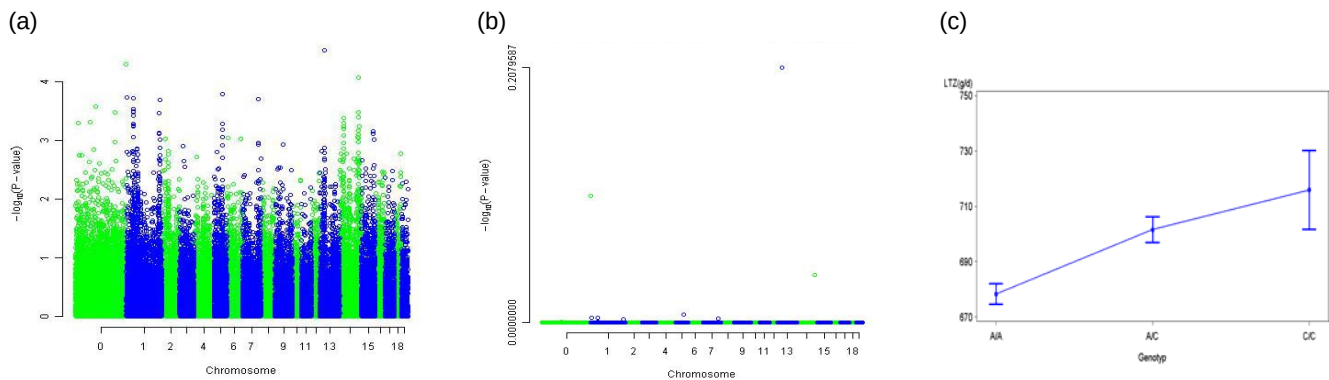


Abbildung 7: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Lebensstagszunahme und genotypische Unterschiede in der Lebensstagszunahme

- Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- Darstellung der Genotypeneffekte bei 266 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=154, AC=101, CC=11, Allelfrequenzen: A=0,77 C=0,23)

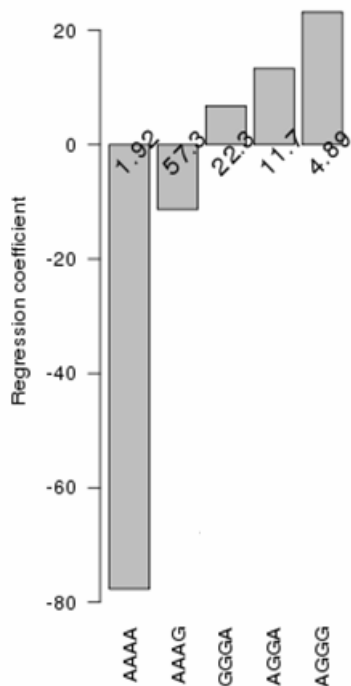


Abbildung 8: Darstellung der Haplotypeneffekte im Merkmal Lebensstagszunahme für den Haplotypenblock, der den signifikantesten SNP enthält

Weil Haplotypenblöcke aus mehreren SNPs gebildet werden, können zwei oder mehr als zwei Haplotypen in einem Haplotypenblock auftreten. Haplotypen, die seltener als 1 % in der Population auftraten, wurden eliminiert. Die Häufigkeiten für jeden Haplotypen sind entlang der x-Achse angegeben.

Über das gesamte Genom betrachtet sind auf den Chromosomen 13 und 14 SNPs lokalisiert, die ohne Berücksichtigung des multiplen Testens p-Werte unter 10^{-4} aufweisen und damit für die Selektion auf die Lebensstagszunahme von Bedeutung sein können. Darüber hinaus sind aber auch auf den Chromosomen 1, 5 und 7 Anhäufungen von SNPs mit geringem p-Wert zu finden, die ein Hinweis darauf sein können, dass auch in diesen Chromosomenregionen Gene mit Einfluss auf die Lebensstagszunahme liegen (Abbildung 7). Lebens- und Prüftagszunahme sind Merkmale, die sehr eng miteinander korrelieren

(Tabelle 3, $r=0,80$) und nicht als unabhängig voneinander betrachtet werden können. Aus den jeweiligen Manhattan-Plots für die Prüftagszunahme und Lebenstagszunahme (Abbildungen 6a und 7a) lässt sich eine gemeinsame genetische Basis erkennen. So zeigen die Muster der aufgetragenen p-Werte auf den Chromosomen 1 und 14 auffällige Übereinstimmungen. SNPs auf dem Chromosom 4, die für die Prüftagszunahme von Bedeutung waren, fehlen dagegen bei der Lebenstagszunahme. Möglicherweise sind hier Gene lokalisiert, die außerhalb des Prüfabschnitts am stärksten aktiv sind.

Am Beispiel des Merkmals Lebenstagszunahme sollen neben den SNP- auch Haplotypeneffekte vorgestellt werden (Abbildung 8). Die Analyse der Haplotypenblockstruktur hat gezeigt, dass die beiden signifikanten SNPs auf Chromosom 13 tatsächlich im Kopplungsungleichgewicht stehen. Zusammen mit zwei weiteren SNPs bilden sie einen Haplotypenblock, für den fünf verschiedene Haplotypen mit einer Frequenz größer 0,01 in der untersuchten Population nachgewiesen wurden (Abbildung 8). Die fünf Haplotypen kommen mit Häufigkeiten von 57,3; 22,8; 11,7; 4,88 und 1,92 vor. Der Effekt des Haplotypenblocks war hochsignifikant ($p = 1,1 \cdot 10^{-6}$). Die Effekte der einzelnen Haplotypen auf die Lebenstagszunahme variieren zwischen +20g und -80 g pro Tag. Der am häufigsten vorkommende Haplotyp reduziert die Lebenstagszunahme um 10 g im Vergleich zum Populationsdurchschnitt. Der selten auftretende Haplotyp AAAA bremst die Lebenstagszunahme drastisch um -80 g pro Tag. Die anderen drei Haplotypen tragen zu einer überdurchschnittlichen Lebenstagszunahme bei. Ebenso wie bei Allelen muss auch bei den Haplotypen berücksichtigt werden, dass die Schätzwerte für seltene Allele und Haplotypen ungenau sind.

3.4.3 Futterverzehr

Neben der Prüftags- und Lebenstagszunahme ist auch der Futterverzehr während der Eigenleistungsprüfung im Teilzuchtwert Mastleistung integriert. Im Durchschnitt nahmen die Tiere hier 200 kg Futter auf, mit Extremwerten zwischen 69 und 330 kg. Die Standardabweichung betrug 43 kg. Von 265 Tieren lagen Daten vor.

Es fällt auf, dass unter den bislang nicht genomisch lokalisierten SNPs und die SNPs auf Chromosom 2 und auf Chromosom 17 in der genomweiten Assoziationsanalyse p-Werte unter 10^{-4} aufweisen. Die Genomregionen, die für die Prüftagszunahme und die Lebenstagszunahme eine Rolle spielen, besitzen für die Futteraufnahme eher eine untergeordnete Bedeutung. Auch zeigen die Muster der abgebildeten p-Werte keine markanten Übereinstimmungen mit vorangegangenen Mastleistungsmerkmalen. Daher ist es wahrscheinlich möglich, den Futterverzehr züchterisch weitgehend unabhängig von der Prüftagszunahme und Lebenstagszunahme zu bearbeiten. Wenngleich die phänotypischen Korrelationskoeffizienten (Tabelle 2) nicht der wirkliche Maßstab für eine solche Bewertung sind, geben auch sie Hinweise auf eine relative lose Beziehung zwischen den Körpermassezunahmen und der Futteraufnahme.

Der SNP mit dem kleinsten p-Wert in der genomweiten Assoziationsanalyse wurde auf Chromosom 17 identifiziert. Homozygote Anlagenträger des Allels A haben im Prüfabschnitt 184 kg Futter aufgenommen, homozygote Träger des alternativen Allels G 213 kg und damit 29 kg mehr. Übereinstimmend mit einer ausgeglichenen Frequenz beider Allele (Minderheitenallelfrequenz = 0,47) treten heterozygote Genotypen am häufigsten auf (0,45). Heterozygote Genotypen realisierten im Prüfabschnitt eine Futteraufnahme von 201 kg. Damit lagen sie sehr nahe am Mittel zwischen den homozygoten Gruppen. Folglich kann für diesen SNP eine additive Allelwirkung angenommen werden.

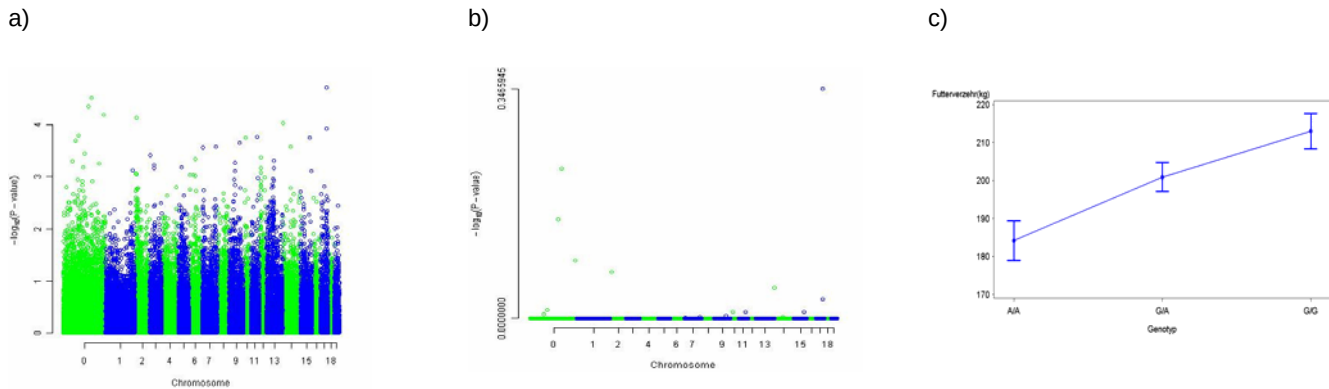


Abbildung 9: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Futterverzehr und genotypische Unterschiede im Futterverzehr

- a) Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- b) wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- c) Darstellung der Genotypeneffekte bei 265 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=64, GA=120, GG=81; Allelfrequenzen: A=0,47 G=0,53)

3.4.4 Ultraschall-Fettdicke

Im Mittel betrug die Ultraschall-Fettdicke 13 mm. Sie wies eine recht hohe Variabilität mit einer Standardabweichung von 3 mm auf. Aus Tabelle 1 ist ersichtlich, dass sie über einen weiten Bereich von 5 bis 23 mm variierte. Messwerte für die Ultraschall-Fettdicke lagen von 266 Tieren vor.

Dem Chromosom 1 ist bei der phänotypischen Ausprägung der Ultraschall-Fettdicke eine besondere Bedeutung beizumessen (Abbildung 10). Auf diesem Chromosom wurden fünf SNPs mit den kleinsten p-Werten ($p < 10^{-4}$) nachgewiesen. Die Schätzwerte für die Effekte des SNP mit der höchsten Signifikanz auf Chromosom 1 waren innerhalb eines engen Bereiches angesiedelt. Der AA-Genotyp wies eine mittlere Ultraschall-Fettdicke von 14 mm auf. Die AG- und GG-Genotypen hatten 12,8 und 11,9 mm Ultraschall-Fettdicken, was wiederum auf einen additiv genetischen Alleleffekt hinweist. Weitere Gene mit Einfluss auf die Ultraschall-Fettdicke könnten auf den Chromosomen 3, 4, 5 und 7 liegen, weil hier in engen Chromosomenregionen SNPs mit geringen p-Werten (aber nicht $< 10^{-4}$) liegen.

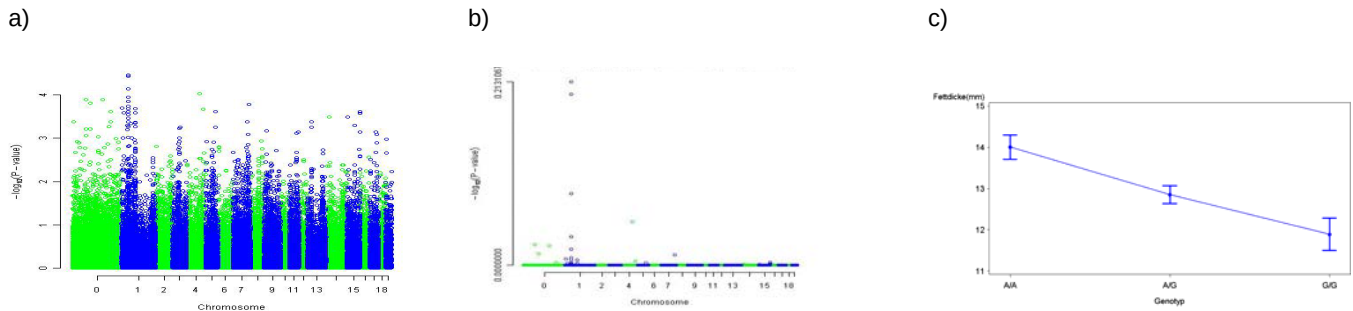


Abbildung 10: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Ultraschall-Fettdicke und genotypische Unterschiede in der Ultraschall-Fettdicke

- a) Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- b) wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- c) Darstellung der Genotypeneffekte bei 266 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=80, AG=141, GG=45; Allelfrequenzen: A=0,57 G=0,43)

3.4.5 Fettfläche

Die Fettfläche steht in enger phänotypischer Beziehung zur Ultraschall-Fettdicke (Tabelle 2, $r=0,58$). Die mittlere Fettfläche aller 236 geschlachteten Eber betrug 17,7 cm². Die Standardabweichung von 5 cm², ein Minimum von 5 cm² und ein Maximum von 36 cm² signalisieren eine hohe Variabilität des Merkmals.

Aus Abbildung 11 geht hervor, dass vor allem die Chromosomen 1,13 und 15 mit hoher Wahrscheinlichkeit Gene tragen, die mit der Fettfläche eng assoziiert sind. Auf Chromosom 1 ergaben sich in allen Modellvarianten (mit und ohne Stratifikation) die stabilsten Ergebnisse mit den geringsten p-Werten. Von den zehn signifikantesten Marker-SNPs aus der genomweiten Assoziationsanalyse lagen allein sieben auf diesem Chromosom (Tabelle 4). Darüber hinaus weisen offenbar auch eine Reihe von SNPs auf nicht zuordenbaren Chromosomen (Chromosom 0) einen engeren Bezug zu dem Merkmal auf.

Der SNP mit der höchsten Signifikanz lag auf Chromosom 15. Der günstigste Genotyp GG besaß im Durchschnitt eine Fettfläche von 16 cm², der ungünstigste Genotyp AA eine Fettfläche von 19 cm² (Abbildung 14c). Der Alleleffekt ist additiv.

Auch im Falle der Fettfläche ließ sich ein deutlicher Effekt des Haplotypenblocks nachweisen. Am signifikantesten war der Haplotypenblock ($p=7,4 \cdot 10^{-6}$), in dem sich der signifikanteste SNP befindet. Der Haplotypenblock wird von sieben SNPs gebildet. Sechs Haplotypen mit Frequenzen größer 1 % waren in der untersuchten Population vorhanden. Der günstigste Haplotyp (GAGGAAG) hatte einen reduzierenden Effekt auf die Fettfläche von -2,5 cm² und der ungünstigste Haplotyp erhöhte die Fettfläche um +2,5 cm². Der Haplotyp, der die Fettfläche reduziert, kommt in der Population mit der hohen Frequenz von 34,3 % vor, der Haplotyp mit der höchsten Fettfläche kommt mit einer Frequenz von 1,94 % vor, wird aber von dem ebenfalls über der durchschnittlichen Fettfläche liegenden häufigsten Haplotyp gefolgt (49,5 %).

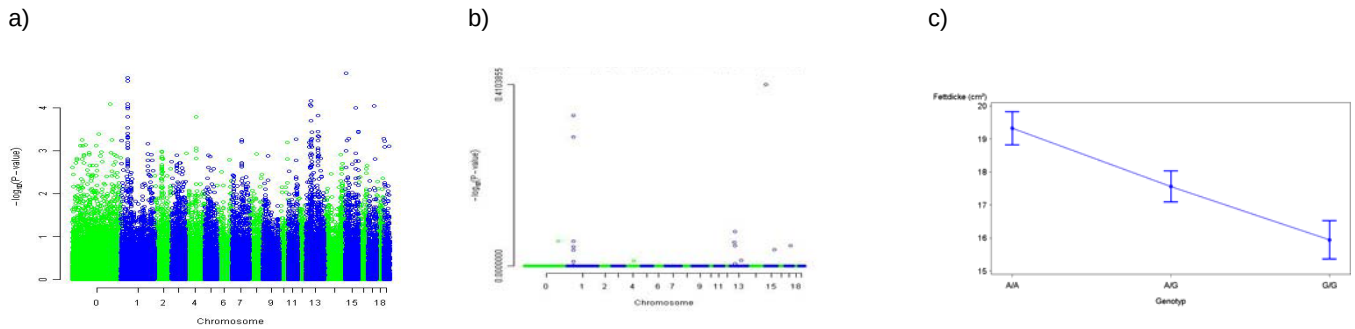


Abbildung 11: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Fettfläche und genotypische Unterschiede in der Fettfläche

- a) Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- b) wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- c) Darstellung der Genotypeneffekte bei 235 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=81, AG=94, GG=60; Allelfrequenzen: A=0,54 G=0,46)

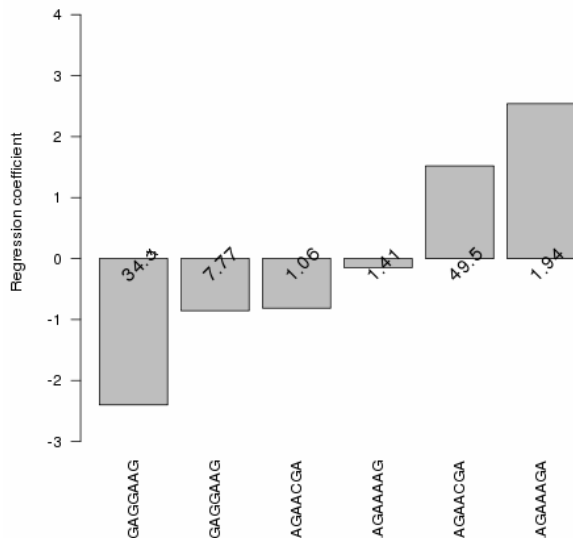


Abbildung 12: Darstellung der Haplotypeneffekte im Merkmal Fettfläche für den Haplotypenblock, der den signifikantesten SNP enthält

Weil Haplotypenblöcke aus mehreren SNPs gebildet werden, können zwei oder mehr als zwei Haplotypen in einem Haplotypenblock auftreten. Haplotypen, die seltener als 1 % in der Population auftraten, wurden eliminiert. Die Häufigkeiten für jeden Haplotypen sind entlang der x-Achse angegeben.

Am Beispiel der Fettfläche kann sehr gut gezeigt werden, dass die Populationssubstruktur für die untersuchten Tiere kaum einen Einfluss auf die Assoziationsanalysen hat. Einige SNPs sind sowohl im Modell mit als auch ohne Stratifikation von besonderer Bedeutung für die Merkmalsausprägung. In Tabelle 4 sind die zehn signifikantesten SNPs hinsichtlich ihrer Rangfolge für den p-Wert bei Verwendung verschiedener Modelle mit und ohne Populationsstratifikation aufgelistet. So hatte SNP1, der im Modell mit Stratifikation den niedrigsten p-Wert aufwies, im Modell ohne Stratifikation den zweitniedrigsten p-Wert. Für SNP2, der im Modell mit Stratifikation den zweitniedrigsten p-Wert hatte, wurde unter Berücksichtigung der genetischen Verwandtschaftsstruktur der drittniedrigste p-Wert ermittelt.

Tabelle 4: p-Werte und Rangfolge von SNPs in Modellen mit und ohne Berücksichtigung der Populationsstruktur (Stratifikation)

SNP (Chromosom)		p-Wert Modell mit Stratifikation	Rang Modell mit Stratifikation	p-Wert Modell ohne Stratifikation	Rang Modell ohne Stratifikation
SNP1	(Chr 15)	0,0000194	1	0,0000201	2
SNP2	(Chr 1)	0,0000207	2	0,0000237	3
SNP3	(Chr 1)	0,0000207	2	0,0000237	9
SNP4	(Chr 1)	0,0000213	3	0,0000912	14
SNP5	(Chr 1)	0,0000262	4	0,0000238	5
SNP6	(Chr 1)	0,0000262	4	0,0000238	4
SNP7	(Chr 13)	0,0000620	5	0,0000159	1
SNP8	(Chr 0)	0,0000194	6	0,0002928	17
SNP9	(Chr 1)	0,0000207	7	0,0005666	37
SNP10	(Chr 1)	0,0000826	8	0,0001385	7

3.4.6 Ultraschall-Muskeldicke

Für die Analyse der Ultraschall-Muskeldicke standen nur 148 Tiere zur Verfügung, sodass die Sicherheiten der vorliegenden Ergebnisse im Vergleich zu den vorangegangenen Merkmalen noch geringer sind. Für die Muskeldicke wurde ein Mittelwert von 52 mm ermittelt. Die Extremwerte betrugen 43 und 59 mm, die Standardabweichung 2 mm. Damit war die Ultraschall-Muskeldichte im vorliegenden Datenmaterial deutlich weniger variabel als die beiden Fettmerkmale Ultraschall-Fettdicke und Fettfläche.

Die Ergebnisse der genomweiten Assoziationsanalyse (Abbildung 13) zeigten, dass mit den Chromosomen 6, 7 und 13 SNPs liegen, die p-Werte unter 10^{-4} aufwiesen. Auf diesen Chromosomen könnten Gene lokalisiert sein, die bei der Ausprägung des Merkmals Ultraschall-Muskeldicke von Bedeutung sein könnten. Der niedrigste p-Wert wurde für einen SNP auf einem nicht zuordenbaren Chromosom gefunden. Für diesen SNP hatte der Genotyp AA eine Ultraschall-Muskeldichte von 51,6 mm, die heterozygoten AG Genotypen hatten eine Dicke von 52,9 mm und die GG-Genotypen von 54,1 mm. Nur 18 Tiere (12 %) besaßen den vorteilhaften GG-Genotyp. Der Alleleffekt war additiv.

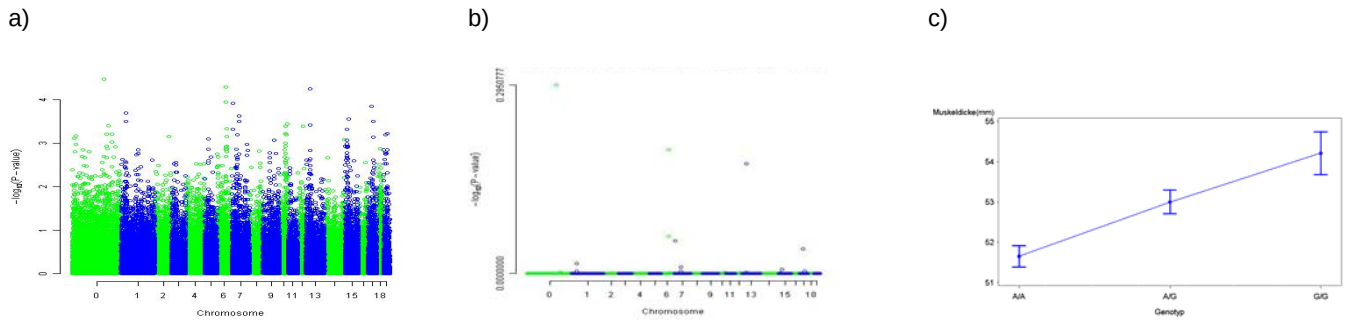


Abbildung 13: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Ultraschall-Muskeldicke und genotypische Unterschiede in der Ultraschall-Muskeldicke

- a) Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifizierung). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- b) wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- c) Darstellung der Genotypeneffekte bei 148 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=73, AG=57, GG=18; Allelfrequenzen: A=0,69 G=0,31)

3.4.7 Intramuskulärer Fettgehalt

Der intramuskuläre Fettgehalt lag von 158 Ebern vor. Er betrug im Durchschnitt 1 % und variierte zwischen 0,01 und 1,8 %. Die Standardabweichung betrug 5 %. Das Merkmal musste logarithmiert werden, um annähernd normal verteilte Werte für die Assoziationsanalysen zu erhalten.

Die Ergebnisse der genomweiten Assoziationsanalyse zeigten, dass zwei SNPs auf Chromosom 8 die niedrigsten p-Werte unter 10^{-4} aufwiesen (Abbildung 14). Weitere niedrige p-Werte wurden auch für SNPs auf nichtzuordenbaren Chromosomen (Chromosom 0) und auf dem X-Chromosom gefunden.

Der signifikanteste SNP liegt auf Chromosom 8. Für AA-Genotypen wurde ein intramuskulärer Fettgehalt von 1,5 % ermittelt, für AG-Genotypen waren es 0,9 % und für GG-Genotypen 0,8 %. Weil der heterozygote Genotyp etwa dem Mittelwert aus den beiden homozygoten Genotypen entsprach, kann wiederum ein überwiegend additiver Erbgang unterstellt werden. Die Selektion auf den AA Genotyp könnte zur Erhöhung des intramuskulären Fettgehalts beitragen.

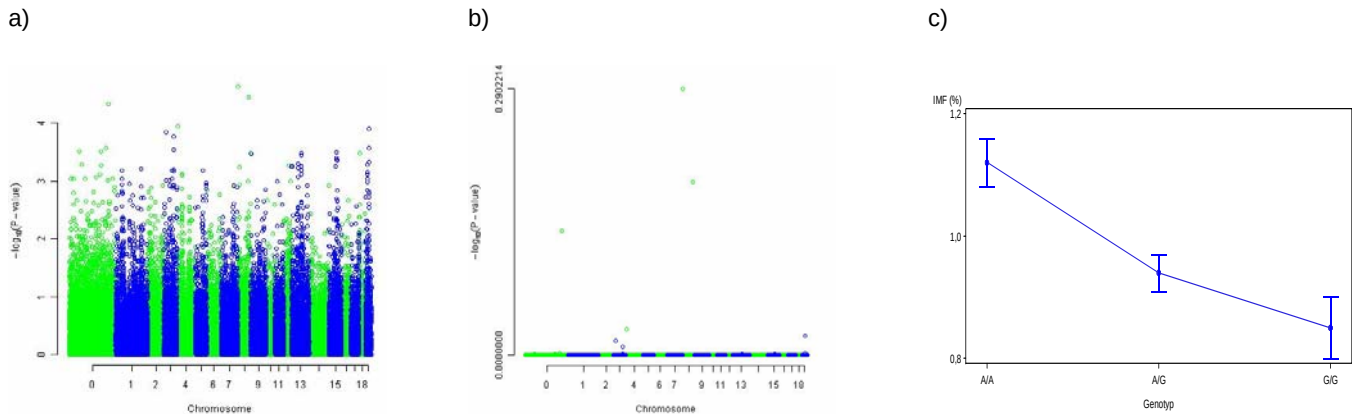


Abbildung 14: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal intramuskulärer Fettgehalt und genotypische Unterschiede im intramuskulären Fettgehalt

- Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- Darstellung der Genotypeneffekte bei 158 Tieren am SNP mit dem signifikantesten p-Wert (Anzahl Tiere in den Genotypenklassen: AA=57, AG=78, GG=23; Allelfrequenzen: A=0,61 G=0,39)

3.4.8 Dripverlust

Dripverlust und pH-Wert sind wichtige Eigenschaften, die die Muskel- und Fleischqualität bestimmen. Der geringste Dripverlust betrug 0,6 %, der höchste 9,1 %. Über alle Tiere wurde ein durchschnittlicher Dripverlust von 4 % ermittelt. Mit 1,7 % war die Standardabweichung sehr hoch.

Die signifikantesten SNPs mit p-Werten unter 10^{-4} waren auf Chromosom 17 zu finden (Abbildung 15). Am signifikantesten SNP-Lokus auf Chromosom 17 betrug die Minderheitenallelfrequenz 19,9 %. Daraus resultiert, dass nur elf homozygote AA-Tiere (5,3 %) in der Population vorhanden waren. Diese Tiere weisen den höchsten Dripverlust von durchschnittlich 5,3 % auf. Die geringe Anzahl homozygote AA-Träger führt dazu, dass der phänotypische Effekt dieser Genotypenklasse mit einem hohen Standardfehler behaftet ist. Der Alleleffekt ist additiv (Abbildung 15c). Jedes negative Allel A führt zu einer Verschlechterung der Muskelqualität durch erhöhten Dripsaftverlust von 0,8 %. Die signifikante Differenz zwischen den Genotypengruppen lässt unter Einbeziehung dieses SNPs jedoch gute Selektionschancen für die weitere Verbesserung der Fleischqualität erwarten.

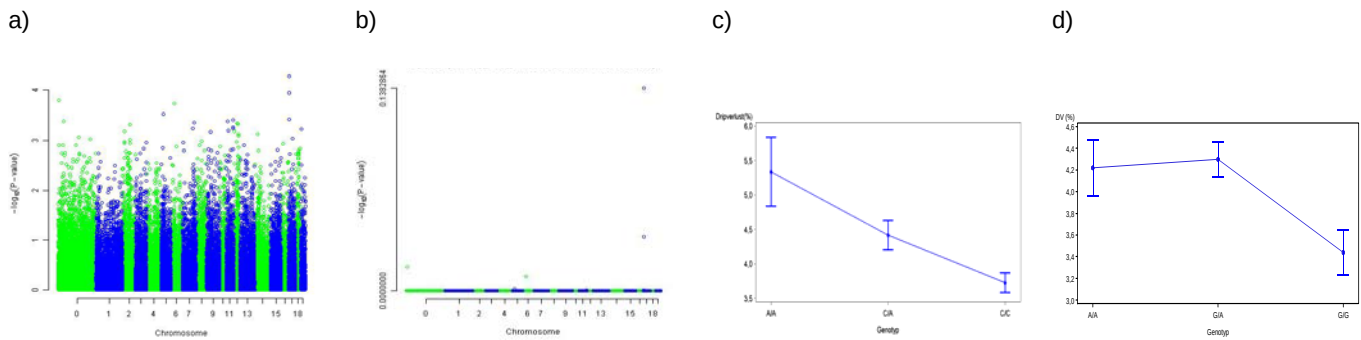


Abbildung 15: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal Dripverlust und genotypische Unterschiede im Dripverlust

- a) Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- b) wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- c) Darstellung der Genotypeneffekte für Dripverlust bei 209 Tieren am SNP mit dem signifikantesten p-Wert für Dripverlust (Anzahl Tiere in den Genotypenklassen: AA=11, CA=61, CC=137; Allelfrequenzen: A=0,20 C=0,80)
- d) Darstellung der Genotypeneffekte für Dripverlust bei 209 Tieren am SNP mit dem signifikantesten p-Wert im Merkmal pH-Wert (Anzahl Tiere in den Genotypenklassen: AA=41, GA=102, GG=66; Allelfrequenzen: A=0,44 C=0,56)

3.4.9 pH-Wert 1. Kontrolle

Für die Analyse von SNP-Effekten auf den pH-Wert standen 128 Tiere zur Verfügung. Der pH-Wert im Kotelett zur ersten Kontrolle 45 Minuten nach dem Schlachten betrug im Durchschnitt 6,40 mit einer Standardabweichung von 0,18. Der kleinste pH-Wert unter den untersuchten Tieren lag bei 6, der größte bei 6,9.

Auf Chromosom 15 lagen drei SNPs mit p-Werten unter 10^{-4} , wenn keine Korrektur auf multiples Testen erfolgte (Abbildung 16). Für die Darstellung der genetischen Effekte wurde der markanteste SNP auf Chromosom 15 ausgewählt (Abbildung 16). Im Vergleich der Schätzwerte für die drei Genotypen an diesem signifikantesten SNP zeigte sich, dass der pH-Wert von 6,58 im Genotyp GG um 0,16 bzw. 0,22 Einheiten höher liegt und damit aus der Sicht der Fleischqualität besser ist als die beiden anderen Genotypen GA und AA. Der Genotyp GG wäre daher bei der Selektion zu bevorzugen. Die Allelfrequenzen sind ausgewogen (0,5). Das widerspiegelt, dass diese Genomregion bislang offenbar nicht unter Selektionsdruck stand. Chromosom 15 könnte für nachfolgende Forschungsarbeiten zur Verbesserung des Wasserbindevermögens im Muskel, insbesondere im Kotelett, aber auch wie für die praktische Nutzenanwendung zur Selektion bevorzugter Genotypen von besonderem Interesse sein.

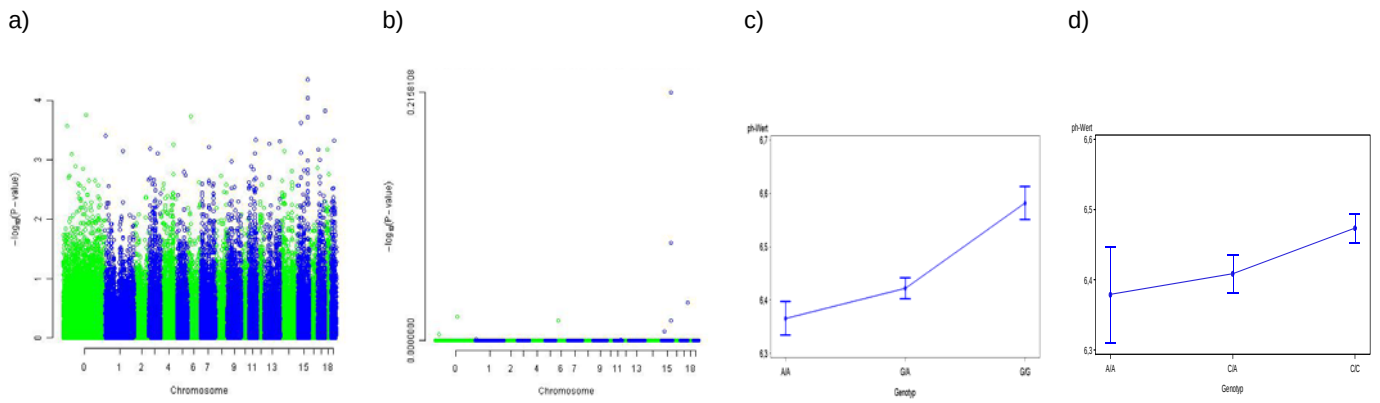


Abbildung 16: Ergebnisse der genomweiten Assoziationsanalyse für das Merkmal pH-Wert und genotypische Unterschiede im pH-Wert

- Darstellung der p-Werte für alle Marker im Genom im Manhattan-Plot (Modell mit Populationsstratifikation). Die p-Werte sind entsprechend der genomischen Positionen der SNPs auf den Chromosomen 1 bis 19 sowie X angeordnet. „0“ repräsentiert nicht-kartierte SNPs.
- wie (a), aber nach Korrektur der p-Werte auf multiples Testen
- Darstellung der Genotypeneffekte für den pH-Wert bei 128 Tieren am SNP mit dem signifikantesten p-Wert für den pH-Wert (Anzahl Tiere in den Genotypenklassen: AA=28, GA=71, GG=29; Allelfrequenzen: A=0,50 G=0,50)
- Darstellung des Genotypeneffekts bei 128 Tieren am SNP mit dem signifikantesten p-Wert im Merkmal Dripverlust (Anzahl Tiere in den Genotypenklassen: AA=7, CA=44, GG=77; Allelfrequenzen: A=0,23 G=0,77)

Für die mit dem pH-Wert assoziierte Region auf Chromosom 15 wurde ein Haplotypenblock gefunden, in dem der signifikanteste und zwei weitere signifikante ($p < 10^{-4}$) SNPs liegen. Der Haplotypenblock zeigte in der Assoziationsanalyse ebenfalls p-Werte unter 10^{-4} . In diesem Haplotypenblock treten in der Population vier Haplotypen mit einer Frequenz größer 1 % auf. Diese vier Haplotypen haben im Kotelett 45 Minuten nach der Schlachtung pH-Werte, die im Mittel um 0,25 pH-Einheiten variieren (Abbildung 18). Der am häufigsten auftretende Haplotyp (AAG) zeigt pH-Werte unter dem Durchschnitt der Population. Zwei Haplotypen (AGA und GGA) sind mit den höheren pH-Werten assoziiert. Die Haplotypen geben ein differenzierteres Bild für genetische Effekte dieser Genomregion auf den pH-Wert wieder als jeder SNP einzeln.

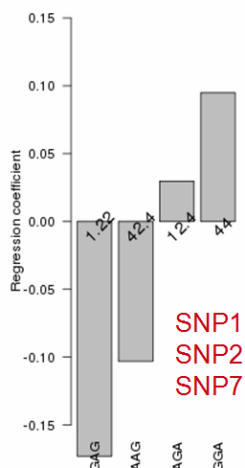


Abbildung 17: Darstellung der Haplotypeneffekte im Merkmal ph-Wert für den Haplotypenblock auf Chromosom 15, der den signifikantesten SNP enthält

Weil Haplotypenblöcke aus mehreren SNPs gebildet werden, können zwei oder mehr als zwei Haplotypen in einem Haplotypenblock auftreten. Haplotypen, die seltener als 1 % in der Population auftraten, wurden eliminiert. Die Häufigkeiten für jeden Haplotypen sind entlang der x-Achse angegeben.

Zwischen Dripverlust und dem pH-Wert des Fleisches wurde eine negative Beziehung nachgewiesen (Tabelle 2: $r=-0.52$). Hohe pH-Werte gehen mit einem geringeren Dripsaft einher. Diesbezüglich haben wir die beiden signifikantesten Marker für Dripsaft und pH-Wert angesehen. In der Tat ist der Genotyp GG am signifikantesten SNP für den pH-Wert mit hohem pH-Wert (Abbildung 16c) und geringem Dripsaftverlust (Abbildung 15d) assoziiert. Der Effekt vom Allel A ist für den pH-Wert überwiegend additiv, für Dripsaftverlust jedoch dominant. Ebenfalls eine negative Beziehung zwischen hohem Dripsaftverlust und geringem pH-Wert wurde für den Genotyp AA am signifikantesten SNP für den Dripverlust nachgewiesen (Abbildungen 15 c und 16 d). Hier spiegeln die genetischen Effekte auf die verschiedenen Merkmale sehr gut die beobachtete Korrelation wider. Zwischen Dripverlust und dem pH-Wert des Fleisches wurde eine negative Beziehung nachgewiesen (Tabelle 2: $r=-0.52$). Trotz der engen Korrelation zwischen dem Dripverlust und dem pH-Wert der ersten Kontrolle war keiner der fünf SNPs mit dem niedrigsten p-Wert für den Dripverlust auch unter den fünf niedrigsten für den pH-Wert. Entweder sind für diese Merkmalsbeziehung nichtgenetische Faktoren verantwortlich zu machen oder eine Reihe eher unauffälliger bzw. nicht entdeckter Loci war an der Ausprägung beider Merkmale beteiligt.

4 Diskussion

Die vorliegenden Untersuchungen verfolgten das Ziel, anhand einer Stichprobe der Deutschen Landrasse die genetische Variabilität zu charakterisieren und die molekulargenetischen Voraussetzungen für die genomische Selektion auf Mast- und Schlachtleistungsmerkmale zu analysieren. Neben der Charakterisierung von SNPs und Haplotypenblöcken wurden Assoziationsanalysen sowohl zwischen SNP-Genotypen als auch Haplotypenblöcken und phänotypischen Merkmalen der Leistungsprüfung durchgeführt.

Die hier präsentierten Ergebnisse stellen erstmals genomweite Untersuchungen in der Deutschen Landrasse vor. Bei der Bewertung der erzielten Ergebnisse muss berücksichtigt werden, dass bis heute beim Schwein nur wenig vergleichbare Untersuchungen durchgeführt oder gar veröffentlicht wurden und deshalb nur sehr begrenzt Referenzwerte bestehen, mit denen die erhaltenen Ergebnisse verglichen werden können.

Die Untersuchungen zur genetischen Diversität zeigen, dass die untersuchte Population eine sehr gute Vielfaltigkeit an segregierenden Allelen in der Population aufweist. Der durchschnittliche Heterozygotiegrad von 0,36 pro Tier für alle untersuchten SNPs entspricht Beobachtungen, die auch beim Rind gefunden wurden. Auch die Allelfrequenzen für das Minderheitenallel über alle untersuchten SNPs mit einer Schwankungsbreite von 0 bis 0,5 und einem mittleren Wert von 0,27 entsprechen den Erwartungen. In Holland wurden z. B. bei Untersuchungen von 234 Ebern der Vatterrasse Pietrain und 288 Sauen der Landrasse als Mutterrasse mittlere Minderheitenallelfrequenzen von 0,274 bzw. 0,287 gefunden (HUISMAN et al. 2010). Obwohl sich die veröffentlichten Daten nur auf Chromosom 1 beziehen, stimmen sie sehr gut mit den beobachteten Minderheitenallelfrequenzen in der sächsischen Deutschen Landrasse überein. Die mittlere Minderheitenallelfrequenz hängt maßgeblich von der Populationsgröße ab und dem Selektionsdruck auf das gesamte Genom. Die gute Übereinstimmung zwischen der holländischen Landrasse und der sächsischen Landrasse weisen auf ähnliche Populationsgrößen und Selektionsdrücke in beiden Populationen hin.

Einen detaillierten Blick auf die Genomstruktur gestattet die Verteilung von Allelfrequenzen der verschiedenen SNPs im Genom. Diese Untersuchungen haben gezeigt, dass im Genom Regionen unterschiedlicher Länge vorhanden sind, in denen SNPs mit geringen Frequenzen des Minderheitenallels auftraten. Interessanterweise waren diese Regionen von scharfen Grenzen umgeben, hinter denen sich Regionen mittlerer Allelfrequenzen anschlossen. Genombereiche mit geringer Allelfrequenz des Minderheitenallels können durch Selektion, aber auch Drift verursacht sein. Die scharfe Abgrenzung der Regionen geringer Minderheitenallelfrequenz zu Nachbarregionen mit durchschnittlicher Allelfrequenz ist jedoch ein sicherer Hinweis darauf, dass eine Basiszuchtpopulation bestand, in die bereits vor vielen Generationen weitere Tiere eingekreuzt wurden. In der

Basiszuchtpopulation haben diese markanten Genomregionen mit geringer Minderheitenallelfrequenz möglicherweise unter Selektionsdruck gestanden.

Einen weiteren Hinweis auf die Geschichte der Population gibt die Haplotypenblockstruktur. In der Population treten sehr viele sehr kurze Haplotypenblöcke auf. Diese entstehen, wenn innerhalb einer relativ geschlossenen Population unter Vermeidung von Inzucht verpaart wird. Das wird für Reinzuchten wie der untersuchten Deutschen Landrassepopulation erwartet. Die Haplotypenblockstruktur weist aber auch eine höhere Anzahl sehr langer Haplotypenblöcke auf. Es wurden Haplotypenblöcke bis zu 500 kb Länge gefunden. Das wäre innerhalb einer geschlossenen Population nicht zu erwarten. Die langen Haplotypenblöcke sind Ausdruck eines hohen Verwandtschaftsgrades. Sie könnten auf Einkreuzungen durch Zukauf aus anderen Zuchtgebieten in der jüngeren Zuchtgeschichte zurückzuführen sein. Unter den Haplotypenblöcken wurden auch solche gefunden, in denen geringe Minderheitenallelfrequenzen für bestimmte Haplotypen auftraten. Minderheitenallelfrequenzen unter 10 % wurden vor allem auf den Chromosomen 1, 7 und 17 nachgewiesen. Solche Regionen könnten durch Selektion oder Drift nahezu fixiert worden sein. Weil auf den genannten Chromosomen auch SNPs detektiert wurden, die in den betroffenen Haplotypenblöcken liegen und die Merkmalsentwicklung beeinflussen, liegt es nahe, dass die betroffenen Chromosomenregionen unter Selektionsdruck stehen.

Hinweise auf Einkreuzungen aus anderen Zuchtgebieten lieferten auch genomische Untersuchungen zu Populationssubstrukturen. Die genetische Verwandtschaft („genetic kinship“) zwischen den Tieren hat Substrukturen aufgedeckt, die in der untersuchten Gesamtpopulation aber keine dramatischen Auswirkungen auf nachfolgende Assoziationsanalysen, in denen die Beziehung zwischen Genotypen und Merkmalsausprägungen untersucht werden, gezeigt haben. In Abhängigkeit davon, wie viele Achsen der genetischen Variation zugelassen werden, ist die Population in unterschiedlich viele Substrukturen unterteilt. In der untersuchten Population war eine kleinere Gruppe deutlich von einer größeren Gruppe von Tieren zu unterscheiden. Wenn gerade die Subpopulationen zu identifizierten langen Haplotypenblöcken beitragen, könnten in den langen Haplotypenblöcken auch solche Gene liegen, die zu den vorteilhaften Eigenschaften der Subpopulation beitragen. Weil wir in den Assoziationsanalysen für die verschiedenen Merkmale sowohl Modelle mit als auch ohne Berücksichtigung der Populationssubstruktur verwendet und nur geringe Unterschiede und keine konträren Ergebnisse erhalten haben, kann davon ausgegangen werden, dass in den Subpopulationen solche Allelvarianten vorhanden sind, die zur Merkmalsausprägung entweder keinen oder nur kleine Effekte beitragen. Wenn in den in jüngster Zuchtgeschichte eingekreuzten Tieren vorrangig Allelvarianten mit geringen Effekten segregieren, ist auch zu erwarten, dass eine Leistungsverbesserung durch diese Allele nur über mehrere Generationen erreicht werden kann.

Auch hinsichtlich der Variabilität von leistungsbeeinflussenden Genombereichen zeigen die vorliegenden Ergebnisse der Assoziationsanalysen, dass in der untersuchten Population der Deutschen Landrasse noch sehr viele Loci segregieren, an denen Allele mit unterschiedlichem Einfluss auf die verschiedenen Merkmale vorkommen. Loci mit Hauptgeneffekten wurden nicht gefunden. Die Ergebnisse weisen für alle untersuchten Merkmale auf die Wirkung mehrerer bis vieler Loci und damit auf komplexe genetische Determinanten hin. Die Frequenzen der positiven bzw. negativen Allelvarianten variieren in der gesamten Spannbreite. In Abhängigkeit vom spezifischen Zuchtziel sind unter Einbeziehung der genomischen Information mehrere Selektionsvarianten denkbar: Selektion von vorteilhaften Varianten, Selektion gegen nachteilige Varianten und Selektion über die Summe aller SNP-Effekte als genomischer Zuchtwert. Die genomische Variabilität in der Population bietet ein sehr gutes Potenzial für die weitere Verbesserung der Rasse durch Selektion innerhalb der Population.

Die Ergebnisse der Assoziationsanalysen unter Einbeziehung der vorkorrigierten phänotypischen Daten lassen eine Reihe von SNP-Regionen auf verschiedenen Chromosomen erkennen, die einen Bezug zu den untersuchten Merkmalen haben. Die untersuchte Stichprobe war aber relativ klein, sodass signifikante Ergebnisse unter Berücksichtigung der stringenten Bedingungen der Populationsstratifikation und des multiplen Testens nicht nachgewiesen werden konnten. Unter Verwendung des weniger stringenten Signifikanzkriteriums von p-Werten unter 10⁻⁴, für das biologisch relevante Effekte angenommen werden können (SPELIOTES et al. 2010), wurden für alle untersuchten Merkmale relevante SNPs gefunden. Neben SNPs mit p-Werten unter 10⁻⁴ sind aber auch Chromosomenregionen von Interesse, in denen mehrere benachbarte SNPs mit niedrigen p-Werten auftreten, auch wenn diese nicht die Grenze von 10⁻⁴ erreichen. Wenn diese SNPs im Kopplungsungleichgewicht stehen und die resultierenden Haplotypenblöcke ebenfalls signifikante Assoziationen zeigen, liegen in diesen Chromosomenregionen mit großer Sicherheit ebenfalls Genvarianten, die zur unterschiedlichen Merkmalsausprägung beitragen. Solche Regionen haben wir auch mit der relativen kleinen Tierzahl gefunden. Damit wird bestätigt, dass die

erarbeiteten Ergebnisse der Assoziationsanalysen für die untersuchten Merkmale der Mast- und Schlachtleistung durchaus relevant sind und gute Anhaltspunkte für genetische Effekte geben, die in der Selektion genutzt werden können. Die Sicherheit der Ergebnisse der Assoziationsanalyse lassen sich auch dadurch erhärten, dass für die SNP mit einem p-Wert unter 10^{-4} in verschiedenen Populationen wiederholt QTL kartiert wurden, die in der QTL-Datenbank für das Schwein PigQTLdb (www.animalgenome.org) hinterlegt sind. Die signifikantesten SNPs, die hier für die verschiedenen Merkmale im Einzelnen vorgestellt wurden, sind die besten Kandidaten, um die SNP-Effekte in einer größeren Stichprobe zu überprüfen.

Die für die verschiedenen Merkmale identifizierten SNPs mit der signifikantesten Assoziation sind merkmalspezifisch; das heißt, keiner der SNPs hat gleichzeitig signifikante Effekte auf zwei oder mehr Merkmale. Dennoch wurden in bestimmten Chromosomenregionen mehrere Assoziationen mit verschiedenen Merkmalen gefunden. Die Lebendmasse, die eng korreliert ist mit dem Fettansatz, hat z. B. auch in der Assoziationsanalyse gezeigt, dass bestimmte Chromosomenregionen nicht nur die Lebendmasse, sondern auch den Fettansatz beeinflussen. Für das Merkmal Ultraschall-Fettdicke konnte gezeigt werden, dass die fünf SNPs mit den kleinsten p-Werten alle auf Chromosom 1 in unmittelbarer Nachbarschaft lagen. Damit scheint diese Region für weiterführende molekularbiologische und züchterische Untersuchungen zur Beeinflussung des Fettansatzes in der Deutschen Landrasse besonders prädestiniert zu sein. Wenn nicht nur einzelne SNPs betrachtet werden, sondern auch Cluster von SNPs mit geringen p-Werten in bestimmten Chromosomenregionen, die im Manhattan-Plot als „Türme“ erkennbar sind, können in solchen Chromosomenregionen ebenfalls Gene vermutet werden, die das Merkmal beeinflussen und eine entsprechende Bedeutung für das züchterisch zu bearbeitende Merkmal besitzen. Solche Chromosomenregionen sind für weitere Untersuchungen interessant, weil die ursächlichen Gene durch eine höhere Markerdichte weiter eingegrenzt werden können. Als Beispiele hierfür seien die Chromosomen 1, 5 und 14 mit Einfluss auf die Lebensstagszunahme, die Chromosomen 1 und 13 für die Fettfläche und das Chromosom 1 für die Fettdicke aufgeführt. Schon aus dieser Aufzählung lässt sich ableiten, dass das Chromosom 1 eine Reihe von Genen trägt, die an den kausalen Zusammenhängen zwischen Lebensstagszunahme und Verfettung in der Deutschen Landrasse beteiligt sind. Der entsprechende phänotypische Zusammenhang zeigte sich auch in der Korrelationsanalyse, in der zwischen Lebensstagszunahme und Ultraschall-Fettdicke sowie zwischen Lebensstagszunahme und Fettfläche hohe positive Korrelationen ermittelt wurden. Für die Futteraufnahme während der Eigenleistungsprüfung hatte Chromosom 1, anders als bei der Lebensstagszunahme, eher eine untergeordnete Bedeutung. Auch die Muster der abgebildeten p-Werte zeigten keine markanten Übereinstimmungen mit den Mastleistungsmerkmalen. Dies kann als ein vorsichtiger Hinweis darauf gewertet werden, dass eine Selektion auf hohe Lebensstagszunahme bei geringem Futterverzehr in begrenztem Maße möglich ist.

Für einige Merkmale zeigten neben den SNPs auch die Haplotypen, die die signifikantesten SNPs enthalten, einen engen Bezug zur Merkmalsausprägung. Beispiele hierfür wurden für die Lebensstagszunahme, die Fettfläche und den pH-Wert vorgestellt. Die Haplotypen zeigen in jedem Fall ein variables Muster der auftretenden genetischen Effekte als die einzelnen SNPs. Das Vorhandensein von Haplotypen mit verschiedenen Effekten weist darauf hin, dass in der Region des Haplotypenblocks nicht nur eine Mutation, sondern mehrere vorhanden sind, die das Merkmal beeinflussen. Damit wird die Komplexität der Merkmalsdetermination in der eingegrenzten Chromosomenregion sichtbar gemacht. Die ursächlichen Mutationen können in einem Gen, aber auch in verschiedenen Genen im Haplotypenblock liegen. Der Effekt einer einzelnen Mutation kann additiv oder dominant sein. Die Effekte, die die Autoren real an einzelnen SNP-Loci gefunden haben, sind überwiegend additiv, folgen aber nicht immer diesem idealen Muster. Sie haben häufig additive und dominante Varianzanteile. Das Zusammenwirken mehrerer Mutationen an einem oder mehreren Genen würde das erklären. Interaktionen mit anderen Genen im Genom können aber auch die Alleleffekte modifizieren. Noch mehr als bei den SNPs sind die geringen Tierzahlen, die hinter einem Haplotyp stehen, kritisch zu beurteilen. Zukünftige Untersuchungen machen daher wesentlich größere Stichprobenumfänge erforderlich.

Die in der Population vorgefundene Biodiversität und die trotz der geringen Stichprobe identifizierten SNP-Effekte ermöglichen die Nutzung der gewonnenen genomischen Information zur Erhaltung der Allelvielfalt als auch zur gerichteten Verbesserung der Population durch Selektion. Die Ergebnisse sind damit auch ein wichtiger Schritt, um Grundlagen für die genomische Selektion zu legen. Um die Aussagen zu SNP-Effekten zu bestätigen und deren Sicherheit zu erhöhen, sind jedoch weitere Untersuchungen in einer größeren Population notwendig.

5 Quellenverzeichnis

- AULCHENKO, Y.S., RIPKE, S., ISAACS, A., VAN DUIJN, C.M. GenABEL: an R package for genome-wide association analysis. *Bioinformatics*. 2007 23(10):1294-6.
- BARRETT, J.C., FRY, B., MALLER, J. & DALY, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263-5.
- HUISMAN, A.E., CHEREL, P., VAN HAANDEL, B.: Linkage disequilibrium and signatures of selection on chromosome 1 in a commercial sire and dam line. World Congress for Genetics Applied to Livestock Production, Leipzig, Germany, 01.-06.08.2010
- PRICE, A., PATTERSON, N., PLENGE, R., WEINBLATT, M., SHADICK, N., REICH, D. (2006): Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*.38, 904-9.
- PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M.A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P.I., DALY, M.J. et al. (2007): PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81, 559-75.
- SCHEET, P. & STEPHENS, M. (2006): A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 78, 629-44.
- SCHMITT, A.O., BORTFELDT, R.H. & BROCKMANN, G.A. (2010): Tracking chromosomal positions of oligomers - a case study with Illumina's BovineSNP50 beadchip. *BMC Genomics* 11, 80.
- SPELIOTES, E.K., WILLER, C.J., BERNDT, S.I., MONDA, K.L., THORLEIFSSON, G., JACKSON, A.U., ALLEN, H.L., LINDGREN, C.M., LUAN, J., MAGI, R. et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 42, 937-48.

Herausgeber:

Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG)
Pillnitzer Platz 3, 01326 Dresden
Telefon: + 49 351 2612-0
Telefax: + 49 351 2612-1099
E-Mail: lfulg@smul.sachsen.de
www.smul.sachsen.de/lfulg

Autoren:

Prof. Dr. Gudrun A. Brockmann, Dr. Uwe Müller, Dr. Armin O. Schmidt
Humboldt-Universität zu Berlin
Prof. Dr. Ottmar Distl
Tierärztliche Hochschule Hannover
Dr. Uwe Bergfeld, Dr. Ulf Müller
Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie

Redaktion:

Dr. Ulf Müller
LfULG, Abteilung Tierische Erzeugung/Referat Tierzucht, Tierhygiene
Telefon: + 49 34222 46-2106
Telefax: + 49 34222 46-2199
E-Mail: ulf.mueller@smul.sachsen.de

Redaktionsschluss:

31.01.2012

ISSN:

1867-2868

Hinweis:

Die Broschüre steht nicht als Printmedium zur Verfügung, kann aber als PDF-Datei unter <http://www.smul.sachsen.de/lfulg/6447.htm> heruntergeladen werden.

Verteilerhinweis

Diese Informationsschrift wird von der Sächsischen Staatsregierung im Rahmen ihrer verfassungsmäßigen Verpflichtung zur Information der Öffentlichkeit herausgegeben. Sie darf weder von Parteien noch von deren Kandidaten oder Helfern im Zeitraum von sechs Monaten vor einer Wahl zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für alle Wahlen.

Missbräuchlich ist insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist auch die Weitergabe an Dritte zur Verwendung bei der Wahlwerbung. Auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl darf die vorliegende Druckschrift nicht so verwendet werden, dass dies als Parteinahme des Herausgebers zugunsten einzelner politischer Gruppen verstanden werden könnte.

Diese Beschränkungen gelten unabhängig vom Vertriebsweg, also unabhängig davon, auf welchem Wege und in welcher Anzahl diese Informationsschrift dem Empfänger zugegangen ist. Erlaubt ist jedoch den Parteien, diese Informationsschrift zur Unterrichtung ihrer Mitglieder zu verwenden.