

Aus dem Department für Veterinärwissenschaften
der Tierärztlichen Fakultät
der Ludwig-Maximilians-Universität München
Arbeit angefertigt unter Leitung von Univ.-Prof. Dr. Dr. habil. Martin Förster

**Genomweite Detektion von
Selektionssignaturen
in divergent selektierten Rinderpopulationen
mit anschließender Identifikation eines
möglichen kausalen Gens**



Inaugural-Dissertation
zur Erlangung der tiermedizinischen Doktorwürde
der Tierärztlichen Fakultät der Ludwig-Maximilians-Universität München

von
Sophie Maria Rothammer
aus Dachau

München 2011

Gedruckt mit der Genehmigung der Tierärztlichen Fakultät
der Ludwig–Maximilians–Universität München

Dekan: Univ.-Prof. Dr. Joachim Braun

Berichterstatter: Univ.-Prof. Dr. Dr. habil. Martin Förster

Korreferent: Univ.-Prof. Dr. Joris Peters

Tag der Promotion: 30. Juli 2011

Für meine Familie

Inhaltsverzeichnis

INHALTSVERZEICHNIS	I
ABKÜRZUNGSVERZEICHNIS	IV
ABBILDUNGSVERZEICHNIS	VIII
TABELLENVERZEICHNIS	IX
ANHANGSÜBERSICHT	IX
1 EINLEITUNG	1
2 LITERATURÜBERSICHT	2
2.1 Das Rind.....	2
2.1.1 Geschichte des Rindes.....	2
2.1.1.1 Abstammung des Hausrindes.....	2
2.1.1.2 Domestikation des Rindes.....	4
2.1.1.3 Entwicklung der Hausrinder.....	5
2.1.2 Populationsgeschichte der untersuchten Rinderrassen.....	7
2.1.2.1 Weitgehend unselektierte Rinderrassen.....	7
2.1.2.1.1Anatolisches Schwarzvieh.....	8
2.1.2.1.2Illyrisches Rind	9
2.1.2.2 Zweinutzungsrasen.....	10
2.1.2.2.1Deutsches Fleckvieh	10
2.1.2.2.2Braunvieh alte Zuchtrichtung (Original Braunvieh)	11
2.1.2.2.3Fränkisches Gelbvieh	13
2.1.2.2.4Murnau-Werdenfelser.....	14
2.1.2.3 Milchrasen	15
2.1.2.3.1Braunvieh neue Zuchtrichtung.....	15
2.1.2.3.2Red Holstein.....	16
2.1.2.4 Fleischrasen	17
2.1.2.4.1Weißblauer Belgier.....	18
2.1.2.4.2Galloway.....	19
2.2 Populationsgenetische Prozesse	20
2.2.1 Mutation.....	20
2.2.2 Rekombination	21
2.2.3 Migration.....	22
2.2.4 Selektion.....	23
2.2.5 Gendrift.....	24
2.2.6 Nicht zufällige Paarung.....	25
2.3 Molekulare Marker.....	26
2.3.1 Historische Entwicklung molekularer Marker.....	27
2.3.2 Single Nucleotide Polymorphisms	28

2.4	Signaturen Positiver Selektion.....	29
2.4.1	Verschiedene Selektionssignaturen und ihre Analysemethoden.....	29
2.4.1.1	Hohe Anzahl funktionsverändernder Mutationen.....	30
2.4.1.2	Reduktion der genetischen Diversität.....	31
2.4.1.3	Neue Allele mit hoher Frequenz.....	32
2.4.1.4	Unterschiede zwischen Populationen.....	33
2.4.1.5	Lange Haplotypen.....	34
2.4.2	Bisher durchgeführte Selektionssignaturstudien im Rindergenom.....	35
3	MATERIAL UND METHODEN.....	38
3.1	Material.....	38
3.1.1	Probenmaterial.....	38
3.1.2	Verwendete Software.....	40
3.2	Methoden.....	41
3.2.1	Genotypisierung.....	41
3.2.2	Haplotypisierung.....	43
3.2.3	Auswahl geeigneter Marker.....	44
3.2.4	Schätzung der <i>identity-by-descent</i> -(<i>IBD</i>)-Wahrscheinlichkeiten.....	44
3.2.5	Methoden basierend auf Haplotypenlänge und –diversität.....	45
3.2.5.1	Extended Haplotype Homozygosity (<i>EHH</i>).....	45
3.2.5.2	Integrated Haplotype Score (<i>iHS</i>).....	47
3.2.5.2.1	Zuteilung des Allelstatus.....	47
3.2.5.2.2	Berechnung der unstandardisierten <i>iHS</i> -Werte.....	48
3.2.5.2.3	Standardisierung der <i>iHS</i> -Werte.....	49
3.2.5.3	Cross Population Extended Haplotype Homozygosity (<i>XP-EHH</i>).....	50
3.2.5.3.1	Berechnung der unstandardisierten <i>XP-EHH</i> -Werte.....	50
3.2.5.3.2	Standardisierung der <i>XP-EHH</i> -Werte.....	51
3.2.6	Auswertung der Daten.....	52
3.2.7	Sequenzierung.....	54
3.2.7.1	Paired-End-Sequenzierung.....	55
3.2.7.1.1	Auswahl des zu sequenzierenden Bereichs.....	55
3.2.7.1.2	Durchführung der Paired-End-Sequenzierung.....	56
3.2.7.2	DNA-Sequenzierung (Sanger).....	59
3.2.7.2.1	Erstellen der Primer.....	59
3.2.7.2.2	Durchführung der DNA-Sequenzierung.....	59
4	ERGEBNISSE.....	61
4.1	Validierung der Methode.....	62
4.1.1	Phänotyp Doppellender.....	62
4.1.2	Phänotyp Rotfärbung.....	64
4.1.3	Phänotyp Hornlosigkeit.....	67
4.1.4	Zusammenfassende Beurteilung der Methode.....	69
4.2	Genomweite Ergebnisse.....	70
4.2.1	<i>iHS</i> -Ergebnisse.....	70
4.2.2	<i>XP-EHH</i> -Ergebnisse.....	74
4.3	Deutsches Fleckvieh.....	78

5	DISKUSSION	98
6	ZUSAMMENFASSUNG.....	111
7	SUMMARY	113
8	LITERATURVERZEICHNIS	115
9	ANHANG	123

Abkürzungsverzeichnis

3'U	3prime UTR (in der 3' Untranslated Region liegend)
Abb.	Abbildung
ABCG2	ATP-binding cassette, sub-family G (WHITE), member 2
AFLP	amplified fragment length polymorphisms
α -MSH	Melanozyten-stimmulierendes Hormon alpha
ASV	Anatolisches Schwarzvieh
BBB	Weißblauer Belgier
BBV	Braunvieh (neue Zuchtrichtung)
bp	Basenpaare
BS	Brown Swiss
BTA	Rinderchromosom (Bos <i>Taurus</i> Autosome)
CHIC2	cysteine-rich hydrophobic domain 2
CL	Corpus Luteum, Gelbkörper
CLL	parametric composite log likelihood
cm	Zentimeter
cSNP	SNP der im codierenden oder Codierung-beeinflussenden (regulierenden) Bereich liegt
dAF	derived-Allelfrequenz
ddNTPs	2'-3'-dideoxy-Nucleosid-Triphosphate
DFG	Deutsche Forschungsgemeinschaft
DFV	Deutsches Fleckvieh
DHV	Deutscher Holstein Verband e. V.
DNA	Desoxyribonukleinsäure
dNTP	Desoxynucleosid-Triphosphate
Down	Downstream (innerhalb 5 kb distal des 3' Endes eines Transkripts)
EFABIS	European Farm Animal Biodiversity Information System
EHH	Extended Haplotype Homozygosity
eP	empirischer <i>P</i> -Wert
E _p	Erwartungswert
E.S	Essential Splice Site (die ersten zwei oder letzten zwei Basenpaaren eines Introns betreffend)
ET	Embryotransfer

e. V.	eingetragener Verein
F	Fenster(nummer)
FAO	Food And Agriculture Organization Of The United Nations
FGV	Fränkisches Gelbvieh
FRYL	furry homolog-like
GEH	Gesellschaft zur Erhaltung alter und gefährdeter Haustierrassen e. V.
GLW	Galloway
IBD	identity by descent
IBHM	International Bovine HapMap
IGV	Integrative Genomics Viewer
iH _{H_A}	integrierter EHH des " <i>ancestral</i> " Allels
iH _{H_D}	integrierter EHH des " <i>derived</i> " Allels
iHS	integrated Haplotype Score
IL	Illyrisches Rind
INDEL	Insertionen und Deletionen
Inter	Intergenic (mehr als 5 kb distal oder proximal eines Transkripts)
Intro	Intronic (im Intron liegend)
ISSR	inter simple sequence repeats
KB	Künstliche Besamung
kb	Kilobasen
KCNIP4	Kv channel interacting protein 4
KDR	kinase insert domain receptor (type III receptor tyrosine kinase)
kg	Kilogramm
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
LAFUGA	Laboratory for Functional Genome Analysis
LAP3	leucine aminopeptidase 3
LCORL	ligand dependent nuclear receptor corepressor-like
LD	Linkage Disequilibrium (Kopplungsungleichgewicht)
LDL	low density lipoprotein
LfL	Bayerische Landesanstalt für Landwirtschaft
LMU	Ludwig-Maximilians-Universität
LOC540334	clock (Pseudogen)
LRH	long-range-haplotype
MAF	minor allele frequency

Mb	Megabasen
Mc1r	Melanocortin-1-Rezeptor
mh	muscular hypertrophy
µl	Mikroliter
mind.	mindestens
MWF	Murnau-Werdenfelser
n. Chr.	nach Christus
N _e	effektive Populationsgröße
Non	Non-synonymous SNP (in codierender Sequenz, Veränderung der Aminosäure in der codierten Peptidsequenz)
OBV	Original Braunvieh
ORF	open reading frame
PCR	Polymerase-Kettenreaktion, polymerase chain reaction
PDGFRA	platelet-derived growth factor receptor, alpha polypeptide
QTL	Quantitative Trait Locus/Loci
RAPD	randomly amplified polymorphic DNA
RefSeq	Referenzsequenz
RFLP	restriction fragment length polymorphism
RH	Red Holstein
RSP	restriction site polymorphism
RTK	receptor tyrosin kinase
s	Selektionsvorteil
SBS	sequencing by synthesis
SCFD2	sec1 family domain containing 2
SD _P	Standardabweichung
SNP	Einzelnukleotid-Polymorphismus, Single Nucleotide Polymorphism
Sp.S	Splice Site (erste ein bis drei Basenpaare innerhalb eines Exons oder erste drei bis acht Basenpaare innerhalb eines Introns)
SQL	Structured Query Language
SRD5A3	steroid 5 alpha-reductase 3
St	Stop gained (in codierender Sequenz, Zugewinn eines Stopp-Codons, Resultat z. B. verkürzte Peptidsequenz)
STR	short tandem repeats
sXP-EHH	standardisierter XP-EHH-Wert

Syn	Synonymous SNP (in codierender Sequenz, keine Veränderung der Aminosäure; stille Mutation)
Tab.	Tabelle
TARBP1	TAR (HIV-1) RNA binding protein 1
TMEM165	transmembrane protein 165
UCSC	University of California, Santa Cruz
Up	Upstream (innerhalb 5 kb proximal des 5' Endes eines Transkripts)
UTR	untranslated region
v. Chr.	vor Christus
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
XP-EHH	Cross Population Extended Haplotype Homozygosity
z. B.	zum Beispiel

Abbildungsverzeichnis

Abbildung 1: Systematik der Rinder	3
Abbildung 2: Auerochse	4
Abbildung 3: Anatolisches Schwarzvieh (Bulle).....	8
Abbildung 4: Illyrisches Rind (Kuh und Kalb verschiedener Schläge)	9
Abbildung 5: Deutsches Fleckvieh (Bulle Holzmichl).....	11
Abbildung 6: Original Braunvieh (Bulle Singau).....	12
Abbildung 7: Fränkisches Gelbvieh (Bulle Mongan).....	13
Abbildung 8: Murnau-Werdenfelser (Bulle Bergfex)	15
Abbildung 9: Braunvieh neue Zuchtrichtung (Bulle Husant)	16
Abbildung 10: Red Holstein (Bulle Malvoy)	17
Abbildung 11: Weißblauer Belgier (Bulle Graphite de Dessous la Ville).....	18
Abbildung 12: Galloway (Bulle Dean).....	19
Abbildung 13: Funktionsweise eines BeadChips.....	42
Abbildung 14: <i>EHH</i> -Verlauf der Beispielhaplotypen aus Tabelle 3.....	46
Abbildung 15: Paired-End-Sequenzierung	58
Abbildung 16: Signal der positiven Selektion auf BTA2 (Fenster 66) der Rasse BBB im Vergleich mit neun weiteren Rinderrassen.	64
Abbildung 17: Signal der positiven Selektion auf BTA18 (Fenster 731-735) der Rasse RH im Vergleich mit neun weiteren Rinderrassen.	66
Abbildung 18: Signal der positiven Selektion auf BTA1 (Fenster 1) der Rasse GLW im Vergleich mit neun weiteren Rinderrassen.	68
Abbildung 19: Übersicht über die Verteilung der 94 verschiedenen Fenster mit signifikanter <i>iHS</i> -Signatur auf die 29 Autosome des Rindes.	71
Abbildung 20: Genomweite Verteilung der 100 signifikanten <i>iHS</i> -Fenster der zehn verschiedenen Rinderrassen	72/73
Abbildung 21: Übersicht über die Verteilung der 90 verschiedenen Fenster mit signifikanter <i>XP-EHH</i> -Signatur auf die 29 Autosome des Rindes..	74
Abbildung 22: Genomweite Verteilung der 103 signifikanten <i>XP-EHH</i> -Fenster der zehn verschiedenen Rinderrassen	76/77
Abbildung 23: Signal der positiven Selektion in den Fenstern 282-284 (BTA6) der Rasse DFV im Vergleich mit neun weiteren Rinderrassen.....	80
Abbildung 24: Signal der positiven Selektion in den Fenstern 294-296 (BTA6) der Rasse DFV im Vergleich mit neun weiteren Rinderrassen.....	82
Abbildung 25: <i>IGV</i> -Ansicht des Sequenzabgleichs im Übergang des im DFV homozygoten Bereichs zum <i>KDR</i> -Gen.....	87
Abbildung 26: Ausschnitt von BTA6 mit Lage der Gene sowie der Varianten	92
Abbildung 27: Lage der Varianten bezüglich dem Signal der positiven Selektion auf BTA6 im Bereich 71.900.000-73.500.000.....	93
Abbildung 28: <i>IGV</i> -Ansicht der drei Varianten im codierenden Bereich des <i>KDR</i> -Gens	94
Abbildung 29: Elektropherogramm-Ausschnitt eines homozygoten Tieres der Fallgruppe sowie eines heterozygoten Tieres der Kontrollgruppe.	96
Abbildung 30: Verteilung der <i>XP-EHH</i> -Werte zweier Vergleiche von je zwei Ras- sen mit ähnlichen respektive stark unterschiedlichen N_e und LD..	103

 Tabellenverzeichnis

Tabelle 1: Übersicht über die zehn ausgewählten Teilpopulationen.....	39
Tabelle 2: Verwendete Software: Programmname, Verwendungszweck und Quelle	41
Tabelle 3: Beispielhaplotypen für <i>EHH</i> -Berechnung	46
Tabelle 4: <i>eP</i> -Werte der <i>XP-EHH</i> -Vergleiche mit BBB in Fenster 66 (5.119.264- 7.676.879)	63
Tabelle 5: <i>eP</i> -Werte der <i>XP-EHH</i> -Vergleiche mit RH in den Fenstern 731-735 (10.173.641-22.887.955)	66
Tabelle 6: <i>eP</i> -Werte für <i>XP-EHH</i> -Vergleiche mit GLW in Fenster 1 (483- 2.557.020)	68
Tabelle 7: Die 19 Selektionssignaturen-beinhaltenden Fenster des Deutschen Fleckviehs	78
Tabelle 8: Übereinstimmung der signifikanten Fenster des DFV mit QTL und möglichen kausalen Genen anderer Studien.....	79
Tabelle 9: Ausschnitt der Haplotypen-Matrix um den für das DFV typischen Chromosomenabschnitt.....	84
Tabelle 10: Die sieben Gene im zu sequenzierenden Bereich auf BTA6.....	85
Tabelle 11: Ermittelte Varianten und ihre Auswirkungen.....	89
Tabelle 12: Erklärung der durch den <i>Variant Effect Predictor</i> ermittelten Auswirkungen der Varianten	89
Tabelle 13: Die 23 Varianten im die Codierung betreffenden sowie im codierenden Bereich.....	90
Tabelle 14: Die Auswirkungen der drei näher untersuchten Varianten auf die Codierung.	97

Anhangsübersicht

Anhang 1: Tierliste	123
Anhang 2: Fensterliste	126
Anhang 3: signifikante <i>iHS</i> -Fenster ($eP \leq 0,01$)	136
Anhang 4: signifikante <i>XP-EHH</i> -Fenster	136
Anhang 4. 1: Illyrisches Rind.....	136
Anhang 4. 2: Anatolisches Schwarzvieh	137
Anhang 4. 3: Deutsches Fleckvieh.....	137
Anhang 4. 4: Original Braunvieh.....	137
Anhang 4. 5: Fränkisches Gelbvieh.....	138
Anhang 4. 6: Murnau-Werdenfelser	138
Anhang 4. 7: Braunvieh (neue Zuchtichtung).....	138
Anhang 4. 8: Red Holstein.....	139
Anhang 4. 9: Galloway	139
Anhang 4. 10: Weißblauer Belgier.....	140
Anhang 5: Alignment des Teilbereichs des <i>KDR</i> -Gens der acht sequenzierten Tiere	141

1 Einleitung

Rinder wurden, wie alle Spezies, durch positive natürliche Selektion geformt. Zusätzlich ist der Einfluss künstlicher Selektion bei dieser Art, wie auch bei den meisten anderen domestizierten Arten, sehr hoch. Dies wird in jüngster Vergangenheit insbesondere durch den Einsatz von anspruchsvollen quantitativ-genetischen und reproduktiven Techniken begünstigt. Aufgrund strikter Verfolgung teilweise sehr einseitiger Zuchtziele sowie Verwendung vergleichsweise weniger Vätertiere, ist davon auszugehen, dass sich insbesondere die in der Populationsgröße effektiv kleinen Hochleistungsrassen stark in den selektierten Bereichen ihres Genoms unterscheiden.

Diese chromosomalen Bereiche sondern sich außerdem durch charakteristische Merkmale vom unselektierten Teil des Genoms ab und werden als Selektionssignaturen bezeichnet. Sie aufzufinden zu machen, ist aus mehreren Gründen erstrebenswert. Zum Einen wird das Auffinden positiver Selektion langfristig dabei helfen die Rolle von Gendrift und Selektion während evolutionären Prozessen zu klären. Zum Anderen grenzen Selektionssignaturen Genomregionen ein, die funktionell wichtig waren oder immer noch sind, und erleichtern daher die Identifikation von genetischen Variationen, welche die Grundlage phänotypischer und funktioneller Diversität darstellen (BISWAS & AKEY, 2006).

Ziel dieser Arbeit ist es, bei zehn verschiedenen Rinderrassen, welche vier unterschiedlichen Nutzungsgruppen angehören, genomweit (Autosome) rassespezifische Selektionssignaturen zu detektieren und entsprechend zu kartieren. Dabei soll diese Kartierung mit den Methoden *integrated Haplotype Score (iHS)*, VOIGHT *et al.*, 2006) und *Cross Population Extended Haplotype Homozygosity (XP-EHH)*, SABETI *et al.*, 2007) erfolgen, welche auf dem Auffinden von Genombereichen mit ausgedehnter Haplotypenhomozygotie basieren.

Beim im süddeutschen Raum dominierenden Rind, dem Deutschen Fleckvieh, soll anschließend versucht werden, für einen der so gefundenen Bereiche ein mögliches kausales Gen sowie die zugrundeliegende Mutation zu identifizieren.

2 Literaturübersicht

2.1 Das Rind

Rinder werden heutzutage vornehmlich zur Erzeugung von Milch und Fleisch gehalten (ELFRICH & ROESICKE, 2009). Da dies nicht immer so war und auch gegenwärtig nicht uneingeschränkt auf alle Rinder zutrifft, wird zunächst ein historischer Überblick über die Entwicklung dieser Tierart gegeben, bevor dann auf die einzelnen Rassen (Teilpopulationen), welche in dieser Arbeit Verwendung fanden, eingegangen wird.

2.1.1 Geschichte des Rindes

Laut Statistischem Bundesamt Deutschland betrug die Anzahl der in der Bundesrepublik gehaltenen Rinder am 3. Mai 2010 12.809.492 (http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2010/07/PD10__233__413,templateld=renderPrint.psm). Diese imposante Zahl veranschaulicht die gegenwärtige Bedeutung dieser Nutztiere.

Um die Entwicklung vom wildlebenden Vorfahren zu den heute gehaltenen Rassen besser nachvollziehen zu können, wird diese im Folgenden skizziert.

2.1.1.1 Abstammung des Hausrindes

Der Auerochse (*Bos primigenius*) gilt als Urahn aller domestizierten Rinder, welche sich wiederum in die zwei Gruppen der buckellosen taurinen Rinder (*Bos taurus*) und der Buckelrinder (*Bos indicus*) einteilen lassen. Von dieser bereits ausgestorbenen Wildart existierten drei verschiedene kontinentale Subpopulationen, nämlich das *Bos primigenius namadicus* in Asien, das *Bos primigenius opisthonomus* in Nordafrika sowie das *Bos primigenius primigenius* in Europa (LOFTUS *et al.*, 1994). Von dem ursprünglich wohl während des Pleistozäns (früher Diluviums) im indischen Raum entstandenen Auerochsen

(KRÄUßLICH, 1994), überlebte nur das *Bos primigenius primigenius* bis in die christliche Zeitrechnung (LOFTUS *et al.*, 1994). Jedoch dezimierte sich auch dieser Bestand durch Jagd und Habitatzerstörung allmählich, bis schließlich 1627 die letzte Urkuh in Polen einem Wilderer zum Opfer fiel (KRÄUßLICH, 1994; LOFTUS *et al.*, 1994).

Der Auerochse auch Ur genannt bildet zusammen mit Gaur, Banteng, Yak und Kouprey die Gattung der „eigentlichen Rinder“ (*Bos*). (FAO, 2000). Diese Systematik der Unterfamilie *Bovinae* ist in Abbildung 1 (Abb.1) dargestellt.

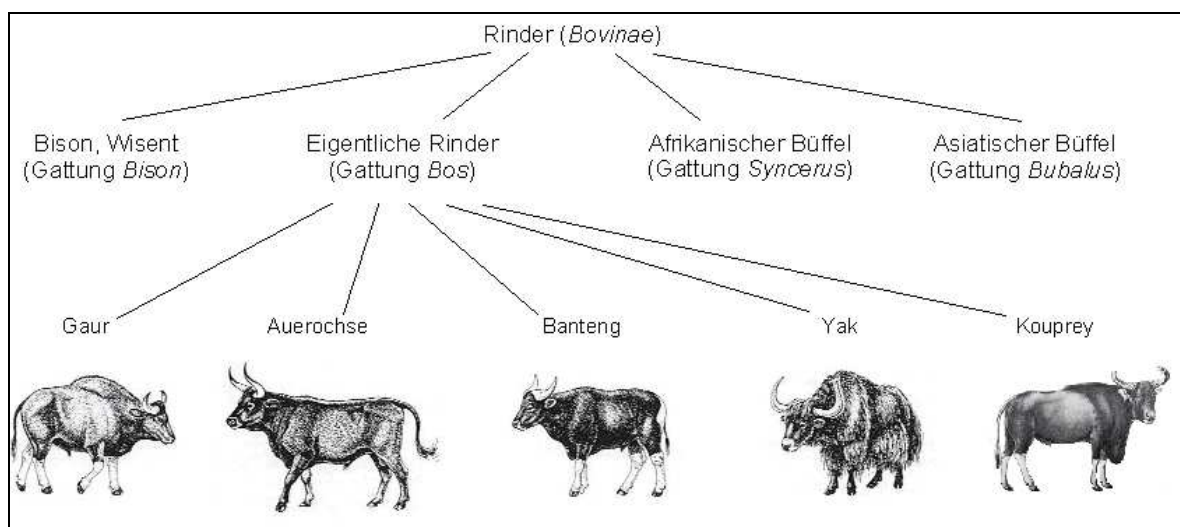


Abbildung 1: Systematik der Rinder

(nach: Ahne *et al.* „Zoologie“, Schattauer, 2000; Quelle Abbildung Kouprey: http://www.nytimes.com/imagepages/2006/09/11/science/12kouprey2_ready.html)

Im Vergleich zu den heutigen Hausrindern wies der Auerochse (Abb. 2) eine gegenüber der Nachhand deutlich kräftiger entwickelte Vorhand auf (FRAHM, 1982) und zeigte einen stark ausgeprägten Geschlechtsdimorphismus. So erreichten Stiere eine Widerristhöhe von bis zu 200 cm, Kühe bis maximal 170 cm. Darüber hinaus fiel die dunkelbraune Färbung des glatten Haarkleides, welches auf dem Rücken einen hellen, fast gelblichen Aalstrich besaß, bei Bullen dunkler (schwarzbraun) aus als bei Kühen (rotbraun) (KRÄUßLICH, 1994; MACKROTT, 1994).



Abbildung 2: Auerochse

(Quelle: M. Haller, Seltene Haus- & Nutztierassen, Leopold Stocker Verlag, 2000)

2.1.1.2 Domestikation des Rindes

Generell bedarf es verschiedener Voraussetzungen, um die Domestikation einer Art möglich und für den Menschen erstrebenswert zu machen. Zum Einen muss sich die auserwählte Art auf eine andere Art, in diesem Fall den Menschen, prägen lassen, damit sie die Scheu vor dem Menschen verliert. Zum Anderen muss der Mensch aus den Produkten oder Leistungen dieser Tiere einen Nutzen ziehen können, der den mit der Haltung der Tiere verbundenen Aufwand rechtfertigt.

Die Domestikation der ersten Tiere begann während des Neolithikums, also der Jungsteinzeit. Damals war von großer Bedeutung, dass die entsprechende Art dem Menschen nicht zu gefährlich werden konnte, und dass es sich bei ihr nicht um einen Nahrungskonkurrenten handelte. Waren diese Voraussetzungen erfüllt, mussten die Tiere am Entkommen und an der Paarung mit der Stammform gehindert werden (SAMBRAUS, 2001). Die daraus folgenden Veränderungen in Sozialstruktur, Fortpflanzungsverhältnissen, Ernährung und Umwelt, ebenso wie die zunehmende künstliche Selektion, führten über viele Generationen zur Anpassung an die neuen ökologischen Bedingungen.

Aufgrund genetischer Anzeichen geht man davon aus, dass es vor etwa 10.000 Jahren zwei separate Domestikationsvorgänge gab, welchen divergente Populationen des lokalen Auerochsens zugrunde lagen. So wurden taurine Rinder

in der Region des Fruchtbaren Halbmonds, Zebus in der Region des Indus-Tals domestiziert (LOFTUS *et al.*, 1994; FAO, 2000).

Dass sich diese Veränderungen vom Ur zum Hausrind tatsächlich bereits vor mindestens 8.500 Jahren vollzogen, beweisen Funde aus Südosteuropa. Zunächst dienten die Hausrinder lediglich als Fleisch-, Fett- und Rohstofflieferanten sowie in einigen Kulturen zu Opferzwecken. Der Einsatz als Zug- und Reittiere erfolgt dagegen erst seit etwa 5.000 Jahren (KRÄUßLICH, 1994). Auch wenn die älteste überlieferte Darstellung, welche eine Kuh zeigt, die gemolken wird, um 2.000 v. Chr. in Ägypten entstand, blieb die Milchgewinnung bis weit in die Neuzeit von untergeordneter Bedeutung (MACKROTT, 1994).

Sicher ist, dass die neolithische Transition, also der Übergang des Lebensstils vom Jäger und Sammler zu einer landwirtschaftlichen Gesellschaft (unter anderem MACHUGH *et al.*, 1998), und die damit verbundene Sesshaftwerdung enormen Einfluss auf die menschliche Geschichte hatten, die allenfalls mit den Auswirkungen der technischen Revolution vergleichbar sind (KRÄUßLICH, 1994).

2.1.1.3 Entwicklung der Hausrinder

Obwohl das Wissen über die Entwicklung in frühgeschichtlicher Zeit sehr lückenhaft ist, lässt sich prähistorischen Dokumenten entnehmen, dass sich die Vielfalt der phänotypischen Formen schnell erweiterte. Dies ist vor allem auf verborgene biologische Fähigkeiten der Stammart sowie natürliche und künstliche Selektion zurückzuführen und wurde mitunter durch die Tatsache gefördert, dass die domestizierten Tiere mit dem Menschen in Gebiete vordrangen, in denen die Wildart nicht vorkam und sie sich somit an völlig neue ökologische Bedingungen anpassen mussten (KRÄUßLICH, 1994). Generell kam es bei den Hausrindern nach der Domestikation zu einer Größenabnahme, gefördert durch die Beschränkung des Lebensraums und die Abschwächung des Geschlechtsdimorphismus sowie vermutlich durch bevorzugte Selektion kleinerer Individuen aufgrund der leichteren Handhabbarkeit. Lediglich die Römer und die alten Griechen, die verhältnismäßig hochstehende Haustierzucht betrieben, züchteten größere Tiere als alle anderen Völker. Mit dem Ende der griechisch-römischen Epoche verschwand jedoch auch ihr tierzüchterischer Einfluss und die

Größe der Tiere nahm wieder ab, während die phänotypische Formenvielfalt erneut zunahm (FRAHM, 1982; KRÄUßLICH 1994). Etwa 4.000 Jahre alte Funde lassen auf eine Widerristhöhe damaliger Hausrinder von 105 - 145 cm schließen, 1.000 n. Chr. erreichte ein weibliches Tier in Deutschland sogar nur 90 - 110 cm bei einem Gewicht von 250 kg (MACKROTT, 1994). Dies entspricht auch der aktuellen Größe autochthoner Buša Rinder in einigen Regionen des Balkans und Kleinasiens (MEDUGORAC *et al.*, 2009; RAMLJAK *et al.*, 2011).

Bis ins 18. Jahrhundert stand die Zucht von Rindern im Vergleich zu anderen Haustierarten, wie Pferd oder Schaf, weit hinten an (KRÄUßLICH, 1994). So lag die durchschnittliche Milchleistung um 1740 lediglich bei drei Litern. Zu dieser Zeit war die Ochsenmast von weit aus größerer Bedeutung als die Milchwirtschaft und in einigen Ackerbaugebieten stand sogar die Lieferung von Dung an erster Stelle. Erst durch die steigende Nachfrage nach tierischen Produkten durch die im Zuge der Industriellen Revolution wachsende Stadtbevölkerung sowie mit Einführung der planmäßigen Futterwirtschaft, vollzog sich die Entwicklung von gut angepassten Landrassen hin zu Leistungsrassen (KRÄUßLICH, 1994; MACKROTT, 1994; FELIUS, 1995). 1880 erreichte eine schwarzbunte Kuh in Friesland erstmals eine Tagesleistung von 30 kg Milch und mit Beginn der Künstlichen Besamung (KB) - die erste deutsche Besamungsstation wurde 1942 in Pinneberg/Holstein eröffnet - stieg der züchterische Erfolg schließlich enorm an (MACKROTT, 1994).

2.1.2 Populationsgeschichte der untersuchten Rinderrassen

Von den weltweit gegenwärtig 953 lokal existierenden Rinderrassen (FAO, 2010) flossen zehn Rassen in vorliegende Untersuchungen mit ein. Über deren Zuchtgeschichte, Nutzungsrichtung und Exterieurmerkmale wird im Folgenden eine Übersicht gegeben.

Auch wenn die Rassen nachfolgend getrennt betrachtet werden, sollte man nicht außer Acht lassen, dass sich insbesondere geographisch naheliegende Rassen auch genetisch ähnlicher sind als solche, die sich durch weite Distanzen getrennt voneinander entwickelten. So konnten Medugorac *et al.* (2009) nachweisen, dass die genetische Distanz innerhalb sowohl alpiner als auch nordwest-europäischer Rassen deutlich geringer ist, als zwischen diesen beiden Gruppen. Darüber hinaus deckten die alpinen sowie nordwest-europäischen Rassen jeweils nur einen kleinen, sich kaum überschneidenden Teil des Diversitätsspektrums unselektierter Rassen ab. Da sich die zehn Rassen, welche in dieser Arbeit Verwendung fanden, ebenfalls nach diesen Kriterien unterteilen lassen, wird dieser geographische Zusammenhang nachfolgend ebenfalls Erwähnung finden.

2.1.2.1 Weitgehend unselektierte Rinderrassen

Für viele Rassen existieren weder gemeinsames Zuchtziel noch Herdbuch, weshalb sie nicht gezielt auf bestimmte Leistungsmerkmale selektiert werden. Meist handelt es sich dabei um Landrassen, die auch derzeit gelegentlich noch als echte Dreinutzungsrassen gehalten werden. Das bedeutet, dass diese Tiere neben der Milch- und Fleischleistung auch Arbeitsleistung erbringen. Arbeitsleistung spielt zwar im westlichen Teil Europas sowie in Nordamerika und anderen hochtechnisierten Regionen keine Rolle mehr, ist jedoch weltweit noch von erheblicher Bedeutung (SAMBRAUS, 2001).

Auch bei den Vertretern des Anatolischen Schwarzviehs und des Illyrischen Rindes werden Tiere nicht gezielt selektiert sondern hauptsächlich nach individuellen Vorstellungen einzelner Züchter ausgewählt. Übergeordnete Zuchtprogramme und Leistungskontrollen hingegen fehlten zum Zeitpunkt der Probennahme völlig, so dass sie daher im Weiteren als weitgehend unselektiert betrachtet werden.

2.1.2.1.1 Anatolisches Schwarzvieh

In Anatolien, also dem asiatischen Teil der Türkei, wurden seit der Zeit der Hethiter vor mehr als 4.000 Jahren kurzgehörnte Rinder bevorzugt. Diese Favorisierung führte letztendlich zum Anatolischen Schwarzvieh (ASV, Abb. 3), auch Anatolian Black oder Anadolu Yerli Kara genannt (FELIUS, 1995; Department of Animal Science – Oklahoma State University, 2001). Das Hauptverbreitungsgebiet dieser Rasse ist auch heute noch die Türkei. Diese relativ kleinen Rinder (Höhe weiblicher Tiere: 100 bis 115 cm; Gewicht: 150 bis 300 kg) zeichnen sich vor allem durch ihre Genügsamkeit und geringe Krankheitsanfälligkeit sowie der Fähigkeit aus, sich auch an widrige klimatische Bedingungen anzupassen. Anatolisches Schwarzvieh wird vorwiegend auf kleinen Bauernhöfen als Milchlieferant und gelegentlich auch heute noch als Arbeitstier gehalten, kann jedoch auch gemästet werden. So verschieden wie die Nutzungsrichtungen, ist auch das äußere Erscheinungsbild, so dass das Exterieur einiger Vertreter mehr an Fleischrinder, das anderer jedoch an Milchrinder erinnert (FELIUS, 1995). Gemeinsam sind jedoch allen die kurzen und kräftigen Beine, der lange Nacken und die große Wamme (Department of Animal Science – Oklahoma State University, 2001). Ihre Haut ist rau und die Fellfarbe reicht von braunschwarz bis zu dunklem Grau (FELIUS, 1995).



Abbildung 3: Anatolisches Schwarzvieh (Bulle)
(Quelle: Department of Animal Science – Oklahoma State University,
[http:// www.ansi.okstate.edu/breeds/cattle/](http://www.ansi.okstate.edu/breeds/cattle/))

2.1.2.1.2 Illyrisches Rind

Das Illyrische Rind (IL, Abb. 4), auch Buša genannt, ist vermutlich schon seit der Zeit des Neolithikums (Jungsteinzeit) ein heimisches Rind der Balkanhalbinsel. Ein weibliches Rind erreicht im Durchschnitt eine Widerristhöhe von 104 cm und ein Gewicht zwischen 230 und 270 kg (KUGLER, 2005). Es handelt sich bei den Illyrischen Rindern also um eine relativ kleine Rasse, die in verschiedensten Schlägen in Serbien, Montenegro, Kroatien, Bosnien, Mazedonien, dem Kosovo sowie in Albanien vorkommt (FELIUS, 1995). So sind Bušas in der Region um Metohija im Kosovo beispielsweise rot gefärbt, während solche in Mazedonien blaugrau sind (KUGLER, 2005). Weiß ausgenommen, kommen fast alle Farbschläge vor. Charakteristisch für alle Buša-Schläge sind dabei die Einfarbigkeit der Tiere, wobei die Eutergegend immer eine helle Färbung aufweist, der Aalstrich, das Rehmaul sowie das einfarbige schwarze Flotzmaul. Vertreter dieser Rasse tragen kurze nach vorne und innen gebogene Hörner (FINCI, 1932; RAMLJAK *et al.*, 2011; persönliche Mitteilung Dr. Ivica Međugorac). Im Jahre 1984 vertrat diese Rasse oder vielmehr die Gesamtheit der Buša-Schläge etwa 20 Prozent der Rinderpopulation in einzelnen Ländern des Balkans (Department of Animal Science – Oklahoma State University, 2001). Die nominelle Populationsgröße nahm dann jedoch stetig ab. Allerdings konnte dieser Populationsschwund in einzelnen Ländern wieder gestoppt werden, was vorwiegend Programmen zur Erhaltung der Illyrischen Rinder beziehungsweise einzelner Buša-Schläge zu verdanken ist (EFABIS, 2008).



**Abbildung 4: Illyrisches Rind
(Kuh und Kalb verschiedener Schläge)**
(Quelle: Dr. Ivica Međugorac)

2.1.2.2 Zweinutzungsrasen

Bei Zweinutzungsrasen wird annähernd gleich viel Wert auf Fleisch- und Milchleistung gelegt. Steht einer der beiden Nutzungszwecke im Vordergrund, so spricht man auch von fleisch- respektive milchbetonten Zweinutzungsrasen (SAMBRAUS, 2001). Die im Anschluss vorgestellten Vertreter dieser Nutzungsrichtung stammen alle aus dem alpinen Raum. Es handelt sich dabei um das Deutsche Fleckvieh, das Original Braunvieh, das Fränkische Gelbvieh sowie das Murnau-Werdenfelser Rind.

2.1.2.2.1 Deutsches Fleckvieh

Das Deutsche Fleckvieh (DFV) geht auf Tiere aus dem schweizerischen Simmental zurück, weshalb diese Rasse auch unter dem Namen „Simmentaler“ bekannt ist. Nach etlichen Einkreuzungen in einige deutsche Rassen während des 18. Jahrhunderts, kamen sie erstmals am 8. Oktober 1837 in das jetzige Verbreitungsgebiet Deutschlands. Dieser erste Transport von 18 Stück Original-Simmentaler Fleckvieh wurde von Max Obermayer und Johann Fischbacher aus Gmund am Tegernsee unternommen und setzte den Grundstein für die Zucht des Deutschen Fleckviehs, das gegenwärtig über 80 Prozent des bayerischen Rinderbestands ausmacht (ROSENBERGER, 2007). Zunächst wurde die Kreuzung mit einheimischen Rindern „neuer Miesbacher Schlag“ genannt (SAMBRAUS, 2001). Im Anschluss kam es jedoch zu einer Verdrängungskreuzung aus der schließlich die Rasse „Deutsches Fleckvieh“ hervorging (FRAHM, 1982). Die Gründung eines Zuchtverbands und die Eröffnung eines Herdbuchs vollzogen sich gegen 1891. Seit Ende der 60er Jahre des 20. Jahrhunderts erfolgt die Einkreuzung von Rotbunten (Red Holstein) zur Verbesserung von Frühreife, Milchleistung und Melkbarkeit unter Erhaltung der guten Fleischleistungseigenschaften sowie der Eutergesundheit (BREM *et al.*, 1990). Beim Deutschen Fleckvieh handelt es sich um eine Zweinutzungsrasse mit gleicher Betonung von Milch- und Fleischleistung, neben welcher aber auch eine reine Fleischlinie besteht. Tiere der Zweinutzungslinie (Abb. 5) sind mittelgroß bis großwüchsig und gescheckt, wobei die Farbe von hellgelb bis zu dunklem

rotbraun variiert. Ein besonderes Rassemerkmal ist, dass der Kopf stets bis hinter die Augen weiß ist. Auch der distale Teil der Beine, Unterbauch und Schwanzquaste sind weitgehend weiß. Der Großteil dieser Tiere ist behornt (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009).

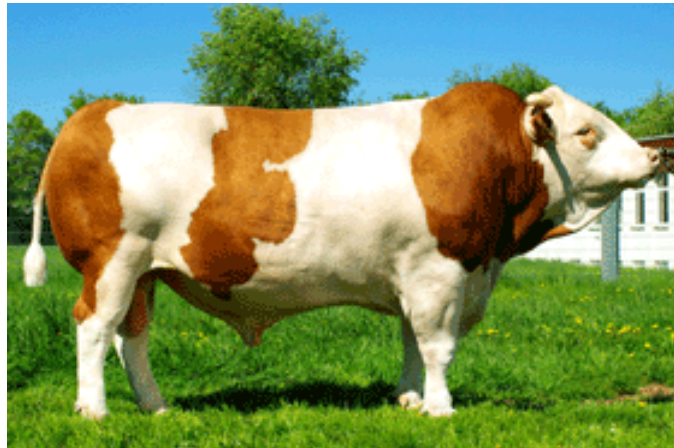


Abbildung 5: Deutsches Fleckvieh (Bulle Holzmichl)

(Quelle: Besamungsstation Greifenberg; http://www.greifenberg-genetics.de/index.php?pid=15&subpage=detail&tier_id=276000937169790)

2.1.2.2.2 Braunvieh alte Zuchtrichtung (Original Braunvieh)

In der Zentralschweiz begann man vor etwa 600 Jahren mit der Zuchtarbeit an einfarbigen vorwiegend braunen Schlägen, wofür insbesondere das Kloster Einsiedeln bekannt war. Der Export einiger Tiere nach Nordamerika erfolgte schon um die Zeit der ersten Milchleistungsprüfung gegen 1870. Aus diesen Rindern entwickelte sich in den Vereinigten Staaten die eigenständige Rasse Brown-Swiss (BS), welche seit Beginn der 60er Jahre des 20. Jahrhunderts in das europäische Braunvieh eingekreuzt wurde (SAMBRAUS, 2001). Ziel dieser Kreuzungen war die Verbesserung der Sicherheit der Melkleistung, des Formats, des Fundaments und der Euterform (BREM *et al.*, 1990). Im Zuge dieser Einkreuzungen drängte man jedoch das Braunvieh der ursprünglichen alten Zuchtrichtung an den Rand des Aussterbens. Um den Erhalt der Rasse zu sichern, wurde 1988 in Bayern die „Arbeitsgemeinschaft zur Erhaltung und Züchtung des Original Braunviehs im Allgäu“ gegründet, welche 1995 zum „Allgäuer Original Braunviehzuchtverein e.V.“ wurde. Im Rahmen von geförderten Erhaltungszuchtprogrammen wird nun versucht, die genetische Vielfalt zu bewahren, wobei die Zuchttiere in zwei

Kategorien, „A“ für Tiere mit 0 % BS-Blutanteil und „B“ für solche mit bis zu 12,5 % BS-Blutanteil, eingeteilt werden (SCHICK, 2008). Trotz dieser Bemühungen gilt das Braunvieh alter Zuchtrichtung nach der Roten Liste der Gesellschaft zur Erhaltung alter und gefährdeter Haustierrassen e. V. (GEH) immer noch als stark gefährdet (ELFRICH & ROESICKE, 2009). Während in Deutschland nur rund 350 Herdbuchtiere existieren (GEH, 2009), befinden sich in der Schweiz, in der die Erhaltung des Original Braunviehs frühzeitig begann (1981), etwa 12.000 Tiere im Herdbuch des Schweizer Braunviehzuchtverbands, so dass das schweizer Original Braunvieh für die weitere Züchtung zahlenmäßig gut genug vertreten ist (http://homepage.braunvieh.ch/index.html?page_id=140&l=2). Beim Original Braunvieh (Abb. 6) handelt es sich um ein Zweinutzungsrind mit gleicher Gewichtung von Milch und Fleisch, welches von mittelgroßer und mittelschwerer Statur ist (SCHICK, 2008). Die Färbung dieser behornen Rinder ist einheitlich braun bis graubraun, wobei sie bei Stieren etwas dunkler ausfällt als bei Kühen. Charakteristisch ist, dass sowohl Hornspitzen und Flotzmaul als auch Klauen dunkel pigmentiert sind und die Augen wie auch das Flotzmaul von einem hellen Rand gesäumt sind (SAMBRAUS, 2001). Typisch für Braunviehmilch ist der hohe Kappa-Kasein-B-Anteil, der zu einer überdurchschnittlich hohen Käseausbeute führt (MACKROTT, 1994).



Abbildung 6: Original Braunvieh (Bulle Singau)
(Quelle: Besamungsstation Greifenberg; <http://www.greifenberg-genetics.de/userfiles/file/pruefstiere/276000940438790.pdf>)

2.1.2.2.3 Fränkisches Gelbvieh

Das Fränkische Gelbvieh (FGV), auch Frankenvieh genannt, entwickelte sich vermutlich aus dem Anfang des 19. Jahrhunderts nach Franken eingeführten Heilbronner Vieh (SAMBRAUS, 2001). Dies geschah unter anderem durch Einkreuzung einfarbiger Simmentaler Bullen sowie einzelner Shorthorn- und South-Devon-Tiere zur Verbesserung von Mast- und Schlachtleistung Mitte des 19. Jahrhunderts (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009). 1872 wurde erstmals ein eindeutiges Zuchtziel für diese ursprüngliche Dreinutzungsrasse formuliert und 1897 schließlich der Zuchtverband gegründet. In den darauffolgenden Jahren erfolgte genetischer Austausch lediglich mit anderen Gelbviehrassen, bis schließlich 1963 vier Spitzenbullen der Rasse Dänisches Milchvieh, sowie 1980 weitere vier Bullen des Roten Flämischen Rindes zur Verbesserung der Milchleistung eingekreuzt wurden (BREM *et al.*, 1990; SAMBRAUS, 2001). Gegenwärtig unterscheidet man beim Gelbvieh Fleischrinder, die vor allem in anderen Ländern gehalten werden, sowie die hierzulande fast ausschließlich in kleinbäuerlichen Betrieben Frankens gezüchteten fleischbetonten Doppelnutzungsrinder. Vertreter dieser Zweinutzungsrasse sind mittel- bis großrahmig und gut bemuskelt. Sie sind einfarbig hellgelb bis rotgelb und weisen in der Regel ein helles Flotzmaul auf. Es gibt jedoch vereinzelt auch Tiere mit dunklem Flotzmaul. Darüber hinaus trägt das mehrheitlich behornte Fränkische Gelbvieh (Abb. 7) im Allgemeinen helle Hörner (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009).



Abbildung 7: Fränkisches Gelbvieh (Bulle Mongan)

(Quelle: SPERMEX GmbH; http://www.spermex.de/102.0.html?&no_cache=1&L=1%3Fref%3DSex%C5%9Ehop.com)

2.1.2.2.4 Murnau-Werdenfelser

Die Vorfahren des Murnau-Werdenfelser (MWF) Rindes wurden vermutlich ursprünglich aus Tirol eingeführt und kamen über das Kloster Ettal und andere Klöster schließlich in das jetzige namensgebende Verbreitungsgebiet, nämlich das Werdenfelser Land und das Murnauer Moos (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009). In den Jahren von 1851 bis 1986 wurden laufend Einkreuzungen zur Verbesserung der Leistungsfähigkeit und zur Inzuchtvermeidung durchgeführt (BREM *et al.*, 1990). Besonders stark verwandt ist diese Rasse speziell mit dem Braunvieh, was zum Einen daran liegt, dass sie ab 1901 bis zur Gründung eines eigenen Zuchtverbandes im Jahre 1952 mit diesem im gleichen Zuchtverband untergebracht war. Zum Anderen befand sich das Murnau-Werdenfelser-Rind nach den Tuberkulose- und Brucellosebekämpfungen in den 50er und 60er Jahren des 20. Jahrhunderts so nahe am Zusammenbruch, dass 1968 schließlich weibliche Kreuzungstiere aus Braunvieh und Murnau-Werdenfelser zur Zucht zugelassen wurden (BREM *et al.*, 1990; SAMBRAUS, 2001). Darüber hinaus erfolgte, um Inzucht zu verhindern, die Einkreuzung von Tieren der Rasse Tarentaise (BREM *et al.*, 1990). Gegenwärtig erfreut sich das Murnau-Werdenfelser Rind zwar wieder steigender Tendenz, zählt aber dessen ungeachtet mit lediglich 137 Zuchttiere im Jahre 2007 zu den extrem gefährdeten Rinderrassen. Derzeitig dient daher jedes Bestreben dazu, diese Rasse aus tierzüchterischen und landeskulturellen Gründen in ihrer ursprünglichen Form zu erhalten (ELFRICH & ROESICKE, 2009). Diese Form stellt beim Murnau-Werdenfelser (Abb. 8), welches als einzige deutsche Rasse gut an Moor- und Sumpflandschaften angepasst ist, ein kleinrahmiges Rind dar, das einfarbig hellgelb bis dunkelbraun gefärbt ist und einen hellen Aalstrich aufweist. Das an sich dunkle Flotzmaul wird von einem hellen Saum umgeben. Auch Schwanzquaste, Klauen sowie die Spitzen der ansonsten hellen Hörner dieser genügsamen alten Landrasse sind dunkel (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009).



Abbildung 8: Murnau-Werdenfelser (Bulle Bergfex)

(Quelle: Besamungsstation Greifenberg; <http://www.greifenberg-genetics.de/userfiles/file/pruefstiere/2760000935925385.pdf>)

2.1.2.3 Milchrassen

Bei den unten aufgeführten Rassen Braunvieh (neue Zuchtrichtung) und Red Holstein (Rotbunte) steht das Leistungskriterium Milch im Vordergrund, so dass man hier von Einnutzungsrasen sprechen kann (SAMBRAUS, 2001). Allerdings werden beim Braunvieh weiterhin Fleischleistungsmerkmale erfasst, weshalb diese nicht als reine Milchrasse sondern als milchbetontes Zweinutzungsrind gilt. Während das Braunvieh der neuen Zuchtrichtung auf das Original Braunvieh und somit geographisch auf eine alpine Rasse zurückgeht, ist das Red Holstein ein Vertreter der nordwest-europäischen Rassen (MEDUGORAC *et al.*, 2009).

2.1.2.3.1 Braunvieh neue Zuchtrichtung

Die Rasse American Brown-Swiss entwickelte sich in Nordamerika aus europäischen Braunviehtieren, die gegen 1870 eingeführt und jahrzehntelang einseitig auf Milchmenge gezüchtet wurden (MACKROTT, 1994). Diese Brown-Swiss Tiere wurden ab den 60er Jahren des 20. Jahrhunderts in das europäische Braunvieh eingekreuzt. Aus dieser Verdrängungskreuzung mit dieser alpinen Rasse ging das Braunvieh in der gegenwärtig üblichen neuen Zuchtrichtung

hervor, welches heute gemeinhin unter der Bezeichnung „Braunvieh“ verstanden wird. Rinder der neuen Zuchtrichtung unterscheiden sich von denen der alten Zuchtrichtung insbesondere durch ihr größeres Format (138 bis 152 cm statt 130 bis 140 cm Widerristhöhe) und die verbesserte Milchleistung (mind. 8.000 bis 9.000 kg anstelle von etwa 6.000 kg Milch). Beim mehrheitlich behornten Braunvieh der neuen Zuchtrichtung (BBV, Abb. 9) handelt es sich um ein milchbetontes Zweinutzungsrind von mittelgroßer und mittelschwerer Statur, welches bezüglich Färbung dem Original Braunvieh gleicht (SAMBRAUS, 2001; SCHICK, 2008).



Abbildung 9: Braunvieh neue Zuchtrichtung (Bulle Husant)

(Quelle: Besamungsstation Greifenberg; http://www.greifenberg-genetics.de/index.php?pid=16&subpage=detail&tier_id=276000812891894)

2.1.2.3.2 Red Holstein

Red Holsteins (RH) lassen sich auf die Deutschen Rotbunten zurückführen, die ihren Ursprung in den deutschen Küstenregionen der Nord- und Ostsee haben. In diese schon sehr lange existierenden Rinder wurden Anfang des 19. Jahrhunderts Shorthorn-Tiere eingekreuzt. Außerdem erfolgte eine Teilung in unterschiedliche Schläge. Im Jahre 1876 wurde in Fischbek bei Altmark die erste deutsche offizielle Stammzuchtgenossenschaft gegründet, in deren Folge sich immer mehr Vereinigungen und Genossenschaften etablierten, wovon die meisten schwarzbunte und rotbunte Tiere zunächst im gleichen Zuchtbuch führten (FELIUS, 1995; ELFRICH & ROESICKE, 2009). Bereits um 1875 wurden schwarzbunte Rinder nach Nordamerika exportiert, wo man dann größere Tiere

mit hoher Milchleistung und schwächerer Bemuskelung (Holstein Friesian) züchtete. Ab den siebziger Jahren des letzten Jahrhunderts wurden diese Holstein Friesians dann zunehmend in die Zweinutzungsrasse eingekreuzt, so dass eine reine Milchrasse entstand (Deutsche Holstein). Diese Entwicklung vollzog sich separat zunächst für die schwarzbunten und etwas später auch für die rotbunten Rinder, die zu dieser Zeit von den Zuchtverbänden getrennt erfasst wurden. Da sich aber die Zuchtziele immer mehr annäherten, wurde 1996 für rotbunte und schwarzbunte Holsteins wieder ein gemeinsames Zuchtziel definiert und der Deutsche Holstein Verband e. V. (DHV) gegründet. Die rotbunten Vertreter der Holstein-Rasse, die sogenannten Red Holstein (Abb. 10), sind großrahmige, hochbeinige, flach bemuskelte Tiere mit rot-weißer Scheckung, selten auch fast weiß oder komplett gefärbt mit wenigen Abzeichen. Tiere dieser Rasse sind im Allgemeinen behornt (ELFRICH & ROESICKE, 2009).



Abbildung 10: Red Holstein (Bulle Malvoy)

(Quelle: Besamungsstation Greifenberg; http://www.greifenberg-genetics.de/index.php?pid=17&subpage=detail&tier_id=132414870)

2.1.2.4 Fleischrassen

Bei den Fleischrassen handelt es sich ebenfalls um Einnutzungsrasse, allerdings steht hier wie der Name schon sagt das Leistungskriterium Fleisch im Vordergrund (SAMBRAUS, 2001). Hierzu zählen die nachfolgend beschriebenen Rassen Weißblauer Belgier und Galloway, welche beide Vertreter der nordwest-europäischen Gruppe sind (MEDUGORAC *et al.*, 2009).

2.1.2.4.1 Weißblauer Belgier

Die Weißblauen Belgier (BBB) gehen auf zwei im Typ unterschiedliche Landschläge zurück. In der zweiten Hälfte des 19. Jahrhunderts erfolgte zusätzlich die Einkreuzung von Charolais- und Shorthornrindern. Ursprüngliches Zuchtziel war ein Zweinutzungsrind (FELIUS, 1995). Doch gegen 1950 wurde schließlich bei einem Teil der Population die Zucht auf Muskelfülle in den Vordergrund gestellt, so dass gegenwärtig sowohl Zweinutzungstyp als auch Fleischtyp, dem die Mehrzahl der Zuchttiere angehört, existieren (FELIUS, 1995; SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009). Diese Selektion auf Muskelhypertrophie führte jedoch zu einem Anstieg außergewöhnlich bemuskelter Tiere und damit assoziierten Problemen wie beispielsweise reduzierter Fruchtbarkeit und Dystokie (ARTHUR, 1995, KAMBADUR *et al.* 1997). Grund für diese extreme Muskelentwicklung ist eine Mutation in dem Gen, welches für die Bildung von Myostatin zuständig ist. Dieses Protein verhindert normalerweise unkontrolliertes Muskelwachstum (GROBET *et al.*, 1997; KAMBADUR *et al.* 1997). Bei seinem Fehlen kommt es daher zum Phänotyp des schwerbemuskelten „Doppellenders“. Daneben hemmt diese Mutation auch den Fettansatz, was zu sehr magerem Fleisch führt. Der Weißblaue Belgier (Abb. 11) ist ein mittelgroßes bis großrahmiges Rind mit relativ kleinem Kopf und kurzen, stumpfen Hörnern, das in den Färbungen weiß und blauweiß-gescheckt sowie gelegentlich in schwarz-weiß vorkommt (SAMBRAUS, 2001; ELFRICH & ROESICKE, 2009). Während diese Rasse mit einem Anteil von 45 Prozent die populärste Rasse Belgiens ist, gibt es im Bundesgebiet nur wenige Reinzuchtherden. Jedoch erfolgen zum Teil Gebrauchskreuzungen mit körperlich ausgereiften Kühen der Zweinutzungsrassen (HAMPEL, 1994).



Abbildung 11: Weißblauer Belgier (Bulle Graphite de Dessous la Ville)

(Quelle: Hr. A. Osio, Belgian Blue Group, Ciney)

2.1.2.4.2 Galloway

Ihren Namen verdanken diese Rinder der Region Galloway, die sich im Südwesten Schottlands befindet und in welcher diese Rasse vor Jahrhunderten entstand (SAMBRAUS, 2001). Dabei stammt das Galloway, welches die älteste Rasse Großbritanniens ist, von überwiegend behornen Tieren ab, die seit der Zeit der Kelten in Schottland anzutreffen waren. Erst während des 18. Jahrhunderts wurden hornlose Tiere bevorzugt. 1780 begann dann die Zucht eines reinen Einnutzungsrindes vom Fleischtyp, was nach einem Jahrhundert selektiver Zucht in der Rasse Galloway resultierte (FELIUS, 1995). Die Gründung der Zuchtgesellschaft erfolgte im Jahre 1878 und bereits ein Jahr später wurde das erste Herdbuch herausgegeben. Rinder dieser Rasse sind dominant hornlos und besitzen einen breiten aber kurzen Kopf. Sie weisen einen kleinen bis mittleren Rahmen auf und haben langes, weiches, welliges Fell mit dichtem Unterhaar (SAMBRAUS, 2001). Galloways (GLW, Abb. 12) kommen hauptsächlich in den Farben schwarz mit rotbraunem Unterhaar, blond (dun) und rot, selten auch „parkfarben“ mit weißem Fell sowie schwarzer Pigmentierung des Flotzmauls und der Ohren vor. Tiere mit schwarzer, roter oder blonder Grundfarbe und weißer „Leibbinde“ hingegen bilden eine eigene Rasse namens „Belted Galloways“ (HAMPEL, 1994; ELFRICH & ROESICKE, 2009).



Abbildung 12: Galloway (Bulle Dean)

(Quelle: Rinderzucht Schleswig-Holstein eG; <http://www.rsheg.de/bullen/zucht-und-bullenangebote/top-fleischrinder.html>)

2.2 Populationsgenetische Prozesse

In einer gegebenen Population können drei verschiedene Ereignisse die genetische Variation erhöhen. Dabei handelt es sich um Mutation, Rekombination und Immigration von Genen. Wobei als Voraussetzung für eine Änderung der genetischen Zusammensetzung mittels Rekombination bereits eine Mutation, die ohnehin die Quelle aller genetischen Variation darstellt, stattgefunden haben muss. Ebenso müssen in der immigrierenden Population schon Allele vorhanden sein, die in der gegebenen Population fremd sind, damit eine solche Migration überhaupt einen Effekt auf die genetische Vielfalt haben kann (GRIFFITHS *et al.*, 2008). Die Genotyphäufigkeiten werden darüber hinaus sowohl von Migration und Selektion als auch durch genetische Drift, also zufälligen Genotypfrequenzschwankungen in kleinen Populationen, beeinflusst (KRÄUßLICH, 1994). Zusätzlich ändern nicht zufällige Paarungen, wie assortative und disassortative Paarung sowie Inzucht, die Genotypenfrequenz der Nachkommen (KRÄUßLICH & BREM, 1997). Im Folgenden wird ein kurzer Überblick über diese wichtigsten sechs das Genom beeinflussenden Prozesse, nämlich Mutation, Rekombination, Migration, Selektion, Gendrift und nicht zufällige Paarung, gegeben.

2.2.1 Mutation

Mutationen werden durch Fehler bestimmter molekularer Vorgänge, wie Replikation oder Rekombination, sowie durch Mutagene verursacht und stellen relativ seltene Ereignisse dar (SENDLINGER *et al.*, 2006). Generell gibt es verschiedene Klassen von Mutationen. Dabei sind die strukturelle und numerische Chromosomenmutation sowie die Genommutation der Zytogenetik zuzuordnen. Im Arbeitsbereich der Molekulargenetik (KRÄUßLICH, 1994) lassen sich entsprechend ihrer Auswirkung auf die DNA-Sequenz Basensubstitutionen, Deletionen und Insertionen unterscheiden. Erreicht eine so entstandene neue Variante eine Häufigkeit von über einem Prozent, spricht man üblicherweise von einem DNA-Polymorphismus, da sich eine solche Frequenz nicht mehr durch das alleinige wiederholte Auftreten von Neumutationen erklären lässt. Solche

Polymorphismen können dann, wie in Kapitel 2.3.2 beschrieben, als DNA-Marker genutzt werden (STRACHAN & READ, 2005). Auch wenn Mutationen die Voraussetzung für jegliche genetische Variabilität und daher jegliche evolutionäre Entwicklung darstellen, würde sich eine Veränderung in der Genfrequenz ohne zusätzliche Prozesse nur sehr langsam einstellen (GRIFFITHS *et al.*, 2008). Dies liegt daran, dass die spontane Mutationsrate der Mehrzahl der Loci der meisten Organismen mit etwa 10^{-5} bis 10^{-6} pro Generation sehr gering ist (FALCONER & MACKAY, 1996).

Nach Auftreten einer Mutation spricht man bei dem ursprünglichen Allel auch von „*ancestral*“, bei dem davon abgeleiteten neuen Allel von „*derived*“ (VOIGHT *et al.*, 2006). Die meisten der neu entstehenden Mutationen haben keine offensichtlichen Auswirkungen oder sind sogar nachteilig, so dass sie im Allgemeinen selektionsneutral sind oder negativer Selektion unterliegen (siehe 2.2.4). Manche Mutationen sind jedoch vorteilhaft und somit in der Lage die Anpassungsfähigkeit einer Population an bestimmte Selektionsdrücke zu erhöhen (RADMAN, 1999).

2.2.2 Rekombination

Unter Rekombination versteht man den zufälligen Austausch von Abschnitten homologer Chromosomen. Dieser Vorgang ist bei der Entwicklung der Keimzellen als Crossing-over bekannt (STRACHAN & READ, 2005). Im Gegensatz zur Mutation entstehen bei der Rekombination keine neuen Allele, sondern neue Kombinationen bereits vorhandener Allele. Allerdings kann eine fehlerhafte Rekombination auch zu Deletionen und Insertionen führen. Angenommen auf einem Chromosom sind zwei benachbarte Loci A und B monomorph für die Allele „a“ und „b“ und an beiden Loci entstehen durch Mutation zwei günstige Allele. Da Mutation ein sehr seltenes Ereignis darstellt, entstehen hierdurch höchstwahrscheinlich zwei neue Gameten respektive Haplotypen „Ab“ und „aB“. Ohne Rekombination bliebe es in der Population bei dem ursprünglichen Haplotypen „ab“ sowie den beiden neuen durch Mutation entstandenen „Ab“ und „aB“. Erst durch Rekombination kommt es auch zur Kombination der beiden neuen („*derived*“) günstigen Allele, also zu „AB“. Entspricht nach wiederholten Rekombinationsereignissen die Frequenz der jeweiligen Allelkombinationen dem

Produkt der Frequenzen der beteiligten Allele, befinden sich die Loci im sogenannten „*linkage equilibrium*“. Umso stärker gekoppelt die betreffenden Gene sind, desto mehr Generationen sind letztendlich nötig, um diesen Zustand des Gleichgewichts zu erlangen. Das ursprüngliche Ungleichgewicht, genannt „*linkage disequilibrium*“ (LD), verringert sich dann von Generation zu Generation proportional zu der Anzahl an Rekombinationen zwischen diesen Genen. Dieser Umstand ermöglicht es die Position von unbekanntem Genen ausfindig zu machen und den Beweis zu führen, dass ein bestimmter Phänotyp von einem unbekanntem Gen beeinflusst wird (GRIFFITHS *et al.*, 2008).

Die enorme Bedeutung der Rekombination für die Genomentwicklung und den Selektionsfortschritt zeigten Zhang *et al.* (2002) beispielhaft mittels Selektionsexperimenten an Mikroorganismen.

2.2.3 Migration

Migration beschreibt das Ein- oder Auswandern von Genen in Form von Individuen und Gameten von einer Population in eine Andere. Dabei treten Änderungen in der Genotypfrequenz immer dann auf, wenn sich Donor- und Empfängerpopulation in ihren Allelhäufigkeiten unterscheiden (KRÄUßLICH, 1994). Die Frequenz der daraus entstehenden gemischten Population liegt dann irgendwo zwischen dem originalen Wert und dem der Donorpopulation. Dabei ist die Änderung in der Genotypfrequenz proportional zur Migrationsrate sowie der Differenz der Frequenzen der beteiligten Populationen. Die Migrationsrate kann dabei im Gegensatz zur Mutationsrate sehr hoch sein (GRIFFITHS *et al.*, 2008). Die Allelfrequenz der vermischten Population (q_1) beträgt nach einer Generation in Abhängigkeit von der Donorfrequenz (q_m) der Immigranten, der Frequenz der ursprünglichen Population (q_0) und dem Anteil der immigrierenden Individuen, bezogen auf die neue Gesamtpopulation (m), folgenden Wert:

$$q_1 = mq_m + (1 - m)q_0 = m(q_m - q_0) + q_0$$

Migration ist daher ein systematischer Prozess, bei dem Änderungen in der Genfrequenz in einer Art erfolgen, die es zulässt sowohl Größe als auch Richtung vorherzusagen (FALCONER & MACKAY, 1996).

Die Migration von Genmaterial ist in der Haustierpopulation von entscheidender Bedeutung, wenn Populationen rasch verändert respektive den aktuellen Marktansprüchen angenähert werden sollen, oder um eine schnellere Anpassung an veränderte Umweltbedingungen zu ermöglichen. Erleichtert wird dieser Genaustausch durch moderne Reproduktionstechniken, wie Künstliche Besamung (KB) oder Embryotransfer (ET; KRÄUßLICH & BREM, 1997).

2.2.4 Selektion

Wirkt sich eine Mutation auf die betroffenen Individuen insofern aus, dass sich ihre Nachkommenzahl im Mittel von denen, die nicht Träger dieser Mutation sind unterscheidet, spricht man von Selektion. Handelt es sich hierbei um einen Selektionsvorteil, so wird das jeweilige Gen häufiger (positive Selektion), bei einem Nachteil hingegen seltener (negative Selektion) werden (TARIVERDIAN & BUSELMAIER, 2004). Selektion verändert daher die Genfrequenz in Populationen. Geschieht dies in vom Menschen unbeeinflusster Natur, so spricht man von natürlicher Selektion. In deren Folge werden meist die relativen Häufigkeiten jener Allele verändert, die die Überlebensfähigkeit, die Fruchtbarkeit oder beide zusammen unter den aktuellen Lebensbedingungen, also der Umwelt, positiv beeinflussen. In diesem Zusammenhang ist häufig von „reproduktiver Fitness“ die Rede (KRÄUßLICH & BREM, 1997). Dabei ist „Fitness“ definiert als die individuelle Fähigkeit sich zu entwickeln und vitale Nachkommen zu zeugen (FALCONER & MACKAY, 1996). Sowohl natürliche als auch künstliche Selektion basieren daher auf Phänotypen und nicht auf Genotypen (WEISS & FULLERTON, 2000). Deswegen ist die „Fitness“ nicht vom Genotypen allein abhängig, sondern vielmehr die Konsequenz aus der Beziehung zwischen dem entsprechenden Phänotypen und der Umwelt, in der er lebt. So kann ein und derselbe Genotyp, wenn er verschiedenen Umweltsituationen ausgesetzt wird, unterschiedliche Fitnesswerte aufweisen (GRIFFITHS *et al.*, 2008). Im Zuge der Zucht greift der Mensch jedoch aktiv in das Fortpflanzungsgeschehen ein und wählt nach festgelegten Kriterien Zuchttiere aus. Diese künstliche oder artifizielle Selektion, die mit dem Ziel der Leistungsverbesserung durchgeführt wird, ist die wichtigste züchterische Maßnahme zur Veränderung der Genhäufigkeiten. Jedoch können

auch bei der künstlichen Selektion die für die natürliche Selektion wichtigsten Merkmale, die Vitalität und die Fertilität, nicht gänzlich außer Acht gelassen werden, da dies die Nachhaltigkeit der Tierzucht und die züchterische Nutzung von vielversprechenden leistungsstarken Tieren verhindern könnte (KRÄUßLICH & BREM, 1997).

Der Suche nach Selektionssignaturen wird im Allgemeinen ein einfaches Model der positiven Selektion zugrunde gelegt, indem Mutationen sobald sie entstehen bevorzugt werden. Diese Annahme favorisiert also immer das „*derived*“ Allel. Darüber hinaus setzt hier Selektion immer bei Allelen ein, die nur in sehr niedriger Frequenz vorliegen. Allerdings kommt in realen Populationen nicht nur Selektion von „*derived*“ Allelen vor. Zum einen können Veränderungen der Umwelt dazu führen, dass zuvor neutrale oder sogar schwach nachteilige Allele plötzlich vorteilhaft sind, was bedeuten kann, dass auch „ancestrale“ Allele unter Selektion stehen können. Zum Anderen können solche neutralen respektive schwach nachteiligen Allele auch allein durch genetische Drift (siehe 2.2.5) hohe Frequenzen erlangen und neben der ursprünglichen Variante bestehen, so dass bei einer dann einsetzenden Selektion auf eine solche „*standing genetic variation*“ die Initialfrequenz bereits sehr hoch sein kann (PRZEWORSKI *et al.*, 2005).

2.2.5 Gendrift

Gendrift, auch genetische oder zufällige Drift genannt, bezeichnet die zufälligen Veränderungen von Genotypfrequenzen, welche von Generation zu Generation zu beobachten sind (FALCONER & MACKAY, 1996). Dabei liegt der Grund für diese unvorhersehbaren Veränderungen der Allelfrequenzen in dem Vorhandensein kleiner Populationsgrößen. Der Effekt, den die Gendrift dabei auf das Genom ausübt, ist indirekt proportional zur effektiven Populationsgröße (N_e). Dies bedeutet, dass in komplett isolierten Gruppen der gleichen Population die Gendrift allein zur Fixierung von verschiedenen Allelen in selbigen Gruppen führen würde. Im Unterschied zur (natürlichen) Selektion wirkt sich die Gendrift im Durchschnitt gleichermaßen auf alle Loci aus (SLATKIN, 1987). Jedoch sind auch die hier zufällig fixierten oder angehäuften Allele immer im Kontext von Haplotypen zu verstehen. Deshalb ist Gendrift durchaus in der Lage ein

Kopplungsungleichgewicht zwischen nahegelegenen Loci zu erzeugen. Dies erschwert insbesondere in Hochleistungsrassen mit häufig sehr kleiner effektiver Populationsgröße und ausgewählter („*assortative*“) Paarung (siehe 2.2.6) die Unterscheidung der Auswirkungen von Gendrift und Selektion (OHTA & KIMURA, 1969; QANBARI *et al.*, 2010b).

2.2.6 Nicht zufällige Paarung

Man spricht in einer Population mit gleicher Anzahl beider Geschlechter von Panmixie, wenn jedes Individuum die gleiche Chance hat, sich mit jedem Individuum des anderen Geschlechts mit gleicher Fruchtbarkeit zu paaren, während gleichzeitig weder Mutation, Selektion noch Migration auftreten. Dies ist dem Begriff des „*random mating*“, also der zufälligen Paarung, gleichzusetzen. Darüber hinaus gilt unter diesen Umständen das Hardy-Weinberg-Gesetz, welches besagt, dass die Genhäufigkeiten, also der Anteil der verschiedenen Allele eines Gens in einer Population, und damit die Häufigkeiten der beiden Homozygoten und des heterozygoten Genotyps, von Generation zu Generation konstant bleiben. Diese Beziehung zwischen Genhäufigkeit und Genotypenhäufigkeit bleibt, solange Panmixie herrscht, über alle weiteren Generationen bestehen (TARIVERDIAN & BUSELMAIER, 2004). Da das „*random mating*“ bei der Zucht von reinrassigen Rindern jedoch praktisch nicht vorkommt (THRIFT, 2000), wird nun auf die Effekte des „*nonrandom mating*“, also der nicht zufälligen Paarung, eingegangen. Generell sind hierbei zwei Formen zu unterscheiden, nämlich die Inzucht und die ausgewählte Paarung („*assortative mating*“). Letztere bezieht sich immer auf ausgewählte phänotypische Merkmale und lässt sich in „*positive assortative mating*“, bei welchem phänotypisch möglichst ähnliche Partner bevorzugt werden, und „*negative assortative mating*“, bei dem möglichst unterschiedliche Paarungen favorisiert werden, unterteilen. Dem „*positive assortative mating*“ und der Inzucht ist gemein, dass beide zu einem Anstieg der Homozygotie über das, durch das Hardy-Weinberg-Gesetz, vorausgesagte Level führen. Von Inzucht spricht man, wenn Verpaarungen zwischen Verwandten häufiger vorkommen, als dies durch reinen Zufall geschehen würde, wobei „verwandt“ bedeutet, dass ihnen mindestens ein

Vorfahre gemein ist. Sie betrifft immer das gesamte Genom und erhöht darüber hinaus die Chance der „*homozygosity by descent*“. Diese erhöhte Wahrscheinlichkeit, dass homozygote Allele auf den gleichen Ahnen zurückzuführen sind, entspricht dem „*inbreeding coefficient*“ F . Inzucht ist in der Lage die genetische Variation innerhalb der ursprünglichen Population in Variationen zwischen homozygoten Inzuchtlinien zu konvertieren (GRIFFITHS *et al.*, 2008). Theoretisch würde sich bei vollständiger Inzucht ($F = 1$) die genetische Varianz der Gesamtpopulation verdoppeln und sich außerdem vollständig zwischen den Inzuchtlinien befinden, während die Varianz innerhalb der Inzuchtlinien Null erreichen würde (FALCONER & MACKAY, 1996). Praktisch führt Inzucht in der Tierzucht jedoch zu unwiderruflichen Verlusten der genetischen Variabilität, da in einzelnen Inzuchtlinien, Teilpopulationen oder Rassen meist die gleichen seltenen genetischen Varianten verloren gehen. Dieser Verlust gleicher Varianten kann insbesondere durch das auf geographisch begrenztem Raum (Alpen und Nordwest Europa) gehäufte Vorkommen von Rinderrassen mit ähnlichen Zuchtzielen verstärkt werden (MEDUGORAC *et al.*, 2009).

2.3 Molekulare Marker

Um die genetische Variation studieren zu können ist man, sofern sich die Allele an dem jeweiligen Locus phänotypisch nicht so eindeutig darstellen, wie beispielsweise die von Johann (Taufname) Gregor (Ordensname) Mendel untersuchten Merkmale der Erbsen (z.B. WEILING, 1966), auf molekulare Marker angewiesen. Im Folgenden wird über die historische Entwicklung dieses noch jungen Arbeitsgebiets der molekularen Forschung ein kurzer Einblick gegeben, um dann anschließend genauer auf die Single Nucleotid Polymorphisms einzugehen, da diese DNA-Marker in der vorliegenden Arbeit Verwendung fanden.

2.3.1 Historische Entwicklung molekularer Marker

Die eigentliche molekulare Forschungsarbeit begann um die 60er Jahre des 20. Jahrhunderts als Blutproteine mit Hilfe der Proteinelektrophorese in Einzelkomponenten aufgetrennt wurden, um anschließend die Aminosäuresequenzen zu bestimmen. Dieses Verfahren war allerdings für manche Proteine sehr aufwendig und hatte den entscheidenden Nachteil, dass mittels der angewendeten Proteinelektrophorese Unterschiede in der Aminosäuresequenz oftmals nicht zu erkennen waren (STORCH *et al.*, 2007).

Die Entdeckung der Restriktionsendonukleasen durch Arber und Dussoix (1962) legte den Grundstein, für die besonders zwischen 1970 und 1990 häufig durchgeführte RFLP-Analyse (Restriktionsfragment-Längen-Polymorphismus). Restriktionsendonukleasen sind spezielle Enzyme, die spezifische Sequenzabschnitte der DNA, meist Palindromstrukturen, erkennen und so fähig sind die DNA in Abschnitte definierter Länge zu zerschneiden, so dass nach Gelelektrophorese komplexe Bandenmuster entstehen (STORCH *et al.*, 2007).

Im Jahre 1977 wurden dann schnelle Methoden zur DNA-Sequenzierung entwickelt. Dies waren zum Einen die chemische Sequenzierung durch Maxam und Gilbert (1977) und zum Anderen die Strangabbruchmethode von Sanger *et al.* (1977), welche sich heute vor allem bei der automatischen Sequenzierung durchgesetzt hat.

Die von Mullis *et al.* (1986) hervorgebrachte Polymerase-Kettenreaktion (PCR), die es ermöglicht einzelne DNA-Abschnitte gezielt *in vitro* zu vermehren, so dass für eine anschließende Sequenzierung genug Material zur Verfügung steht, stellt die Basis für die folgenden Analysemethoden dar.

Die Mikrosatelliten- oder STR-Analyse (short tandem repeats) findet aufgrund ihrer hohen Empfindlichkeit auch heute noch oft Verwendung. Dabei flankieren Primer von circa 20 Nukleotiden Länge polymorphe Mikrosatelliten-Loci. Am Ende bekommt man bei diploiden Zellen pro heterozygoten Locus zwei Banden, die das väterliche und mütterliche Allel darstellen.

Die vor mehr als zehn Jahren entwickelte AFLP-Analyse (amplified fragment length polymorphisms) kombiniert die Selektivität der RFLP mit der Empfindlichkeit der PCR ohne auf das Vorhandensein spezifischer PCR-Primer angewiesen zu sein (STORCH *et al.*, 2007). Daher wurden AFLP Marker besonders häufig zur

genetischen Analyse von Spezies mit weniger bekannter genetischer Information (Sequenz) verwendet (z.B. BONIN *et al.*, 2007).

Natürlich wurde im Laufe der Zeit neben den oben erwähnten Methoden eine Reihe weiterer entwickelt. Zu diesen gehören beispielsweise die DNA-DNA-Hybridisierung (z.B. SIBLEY & AHLQUIST, 1984), die RAPD-Analyse („*randomly amplified polymorphic DNA*“) sowie die ISSR-Analyse („*inter simple sequence repeats*“). Diese haben sich jedoch nicht weiter etabliert oder fließen nur in Teilen in die gegenwärtige Forschung ein (STORCH *et al.*, 2007).

In den letzten Jahren sind darüber hinaus die Single Nucleotide Polymorphisms (SNP) von stark zunehmender Bedeutung (STORCH *et al.*, 2007). Dabei sind SNP Marker, die sich speziell für genetische Studien mit höherem Probendurchsatz besonders gut eignen. Sie werden aufgrund des Stellenwerts, der ihnen in dieser Arbeit zukommt im nächsten Abschnitt ausführlicher vorgestellt.

2.3.2 Single Nucleotide Polymorphisms

Unter einem Single Nucleotide Polymorphism (SNP, ausgesprochen „Snip“), zu Deutsch Einzelnukleotidpolymorphismus, versteht man DNA-Varianten, die nur einzelne Nucleotidpositionen betreffen (GELDERMANN, 2005a). Der Nachweis über Abweichungen einzelner Nucleotide erfolgt in der Regel durch DNA-Sequenzierung oder Primer-Verlängerungstests. Kommt es durch einen solchen Polymorphismus zur Entstehung oder zum Verlust einer Restriktionsschnittstelle, wird dieser besondere SNP auch „*restricion site polymorphism*“ (RSP; Restriktionsschnittstellen-Polymorphismus), oder, sofern er zu einer Längenveränderung des Restriktionsfragments führt, „*restriction fragment length polymorphism*“ (RFLP, Restriktionsfragment-Längenpolymorphismus) genannt (STRACHAN & READ, 2005). Innerhalb einer Spezies sind über 90 Prozent der Sequenzunterschiede auf Ungleichheiten in einzelnen Basen zurückzuführen. Im Rindergenom, welches etwa drei Milliarden Basenpaare (bp) umfasst, sind circa 2,5 Millionen solcher variabler (= polymorpher) Stellen zu finden (BRADE, 2010). Da codierende Regionen nur etwa fünf Prozent des Genoms ausmachen, und

SNPs in diesen und regulierenden Regionen (cSNP) mit höherer Wahrscheinlichkeit funktionelle Genveränderungen mit sich bringen, bleiben die meisten Polymorphismen ohne funktionelle Bedeutung. Dennoch sind viele dieser selbst funktionslosen SNPs durchaus nützlich, um den Phänotyp beeinflussende Veränderungen zu kartieren, da sich einige dieser Allele in „*linkage disequilibrium*“ (LD) mit eng gekoppelten Loci befinden (COLLINS *et al.*, 1998). Gegenüber Mikrosatelliten zeichnen sich Einzelnukleotidpolymorphismen vor allem durch vier Merkmale aus: (1) sie sind häufiger im Genom vertreten, (2) einige haben selbst Einfluss auf die codierte Proteinstruktur oder deren Expression, wodurch es zur Entstehung neuer Phänotypen kommen kann, (3) sie werden beständiger vererbt (deutlich niedrigere Mutationsrate) und (4) sie sind dank bi-allelischer Natur mittels moderner Methoden mit höherem Probendurchsatz leichter nachzuweisen (LANDEGREN *et al.*, 1998). Allerdings beschrieben Weiss und Terwilliger (2000) eindrucksvoll mögliche Nachteile von SNPs und dämpften so übertriebene Erwartungen.

2.4 Signaturen Positiver Selektion

Selektion bewirkt im Genom verschiedene Veränderungen, die eine Unterscheidung dieser Bereiche von neutralen Genomabschnitten möglich machen. Diese Signaturen, sowie ihre Analysemethoden werden nun vorgestellt, bevor dann auf die Ergebnisse bereits durchgeführter Studien eingegangen wird, deren Ziel ebenfalls das genomweite Auffinden von Selektionssignaturen war.

2.4.1 Verschiedene Selektionssignaturen und ihre Analysemethoden

Bis vor Kurzem war die einzige geeignete Methode, Fälle positiver Selektion zu identifizieren, einzelne Kandidatengene zu untersuchen. Dies änderte sich jedoch mit dem Aufkommen genomweiter Sequenz- und Polymorphismendaten (SABETI *et al.*, 2006), was dazu führte, dass seit Ende der 80er Jahre des letzten Jahrhunderts laufend neue statistische Tests zum Nachweis positiver Selektion

auf DNA-Ebene in Populationsproben entwickelt wurden und immer noch werden (KREITMAN, 2000). Diese neuen populations-genomischen Verfahren ermöglichen das Ausfindigmachen der stärksten Kandidatenloci unter Selektion sowie die Neubewertung bereits bekannter Regionen. Dabei beruhen diese neuen Methoden darauf, charakteristische Signaturen genetischer Variation in der DNA-Sequenz aufzuspüren, indem simultan eine Vielzahl von Loci respektive Genomregionen untersucht werden (LUIKART *et al.*, 2003; SABETI *et al.*, 2006). Diese entstehen während eines sogenannten „*selective sweep*“, also dem Prozess bei dem eine neue vorteilhafte Mutation die Variation mit ihr gekoppelter neutraler Regionen reduziert oder eliminiert, während ihre Frequenz in der Population ansteigt (MAYNARD-SMITH & HAIGH, 1974; NIELSEN, 2005). Diese Selektionssignaturen lassen sich durch Vergleich mit der Hintergrundverteilung der genetischen Variation identifizieren, wobei der Großteil der beobachteten Hintergrundvariation neutral ist, was bedeutet, dass sich die Prävalenz in der Population meist allein durch zufällige Drift ändert (siehe 2.2.5). Mittlerweile gibt es eine große Anzahl spezifischer statistischer Tests auf Selektion, die jedoch im Allgemeinen auf den Nachweis von lediglich fünf verschiedenen typischen Signaturen zurückgreifen und sich daher in fünf entsprechende Gruppen unterteilen lassen. Diese Gruppen, jeweils benannt nach der entsprechenden nachzuweisenden charakteristischen Signatur sowie ihre methodischen Stärken und Schwächen, werden im Folgenden dargestellt (SABETI *et al.*, 2006).

2.4.1.1 Hohe Anzahl funktionsverändernder Mutationen

Dieses sehr gebräuchliche Nachweisverfahren für positive Selektion profitiert von dem Umstand, dass es in codierenden Regionen zwei verschiedene Arten von Mutation gibt. Nicht-synonyme Mutationen haben eine Veränderung der resultierenden Aminosäuresequenz zur Folge, während synonyme Mutationen keinerlei Auswirkung auf die jeweilige codierte Aminosäure haben. Es wird angenommen, dass beide Mutationstypen gleich häufig auftreten. Daher stellt ein Übermaß von nicht-synonymen im Vergleich zu synonymen Mutationen ein starkes Indiz für einen früheren Einfluss positiver Selektion auf Proteinebene dar (NIELSEN *et al.*, 2005). Erfasst werden diese Selektionssignaturen durch den

Vergleich der Rate von nicht-synonymen zu synonymen oder mutmaßlich neutralen Veränderungen mit der entsprechenden Rate anderer Abstammungslinien oder anderer Spezies (SABETI *et al.*, 2006). Dass ein solcher Vergleich von synonymen und nicht-synonymen Substitutionen ein nützlicher Weg ist, um positive Selektion zu erforschen, zeigten auch Hughes und Nei (1988). Der Vorteil dieser Methode besteht darin, dass die Signaturen eines großen Bereichs der evolutionären Zeitskala, nämlich viele Millionen Jahre zurückreichend, ausfindig gemacht werden können, und dass der Fokus dieses Nachweises auf den nutzbringenden Allelen selbst beruht. Jedoch ist die Leistung dieser Methode dadurch eingeschränkt, dass viele selektierte Veränderungen nötig sind, bevor sich ein Gen vom Hintergrund der neutralen Mutationsrate unterscheidet. Darüber hinaus ist lediglich anhaltende oder rekurrente Selektion nachweisbar (SABETI *et al.*, 2006).

2.4.1.2 Reduktion der genetischen Diversität

Anders als beim Nachweis einer hohen Anzahl funktionsverändernder Mutationen, wo verschiedene Spezies miteinbezogen werden, wird nach Genombereichen mit reduzierter genetischer Diversität stets innerhalb einer Spezies gesucht. Bei den hierbei üblicherweise verwendeten Varianten handelt es sich um Single Nucleotide Polymorphisms (SABETI *et al.*, 2006).

Die Entstehung der aufzuspürenden Signatur wird nun kurz dargestellt. Steigt die Frequenz eines Allels einer Population an, so beeinflusst dieser Vorgang auch die diesem Allel naheliegenden und mit ihm gekoppelten Varianten desselben Chromosoms. Dieser Effekt, dass eigentlich neutrale Allele ebenfalls eine Frequenzerhöhung erfahren, wird „*hitchhiking*“ genannt und führt zu sogenannten „*selective sweeps*“, also der Umgestaltung der typischen Struktur der genetischen Variation in einer Region (MAYNARD-SMITH & HAIGH, 1974; NIELSEN, 2005). Diese Bereiche stellen dann aufgrund ihrer geringen Gesamtdiversität und dem Übermaß an seltenen Allelen eine Signatur positiver Selektion dar. Von einem „*complete selective sweep*“ spricht man dann, wenn die Frequenz des selektierten Allels so lange ansteigt, bis es schließlich fixiert wird. Dabei werden eng gekoppelte Varianten auf die gleiche Ebene gebracht, was zur Eliminierung der

Diversität in der unmittelbaren Umgebung und zu ihrem Abfallen in einer ausgedehnteren Region führt (SABETI *et al.*, 2006). Wie groß dabei die Genomregion ist, die von solch einem Sweep beeinflusst wird, hängt von der Selektionsintensität, also der Geschwindigkeit mit der ein selektiertes Allel hohe Frequenzen erreicht, ab. Geschieht dies sehr schnell (hohe Intensität), sind die betroffenen Bereiche entsprechend groß. Dabei steigt die Größe des Intervalls in etwa mit dem durch die Mutation bedingten Selektionsvorteil (s). Angenommen ein Allel hätte einen Selektionsvorteil von einem Prozent (gleichzusetzen mit moderater Selektionsintensität), so entspricht die Modelgröße der beeinflussten Region nach Schätzungen etwa 600.000 Basenpaaren (600 kb; The Chimpanzee Sequencing and Analysis Consortium, 2005). Diese doch beträchtlichen Größen der Selektionssignaturen erleichtern einerseits natürlich deren Detektion, andererseits erschwert es die Aufgabe die kausale Variante ausfindig zu machen. Von Vorteil ist, dass „*selective sweeps*“ über einen längeren Zeitraum bestehen bleiben, als andere populationsgenetische Effekte. Neue Mutationen, welche die Diversität wiederherstellen, sind sehr rar, und die charakteristische Zeit, die eine solche neue Mutation benötigt, um bei neutraler Evolution nur durch genetische Drift zu hohen Frequenzen zu gelangen ist lang. In der Humanpopulation beträgt sie etwa eine Million Jahre, so dass statistisch signifikante Signale hier selbst nach einigen hunderttausenden von Jahren noch ausfindig zu machen sind. Allerdings lassen sich diese Signaturen nur schwer von demographischen Effekten, wie beispielsweise stark expandierenden Populationen, abgrenzen (SABETI *et al.*, 2006).

2.4.1.3 Neue Allele mit hoher Frequenz

Da neue (auch „*derived*“ oder „*nonancestral*“ genannt) Allele durch Mutation aus alten („*ancestral*“) Allelen entstehen, besitzen diese typischerweise auch niedrigere Frequenzen in der jeweiligen Population (SABETI *et al.*, 2006). Um nun Signaturen, welche viele nicht-ancestrale Allele hoher Frequenz enthalten ausfindig machen zu können, ist zuerst das Wissen nötig, welche Allele Ursprüngliche und welche Neue sind. Dieser Allelstatus wird üblicherweise durch Zuhilfenahme von sogenannten „*outgroups*“, also nah verwandten Spezies oder

möglichst ursprünglichen Vertretern der zu untersuchenden Spezies, bestimmt. Dabei werden zwei Annahmen gemacht: (1) das in der Vergleichsgruppe häufigere Allel stellt das Ältere dar, (2) jede Mutation ist ein einmaliges Ereignis, das nach der Trennung der gemeinsamen Evolution der beiden Gruppen oder Spezies stattgefunden hat (FAY & WU, 2000; SABETI *et al.*, 2006). Diese neuen Allele erlangen (ähnlich der Reduktion der genetischen Diversität; siehe 2.4.1.2), sofern sie an ein vorteilhaftes Allel gekoppelt sind, während eines „*selektive sweeps*“ durch „*hitchhiking*“ unüblich hohe Frequenzen (SABETI *et al.*, 2006). Der Unterschied dieser Signaturen zu Bereichen reduzierter genetischer Diversität, besteht in erster Linie aus zwei Punkten. Zum Einen bleiben neue Allele hoher Frequenz lediglich über eine kürzere Zeit (beim Menschen etwa 80.000 Jahre) nachweisbar, da sie der Fixation durch genetische Drift schnell sehr nahe kommen, oder diese sogar erreichen, und daher nicht mehr zur Polymorphismenanzahl beitragen (PRZEWORSKI, 2002; SABETI *et al.*, 2006). Zum Anderen wird ihr Vermögen Selektion zu identifizieren von verschiedenen demographischen Effekten unterschiedlich stark beeinflusst. So ist eine starke Populationsexpansion hauptsächlich für den Nachweis von Selektion anhand reduzierter genetischer Diversität ein möglicher Störfaktor, während Populationsunterteilungen eher ein Problem beim Nachweis von Selektion mittels neuer Allele darstellen. Auch durch Flaschenhälse können analoge Bereiche entstehen, so dass zu beachten ist, dass solche Allele keine eindeutige Signatur positiver Selektion sind (PRZEWORSKI, 2002).

2.4.1.4 Unterschiede zwischen Populationen

Werden Populationen teilweise oder vollständig reproduktiv getrennt und wirken auf diese dann unterschiedliche Selektionsdrücke, so ist es möglich, dass sich bestimmte Allelfrequenzen in nur einer Population entsprechend dem jeweiligen Selektionseinfluss verändern. Findet man also im Genom Allele mit relativ großen Frequenzunterschieden zwischen einzelnen Teilpopulationen, so kann man davon ausgehen, dass dieses Allel selbst oder ein eng gekoppeltes Allel positiver Selektion ausgesetzt war (SABETI *et al.*, 2006). Natürlich ergeben sich auch während evolutionär neutralen Phasen nur durch genetische Drift Unterschiede in

den Allelfrequenzen voneinander getrennter Populationen. Jedoch wird hierbei das ganze Genom gleichermaßen beeinflusst, während es sich bei der Selektion um eine systematische und Locus-spezifische Kraft handelt und sich folglich Auffälligkeiten lokal anhäufen (AKEY *et al.*, 2002). Dennoch fällt es auch bei dieser Nachweisart schwer, Signaturen wirklicher Selektion von genomischen Auswirkungen von Flaschenhälsen zu unterscheiden (SABETI *et al.*, 2006). Ein häufig angewendeter Vertreter dieser Nachweisgruppe ist der von Weir und Cockerham (1984) eingeführte F_{ST} -Wert.

2.4.1.5 Lange Haplotypen

Diese Methode basiert auf der Beziehung zwischen der Frequenz eines Allels und dem Ausmaß an Kopplungsungleichgewicht (LD), das es umgibt (SABETI *et al.*, 2002). Haben mehrere Allele einer chromosomalen Region die Tendenz in Individuen immer gemeinsam aufzutreten, sind sie also miteinander gekoppelt, so nennt man diese bestimmte Allelabfolge einen Haplotypen. Verfahren die zu dieser Nachweisgruppe gehören suchen nach der Kombination zweier bestimmter Kriterien. Zum Einen sollen die Allele eine hohe Frequenz aufweisen, was eigentlich für alte Allele typisch ist, und zum Anderen eine sogenannte „*long-range association*“, also eine weiträumige Gemeinschaft bilden, die wiederum bei jungen Allelen üblich ist. Diese „*long-range associations*“ stellen sich als ausgedehnte Haplotypen dar, die, bei entsprechend starker Selektion, so schnell an Prävalenz gewonnen haben, dass sie noch nicht durch Rekombination aufgelöst werden konnten. Die Vorteile dieses Nachweises sind, dass sie auch unvollständige „*sweeps*“, bei denen die Allelfrequenz gerade einmal zehn Prozent beträgt, erfassen können, sie den selektierten Bereich sehr gut eingrenzen, teilweise sogar auf nur ein Gen, und relativ robust bezüglich der Wahl der genetischen Marker sind (SABETI *et al.*, 2006). Von Nachteil ist jedoch, dass diese langen Haplotypen nur vergleichsweise kurze Zeit nachweisbar sind, da sie schnell durch Rekombination aufgespalten werden. Selektive Ereignisse, die weniger als 400 Generationen zurückliegen, sollten aber eine deutlich zu erkennende Spur im Genom hinterlassen (SABETI *et al.*, 2002). Ein weiterer kritischer Punkt ist, dass es scheint, dass Genome von Säugetieren aus einem Mosaik an Regionen mit

unterschiedlichen Rekombinationswahrscheinlichkeiten bestehen, diese daher zwischen den einzelnen chromosomalen Regionen variieren können (YAUK *et al.*, 2003). Vertreter dieser Nachweisart sind unter anderem der *long-range-haplotype Test (LRH)* von Sabeti *et al.* (2002) und der von Voight *et al.* (2006) eingeführte *integrated Haplotype Score (iHS)*.

2.4.2 Bisher durchgeführte Selektionssignaturstudien im Rindergenom

Die Suche nach Selektionssignaturen beim Rind wurde sehr lange durch die beschränkte Markerdichte gehemmt (HAYES *et al.*, 2009). Erst in jüngster Vergangenheit wurde es mit dem Aufkommen der kosteneffektiven Hochdurchsatzgenotypisierung möglich, nicht nur, wie beispielweise 2008 von Prasad *et al.* durchgeführt, ausgewählte Loci oder Chromosomen zu untersuchen, sondern auch genomweite Studien durchzuführen. Da dieses Forschungsgebiet folglich noch relativ jung ist, ist die Zahl der bereits veröffentlichten Studien noch vergleichsweise gering.

Hayes *et al.* (2009) verglichen die Allelfrequenzen von 9.323 SNPs von Milch- (Holstein) und Fleischrindern (Angus, Hereford und Limousin), um so divergent selektierte Regionen ausfindig zu machen. Dabei fanden sie die größten Frequenzunterschiede auf Chromosom 20 (BTA20) in einer Region nahe dem Gen für den Rezeptor des Wachstumshormons. Dass eine Mutation in diesem Gen Auswirkungen auf Milchmerkmale hat, wurde bereits für mehrere Milchrinderpopulationen bestätigt (BLOTT *et al.*; 2003). Weitere signifikante Regionen fanden Hayes *et al.* (2009) ferner auf den Chromosomen 2, 6, 14 und 24.

Flori *et al.* (2009) suchten in den drei häufigsten Milchrassen Frankreichs (Holstein, Normande und Montbéliarde) nach Regionen, die durch artifizielle Selektion beeinflusst wurden. Anhand von F_{ST} -Berechnungen innerhalb und zwischen den Rassen konnten sie 13 hoch signifikante Regionen identifizieren. Einige dieser Regionen beinhalten Gene, für deren kausale Varianten bereits ein starker Einfluss auf Milchmerkmale und Färbung gezeigt wurde.

Durch einen Vergleich der Selektionssignaturen, die im Bovinen HapMap Projekt (GIBBS *et al.*, 2009) gefunden wurden, mit denen aus Genotypen eines

unabhängigen Tiersets, gelang es Barendse *et al.* (2009) 129 SNPs ausfindig zu machen, die in beiden Datensets hoch divergente F_{ST} -Werte aufwiesen. Mit Hilfe dieser SNPs und bekannten Phänotypen konnten zwölf Regionen identifiziert werden, denen additive genetische Effekte auf Merkmale wie Restfutteraufnahme („*residual feed intake*“, ein Maß der Futtermittelverwertung), Fleischanteil oder intramuskulärer Fettanteil zugeschrieben werden.

MacEachern *et al.* (2009) untersuchten die Frequenzverteilung von 7.500 Polymorphismen in einer Fleisch- und einer Milchrasse (Angus und Holstein) mittels Fay und Wu's H Test (FAY & WU, 2000) und fanden so die größten Auffälligkeiten auf BTA2 und BTA13 für beide Rassen gemeinsam sowie auf den Chromosomen 25, 26 und 29 für jeweils eine der beiden Rassen. Anhand überlagerter graphischer Darstellung von Allelfrequenzen und F_{ST} -Werten, konnten darüber hinaus auf den Chromosomen 8, 20 und 24 Selektionssignaturen in der Rasse Angus ausgemacht werden. Auch wenn für diese Regionen kein bestimmtes Kandidatengen ausfindig gemacht werden konnte, gibt es deutliche Hinweise, dass es sich um QTL handelt, die die Schlachtkörperzusammensetzung (BTA8 und BTA24) sowie die Schlachtausbeute (BTA20) beeinflussen.

Im Jahr 2010 wurde schließlich ein systematischer genomweiter Scan mittels *Extended Haplotype Homozygosity (EHH)* bei Holstein-Friesians durchgeführt (QANBARI *et al.*, 2010a). Zum Einen überprüften Qanbari *et al.* Regionen auf Selektionssignaturen, für die bereits funktionelle Kandidatengene respektive QTL nachgewiesen waren, und konnten zeigen, dass diese Bereiche größtenteils tatsächlich eine solche Signatur aufweisen. Zum Anderen kartierten sie 161 mutmaßlich unter Selektion stehende Regionen, welche sie anhand der genomweit extremsten P -Werte ($< 0,01$) identifizierten. Für die meisten davon kann eine biologische Verbindung zu Merkmalen wie Milchmenge und Milchzusammensetzung, Reproduktion und Verhalten zumindest angenommen werden.

Stella *et al.* (2010) entwickelten einen Ansatz um Selektionssignaturen zu detektieren, indem sie die SNP-Diversität zwischen phänotypisch ähnlichen Rassen der „*International Bovine HapMap*“-Studie (IBHM) und einer größeren Population (Gesamtheit der IBHM Tiere) verglichen. Dazu wurden zunächst die Allelfrequenzen in der Gesamt- sowie der Teilpopulation getrennt bestimmt. Anschließend wurde mit den so erhaltenen Frequenzen die parametrische

zusammengesetzte logarithmische Wahrscheinlichkeit („*parametric composite log likelihood*“, *CLL*) der Allelfrequenzen von SNPs benachbarter Loci berechnet. Auf diese Weise konnten 699 vermeintliche Selektionssignaturen ausfindig gemacht werden, wobei sich der größte *CLL*-Wert auf BTA6 befand und dem *KIT*-Gen zugehörig ist, welches phänotypisch unter anderem für die Scheckung des Fells verantwortlich ist. Des Weiteren waren unter den höchsten *CLL*-Werten Gene die mit Kaliumkanälen in Verbindung gebracht werden (BTA14, 16 und 25), Integrine (BTA18 und 19) sowie serin- und argininreiche Splicingfaktoren (BTA20 und 23).

3 Material und Methoden

3.1 Material

Im Folgenden wird näher auf das verwendete Probenmaterial, also die Anzahl der Tiere, ihre Rassezugehörigkeit sowie ihre Herkunft eingegangen. Im Anschluss wird eine Übersicht über die verwendete Software gegeben, die während dieser Studie zum Einsatz kam.

3.1.1 Probenmaterial

Das für die eigentlichen Selektionsanalysen verwendete Probenmaterial stammt von Vertretern der zehn oben beschriebenen Rassen und wurde der Probensammlung des Lehrstuhls für Tierzucht und Allgemeine Landwirtschaftslehre entnommen. Insgesamt wurden 494 Proben untersucht, die in den Jahren von 1992 bis 2008 gesammelt wurden und deren DNA zu Beginn dieses Projekts bereits extrahiert vorlag, wobei die Extraktion im lehrstuhleigenen Labor durchgeführt wurde.

Die genaue Probenanzahl, -herkunft und -art ist für jede Rasse Tabelle 1 (Tab. 1) zu entnehmen.

Nach *IBD*-basierter Verwandtschaftskorrektur (siehe 3.2.4) flossen von den 494 beprobten Tieren letztlich 380 möglichst unverwandte in die eigentlichen Selektionssignaturalysen mit ein. Die jeweilige Rassezugehörigkeit sowie die Labornummern dieser 380 Tiere sind Anhang 1 zu entnehmen.

Tabelle 1: Übersicht über die zehn ausgewählten Teilpopulationen.

Aufgeführt sind der jeweilige Name der Rasse sowie die in dieser Arbeit verwendete Abkürzung, die gesammelte Probenanzahl mit der Anzahl der davon verwendeten Proben in Klammern, das Land der Probennahme (Probenherkunft) und falls regional begrenzt mit Angabe der Region in Klammern sowie die jeweilige Selektionsrichtung.

Name	Abkürzung	Anzahl	Probenherkunft	Probenart	Selektionsrichtung
Anatolisches Schwarzvieh	ASV	49 (43)	Türkei (Anatolien)	Blutproben	unselektiert*
Illyrisches Rind	IL	45 (41)	Albanien	Blutproben	unselektiert*
Deutsches Fleckvieh	DFV	55 (50)	Deutschland	Spermaproben	Doppelnutzungsring
Original Braunvieh	OBV	48 (32)	Schweiz [#]	Spermaproben	Doppelnutzungsring
Murnau-Werdenfelser	MWF	53 (23)	Deutschland	Blut- oder Spermaproben	Doppelnutzungsring
Fränkisches Gelbvieh	FGV	50 (44)	Deutschland	Haar- oder Spermaproben	Doppelnutzungsring
Braunvieh (neue Zuchtrichtung)	BBV	50 (43)	Deutschland/weltweit	Blut- oder Spermaproben	Milchbetontes Doppelnutzungsring
Red Holstein	RH	50 (46)	Deutschland/weltweit	Spermaproben	Milchrind
Weißblauer Belgier	BBB	47 (28)	Belgien	Spermaproben	Fleischring
Galloway	GLW	47 (30)	Deutschland/weltweit	Blut-, Haar- oder Spermaproben	Fleischring

*wird als unselektiert betrachtet, da bis zum Zeitpunkt der Probennahme keine Leistungskontrolle und keine organisierte Selektion vorgenommen wurde

[#] freundlicherweise zur Verfügung gestellt von Swisegenetics

Darüber hinaus wurden, um die Genauigkeit der Haplotypisierung zu erhöhen (siehe 3.2.2), SNP-Genotypen von 1.318 weiteren Tieren aus der Datenbank des Lehrstuhls für Tierzucht und Allgemeine Landwirtschaftslehre verwendet.

3.1.2 Verwendete Software

Folgende Programme haben im Laufe dieser Arbeit Verwendung gefunden und sind nochmals in Tabelle 2 aufgelistet. Für die Haplotypisierung und zur Auffüllung von fehlenden Allelen wurde ein im Programm *BEAGLE 3.0.4* (BROWNING & BROWNING, 2009) umgesetztes „*Hidden Markov Model*“ angewendet. Die genomweite Abschätzung der paarweisen *IBD*-Wahrscheinlichkeiten („*identity by descent*“) erfolgte mit *PLINK* (PURCELL *et al.*, 2007). Im Anschluss fanden die eigentlichen Selektionssignatur-Analysen *iHS* (VOIGHT *et al.*, 2006) und *XP-EHH* (SABETI *et al.*, 2007), in der Implementierung von Kudravalli (VOIGHT *et al.*, 2006; PICKRELL *et al.*, 2009) respektive Pickrell (PICKRELL *et al.*, 2009) Gebrauch. Alle Daten wurden mittels Microsoft SQL Server 2005 in eine Datenbank abgelegt und mit Hilfe des Microsoft Visual Studios (Programmiersprache: Visual Basic) sowie dem Programm *R* (IHAKA & GENTLEMAN, 1996) in die für die jeweiligen Analysen entsprechende Form gebracht. Die Visualisierung der Daten erfolgte mit dem Programm *MATLAB*. Die Auswertung der Resultate der Paired-End-Sequenzierung wurde auf einem Webserver (*GALAXY*) und ihre optische Darstellung mit dem „*Integrative Genomics Viewer*“ (*IGV*) vorgenommen. Die Lage gefundener DNA-Varianten, sowie gegebenenfalls ihre Auswirkungen auf die Codierung, wurden mit dem Online-Tool „*Variant Effect Predictor*“ bestimmt. Das Erstellen der Primer für eine anschließende Sanger-Sequenzierung erfolgte mit „*Primer3 Input 0.4.0*“ und ihre Überprüfung auf Fehlpaarungsverhalten mittels „*Beacon Designer™ Free Edition*“. Die anschließende Auswertung der Sequenz des DNA-Fragments wurde mit Hilfe des Online-Programms „*ORF Finder*“ („*open reading frame*“) vorgenommen.

Tabelle 2: Verwendete Software: Programmname, Verwendungszweck und Quelle

Programmname und Version	Verwendung und Quelle
BEAGLE 3.0.4	Haplotypisierung und Allel-Imputation http://www.stat.auckland.ac.nz/~browning/beagle/beagle.html
PLINK 1.07	Berechnung der <i>IBD</i> -Wahrscheinlichkeiten http://pngu.mgh.harvard.edu/purcell/plink/
iHS	Berechnung des integrated Haplotype Scores http://hgdp.uchicago.edu/Software/
XP-EHH	Berechnung der Cross Population Extended Haplotype Homozygosity http://hgdp.uchicago.edu/Software/
Microsoft SQL Sever 2005	Datenbank zur Datenverwaltung http://msdn.microsoft.com/de-de/sqlserver
Microsoft Visual Studio 2005	Datenvorbereitung für diverse Analysen http://www.microsoft.com/visualstudio/de-ch/products/2005-editions
R 2.9.0	Datenvorbereitung für diverse Analysen http://www.r-project.org/index.html
MATLAB (R2008b)	Visualisierung der Daten http://www.mathworks.com/products/matlab/
GALAXY	Auswertung der Sequenzierungsergebnisse https://bitbucket.org/galaxy
IGV 1.5	Optische Darstellung der Alignments http://www.broadinstitute.org/software/igv//node/120
Variant Effect Predictor	Bestimmung von Lage und Auswirkung der DNA-Varianten http://www.ensembl.org/tools.html
Primer3 Input 0.4.0	Erstellen der Primer http://frodo.wi.mit.edu/primer3/
Beacon Designer™ Free Edition	Überprüfung der Primer auf Fehlpaarungsverhalten http://www.premierbiosoft.com/qpcr/index.html
ORF Finder	Auswertung der Sequenz des DNA-Fragments http://www.geneinfinity.org/sms/sms_orffinder.html

3.2 Methoden

Im Zuge dieser Arbeit fanden mehrere Methoden Verwendung. Diese waren für die Vorbereitung der Proben (Genotypisierung, Haplotypisierung), die Qualitätssicherung (Auswahl geeigneter Marker, *IBD*-Analyse), für die eigentliche Detektion der Selektionssignaturen (*iHS*, *XP-EHH*) sowie, nach Auswahl einer geeigneten Signatur, für ihre weitere Analyse (Sequenzierung) nötig.

3.2.1 Genotypisierung

Für die Durchführung des Forschungsvorhabens wurden die Probensammlungen der zehn verschiedenen Rinderrassen mittels Illumina Bovine BeadChip 50K

(Illumina, San Diego, USA) von der Tierzuchtforschung e. V. München entsprechend den Herstellerangaben genotypisiert.

Das besondere an diesem BeadChip ist, dass für die so durchgeführte Genotypisierung keine PCR nötig ist. Bis zur Entwicklung dieses Chips, war es üblich das Genom Abschnittsweise zu amplifizieren, was den entscheidenden Nachteil hatte, dass die SNP-Auswahl limitiert war, da nur SNPs gewählt werden konnten, die in diesen amplifizierten Abschnitten vorhanden waren. Mit dem BeadChip ist es jetzt jedoch durch einfache automatisierte Prozesse möglich große Mengen genetischer Information direkt vom Genom abzulesen. Dabei ist die Prozedur in die folgenden vier Schritte unterteilt. Den Anfang macht die Amplifikation des gesamten Genoms, Schritt zwei ist die Hybridisierung an eine spezielle Matrix mit sensitiven Oligonukleotiden. Anschließend wird der SNP mit Hilfe einer Allel-spezifischen Primer-Verlängerungsreaktion identifiziert, da diese Reaktion es möglich macht jede SNP-Kategorie mit einer bestimmten Farbe an der jeweiligen Matrixposition abzufragen. Im letzten Schritt wird dieses sensitive Signal zum Auslesen der Daten verstärkt. In Abbildung 13 ist diese Funktionsweise nochmals schematisch dargestellt (GUNDERSON *et al.*, 2005).

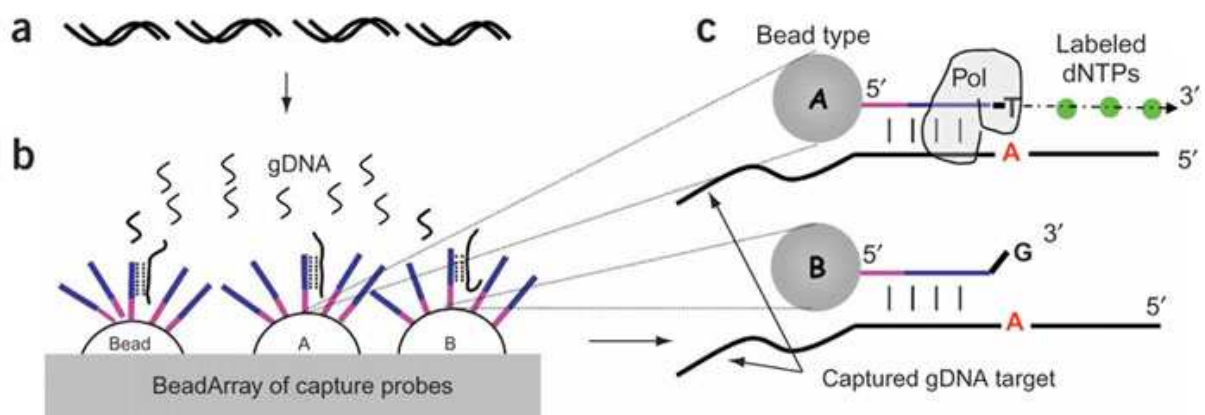


Abbildung 13: Funktionsweise eines BeadChips

a) Amplifikation des gesamten Genoms; b) Hybridisierung der Amplifikationsprodukte an ein spezifisches und sensitives Oligonukleotidarray; c) Allel-spezifische Primer-Verlängerungsreaktion (Einbau von Biotin-markierten dNTP-Nukleotiden in die entsprechende Allelprobe) gefolgt von einer sensitiven Detektion und Signalamplifikation um die eingebauten Labels auszulesen.

Für einen gegebenen SNP werden zwei verschiedenen Bead-Typen verwendet, deren Sequenz abgesehen vom 3'-Ende gleich ist. Dabei ist die letzte Base so gewählt, dass die Proben des einen Bead-Typs das Gegenstück zu Allel A und die des Anderen das zu Allel B bilden. Da die Verlängerung durch die Polymerase vorzugsweise am zusammenpassenden 3'-Ende erfolgt, ist so eine adäquate Identifikation des SNPs möglich.

(Quelle: Gunderson et al. 2005, http://www.nature.com/ng/journal/v37/n5/fig_tab/ng1547_F1.html)

Jede Probe der 494 Tiere wurde auf diese Art an 54.001 der beim Rind vorkommenden SNPs getestet. Dabei beträgt die durchschnittliche Markerdistanz des BeadChips 51,5 Kilobasen. Die Anzahl der SNPs, die so genomweit für alle zehn Rinderrassen erfolgreich genotypisiert werden konnten, lag bei 53.725.

3.2.2 Haplotypisierung

Für die meisten Multilokusanalysen, wie auch für die in dieser Arbeit durchgeführten Untersuchungen, ist es essenziell, die Daten in Form von Haplotypen vorliegen zu haben. Dies bedeutet, dass es erforderlich ist, den bei der sexuellen Fortpflanzung durch Fusion von maternalem und paternalem Haplotyp entstandenen Genotypen wieder entsprechend der paternalen oder maternalen Herkunft der Allele zu trennen. Um dies zu bewerkstelligen entschied man sich für das Programm *BEAGLE*. Diese Software leitet nicht nur die wahrscheinlichsten Haplotypen ab, sondern errechnet auch für nicht-genotypisierte Marker die wahrscheinlichsten Allele und füllt somit die bei der Genotypisierung entstandenen Lücken auf („*imputation*“). Dabei ist dieser Algorithmus bei vergleichbarer Genauigkeit signifikant schneller, als andere Methoden und schon bei seiner Einführung 2007 stellte dieses Programm den besten Algorithmus für die Haplotypisierung von umfangreichen Datensets dar (BROWNING & BROWNING, 2007). Die hier verwendete Version *BEALGE 3.0.4* von 2009 ermöglicht es dem Benutzer verschiedene Datensets in einer Analyse zu kombinieren und dadurch die Genauigkeit der einzelnen Analysen zu erhöhen. Genauer gesagt können sowohl Haplotypen als auch Genotypen von Unverwandten, Eltern-Kind-Paaren und Eltern-Kind-Trios gleichzeitig analysiert werden. Diese aufgrund der Verwandtschaftsverhältnisse zusätzlich gewonnenen Informationen über die abzuleitende Haplotypenstruktur führen zu extrem genauen Schätzungen (BROWNING & BROWNING, 2009).

Um von diesem gewichtigen Vorteil Gebrauch zu machen, wurden zu den 494 für das Projekt ausgewählten Tieren weitere 1318 Tiere aus der Datenbank ergänzt, so dass schließlich 888 Eltern-Kind-Paare und 62 Eltern-Kind-Trios sowie 752 unverwandte Tiere für die Haplotypisierung und die Ableitung nicht-genotypisierter Allele zur Verfügung standen. Die auf diesen Weg erlangten lückenlosen

Haplotypen wurden dann in die SQL-Datenbank eingelesen, so dass auf sie für die nachfolgenden Untersuchungen zugegriffen werden konnte.

3.2.3 Auswahl geeigneter Marker

Von den 53.725 genotypisierten SNPs, wurden im Folgenden für dieses Projekt uninformative Marker ausgeschlossen. So wurden alle SNPs entfernt, die an der jeweiligen Position weniger als zwei heterozygote Tiere in allen zehn Rassen aufwiesen, außerdem alle Einzelnukleotidpolimorphismen, deren Position unbekannt war. Ferner wurden alle Marker ausgenommen, die bei bekannter Verwandtschaft der Tiere Abstammungsprobleme verursachten. Dabei wurde dieser Abstammungstest mit Hilfe einer eigenen Visual-Basic-basierten Anwendung durchgeführt. Letztendlich blieben 47.003 informative SNPs, welche in mindestens 90 % der Tiere erfolgreich genotypisiert wurden (Marker Call-Rate über 0,90), übrig.

3.2.4 Schätzung der *identity-by-descent*-(*IBD*)-Wahrscheinlichkeiten

Unterschiede in der Populationsstruktur der einzelnen Rassen haben Auswirkungen auf die jeweiligen Allelfrequenzen. Dies kann, bei Methoden, welche die Unabhängigkeit der Individuen innerhalb der Teilpopulationen voraussetzen, wie es auch in vorliegender Studie der Fall ist, eine Verfälschung der Ergebnisse zur Folge haben. Es ist daher im Rahmen der Qualitätskontrolle der Stichproben unumgänglich, die einzelnen Proben auf Verwandtschaftsbeziehungen zu untersuchen und gegebenenfalls zu korrigieren (LAURIE *et al.*, 2010). Da für die unselektierten Rassen keine Abstammungsdaten zur Verfügung standen, und die verfügbaren Stammbäume der anderen Stichproben darüber hinaus auch fehlerbehaftet sein können, wurde die realisierte Verwandtschaft jeder möglichen Paarung von Individuen innerhalb einer jeden der zehn Rassen mit Hilfe der *IBD*-Wahrscheinlichkeit geschätzt. Diese gibt die genomweite Wahrscheinlichkeit an, dass die Allele dieser Individuen

herkunftsgleich, also „*identity-by-descent*“ (*IBD*), sind (PURCELL *et al.*, 2007; LAURIE *et al.*, 2010). Die *IBD*-Bestimmung beruht dabei auf einer „*maximum-likelihood*“-Schätzung (MILLIGAN, 2003), welche unter Anderem in dem Programm *PLINK* implementiert ist (PURCELL *et al.*, 2007), das in dieser Arbeit verwendet wurde. Lag bei einem Paar die genomweite *IBD* bei einem Wert von über 0,20, so wurde das Individuum mit der schlechteren Call-Rate von den weiteren Untersuchungen ausgeschlossen.

3.2.5 Methoden basierend auf Haplotypenlänge und –diversität

Obwohl für diese Studie lediglich der *integrated Haplotype Score (iHS)* und die *Cross Population Extended Haplotype Homozygosity (XP-EHH)* verwendet wurden, ist es unumgänglich an dieser Stelle auch auf die *Extended Haplotype Homozygosity (EHH)* einzugehen, da sowohl *iHS* als auch *XP-EHH* auf dieser Methode beruhen.

3.2.5.1 Extended Haplotype Homozygosity (*EHH*)

Die *EHH*-Wahrscheinlichkeit wird immer für zwei Allele eines Kernmarkers berechnet. Diese Kernallele werden in der original Literatur als „*Core*“-Allele und von ihnen ausgehende Haplotypen als „*Core*“-Haplotypen bezeichnet. Diese Bezeichnung wird im Folgenden übernommen.

Definiert wird die *EHH* als die Wahrscheinlichkeit, dass zwei zufällig gewählte Chromosomen, die dasselbe *Core*-Allel tragen, für das gesamte Intervall von der *Core*-Region bis zu einem Punkt *X* identisch sind. Sie kann mit folgender Formel

berechnet werden:

$$EHH = \frac{\sum_{i=1}^G \binom{n_i}{2}}{\binom{N}{2}}$$

Wobei *N* die Anzahl der Haplotypen darstellt, die dasselbe *Core*-Allel tragen, *G* die Zahl der verschiedenen Haplotypen dieses einen *Core*-Typs, die sich bis zum SNP *X* ergeben und *n_i* die Anzahl des jeweiligen Haplotyps (*G_i*).

Dabei erfolgt die Berechnung getrennt für beide Core-Allele und innerhalb desselben Core-Allels getrennt für die distale und proximale Seite vom Core-SNP (SABETI *et al.*, 2002).

Um im folgenden Beispiel (Tab. 3) den *EHH*-Wert, ausgehend vom Core-SNP in distale Richtung, also nach rechts, bis zum SNP X, zu berechnen, legt man den gewünschten Core-Typen fest. Wählt man also zur Darstellung Core = 1, dann ergibt sich für $N = 6$, für $G = 4$ (G_1 : 111111, G_2 : 111110, G_3 : 111100, G_4 : 110111) und für die zugehörigen n_r -Werte $n_1 = 1$, $n_2 = 1$, $n_3 = 3$, $n_4 = 1$. Eingefügt in obige Formel erhält man also:

$$EHH = \frac{\left(\binom{1}{2} + \binom{1}{2} + \binom{3}{2} + \binom{1}{2} \right)}{\binom{6}{2}} = 0,2$$

Berechnet man nun alle *EHH*-Werte für beide Core-Allele und in beide Richtungen, wobei nacheinander jeder SNP vom Core aus als SNP X festgelegt wird, ergibt sich der in Abbildung 14 gezeigte Kurvenverlauf.

Tabelle 3: Beispielhaplotypen für *EHH*-Berechnung

(C = Core; X = SNP X;
1 = ancestral Allel; 0 = derived Allel)

X	◀	C						▶	X
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	0
0	0	1	1	1	1	1	0	1	1
0	0	1	1	1	1	1	1	1	0
0	1	1	0	1	1	1	1	1	0
0	0	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	1	0	1	0
0	0	1	0	1	0	0	0	1	0
0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0

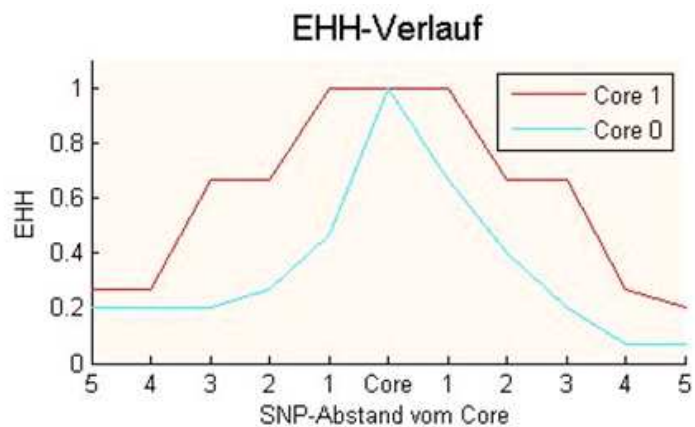


Abbildung 14: *EHH*-Verlauf der Beispielhaplotypen aus Tabelle 3

EHH-Wert in Abhängigkeit vom SNP-Abstand vom Core

Betrachtet man nun den Graphen (Abb. 14), so ist eindeutig ersichtlich, dass die Haplotypen um den Core-SNP mit dem Allel „1“ insgesamt deutlich länger sind als Haplotypen mit einem Core-Allel „0“. Lang zu sein ist jedoch – sofern kein Selektionsdruck wirkt – nur für relativ junge Haplotypen mit noch niedriger Frequenz typisch, da diese noch nicht in hohem Maße durch Rekombination

verkürzt wurden. Und dies wiederum bedeutet, dass ein Kurvenverlauf wie oben dargestellt ein deutlicher Hinweis auf Selektion ist.

Es ist also mit Hilfe der *EHH* möglich, eine untypisch lange Haplotypenhomozygotie als charakteristische Selektionssignatur ausfindig zu machen (siehe 2.4.1.5). Dies bedeutet darüber hinaus, dass auch die auf *EHH* basierenden Methoden *iHS* und *XP-EHH* auf diese Art von Signaturen spezialisiert sind.

3.2.5.2 Integrated Haplotype Score (*iHS*)

Der *iHS* wurde von Voight *et al.* (2006) eingeführt und beruht auf der Berechnung des Verhältnisses der Flächen unter den *EHH*-Kurven zweier Core-Allele. Er lässt daher, wie auch die *EHH*, Rückschlüsse auf Selektion innerhalb einer Rasse zu und wurde daher für jede der zehn Rassen getrennt berechnet.

3.2.5.2.1 Zuteilung des Allelstatus

Bevor der *iHS* berechnet werden kann, bedarf es zuerst der Zuteilung eines jeden Alleles zum jeweiligen Status „*ancestral*“ (sofern das Allel ursprünglich von den Ahnen stammt, also „alt“ ist) oder „*derived*“ (wenn das Allel erst später entstanden, also „neu(er)“ ist) (VOIGHT *et al.*, 2006). Idealerweise würde das Allel, das vom Urrind, also dem Aurochs, überliefert wurde, den Status „*ancestral*“ bekommen. Da dieser aber bereits ausgestorben ist und daher keine Vergleichsproben zur Verfügung stehen, wird dieser Status analog zu der Arbeit von Hayes *et al.* (2009) jeweils demjenigen Allel zugeteilt, welches in mutmaßlich ursprünglichen und weitgehend unbeeinflussten Artgenossen am häufigsten vorkommt und somit vermeintlich das ältere Allel darstellt. Bei Hayes *et al.* wurden hierzu Proben von Yak, Bison und Banteng genotypisiert, während sich bei vorliegender Arbeit dazu entschieden wurde, die beiden unselektierten Rassen, also ASV und IL, der Allelstatuszuteilung zugrunde zu legen.

3.2.5.2.2 Berechnung der unstandardisierten *iHS*-Werte

Ist der Allelstatus dann bekannt, erfolgt die eigentliche Berechnung der *iHS*-Werte, wobei jeder innerhalb der jeweiligen Rasse polymorphe SNP einmal als Core-Marker dient. Für den jeweiligen Core-SNP wird dann der Abfall der *EHH* zu beiden Seiten für beide Allele berechnet und die jeweiligen *EHH*-Werte mit geraden Linien verbunden, so dass eine wie in Abbildung 14 gezeigte Graphik entsteht. Anschließend werden die Flächen (Integral) unter den beiden Kurven, die jeweils zum Core-Allel mit dem Status „*ancestral*“ oder „*derived*“ gehören, getrennt ermittelt. Dabei erstreckt sich das Integral distal und proximal vom Core-SNP per Definition bis zum geringstentferntesten Punkt, an dem ein *EHH*-Wert von 0,05 erreicht wird. Dieser integrierte *EHH* (*iHH*) wird dann abhängig vom Allelstatus als *iHH_A* („*ancestral*“) oder *iHH_D* („*derived*“) bezeichnet, und der unstandardisierte *iHS*-Wert wie folgt berechnet:

$$unst.iHS = \ln\left(\frac{iHH_A}{iHH_D}\right)$$

Durch den Logarithmus wird der unstandardisierte *iHS* einen Wert um Null erlangen, wenn der *EHH*-Abfall bei beiden Allelen in etwa gleich ist, also nichts auf eine selektive Bevorzugung eines Allels hindeutet. Hohe negative Werte dagegen weisen auf unüblich lange homozygote Haplotypen, die das „*derived*“ Allel tragen hin. Umgekehrt stehen hohe positive unstandardisierte *iHS*-Werte für auffällig lange gleichartige Haplotypen um das *ancestrale* Allel.

Auf diese Art wurde für jeden SNP ein *iHS*-Wert berechnet, sofern nicht eines der folgenden Kriterien zutrif. Erreichte die Region, die durch die Bedingung *EHH* > 0,05 bestimmt wird, das Ende eines Chromosoms oder eine Marker-Lücke mit einem SNP-Abstand größer 200 Kilobasen, so wurde kein *iHS*-Wert übermittelt. Bei Abständen von 20 - 200 kb wurde diese genetische Distanz durch einen Multiplikationsfaktor von 20/g, wobei „g“ die Größe der Lücke („*gap*“) in Kilobasen darstellt, reduziert (VOIGHT *et al.*, 2006). Des Weiteren bleibt zu beachten, dass an bereits fixierten Allelen, aufgrund des daraus resultierenden Fehlens eines Vergleichsallels an selber Position, mittels *iHS* keine Selektion nachgewiesen werden kann.

Zur Erlangung dieser unstandardisierten *iHS*-Werte wurde in diesem Projekt die Implementierung von Kudravalli (VOIGHT *et al.*, 2006; PICKRELL *et al.*, 2009)

verwendet. Die so ermittelten Werte wurden anschließend in eine SQL-Datenbank übertragen, um für die nachfolgende Standardisierung leicht zugänglich zu sein.

3.2.5.2.3 Standardisierung der *iHS*-Werte

Da in neutralen Modellen Allele mit geringer Frequenz im Allgemeinen jünger („*derived*“) sind, sind diese auch mit längeren Haplotypen assoziiert als höherfrequente alte („*ancestral*“) Allele. Deshalb muss der unstandardisierte *iHS* so angepasst werden, dass die endgültigen Werte von der Allelfrequenz des Core-SNPs unabhängig und somit vergleichbar sind. Dabei erfolgt die Standardisierung auf einen Mittelwert von Null und eine Varianz von Eins wie folgt:

$$iHS = \frac{\ln\left(\frac{iHH_A}{iHH_D}\right) - E_P\left[\ln\left(\frac{iHH_A}{iHH_D}\right)\right]}{SD_P\left[\ln\left(\frac{iHH_A}{iHH_D}\right)\right]}$$

Um die geforderte Frequenzunabhängigkeit zu erreichen, werden sowohl der Erwartungswert (E_P) als auch die Standardabweichung (SD_P) von $\ln(iHH_A/iHH_D)$ nur anhand der empirischen Verteilung jener SNPs geschätzt, deren Frequenz des Alles mit dem Status „*derived*“ der Frequenz des Core-SNPs entspricht (VOIGHT *et al.*, 2006).

Um dies zu bewerkstelligen wurde ein Programm entworfen, welches für jeden Core-SNP die *iHS*-Werte aller SNPs aus der Datenbank ausliest, deren „*derived*“ Allelfrequenz in einem Bereich von $\pm 2,5$ Prozent um die Core-Allel-Frequenz liegt. Für Standardabweichung und Erwartungswert und folglich standardisierten *iHS*-Wert erfolgte die Ermittlung dann lediglich anhand der empirischen Verteilung der unstandardisierten *iHS*-Werte in diesem fünf Prozent Intervall. Da junge Allele in der Regel eine niedrigere Frequenz aufweisen als ältere, nimmt die SNP-Dichte in diesen fünf Prozent Bereichen mit steigender *derived*-Allelfrequenz (dAF) ab. Daher wurde, um eine sinnvolle SNP-Dichte bei der Standardisierung zu gewährleisten, für SNPs ab einer dAF über 0,70 auf diese fünf Prozent Intervalle verzichtet und stattdessen alle SNPs im Frequenzbereich von 0,70 bis 1,00 gemeinsam standardisiert. Für die Rassen IL und ASV lag diese Grenze bereits bei 0,50, da es aufgrund der Vorgehensweise bei der Allelstatuszuweisung (siehe

Punkt 3.2.5.2.1) hier nur vereinzelt SNPs mit höherer dAF gab. Die anhand der jeweiligen Frequenzbereiche standardisierten *iHS*-Werte wurden schließlich wieder in die SQL-Datenbank eingelesen, um für die spätere Auswertung zugänglich zu sein.

Allerdings stellt der so anhand der genomweiten empirischen Verteilung standardisierte *iHS* keinen eigentlichen Signifikanztest dar, sondern ist ein Maß für die Ungewöhnlichkeit der Haplotypen um einen bestimmten SNP im Bezug auf das Gesamtgenom einer Teilpopulation (VOIGHT *et al.*, 2006).

3.2.5.3 Cross Population Extended Haplotype Homozygosity (*XP-EHH*)

Die *XP-EHH*, entwickelt von Sabeti *et al.* (2007), beruht auf dem Verhältnis der Flächen unter *EHH*-basierten Kurven zweier Rassen und lässt daher Rückschlüsse auf Selektion im Vergleich zwischen zwei Rassen zu, und wurde für jeden der insgesamt 45 möglichen Rassenvergleiche getrennt berechnet.

Die *XP-EHH* ist aufgrund des Haplotypenvergleichs zwischen Populationen so konzipiert, dass sie anders als *EHH* und *iHS* auch imstande ist, selektierte Allele zu ermitteln, die in einer Population nahe an einer Frequenz von 1,00 sind oder diese bereits erreicht haben (SABETI *et al.*, 2007).

3.2.5.3.1 Berechnung der unstandardisierten *XP-EHH*-Werte

Auch bei der Berechnung der *XP-EHH*-Werte dient jeder SNP einmal als Core-SNP. Dafür wird die *EHH* an dem jeweiligen Core-SNP unabhängig vom entsprechenden Core-Allel für zwei Populationen A und B getrennt berechnet, so dass für jede Population nur eine *EHH*-Kurve entsteht. Auf die gleiche Weise wird dann für alle Chromosomen beider Populationen eine gemeinsame *EHH*-Kurve bestimmt. Diese gemeinsame Kurve dient der Grenzfindung für die Flächenberechnung der jeweiligen Populationskurven. So wird der SNP X, der am nächsten an einem gemeinsamen *EHH*-Wert von 0,04 liegt, Grenze für die

Integralberechnung. Für jede populationsspezifische *EHH*-Kurve wird dann die Fläche, die durch diese Grenzen bestimmt ist, berechnet (I_A bzw. I_B). Im Anschluss wird analog zum *iHS* der Logarithmus aus dem Verhältnis von I_A zu I_B ermittelt und so der unstandardisierte *XP-EHH*-Wert ermittelt:

$$\text{unst. } XP - EHH = \ln\left(\frac{I_A}{I_B}\right)$$

Durch den Logarithmus wird, wie auch bei der Berechnung des unstandardisierten *iHS*-Werts, der unstandardisierte *XP-EHH*-Wert einen Wert um Null erlangen, wenn der *EHH*-Abfall beider Populationen A und B gleich ist. Hohe negative Werte weisen auf eine unüblich ausgeprägte Homozygotie der Haplotypen in Population B, hohe Positive auf eine überdurchschnittlich ausgedehnte Haplotypen-homozygotie in Population A hin (SABETI *et al.*, 2007).

Für die Berechnung der unstandardisierten *XP-EHH*-Werte wurde die Implementierung von Pickrell (PICKRELL *et al.*, 2009) verwendet, wobei die ursprüngliche Version, die bei Erreichen von Markerabständen über 200 Kilobasen keinen *XP-EHH*-Wert ermittelte, von J. Pickrell freundlicherweise durch das Entfernen dieses Kriteriums auf die hier im Vergleich zur Humangenetik geringere SNP-Dichte angepasst wurde. Dies bedeutet, dass anders als beim *iHS*, für jeden SNP unabhängig von den Markerabständen ein *XP-EHH*-Wert ermittelt wurde. Die errechneten *XP-EHH*-Werte wurden dann in die SQL-Datenbank eingelesen, um für die folgende Standardisierung zur Verfügung zu stehen.

3.2.5.3.2 Standardisierung der *XP-EHH*-Werte

Insbesondere aufgrund der unterschiedlichen effektiven Populationsgrößen der jeweiligen Rassen unterscheiden sich die durchschnittlichen Haplotypenlängen, je nach Rassenkombination, auch ohne gezielte Selektion mal mehr und mal weniger deutlich, so dass die unstandardisierten Werte der verschiedenen Rassegegenüberstellungen nicht vergleichbar sind. Dies wird vor allem bei Vergleichen zwischen einer Rasse mit sehr kleiner und einer Rasse mit sehr großer effektiver Populationsgröße deutlich. Die Verteilung der unstandardisierten *XP-EHH*-Werte wird sich hier deutlich auf die Seite der effektiv kleinen Population verschieben, da hier die Haplotypen schon unter Neutralität länger sind und auch

geringe Selektion ausreicht, um diesen Effekt zu verstärken. Auf Seiten der effektiv großen Population hingegen bedarf es schon Selektion, um die Haplotypen auch nur auf die gleiche Länge wie die der effektiv kleinen Rasse zu bringen, so dass extrem hohe Werte für die Rasse mit effektiv großer Populationsgröße trotz Selektion eher unwahrscheinlich sind. Da sich also rein aufgrund der Unterschiede in der neutralen Haplotypenlänge die absoluten unstandardisierten *XP-EHH*-Werte um ein Vielfaches unterscheiden können, ist es nötig diese Werte so abzuändern, dass eine Vergleichbarkeit unabhängig von der jeweiligen effektiven Populationsgröße gegeben ist. Um dies zu erreichen wird, wie folgt, jeder einzelne der insgesamt 45 Rassenvergleiche auf einen Mittelwert von Null und eine Varianz von Eins standardisiert (SABETI *et al.*, 2007):

$$XP-EHH = \frac{\ln\left(\frac{I_A}{I_B}\right) - E_P \left[\ln\left(\frac{I_A}{I_B}\right) \right]}{SD_P \left[\ln\left(\frac{I_A}{I_B}\right) \right]}$$

Mithilfe eines hierfür entwickelten Programms wurden sowohl der Erwartungswert (E_P) als auch die Standardabweichung (SD_P) von $\ln(I_A/I_B)$ nur anhand der empirischen Verteilung der unstandardisierten *XP-EHH*-Werte des jeweiligen Vergleichs berechnet. Dies wurde für alle 45 Rassenvergleiche separat durchgeführt und die so erlangten standardisierten *XP-EHH*-Werte (*sXP-EHH*) von dem Programm automatisch für die spätere Auswertung in die SQL-Datenbank eingelesen.

3.2.6 Auswertung der Daten

Um die Ergebnisse der beiden Tests innerhalb und zwischen den Populationen einfacher vergleichen zu können, wurde das Genom in 1.000 etwa gleich große Fenster unterteilt. Die Fenster umfassten im Durchschnitt 2.544.053 Basen, wobei die genaue Größe von der Chromosomengröße abhängig war, so dass sie zwischen 2.497.781 Basen auf BTA24 und 2.598.922 Basen auf BTA29 schwankte (die genaue Unterteilung ist dem Anhang 2 zu entnehmen). Die gewählte Fensteranzahl ergab sich aus dem Kompromiss, sich in etwa an der Markerdichte der Studie von Pickrell *et al.* (2009) zu orientieren (rein rechnerisch

wäre eine Fenstergröße von etwa 2.800.000 Basen nötig) kombiniert mit dem Bestreben bei der anschließenden Vergabe der empirischen P -Werte leichter interpretierbare Ergebnisse zu erzielen.

Auswertung iHS

Für jede der zehn Rassen wurde dann für jedes der 1.000 Fenster der Anteil an SNPs mit einem absoluten iHS -Wert über zwei Standardabweichungen berechnet. Dieses Kriterium wurde gewählt, da bekannt ist, dass die Fraktion hoher iHS -Werte in einer bestimmten Region deutlich aussagekräftiger ist, als die Bestimmung des Maximalwertes (VOIGHT *et al.*, 2006; SABETI *et al.*, 2007; PICKRELL *et al.*, 2009). Dies liegt vor allem daran, dass ein „selective sweep“ dazu tendiert Gruppen extremer iHS -Werte zu erzeugen, während solche Werte in einem neutralen Model gleichmäßiger verteilt sind (VOIGHT *et al.*, 2006). Die Grenze $|iHS| > 2$ wurde, ebenso wie die Tatsache, dass Fenster mit weniger als 20 SNPs nicht berücksichtigt wurden, von Pickrell *et al.* (2009) übernommen.

Den Fenstern wird dann ein empirischer P -Wert (eP) entsprechend ihres durch den Anteil an $|iHS| > 2$ erreichten Rangs zugeteilt ($eP = \text{Rang} / 1.000$), so dass der eP -Wert des besten Fensters immer 0,001 beträgt.

Als unter Selektion stehend werden dann alle Fenster mit einem empirischen P -Wert $\leq 0,01$ erachtet, was jeweils den zehn besten Fenstern jeder Rasse entspricht. Diese so ermittelten Fenster sind im Folgenden Teil dieser Arbeit gemeint, wenn von „signifikant“ die Rede ist.

Die Berechnung des Anteils an SNPs mit $|iHS| > 2$ erfolgte für alle 1.000 Fenster für jede Rasse wieder mit einer eigens dafür konzipierten Datenbank-Anwendung. Die Zuteilung der entsprechenden empirischen P -Werte wurde dann im Anschluss manuell vorgenommen.

Auswertung $XP-EHH$

Im Gegensatz zum iHS , ist bei der $XP-EHH$ der maximale (Population A) beziehungsweise minimale (Population B) Wert innerhalb eines jeden Fensters am aussagekräftigsten. Daher wurde für jedes der 1.000 Fenster für jeden einzelnen der 45 Rassenvergleiche der maximale und minimale $XP-EHH$ -Wert bestimmt und diese Werte anschließend getrennt (minimale Werte für die Suche nach Selektion in Population B und maximale Werte für selbiges in Population A) der Größe nach

geordnet. Dann wird auch hier allen Fenstern entsprechend des Rangs ein empirischer P -Wert zugeteilt ($eP = \text{Rang} / 1.000$), wobei den „Rang 1“ jeweils der $XP\text{-}EHH$ -Wert mit dem größten absoluten Wert bekommt. Es ergeben sich also für jedes Fenster eines jeden Rassenvergleichs zwei verschiedene empirische P -Werte, je nachdem ob nach Selektion in Population A oder Population B gesucht wird. Die 45 Rassenvergleiche entsprechen aufgrund dieser Zweiteilung bezüglich der Auswertung folglich 90 $XP\text{-}EHH$ -Vergleichen, da jede der zehn Rassen im Vergleich mit jeder der neun Anderen separat ausgewertet wird.

Um nun für eine Rasse die unter Selektion stehenden Bereiche ausfindig zu machen, wurden von allen neun diese Rasse beinhaltenden Vergleichen jeweils die zehn Besten Fenster ($eP \leq 0,01$) ermittelt. Da aber einzelne hohe $XP\text{-}EHH$ -Werte unter Umständen nur auf unterschiedlichen effektiven Populationsgrößen oder verschiedenen Rekombinationsraten dieser zwei gegenübergestellten Rassen beruhen, wurden, um die Zahl falsch Positiver zu minimieren, nur solche Fenster als durch Selektion beeinflusst gewertet, bei denen mindestens drei der neun Vergleiche einen empirischen P -Wert von maximal 0,01 aufwiesen. Auch hier gilt, dass die so durch $XP\text{-}EHH$ ermittelten Fenster im Weiteren als „signifikant“ bezeichnet werden.

Die Ermittlung der minimalen und maximalen $XP\text{-}EHH$ -Werte eines jeden Fensters eines jeden Vergleichs wurde wieder durch eine dafür entworfene Datenbank-Anwendung bewerkstelligt. Die anschließende Zuteilung der eP -Werte sowie das Auffinden der unter Selektion stehenden Fenster erfolgte manuell.

3.2.7 Sequenzierung

Da es ein Ziel dieser Arbeit war, das mögliche kausale Gen und die zugrundeliegende Mutation einer Selektionssignatur eines besonders auffälligen Bereiches ausfindig zu machen, musste das Genom wesentlich genauer untersucht werden, als dies durch den BovineSNP50 Genotyping BeadChip von Illumina möglich ist. Diese eingehende Untersuchung wurde mittels der Durchführung sowohl einer Paired-End- als auch einer DNA-Sequenzierung bewerkstelligt.

3.2.7.1 Paired-End-Sequenzierung

Nach Auswertung der *iHS*- und *XP-EHH*-Ergebnisse, wurde aufgrund mehrerer Kriterien eine Signatur für weitere Analysen ausgewählt. Für einen besonders vielversprechenden Bereich der gewählten Signatur wurde dann eine Paired-End-Sequenzierung durchgeführt, um diesen genauer studieren zu können.

3.2.7.1.1 Auswahl des zu sequenzierenden Bereichs

Da im Süden von Deutschland das Deutsche Fleckvieh die dominierende Rinderrasse darstellt und somit von höchster Bedeutung ist, beschloss man die Suche nach einem besonders auffälligen Bereich und dessen Ursache auf diese Rasse zu beschränken.

Um nun die erfolgversprechendste Region ausfindig zu machen, wurden die durch *XP-EHH* und *iHS* ermittelten Fenster auf das Vorhandensein bekannter QTL sowie potenzieller Kandidatengene anderer genomweiter Selektionssignaturstudien überprüft. Darüber hinaus wurde die Anzahl sowohl der Fenster, die die jeweilige Signatur umfasste, als auch der Tests mit einem eP -Wert $\leq 0,01$ berücksichtigt.

Um nun die Region innerhalb der ausgewählten Signatur einzugrenzen, für die eine anschließende Sequenzierung am erfolgversprechendsten ist, wurde innerhalb der Haplotypen-Matrix, in welcher die SNPs an den einzelnen Positionen eines jeden Chromosoms für alle DFV-Tiere ablesbar sind, manuell nach Abschnitten gesucht, in denen möglichst alle Chromosomen der Rasse DFV bei aufeinanderfolgenden SNPs die gleichen Basen und somit den gleichen Haplotypen tragen.

Auf diese Weise konnte eine Region mit auffallend ausgeprägter Homozygotie im DFV ausfindig gemacht werden, welche dann auch bei der Rasse RH genauer studiert wurde. Durch diesen Vergleich wurde sichergestellt, dass sich die RH Haplotypen, in diesem Bereich insbesondere in Frequenz und Länge deutlich vom DFV unterscheiden. Anschließend wurden vier Vertreter des DFVs als Fall- sowie vier des RHs als Kontrollgruppe bestimmt, für welche dann eine Sequenzierung für eine Region von 1.520 kb durchgeführt wurde. Die Entscheidung RH-Tiere als Kontrolle zu nehmen, begründete sich dabei zum Einen auf die

Probenverfügbarkeit (nötig war mindestens eine Menge von einem Mikrogramm in einem maximalen Volumen von 50 µl) und zum Anderen darauf, dass in dem zu sequenzierenden Bereich auch das von Stella *et al.* (2010) als Kandidatengen identifizierte *KIT*-Gen liegt, welches unter anderem für die Scheckung bei Rindern verantwortlich gemacht wird (REINSCH *et al.*, 1999; FONTANESI *et al.*, 2009). Da sowohl DFV als auch RH gescheckt sind, kann diese signifikante Signatur nicht auf der Scheckung selbst beruhen. Darüber hinaus wiesen die einzelnen DFV und RH Tiere verschiedene, nicht im Voraus ausgewählte Grade an Scheckung auf. Somit stellte sich RH als Kontrolle zu verwenden als interessanter dar, als eine ungescheckte Rasse, obwohl auch solche Vergleiche auf signifikante Selektionssignaturen im DFV hindeuteten.

3.2.7.1.2 Durchführung der Paired-End-Sequenzierung

Die Sequenzierung der 1.520 kb umfassenden Region auf BTA6 wurde vom Laboratory for Functional Genome Analysis (LAFUGA, LMU) durchgeführt.

Um aus der Proben-DNA Fragmente unterschiedlicher Größe zu gewinnen, wurde diese zunächst mittels Ultraschall zerkleinert. Für die so gewonnenen DNA-Fragmente erfolgte anschließend die eigentliche Sequenzierung mittels Genome Analyzer Iix mit Paired-End Modul (Illumina, San Diego, USA). Wie in Abbildung 15 gezeigt beruht der Vorgang dabei auf mehreren Einzelschritten, welche sich drei Hauptarbeitsschritten zuordnen lassen. Schritt eins ist die Vorbereitung einer sogenannten DNA-Bibliothek („*Library*“) (Abb. 15 A-D). Hier werden zunächst die Enden der DNA-Fragmente geglättet und phosphoryliert, sowie ein Adenin-Überhang angefügt, an welchen schließlich die Adaptoren ligiert werden. Danach wird eine Amplifikation mittels Polymerase-Kettenreaktion (*PCR*) durchgeführt. Um zu gewährleisten, dass die darauffolgenden Arbeitsschritte nur für DNA-Fragmente durchgeführt werden, die zu der gewünschten Region gehören, erfolgt als Zwischenschritt die Hybridisierung auf einem speziellen Array mit für diese Region spezifischen Oligonukleotiden. Im Anschluss werden die freien DNA-Fragmente durch Waschen entfernt und schließlich die Gebundenen bei hohen Temperaturen eluiert und erneut mittels *PCR* vervielfältigt. Als zweiter Hauptarbeitsschritt folgt dann die Erzeugung der „*Cluster*“ (Abb. 15 E-H), bei dem

die DNA-Fragmente mittels Adapter zunächst an der „*Flow Cell*“ befestigt werden. Nach Anheften der DNA an die „*Flow Cell*“ kommt es zur Brückenbildung und der damit verbundenen Vervielfältigung der DNA, was zur lokal begrenzten Anhäufung von Kopien der einzelnen Fragmente führt, an welche schließlich die Primer angelagert werden. Den letzten Schritt stellt dann die eigentliche Sequenzierung dar (Abb. 15 I-K). Hierzu erfolgt die Zugabe von vier verschiedenen, basenspezifisch Fluoreszenz-markierten und für eine weitere Kettenverlängerung blockierten Desoxynucleosid-Triphosphaten (dNTP). Während jedes Durchgangs der Sequenzierung wird nur ein dNTP an die jeweilige Nukleinsäurekette angefügt, da die Markierungen zum Stoppen der Polymerisation führen. Anschließend wird die Farbe der Fluoreszenz bei allen *Clustern* parallel ermittelt und so die eingebaute Base eines jeden *Clusters* identifiziert. Da diese Technologie so bereits während der Synthese der DNA-Stränge ihre Sequenz bestimmt, wird sie „*sequencing by synthesis*“ oder kurz „*SBS*“ genannt. Um einen weiteren Zyklus der Sequenzierung zu ermöglichen werden dann der Fluoreszenz-Marker und die Schutzgruppe am 3'-OH enzymatisch abgespalten, so dass das nächste Nukleotid eingegliedert werden kann. Dieser Vorgang wird schließlich so oft wiederholt, bis die gewünschte vorher festgelegte Leseweite (hier 80 bp) erreicht ist. Die so gewonnenen Sequenzen werden dann mit der Referenzsequenz abgeglichen. Dabei wird bei der Paired-End-Sequenzierung die Genauigkeit dieses Abgleichs dadurch verbessert, dass für beide Enden der sequenzierten Fragmente die entsprechenden Sequenzen in der Referenzsequenz in einer bestimmten Distanz gefunden werden müssen, und so die Lage der Fragmente im Genom zweifach abgesichert wird. (persönliche Information Dr. Stefan Krebs, LAFUGA; http://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf).

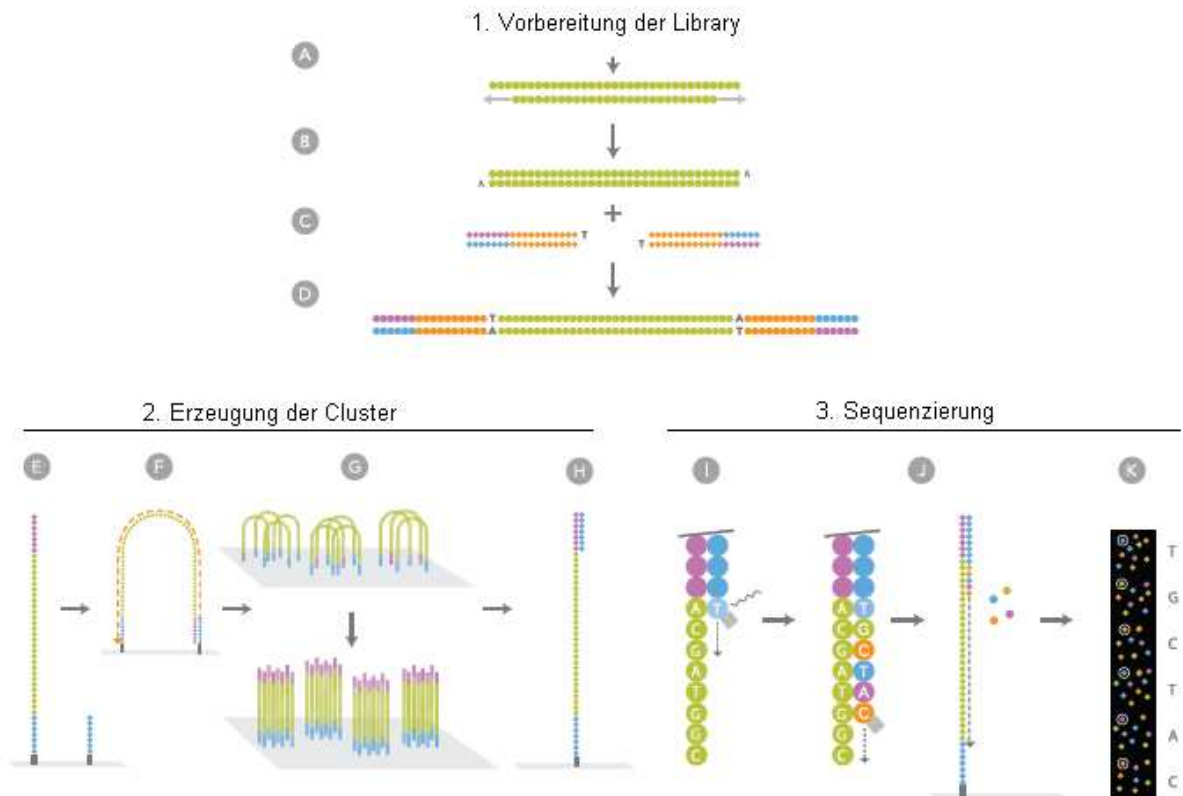


Abbildung 15: Paired-End-Sequenzierung

A) DNA-Fragment; B) Reparatur der DNA-Fragmente und Anfügen des Adenin-Überhangs; C) Adapterbindung; D) Auswahl gebundener DNA; E) Anheften der DNA an die Flowcell; F) Brückenamplifikation; G) Clusterbildung H) Anlagerung des Primers; I) erste Basenverlängerung, Auslesen und Entblocken; J) vorherigen Schritt für Strangverlängerung wiederholen; K) Basen auslesen (Quelle: Illumina; http://www.illumina.com/documents/products/brochures/brochure_genome_analyzer.pdf)

Die Ergebnisse dieser Sequenzierung sowie die verschiedenen Auswertemodule wurden vom LAFUGA auf einem Webserver unter GALAXY (Tab. 2) zur weiteren Analyse zur Verfügung gestellt. Durch Anwendung dieser Module konnte eine Vielzahl an divergenten Mutationen ausgemacht werden. Die Lage dieser Mutationen bezüglich der Gene sowie ihre Effekte auf die Transkription wurden mit dem „Variant Effect Predictor“ (<http://www.ensembl.org/tools.html>) bestimmt. Des Weiteren erfolgte eine visuelle Überprüfung mittels „Integrative Genomics Viewer“ (IGV 1.5, <http://www.broadinstitute.org/software/igv/node/120>).

3.2.7.2 DNA-Sequenzierung (Sanger)

Für einen Teilbereich der Region, die der Paired-End-Sequenzierung unterzogen wurde, erfolgte anschließend eine DNA-Sequenzierung. Dabei wurde dieser Teilbereich aufgrund der Lage der gefundenen Mutationen bezüglich dem Verlauf der Selektionssignatur sowie ihren Effekt auf die Transkription ausgewählt.

3.2.7.2.1 Erstellen der Primer

Bevor eine DNA-Sequenzierung des gewünschten Teilbereichs durchgeführt werden konnte, war zunächst die Erstellung von Primern nötig. Dabei sollten diese die Mutationen flankieren, die laut den Ergebnissen der Paired-End-Sequenzierung dort zu finden sind.

Für die Primererstellung fand das Programm „*Primer3 Input 0.4.0*“, welches vom Whitehead Institute for Biomedical Research online bereitgestellt wird (Tab. 2), Verwendung. Die Vorschläge möglicher Primerkombinationen, die dieses Programm machte, wurden anschließend mittels „*Beacon Designer™ Free Edition*“ (Premier Biosort International, Tab. 2) auf mögliches Fehlpaarungsverhalten, wie die Ausbildung von „*crosspins*“ oder „*hairpins*“ geprüft.

3.2.7.2.2 Durchführung der DNA-Sequenzierung

Die DNA-Sequenzierung, durch welche die genaue Basenabfolge des zu sequenzierenden Teilbereichs ermittelt werden konnte, wurde ebenfalls vom LAFUGA durchgeführt und diente der Verifizierung der Ergebnisse der Paired-End-Sequenzierung. Sie erfolgte mittels ABI 3100-Avant Genetic Analyzer (Applied Biosystems, California, USA).

Dabei beruht diese äußerst präzise DNA-Strukturanalyse auf dem von Sanger *et al.* (1977) entwickelten enzymatischen Sequenzierungsverfahren (Kettenabbruchverfahren). Grundlegendes Prinzip dieses Verfahrens ist, dass 2'-3'-dideoxy-Nucleosid-Triphosphate (ddNTPs) bei Einbau in neugebildete DNA-Moleküle zum

Abbruch der Synthese führen. Dies liegt daran, dass ddNTPs keine 3'-Hydroxylgruppe aufweisen, die mit der 5'-Phosphatgruppe des nachfolgend einzubauenden 5'-desoxy-Nucleosid-Triphosphats (dNTP) reagieren und so zur Verlängerung des DNA-Strangs führen könnte. Inkubiert man folglich Primer und DNA mit einer DNA-Polymerase in Anwesenheit des richtigen Verhältnisses zwischen dNTPs und ddNTPs, entsteht eine Mischung von Fragmenten unterschiedlicher Länge, welche sich jeweils nur um ein Nucleotid unterscheiden. Im Gegensatz zur ursprünglichen Methode ist es für die Sequenzbestimmung allerdings nicht mehr nötig für jedes der vier dNTP einen eigenen Ansatz herzustellen und per Acrylamidgelelektrophorese zu fraktionieren (SANGER *et al.*, 1977; GELDERMANN, 2005b). Vielmehr findet in der heute üblichen automatischen DNA-Sequenzierung die Verwendung von Fluoreszenzmarkierten ddNTPs Gebrauch. Da diese markierten ddNTPs bei Anregung durch Laser für die jeweilige Base spezifisches Licht emittieren, ist die Unterscheidung der eingebauten Basen in nur einem Ansatz möglich (GELDERMANN, 2005b). Darüber hinaus erfolgt die Auftrennung der Fragmente mittlerweile durch Kapillarelektrophorese, welche sich durch einen deutlich höheren Probendurchsatz auszeichnet (DROSSMAN *et al.*, 1990; LUCKEY *et al.*, 1990; WOOLLEY & MATHIES, 1995). Die DNA-Sequenz kann dann aus dem Profil der Fluoreszenzpeaks ermittelt werden, welches sich in Abhängigkeit von Fragmentgröße und emittiertem Licht ergibt.

Die so ermittelte Sequenz des entsprechenden Teilbereichs wurde dann vom LAFUGA an den Lehrstuhl für Tierzucht und Allgemeine Landwirtschaftslehre zurück übermittelt und mittels „ORF Finder“ (von Gene Infinity, siehe Tab. 2) auf das Vorhandensein sowie die Auswirkungen hier liegender Mutationen geprüft.

4 Ergebnisse

Zum Auffinden von Selektionssignaturen in divergent selektierten Rinderrassen, wurden für zehn verschiedene Rassen, die nach *IBD*-basierter Verwandtschaftskorrektur noch insgesamt 380 Tiere umfassten, genomweite (nur Autosome) *iHS*- und *XP-EHH*-Analysen durchgeführt.

Zur Auswertung beider Analysen wurde das Genom - abhängig von der jeweiligen Chromosomengröße - in 1.000 etwa gleich große Fenster unterteilt. Für jedes Fenster wurde im Anschluss sowohl der höchste standardisierte *XP-EHH*-Wert ermittelt als auch das Verhältnis der Markeranzahl mit einem absoluten standardisierten *iHS*-Wert > 2 zur Anzahl aller Marker des Fensters. Diese Werte wurden dann, für jeden der 90 *XP-EHH*-Vergleiche (10 Rassen x 9 Vergleiche) und jede der zehn Rassen (*iHS*) getrennt der Größe der Werte nach geordnet, um schließlich einen empirischen *P*-Wert (*eP*-Wert) zugeteilt zu bekommen. Dabei bekam das jeweils erste, also „beste“ Fenster den *eP*-Wert 0,001. Der *eP*-Wert erhöhte sich dann in Schritten von 0,001, bis schließlich das letzte, also „schlechteste“ Fenster erreicht war (empirischer *P*-Wert = 1).

Als signifikant wurden dann die Fenster deklariert, welche entweder bei der *iHS*-Analyse einen *eP*-Wert $\leq 0,01$ oder mindestens drei *XP-EHH*-Vergleiche mit einem *eP*-Wert $\leq 0,01$ aufwiesen.

Da eine *XP-EHH*-Bestimmung anhand von Rinderhaplotypen in der Literatur bisher nicht beschrieben ist, erfolgt zunächst eine Validierung dieser Methode, mittels bekannter einfach vererbter Merkmale, bevor dann die Lage der gefundenen Selektionssignaturen im bovinen Genom der zehn divergent selektierten Rinderrassen dargestellt wird. Darüber hinaus wird auf die Selektionssignaturen des DFVs, insbesondere auf den sequenzierten Bereich, genauer eingegangen, da das DFV die bedeutsamste Rasse im Süddeutschen Raum darstellt.

4.1 Validierung der Methode

Ähnlich wie Qanbari *et al.* (2010a) und Stella *et al.* (2010) wurde die Zuverlässigkeit der *XP-EHH*-Methode anhand dem Auffinden bekannter einfach vererbter Merkmale abgeschätzt. Hierzu wurden gezielt drei Loci ausgewählt, die sich aufgrund eindeutiger Phänotypen in jeweils einer Rasse deutlich von allen oder zumindest der Vielzahl der verbleibenden neun Rassen unterscheiden sollten. Es handelt sich hierbei um die Mutation im Myostatin-Gen, die beim Weißblauen Belgier den Phänotyp Doppellender zur Folge hat, um den Locus, der für die Rotfärbung des Red Holstein verantwortlichen ist („*extension*“) sowie den „*polled locus*“, welcher der Hornlosigkeit beim Galloway zugrunde liegt. Diese drei Merkmale, sowie ihre Lage im Genom und die dazugehörigen Ergebnisse der Analysen, werden im Folgenden genauer dargestellt.

Auf eine Validierung der *iHS*-Methode wird aufgrund der Tatsache, dass bereits *iHS*-Berechnungen an Rinderdaten durch das Bovine HapMap Consortium durchgeführt wurden, verzichtet (GIBBS *et al.*, 2009).

4.1.1 Phänotyp Doppellender

Der Phänotyp des Doppellenders entsteht aufgrund von Muskelhyperplasie, also einer Erhöhung der Anzahl an Muskelfasern, welche zu einem Mehr an Muskelmasse von etwa 20 Prozent führt. Dieses 1807 erstmals beschriebene Erscheinungsbild tritt weltweit in etlichen Rinderrassen auf (KAMBADUR *et al.*, 1997). Charlier *et al.* gelang 1995 die Kartierung des „*muscular hypertrophy*“ (*mh*) Locus auf das zentromere Ende von BTA2. Smith *et al.* (1997) kartierten schließlich im selben Intervall das Myostatin-Gen. Insbesondere in Anbetracht der Tatsache, dass Knock-out-Mäuse, bei denen das Myostatin-Gen ausgeschaltet wurde, ähnliche Phänotypen aufwiesen, stellte dieses Gen ein sehr aussichtsreiches Kandidatengen dar (McPHERRON & LEE, 1997). Es konnte dann mehrfach bestätigt werden, dass es sich bei der autosomal rezessiv vererbten Mutation, die dem Doppellender beim Weißblauen Belgier zugrunde liegt, um eine elf Basenpaare umfassende Deletion (nt821(del11)) im Myostatin-Gen handelt, welche zu einem vorzeitigen Stopp-Codon führt, was letztendlich zur

Folge hat, dass das Myostatin seiner Rolle als negativer Regulator des Muskelwachstums nicht gerecht wird (GROBET *et al.*, 1997; KAMBADUR *et al.*, 1997; McPHERRON & LEE, 1997).

Um nun zu überprüfen, ob die oben beschriebene Methode dieses Merkmal ausfindig machen kann, ist zunächst die Zuordnung des Myostatin-Gens zu der entsprechenden Fensternummer nötig. Das Gen befindet sich laut UCSC Genome Browser auf BTA2 genau zwischen den Basen 6.532.638 und 6.539.265 und fällt somit in das Fenster mit der Nummer 66 (5.119.264-7.676.879). In diesem Fenster deuten die *XP-EHH*-Werte eine signifikante ($eP < 0,01$) positive Selektion der Rasse BBB in acht von neun möglichen Vergleichen an. Lediglich der Vergleich mit GLW, also der einzigen Fleischrasse außer BBB selbst, weist nur einen indikativen eP -Wert (0,017) auf (Tab. 4).

Tabelle 4: eP -Werte der *XP-EHH*-Vergleiche mit BBB in Fenster 66 (5.119.264-7.676.879)

Exakte eP -Werte aller *XP-EHH* Vergleiche der Rasse BBB mit der jeweiligen Vergleichsrasse für das Fenster (F) 66, welches sich auf dem Rinderchromosom (BTA) 2 befindet.

		<i>XP-EHH</i> der Rasse BBB mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	MWF	FGV	SOBV	RH	BBV	GLW	ASV	IL
2	66	0,008	0,005	0,009	0,006	0,002	0,008	0,017	0,007	0,005

Betrachtet man den in Abbildung 16 dargestellten *XP-EHH*-Verlauf graphisch, so wird ebenfalls deutlich, dass sich der Weißblaue Belgier in diesem Bereich von allen Vergleichsrassen unterscheidet.

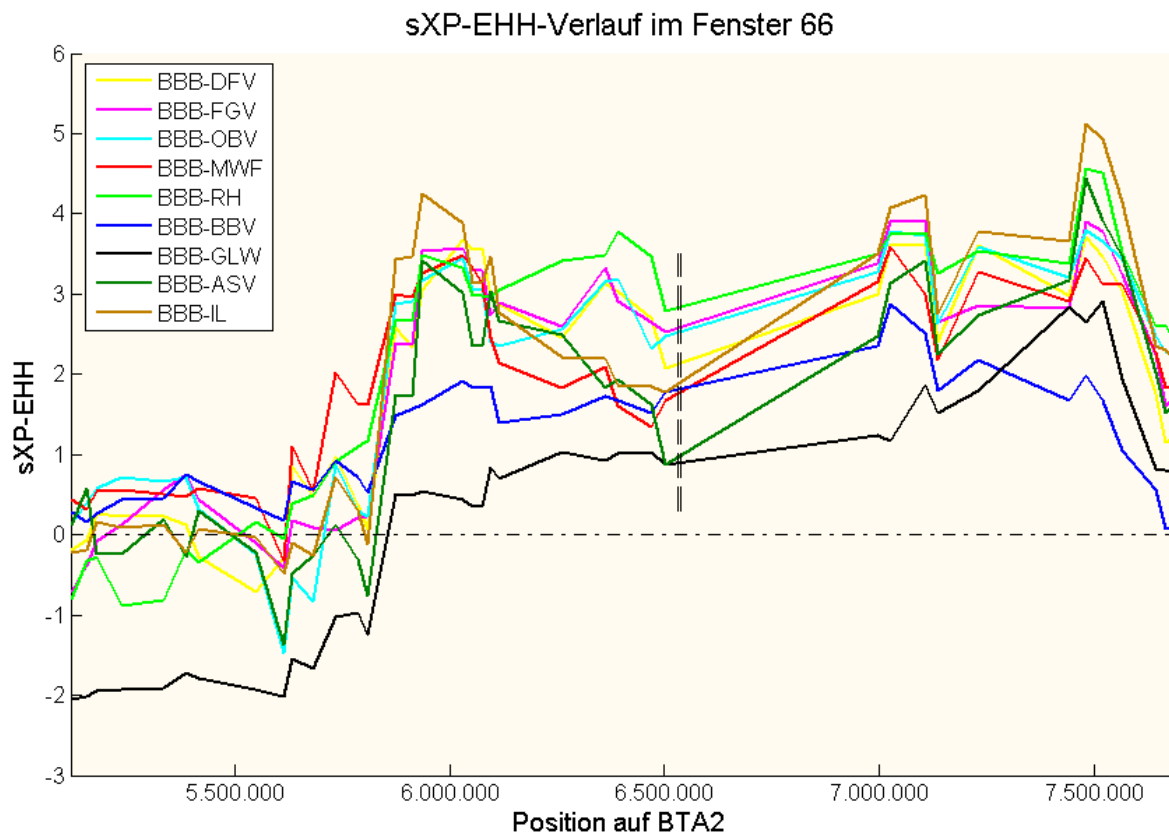


Abbildung 16: Signal der positiven Selektion auf BTA2 (Fenster 66) der Rasse BBB im Vergleich mit neun weiteren Rinderrassen.

Verlauf der standardisierten *XP-EHH*-Werte (*sXP-EHH*; y-Achse) in Abhängigkeit der physikalischen Position (bp) auf dem Rinderchromosom (BTA) 2 (x-Achse). Die Lage des *Myostatingens* wurde durch vertikale gestrichelte Linien gekennzeichnet.

Fenster 66 weist also eine deutliche Selektionssignatur auf, welche aufgrund des Phänotyps Doppellender im BBB zu erwarten war, und bestätigt so die Fähigkeit der Methode im Auffinden von unter Selektion stehenden Bereichen. Dem Anhang 9.4 lässt sich außerdem entnehmen, dass dieses Fenster in keiner der übrigen neun Rassen Signifikanz aufwies, also in diesem Bereich in keiner weiteren Rasse Hinweise auf positive Selektion gefunden werden konnten.

4.1.2 Phänotyp Rotfärbung

Für die Fellfärbung von Säugetieren sind primär zwei Pigmente verantwortlich. Dabei führt das Pigment Phaeomelanin phänotypisch zu rot/gelb, während Eumelanin braun/schwarz bedingt. Die relative Menge beider Pigmente wird im Wesentlichen durch zwei Loci, nämlich „*extension*“ und „*agouti*“, bestimmt (JOERG *et al.*, 1996). Dabei codiert das *Extension*-Gen für den Melanocortin-1-

Rezeptor (*Mc1r*), welchen Klungland *et al.* (1995) auf dem bovinen Chromosom 18 lokalisierten. Darüber hinaus wiesen sie drei verschiedene Allele nach; das dominante Allel E^D , welches zu schwarzer Färbung führt, das rezessive Allel e , das bei homozygoten Tieren zur Rotfärbung führt sowie das Wildtypallel E^+ , welches zu verschiedenen Färbungen führen kann (KLUNGLAND *et al.*, 1995). Joerg *et al.* (1996) zeigten, dass auch in der Holsteinpopulation, die sich bekannterweise in Schwarz- und Rotbunte Tiere unterteilen lässt, alle Red Holsteins homozygot für das rezessive Allel e sind, also für ihre rote Färbung, die bereits durch Klungland *et al.* (1995) beschriebene Leserahmen-Mutation, die dem e Allel zugrunde liegt, verantwortlich ist (JOERG *et al.*, 1996). Bei dieser Mutation handelt es sich um eine Deletion von Guanin an der Position 310, welche eine Verschiebung des Leserahmens zur Folge hat, die wiederum zu einem verfrühten Stopp-Codon führt, so dass das Protein nur aus 155 statt 317 Aminosäuren besteht. Vermutlich ist *Mc1r* deshalb nicht mehr im Stande das Melanozyten-stimulierende Hormon (α -MSH) als Agonisten der Eumelaninsynthese zu binden, während eine Bindung des Agouti-Proteins, welches die Eumelaninsynthese inhibiert, dennoch möglich scheint (ROUZAUD *et al.*, 2000).

Da alle Red Holsteins wie der Name schon sagt, den Phänotyp der Rotfärbung gemein haben, und somit homozygot für e sind, sollte sich dieser Bereich zumindest von den meisten Vergleichsrassen deutlich unterscheiden und daher eine nachweisbare Selektionssignatur aufweisen. Um dies zu prüfen, muss wiederum erneut das zu dieser Region gehörige Fenster ausfindig gemacht werden. Dem UCSC Genome Browser zufolge liegt das *Mc1r*-Gen auf BTA18 zwischen den Basen 13.776.888 und 13.778.639 und somit in Fenster 732 (BTA18: 12.716.504-15.259.366). Die *XP-EHH*-Werte innerhalb dieses Fensters deuteten auf positive Selektion in der Rasse RH hin, da sieben der maximal neun möglichen *XP-EHH*-Vergleiche mit RH-Beteiligung signifikant ($eP \leq 0,01$) waren. Die beiden Vergleiche von RH mit OBV und GLW waren zwar nicht signifikant, zeigten aber indikative eP -Werte auf (0,013 respektive 0,012). Darüber hinaus wiesen auch die umliegenden Fenster (F731-F735) per Definition der Methode mit jeweils mindestens drei signifikanten Vergleichen auf positive Selektion in der Rasse RH hin (Tab. 5).

Tabelle 5: *eP*-Werte der *XP-EHH*-Vergleiche mit RH in den Fenstern 731-735 (10.173.641-22.887.955)

Exakte *eP*-Werte aller *XP-EHH* Vergleiche der Rasse RH mit der jeweiligen Vergleichsrasse für die Fenster (F) 731-735, welche sich auf dem Rinderchromosom (BTA) 18 befinden.

BTA	F	<i>XP-EHH</i> der Rasse RH mit der jeweiligen Vergleichsrasse								
		DFV	MWF	FGV	OBV	BBV	GLW	BBB	ASV	IL
18	731	0,004	0,005	0,005	0,037	0,018	0,052	0,015	0,004	0,008
18	732	0,003	0,001	0,003	0,013	0,003	0,012	0,003	0,003	0,003
18	733	0,001	0,002	0,001	0,002	0,005	0,001	0,001	0,001	0,001
18	734	0,002	0,010	0,002	0,004	0,039	0,002	0,005	0,002	0,002
18	735	0,030	0,017	0,012	0,007	0,022	0,009	0,002	0,009	0,013

Um sicherzustellen, dass es sich bei der fünf Fenster umfassenden Signatur wirklich nur um eine Region handelt, wurde dieser Bereich ebenfalls graphisch dargestellt (Abb. 17).

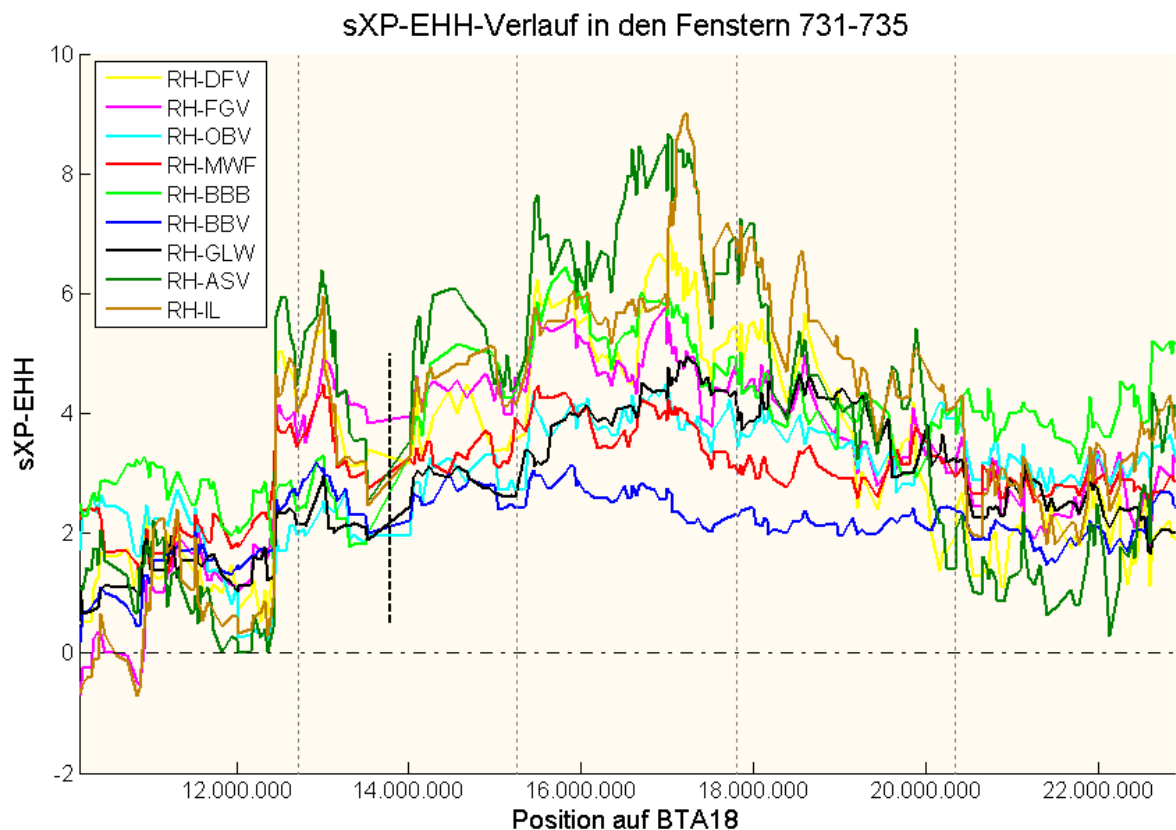


Abbildung 17: Signal der positiven Selektion auf BTA18 (Fenster 731-735) der Rasse RH im Vergleich mit neun weiteren Rinderrassen.

Verlauf der standardisierten *XP-EHH*-Werte (*sXP-EHH*; y-Achse) in Abhängigkeit der physikalischen Position (bp) auf dem Rinderchromosom (BTA) 18 (x-Achse). Die Lage des *Mc1R*-Gens wurde durch eine schwarze vertikale gestrichelte Linie gekennzeichnet. Die grau gestrichelten vertikalen Linien markieren jeweils die Fenstergrenzen.

Zusammenfassend lässt sich feststellen, dass sich der Bereich, der für den Phänotyp der Rotfärbung verantwortlich ist, ebenfalls in einer von der Methode ausfindig gemachten Selektionssignatur befindet. Aufgrund des Umfangs jener

Signatur kann jedoch nicht mit letzter Sicherheit gesagt werden, dass ihr ausschließlich eine Selektion auf *Mc1r* zugrunde liegt und nicht auch andere unter Selektion stehende Gene beteiligt sind.

Im Gegensatz zum RH deutete die *XP-EHH*-Auswertung der verbleibenden neun Rassen weder in Fenster 732 noch in dem umgebenden Fenstern (731-735) auf positive Selektion hin (Anhang 9.4).

4.1.3 Phänotyp Hornlosigkeit

Die Entwicklung von Hörnern ist einzigartig für Bovidae. Jedoch sind generell nicht alle Individuen behornt, so dass es zwischen und teilweise auch innerhalb der Rinderrassen phänotypische Unterschiede gibt (MARIASEGARAM *et al.*, 2010).

Bei taurinen Rindern ist die Hornlosigkeit unter genetischer Kontrolle des autosomal dominanten „*polled-locus*“, so dass Tiere mit dem Genotypen P/P und P/p keine Hörner ausbilden, während p/p-Tiere behornt sind. Bereits 1993 wurde gezeigt, dass dieser Locus an Mikrosatellitenmarker gekoppelt ist, die auf BTA1 sehr nahe dem Centromer positioniert sind (GEORGES *et al.*, 1993), was spätere Studien bestätigten (BRENNEMAN *et al.*, 1996; DRÖGEMÜLLER *et al.*, 2005a). Drögemüller *et al.* (2005b) gelang es zwar, die wahrscheinlichste Lokalisation des „*polled-locus*“ auf ein etwa 1,3 Mb großes Intervall zwischen den Markern *RP42-218J17_MS1* und *BM6438* einzugrenzen, das kausale Gen konnte bisher dennoch nicht identifiziert werden.

Dieser Bereich liegt in etwa zwischen den Basen 661.000 und 2.025.000 und entspricht dem Fenster mit der Nummer eins (483-2.557.020). Da das Galloway die einzige unter den zehn hier verwendeten Rassen darstellt, die durchgehend genetisch hornlos ist, sollte es sich in diesem Bereich von den anderen neun Rassen deutlich unterscheiden. Tabelle 6 zeigt zwar, dass keiner der neun Vergleiche einen empirischen *P*-Wert $\leq 0,01$ aufweist, was zur Folge hat, dass dieses Fenster bei der sehr strengen Methodik nicht als signifikant zählt, dennoch ist der empirische *P*-Wert von sechs Vergleichen unter 0,05.

Tabelle 6: eP -Werte für XP - EHH -Vergleiche mit GLW in Fenster 1 (483-2.557.020)

Exakte eP -Werte aller XP - EHH Vergleiche der Rasse GLW mit der jeweiligen Vergleichsrasse für das Fenster (F) 1, welches sich auf dem Rinderchromosom (BTA) 1 befindet.

		XP-EHH der Rasse GLW mit der jeweiligen Vergleichsrasse									
BTA	F	DFV	MWF	FGV	OBV	RH	BBV	BBB	ASV	IL	
1	1	0.033	0.089	0.013	0.042	0.016	0.054	0.261	0.017	0.033	

Dies wiederum bedeutet, dass anhand der XP - EHH -Berechnungen durchaus Hinweise auf Selektion bestehen, was auch dem Kurvenverlauf in Abbildung 18 zu entnehmen ist.

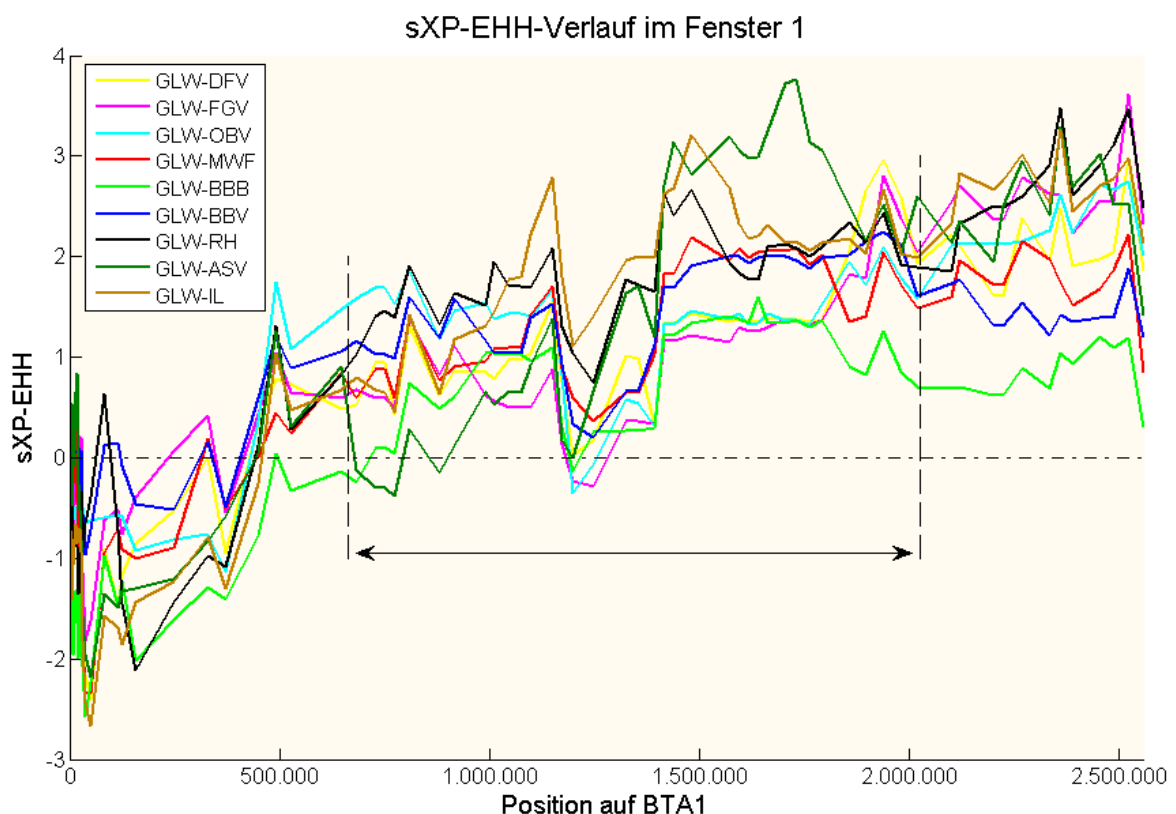


Abbildung 18: Signal der positiven Selektion auf BTA1 (Fenster 1) der Rasse GLW im Vergleich mit neun weiteren Rinderrassen.

Verlauf der standardisierten XP - EHH -Werte (sXP - EHH ; y-Achse) in Abhängigkeit der physikalischen Position (bp) auf dem Rinderchromosom (BTA) 1 (x-Achse). Die Lage des Polled-Locus wurde durch den mit einem Doppelpfeil markierten Bereich gekennzeichnet.

Jedoch wird dieses Fenster aufgrund der sehr stringenten Kriterien, insbesondere des Signifikanzniveaus von $eP = 0,01$, nicht gefunden.

Den Erwartungen entspricht aber, dass dieses Fenster auch für keine der neun anderen Rassen signifikant ist (Anhang 9.4).

Es ist bekannt, dass sich insbesondere in den Anfangsbereichen (von 0 bis etwa 329 kb) des Rinderchromosoms 1 sowohl Markerreihenfolge als auch

Markerposition der aktuelleren Version des *Bos taurus* Genoms *UMD3.1* (http://www.cbcb.umd.edu/research/bos_taurus_assembly.shtml) deutlich von der weniger aktuellen Version *Btau_4.0* (<http://genome.ucsc.edu/cgi-bin/hgGateway>) unterscheiden. Da der verwendete BeadChip jedoch auf der älteren *Btau_4.0* Version basiert, musste ausgeschlossen werden, dass diese falsch zugeordneten Marker erheblichen Einfluss auf die berechneten *XP-EHH*-Ergebnisse hatten. Dazu erfolgte eine neue Berechnung unter Ausschluss dieser problematischen Marker. Da sich der Kurvenverlauf des restlichen Bereichs (über 329 kb) jedoch nur unmerklich änderte, war ein Einfluss der falsch zugeordneten Marker auf den *XP-EHH*-Verlauf im Bereich des Polled-Locus nicht anzunehmen. Dieses Ergebnis war, angesichts der Tatsache, dass die Markerreihenfolge im Bereich des Polled-Locus in der älteren Markerversion bereits korrekt war, zu erwarten.

4.1.4 Zusammenfassende Beurteilung der Methode

Es konnte bei zwei der drei getesteten Phänotypen (Doppellender, Rotfärbung) gezeigt werden, dass die den Phänotypen zugrundeliegenden Gene in Fenstern liegen, die aufgrund dieser Methodik als signifikant deklariert wurden. Lediglich beim Phänotyp Hornlosigkeit, war die Methode in hier angewandter Form nicht in der Lage Selektion nachzuweisen. Würde das Signifikanzniveau allerdings von $eP = 0,01$ (hochsignifikant) auf den weniger stringenten Wert $eP = 0,05$ (signifikant) gesetzt, würde auch dieses Fenster die entsprechenden Kriterien erfüllen.

Für alle drei untersuchten Phänotypen stimmte darüber hinaus die Negativkontrolle mit den Erwartungen überein, da die entsprechenden Fenster in keiner der jeweils neun weiteren Rassen den Status signifikant erreichten.

Alles in allem kann festgestellt werden, dass die durch diese Methode gefundenen vermeintlichen Selektionssignaturen mit hoher Wahrscheinlichkeit tatsächlich auf Selektion zurückzuführen sind, auch wenn mit dem auf $eP = 0,01$ festgelegten Signifikanzniveau einige unter Selektion stehende Bereiche verfehlt werden. Dies wird aber aufgrund der damit verbundenen Herabsetzung falsch positiver Fenster in Kauf genommen.

4.2 Genomweite Ergebnisse

Im Folgenden werden die Ergebnisse der beiden Analysen *iHS* und *XP-EHH* zunächst genomweit für alle zehn Rassen kombiniert dargestellt, bevor anschließend genauer auf die Ergebnisse der Rasse DFV eingegangen wird.

4.2.1 *iHS*-Ergebnisse

Die *iHS*-Berechnung basiert auf einem innerhalb jeder Rasse durchgeführten Vergleich der Ausprägung der Haplotypenhomozygotie entlang der Chromosomen. Die Bestimmung der standardisierten *iHS*-Werte erfolgte daher für jede der zehn Rassen getrennt. Anschließend wurde jedem Fenster für jede Rasse ein empirischer *P*-Wert zugeteilt, mit der Einschränkung, dass Fenster mit weniger als 20 SNPs nicht berücksichtigt wurden. Als signifikant wurden jene Fenster deklariert, die in der jeweiligen Rasse die zehn niedrigsten *eP*-Werte aufwiesen, also $\leq 0,01$ waren.

Von insgesamt 100 (jeweils zehn pro Rasse) der so deklarierten signifikanten Fenstern waren lediglich fünf in mehr als nur einer Rasse signifikant. Dies waren jenes mit der Nummer 22 (OBV, IL) auf BTA1, die Nummer 370 (FGV, OBV) auf BTA8, sowie die Fenster 602 (DFV, RH, GLW), 603 (FGV, RH) und 605 (RH, GLW) auf BTA14.

Die Verteilung der 94 verschiedenen Fenster mit signifikanten *iHS*-Signaturen auf die 29 Chromosomen stellte sich sehr inhomogen dar. So wiesen BTA24 und BTA25 keine Selektionssignatur auf, während auf BTA1 elf Fenster signifikant waren (Abb. 19).

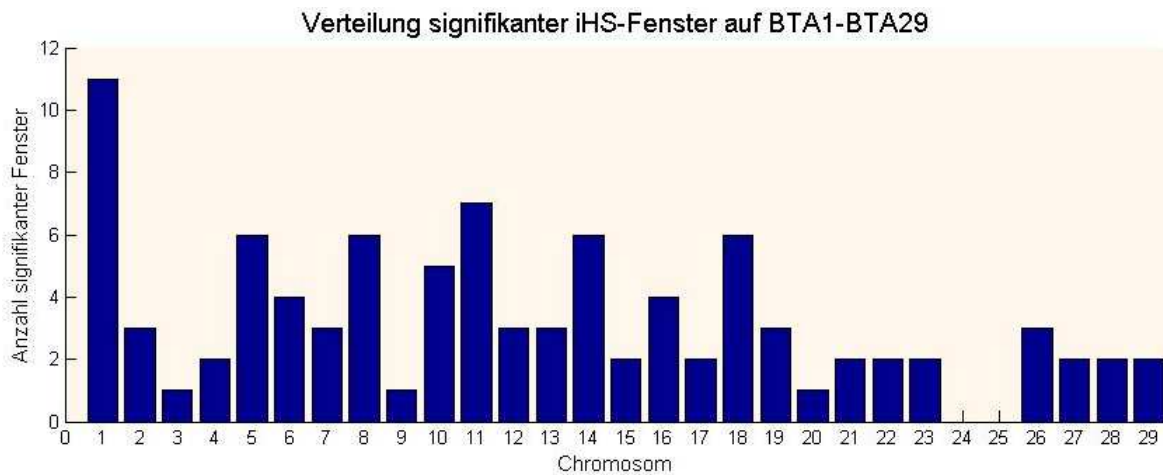
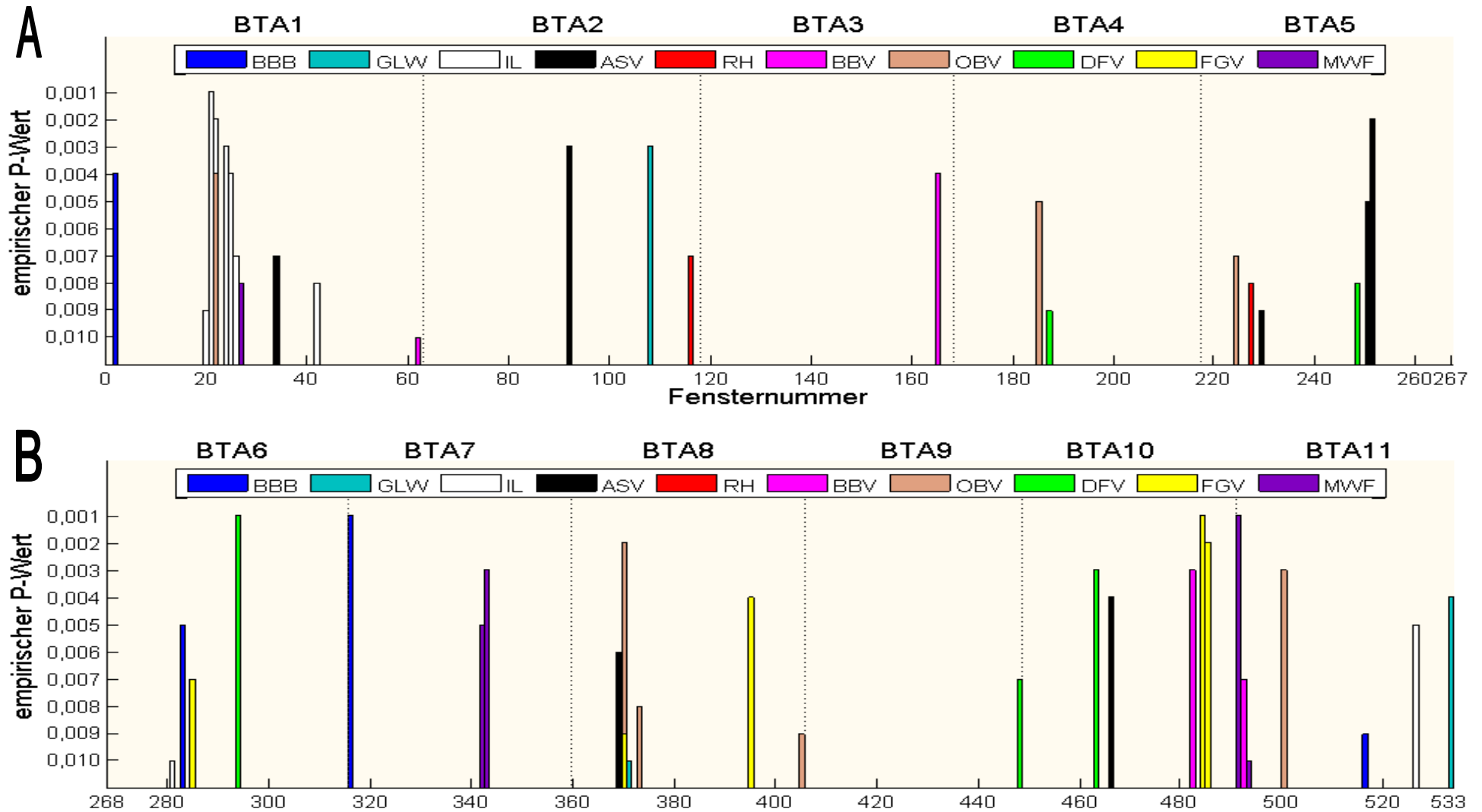


Abbildung 19: Übersicht über die Verteilung der 94 verschiedenen Fenster mit signifikanter *iHS*-Signatur auf die 29 Autosome des Rindes.

Die Anzahl der signifikanten Fenster (y-Achse) ist in Abhängigkeit der Nummer des jeweiligen Rinderchromosoms (x-Achse) dargestellt.

In sechs der zehn Rassen gibt es Gruppen, in denen mehrere direkt aufeinander folgende Fenster signifikant sind. Jedoch lediglich in einer Größe von zwei (FGV, MWF, BBB, ASV) respektive drei Fenstern (RH, IL). In Abbildung 20C ist beispielsweise die drei Fenster umfassende Gruppe der Rasse RH zu erkennen. Die zugehörigen Fensternummern und ihre eP -Werte sind ebenfalls dieser Abbildung zu entnehmen (F601, $eP = 0,002$; F602, $eP = 0,006$; F603, $eP = 0,004$). Für das Fenster 602 ist außerdem ersichtlich, dass dieses auch in zwei weiteren Rassen signifikant ist (GLW, $eP = 0,009$; DFV, $eP = 0,002$). Auch die übrigen Fenstergruppen sowie alle separaten Fenster, welche aufgrund der *iHS*-Ergebnisse als signifikant deklariert wurden und daher Selektionssignaturen beinhalten müssten, sind der genomweiten Übersicht in Abbildung 20 (A-D) zu entnehmen.

Die genauen Fensternummern sowie alle entsprechenden eP -Werte sind außerdem im Anhang 3 aufgeführt.



1. Teil Abbildung 20; Fortsetzung auf nachfolgender Seite

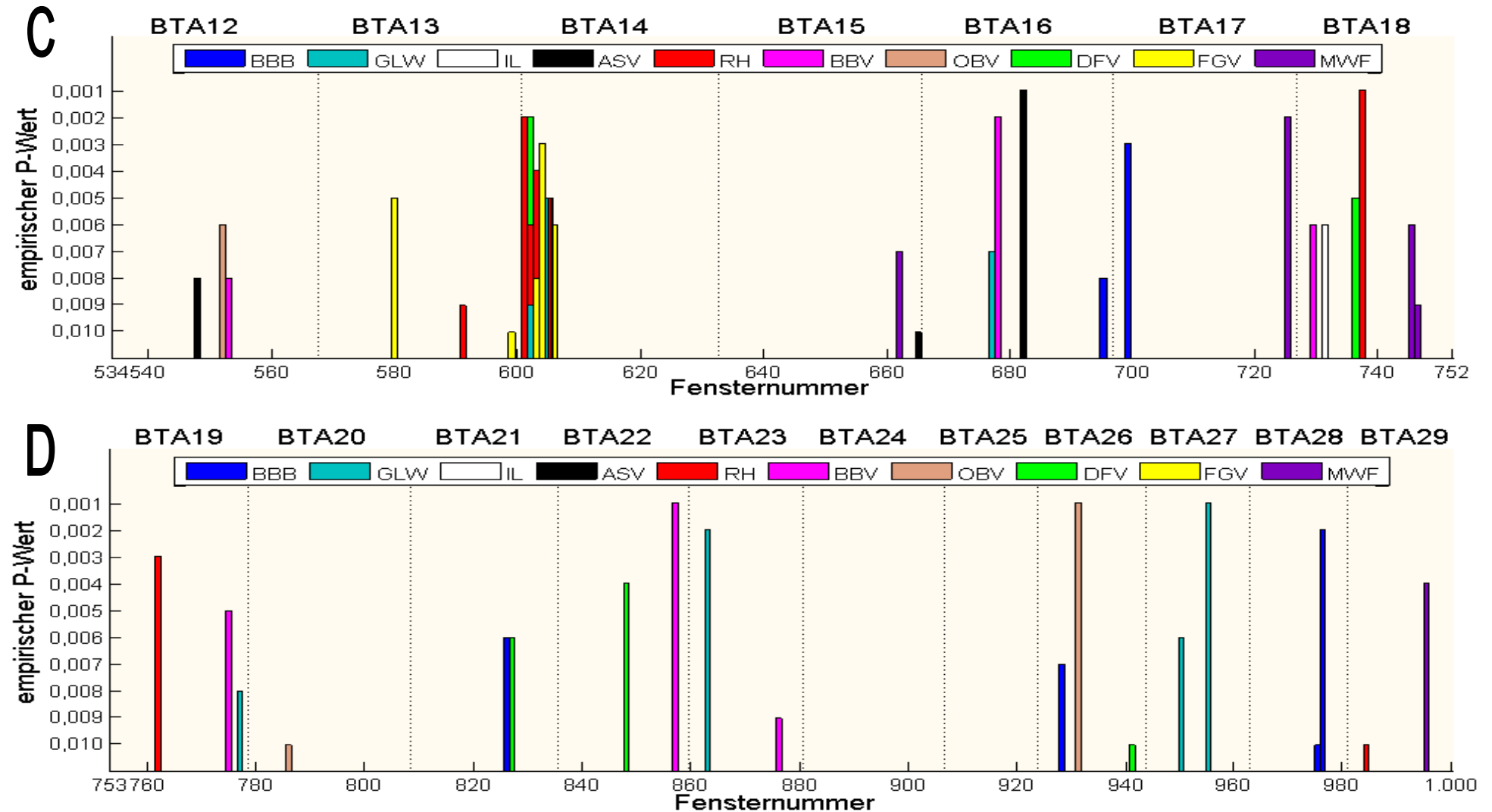


Abbildung 20: Genomweite Verteilung der 100 signifikanten *iHS*-Fenster der zehn verschiedenen Rinderrassen

Dargestellt sind alle Fenster mit empirischen *P*-Wert $\leq 0,01$ (y-Achse) in Abhängigkeit von der Fensternummer (x-Achse). Die Grenzen der jeweiligen Chromosomen (BTA) sind durch vertikale gepunktete Linien gekennzeichnet. A: BTA1-BTA5; B: BTA6-BTA11; C: BTA12-BTA18; D: BTA19-BTA29

4.2.2 XP-EHH-Ergebnisse

Der *XP-EHH*-Berechnung liegt der zwischen zwei verschiedenen Rassen durchgeführte Vergleich der Ausprägung der Haplotypenhomozygotie entlang der Chromosomen zugrunde. Hierfür wurden für jede einzelne der zehn Rassen mit jeder der neun verbleibenden Rassen separat standardisierte *XP-EHH*-Werte bestimmt, so dass insgesamt 90 (zehn Rassen x neun Vergleiche) Rassenvergleiche zu beurteilen waren. Um nun als Selektionssignatur einer bestimmten Rasse deklariert zu werden, muss ein Fenster per Definition der hier angewandten Methode für mindestens drei der jeweils neun Vergleiche dieser speziellen Rasse einen empirischen *P*-Wert $\leq 0,01$ aufweisen.

Diese Kriterien erfüllten in den zehn Rassen insgesamt 103 Fenster, wobei es sich dabei um 90 verschiedene handelte. Die meisten dieser 90 Fenster wiesen also nur für eine der zehn Rassen eine Selektionssignatur auf. Lediglich sieben Fenster erreichten diese Merkmale in zwei oder mehr Rassen. Dies waren das Fenster 119 (DFV, FGV) auf BTA3, die Fenster mit den Nummern 282 (DFV, FGV, OBV, MWF), 283 (DFV, FGV, OBV, MWF, ASV), 284 (DFV, FGV, MWF) und 295 (DFV, FGV) auf BTA6 sowie F932 (OBV, RH) und F933 (OBV, RH) auf BTA26.

Die Verteilung der signifikanten Fenster auf die 29 Chromosomen war, wie auch beim *iHS* (Abb. 19), wenig homogen und schwankte zwischen keiner Signatur (BTA12, BTA17, BTA23, BTA24, BTA27) und zwölf signifikanten Fenstern auf BTA6 (Abb. 21).

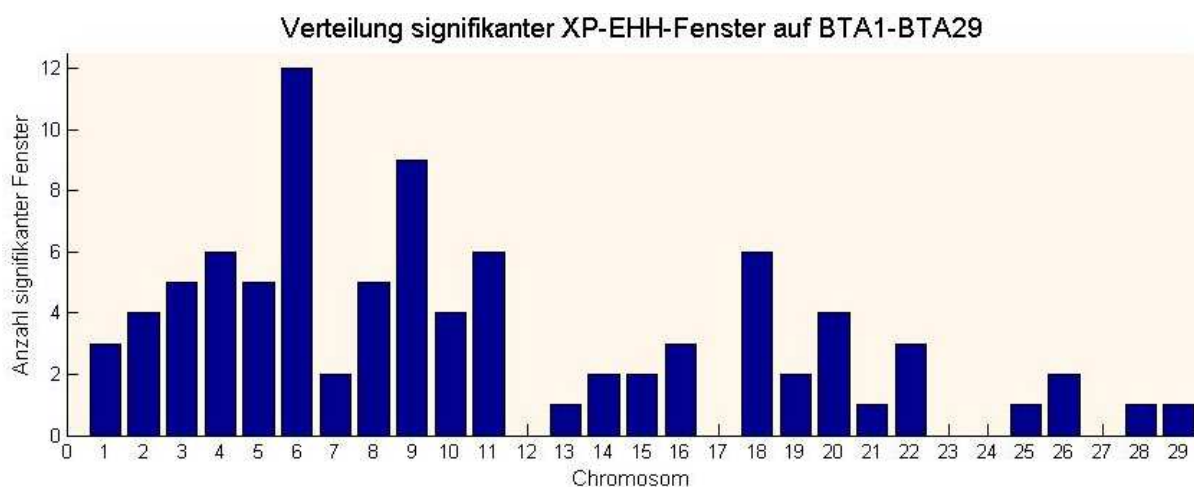
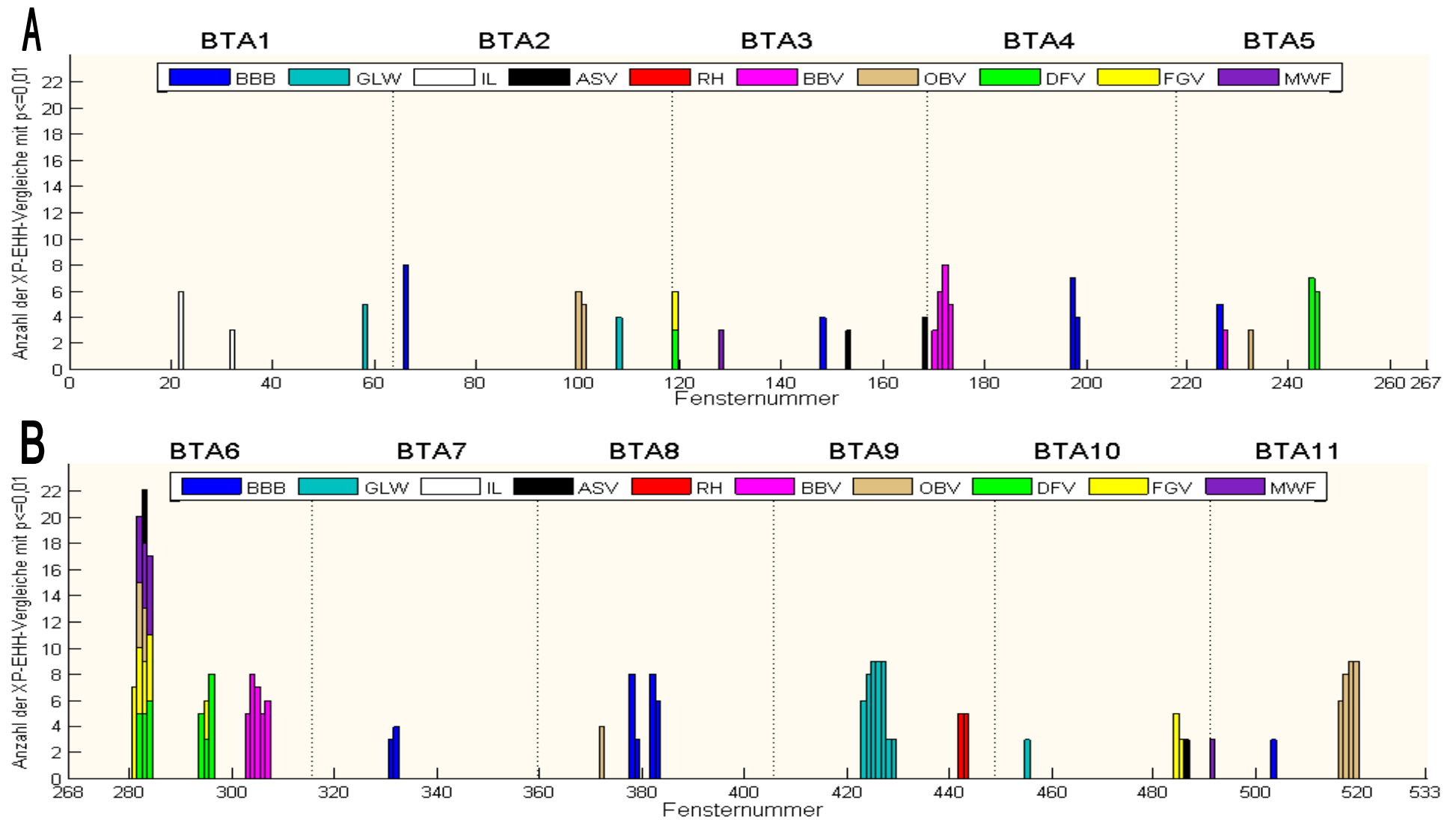


Abbildung 21: Übersicht über die Verteilung der 90 verschiedenen Fenster mit signifikanter *XP-EHH*-Signatur auf die 29 Autosome des Rindes.

Die Anzahl der signifikanten Fenster (y-Achse) ist in Abhängigkeit der Nummer des jeweiligen Rinderchromosoms (x-Achse) dargestellt.

Auch die Anzahl signifikanter Fenster der einzelnen Rassen schwankt mit nur vier beim Illyrischen Rind und 14 beim Weißblauen Belgier deutlich. Abbildung 22 zeigt die genomweite Übersicht aller Fenster, welche aufgrund der *XP-EHH*-Ergebnisse als signifikant definiert wurden und daher Selektionssignaturen beinhalten sollen. Außerdem veranschaulicht diese Abbildung, dass es in acht der zehn Rassen Gruppen gibt, in denen mehrere direkt aufeinander folgende Fenster signifikant sind. Lediglich die unselektierten Rassen ASV und IL weisen nur separate signifikante Fenster auf, während die Größe der Fenstergruppen in den übrigen Rassen zwischen zwei (BBB) und sieben (GLW) benachbarten Fenstern variiert. Besonders auffällige und weitgehend rassespezifische Fenstergruppen lassen die Rassen BBV (F170-173, BTA4; F303-F307, BTA6), DFV (F294-F296, BTA6), GLW (F423-F429, BTA9), OBV (F516-519, BTA11) und RH (F731-735, BTA18) erkennen. Bei der angeführten Fenstergruppe der Rasse RH handelt es sich dabei um den oben (siehe 4.1.2) beschriebenen Bereich, der auch das für die Rotfärbung verantwortliche Extension-Gen, welches *Mc1r* codiert, trägt. Neben diesen rassespezifischen Gruppen ist eine Rasseübergreifende auf BTA6 besonders auffällig. Wie Abbildung 22B zu entnehmen ist, weisen in diesem Bereich (F281-F284) bis zu fünf Rassen mindestens drei *XP-EHH*-Vergleiche mit einem *eP*-Wert $\leq 0,01$ auf und gelten somit als signifikant. In Fenster 283 addieren sich je fünf Vergleiche der Rassen DFV und MWF sowie jeweils vier der Rassen FGV, OBV und ASV zu insgesamt 22 *XP-EHH*-Vergleichen mit *eP*-Wert $\leq 0,01$. Damit stellt F283 das Fenster mit der rasseübergreifend höchsten Anzahl solcher Vergleiche (*eP* $\leq 0,01$) dar.

Die genaue Anzahl an signifikanten Fenstern pro Rasse, die dazugehörigen Fensternummern sowie die entsprechenden *eP*-Werte $\leq 0,01$ sind dem Anhang 9.4 zu entnehmen.



1. Teil Abbildung 22; Fortsetzung auf nachfolgender Seite

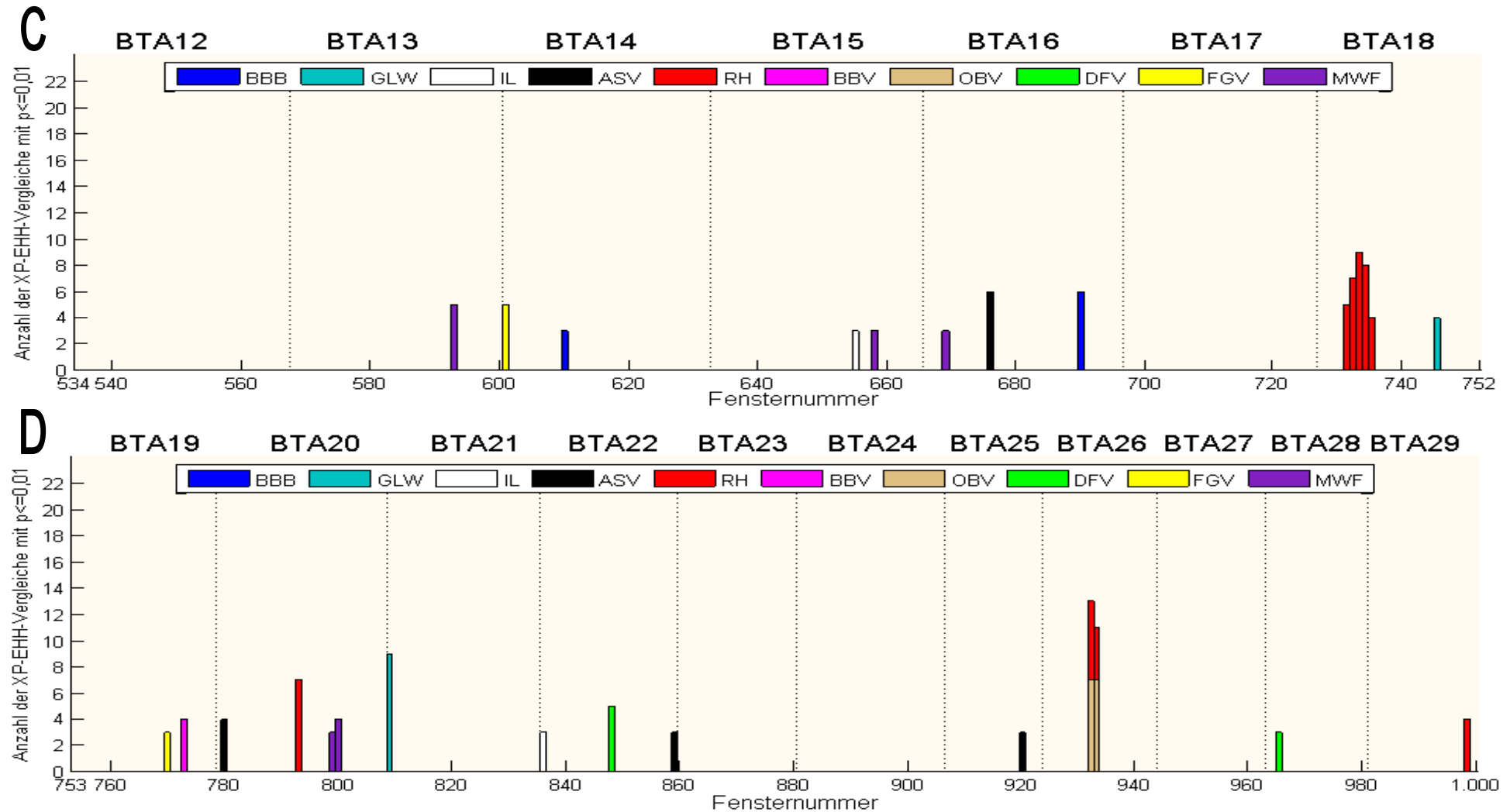


Abbildung 22: Genomweite Verteilung der 103 signifikanten *XP-EHH*-Fenster der zehn verschiedenen Rinderrassen

Dargestellt ist die Anzahl der *XP-EHH*-Vergleiche mit einem empirischen P -Wert $\leq 0,01$ (y-Achse) in Abhängigkeit von der Fensternummer (x-Achse). Für Signifikanz ist dabei eine Mindestanzahl von drei *XP-EHH*-Vergleichen mit $eP \leq 0,01$ nötig. Die Grenzen der jeweiligen Chromosomen (BTA) sind durch vertikale gepunktete Linien gekennzeichnet. A: BTA1-BTA5; B: BTA6-BTA11; C: BTA12-BTA18; D: BTA19-BTA29

4.3 Deutsches Fleckvieh

Nach Angaben des Bayerischen Staatsministeriums für Ernährung, Landwirtschaft und Forsten betrug der Gesamtviehbestand in Bayern nach der Viehzählung vom November 2009 3,40 Millionen Rinder mit einem Fleckviehanteil von 80,9 Prozent (Quelle: <http://www.verwaltung.bayern.de/Anlage1736559/RinderzuchtinBayern-FaktenundDaten.pdf>). Auch in Baden-Württemberg liegt der Anteil am Fleckvieh laut Ministerium für ländlichen Raum, Ernährung und Verbraucherschutz mit 41,0 Prozent deutlich über dem der übrigen Rassen (Quelle: http://www.mlr.baden-wuerttemberg.de/mlr/bro/Bro_Rinderzucht_Rinderhaltung.pdf). Diese Zahlen veranschaulichen die Bedeutung dieser Rasse im süddeutschen Raum und stellen den Grund dar, weshalb die Selektionssignaturen speziell dieser Rasse genauer betrachtet wurden.

Tabelle 7 zeigt die 19 verschiedenen signifikanten Fenster, von welchen neun nur durch *XP-EHH*, acht ausschließlich mittels *iHS* und weitere zwei (294, 848) von beiden Methoden erkannt wurden.

Tabelle 7: Die 19 Selektionssignaturen-beinhaltenen Fenster des Deutschen Fleckviehs

Dargestellt sind alle Fenster, die entweder durch *XP-EHH* oder *iHS* für signifikant erklärt wurden. Die empirischen *P*-Werte (*eP*) wurden für jeden *XP-EHH*-Vergleich des DFV mit einer der neun anderen Rassen (Rasseabkürzungen) sowie für den innerhalb des DFV ermittelten *iHS* aufgeführt. Die *eP*-Werte der einzelnen Fenster auf den zugehörigen Rinderchromosomen (BTA), welche die Signifikanzschwelle erreichten ($eP \leq 0,01$), wurden farblich unterlegt. Fenster in denen weniger als 20 Marker für die Ermittlung des *iHS*-Wertes verfügbar waren wurden mit „<20“, solche die keinen Marker mit $|iHS| > 2$ aufwiesen mit „-“, gekennzeichnet.

BTA	Fenster	<i>XP-EHH</i> der Rasse DFV mit der jeweiligen Vergleichsrasse									<i>iHS</i>
		OBV	MWF	FGV	RH	BBV	GLW	BBB	ASV	IL	
3	119	0,007	0,382	0,634	0,138	0,002	0,006	0,075	0,047	0,081	-
4	187	0,045	0,055	0,058	0,018	0,254	0,171	0,031	0,052	0,021	0,009
5	244	0,004	0,005	0,002	0,008	0,028	0,004	0,006	0,012	0,008	<20
5	245	0,003	0,007	0,076	0,010	0,033	0,012	0,007	0,010	0,009	<20
5	248	0,009	0,019	0,018	0,096	0,856	0,075	0,206	0,082	0,027	0,008
6	282	0,447	0,540	0,262	0,001	0,742	0,002	0,003	0,005	0,003	<20
6	283	0,235	0,968	0,471	0,002	0,202	0,001	0,008	0,002	0,002	<20
6	284	0,002	0,658	0,379	0,003	0,062	0,003	0,004	0,001	0,001	<20
6	294	0,435	0,002	0,029	0,005	0,866	0,924	0,002	0,007	0,007	0,001
6	295	0,017	0,013	0,119	0,017	0,596	0,647	0,010	0,006	0,004	<20
6	296	0,001	0,001	0,001	0,004	0,005	0,071	0,001	0,003	0,005	<20
9	448	0,230	0,128	0,039	0,216	0,014	0,146	0,067	0,049	0,090	0,007
10	463	0,580	0,316	0,568	0,842	0,336	0,854	0,201	0,042	0,034	0,003
14	602	0,053	0,036	0,072	0,368	0,023	0,312	0,016	0,057	0,019	0,002
18	736	0,490	0,466	0,177	0,987	0,359	0,670	0,115	0,427	0,690	0,005
21	827	0,160	0,045	0,110	0,021	0,768	0,186	0,497	0,020	0,031	0,006
22	848	0,008	0,062	0,005	0,009	0,857	0,008	0,009	0,024	0,036	0,004
26	941	0,019	0,087	0,015	0,007	0,092	0,317	0,045	0,011	0,010	0,010
28	965	0,021	0,006	0,004	0,025	0,008	0,046	0,023	0,029	0,074	0,032

Da nur eine Signatur im Verlauf dieser Arbeit weiter untersucht werden sollte, war an dieser Stelle eines der 19 Fenster auszuwählen. Hierzu wurde für die signifikanten Bereiche ein Überschneiden mit bekannten QTL sowie vermeintlichen kausalen Genen aus bereits veröffentlichten Selektionssignaturstudien geprüft und in Tabelle 8 zusammengefasst. Suggestive QTL wurden dabei nur berücksichtigt, wenn innerhalb des jeweiligen Fensters keine signifikanten QTL vorhanden waren.

Tabelle 8: Übereinstimmung der signifikanten Fenster des DFV mit QTL und möglichen kausalen Genen anderer Studien.

Dargestellt sind 19 signifikante Selektionssignaturen, welche auf zwölf Rinderchromosomen (BTA) verteilt sind. In den entsprechenden Genomabschnitten (Fenster) wurden keine (-) bis maximal zwölf verschiedene QTL in der Bovine QTL Datenbank (<http://genomes.sapac.edu.au/bovineqtl/home.php>) gefunden und mit keinem (-) bis maximal drei möglichen kausalen Genen in Verbindung gebracht, von welchen jeweils Gensymbol sowie in Klammern Genbeschreibung angegeben sind.

BTA	Fenster	QTL	kausale Gene
3	119	1	-
4	187	1	#
5	244	8	#
5	245	6	#
5	248	5	#
6	282	8	ABCG2 (ATP-binding cassette, subfamily G, member 2) ³⁾ , LAP3 (leucine aminopeptidase 3) ²⁾ , LCORL (ligand dependend nuclear receptor corepressor-like; Teil) ²⁾
6	283	8	LCORL (ligand dependend nuclear receptor corepressor-like; Teil) ²⁾
6	284	10	KCNIP4 (Kv channel interacting protein 4) ¹⁾
6	294	12	-
6	295	11	FRYL (furry homolog-like) ²⁾ , SCFD2 (sec1 family domain containing 2; Teil) ²⁾
6	296	11	SCFD2 (sec1 family domain containing 2; Teil) ²⁾ , PDGFRA (platelet-derived growth factor receptor, alpha polypeptide) ²⁾ , KIT (v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog) ¹⁾
9	448	-	#
10	463	2	#
14	602	1	-
18	736	2	#
21	827	1*	-
22	848	1	#
26	941	-	#
28	965	-	TARBP1 (TAR (HIV-1) RNA binding protein 1) ¹⁾

¹⁾ STELLA *et al.*, 2010, ²⁾ FLORI *et al.*, 2009, ³⁾ QANBARI *et al.*, 2010a, *) suggestiver QTL, #) laut Supporting Information bei Stella *et al.* (2010) $P < 0,01$, aber keine Nennung des kausalen Gens

Es wird ersichtlich, dass insbesondere die beiden Bereiche auf BTA6 (F282-284 und F294-296) sehr aussichtsreich sind, da diese nicht nur eine Vielzahl an QTL aufweisen, sondern hier auch in verschiedenen genomweiten Studien einige der stärksten Signale gefunden und mögliche Kandidatengene deklariert wurden.

Um abschätzen zu können, ob es sich bei den jeweils drei Fenster umfassenden Bereichen um nur eine Selektionssignatur handelt, oder ob hier mehrere getrennte

Signaturen in unmittelbarer Nähe zueinander liegen sowie um zu beurteilen, inwiefern sich diese Signaturen mit der Lage der in anderen Studien gefundenen vermeintlich kausalen Gene überschneiden, wurden diese Regionen graphisch dargestellt (Abb. 23 und Abb. 24)

Zunächst erfolgt die genaue Betrachtung (Abb. 23) des Areals, das von den Fenstern 282 bis 284, welche BTA6 zugehörig sind (Tab. 8), bestimmt wird.

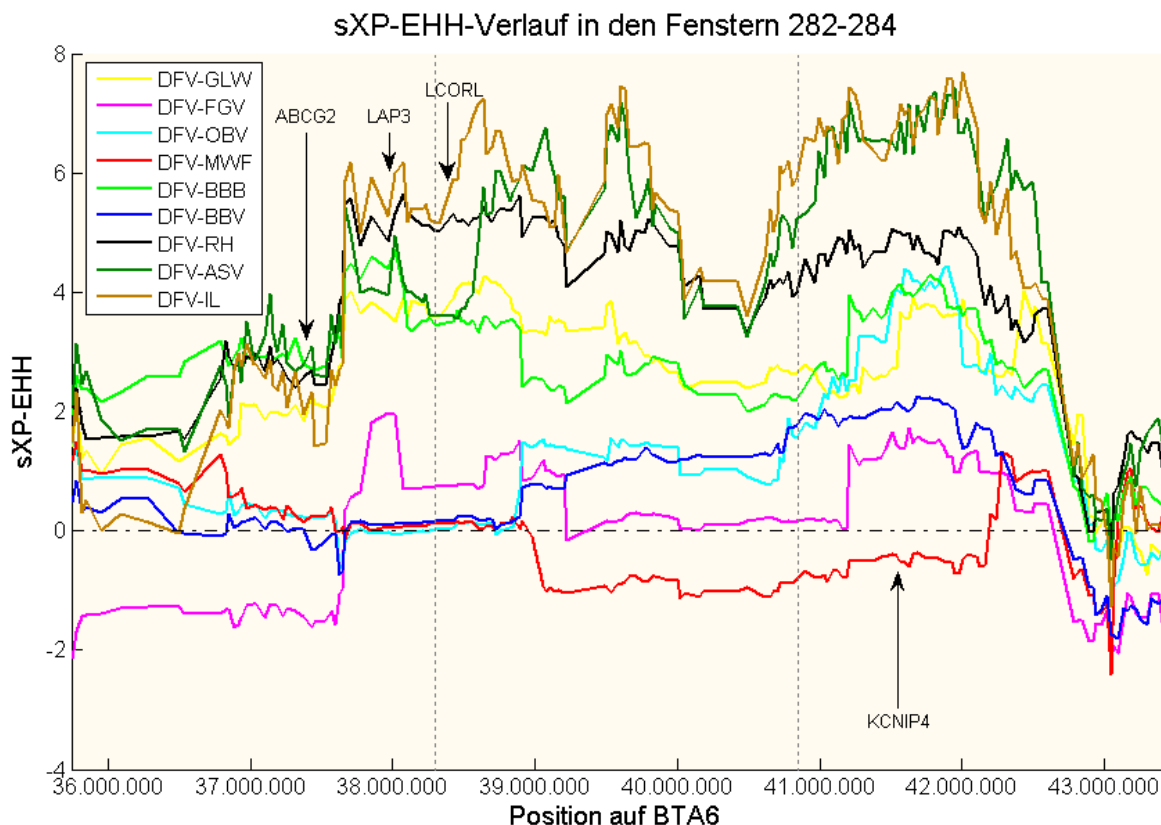


Abbildung 23: Signal der positiven Selektion in den Fenstern 282-284 (BTA6) der Rasse DFV im Vergleich mit neun weiteren Rinderrassen.

Verlauf der standardisierten *XP-EHH*-Werte (*sXP-EHH*; y-Achse) in Abhängigkeit von der physikalischen Position (bp) auf dem Rinderchromosom (BTA) 6 (x-Achse). Die Lage (Genmitte) möglicher Kandidatengene anderer Studien ist mit Pfeilen gekennzeichnet. Die grau gestrichelten vertikalen Linien markieren jeweils die Fenstergrenzen.

Aus Abbildung 23 wird ersichtlich, dass sich auf BTA6 eine deutliche Signatur über einen Bereich von etwa fünf Megabasen (~37,5 bis 42,6 Mb) erstreckt. Jedoch handelt es sich offensichtlich nicht um eine Fleckvieh-spezifische Selektionssignatur, da sich die *sXP-EHH*-Werte der Vergleiche mit den vier anderen Alpenrassen (FGV, OBV, MWF, BBV) insbesondere in dem Anfangsbereich der Signatur, in dem auch die Gene *LAP3* und *LCORL* liegen, kaum von Null unterscheiden. Dies bedeutet, dass sich bei diesen Rassen hier ähnlich lange und frequente Haplotypen befinden. Auch in den folgenden etwa

drei Megabasen erreichen die Werte dieser Rassenvergleiche größtenteils keine Signifikanz. Lediglich der Vergleich DFV-OBV weist im Bereich um *KCNIP4* signifikante *sXP-EHH*-Werte auf, während die Rasse Murnau-Werdenfelser an dieser Stelle, wie bereits die vorhergehenden mehr als zwei Megabasen, sogar negative Werte besitzt, hier also noch etwas stärker selektiert ist, als das DFV. Dies stimmt auch mit den Ergebnissen für diese Rassen überein, was sowohl der genomweiten Darstellung der *XP-EHH*-Ergebnisse (Abb. 22) als auch dem Anhang 9.4 zu entnehmen ist.

Während sich *ABCG2* in keinem stark auffälligen Bereich befindet, liegen sowohl *LAP3*, *LCORL* als auch *KCNIP4* in von Selektion deutlich gekennzeichneten Arealen, und könnten somit der Signatur – oder zumindest einem Teil davon kausal zugrunde liegen. Denn auch wenn die Signatur nicht deutlich unterbrochen ist, so scheint der Verlauf der *sXP-EHH*-Kurven doch darauf hinzuweisen, dass es sich hier um mindestens zwei eng beieinander liegende Selektionssignaturen handelt. Während *LAP3* und *LCORL* von einer Signatur umgeben sind, scheint eine andere *KCNIP4* einzuschließen. Da aber der Bereich um *LAP3* und *LCORL* bereits im Zuge eines ebenfalls am Lehrstuhl laufenden Braunviehprojekts für die Rasse OBV sequenziert wird, wurde für die Sequenzierung im DFV nicht die gleiche Region gewählt, so dass hier lediglich der Bereich um *KCNIP4* in Fenster 284 mit sechs von neun möglichen signifikanten *XP-EHH*-Werten (Tab. 7) in Frage käme, welcher fast zwei Megabasen umfasst.

Um eine endgültige Entscheidung über den zu sequenzierenden Bereich treffen zu können, wurde im Anschluss der *sXP-EHH*-Verlauf in den Fenstern 294 bis 296 (Abb. 24) genauer betrachtet.

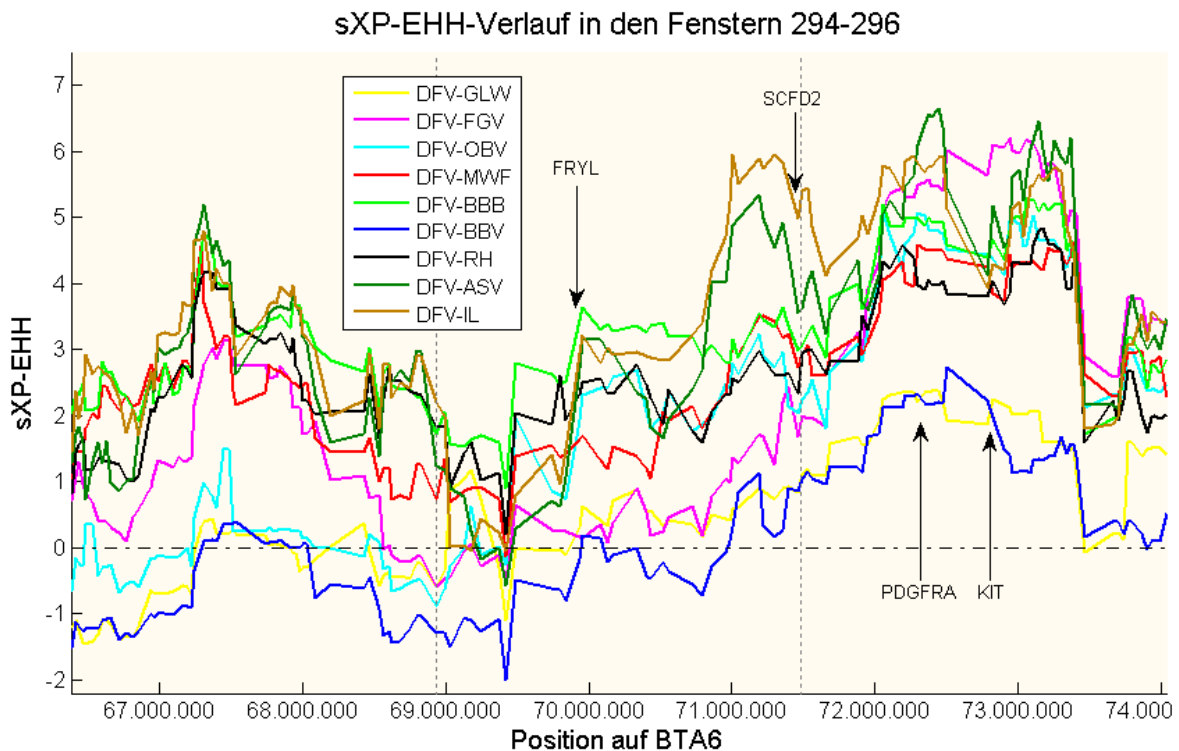


Abbildung 24: Signal der positiven Selektion in den Fenstern 294-296 (BTA6) der Rasse DFV im Vergleich mit neun weiteren Rinderrassen.

Verlauf der standardisierten *XP-EHH*-Werte (*sXP-EHH*; y-Achse) in Abhängigkeit von der physikalischen Position (bp) auf dem Rinderchromosom (BTA) 6 (x-Achse). Die Lage (Genmitte) möglicher Kandidatengene anderer Studien ist mit Pfeilen gekennzeichnet. Die grau gestrichelten vertikalen Linien markieren jeweils die Fenstergrenzen.

In Abbildung 24 wird auf den ersten Blick ersichtlich, dass es sich hier um zwei deutlich getrennte Selektionssignaturen handelt, da alle *XP-EHH*-Vergleiche im Bereich um 69,5 Mb auf Werte um Null liegen, respektive unter Null fallen. Dies bedeutet, dass die Haplotypen im DFV hier nicht länger und frequenter sind, als in den anderen Rassen, also keine Selektion vorliegt. Da in der ersten Signatur zum Einen kein potenzielles kausales Gen liegt, welches bereits in anderen Studien gefunden worden wäre, und zum Anderen Vergleiche mit den Rassen GLW, OBV und BBV nur schwachpositive oder sogar negative *sXP-EHH*-Werte ergeben, entschloss man sich trotz signifikanten *iHS*-Wertes im zugehörigen Fenster 294 gegen die Sequenzierung dieses Bereichs.

Das Gen *FRYL* liegt im Gegensatz zu den anderen möglichen kausalen Genen anderer Studien (*SCFD2*, *PDGFRA* und *KIT*; Tab. 8) in einem eher unauffälligen Areal. Daher scheint der Bereich um *FRYL* (~ 68,6 - 70,6 Mb) für die Sequenzierung ebenfalls als weniger erfolgversprechend.

Um *SCFD2* (Übergang von F295 zu F296) besteht zwar eine erkennbare Selektionssignatur, die sich aber vor allem auf die Vergleiche mit den

unselektierten Rassen ASV und IL zurückführen lässt. Von den sieben übrigen *XP-EHH*-Vergleichen erreicht in Fenster 295 ansonsten nur der DFV-BBB-Vergleich einen signifikanten *eP*-Wert, was Tabelle 7 zu entnehmen ist.

In der Signatur um *PDGFRA* und *KIT* hingegen liegt bei allen neun Rasse-Vergleichen die Selektion auf Seiten des Deutschen Fleckviehs, da hier alle *sXP-EHH*-Werte positiv sind. Acht Vergleiche erreichen dabei im zugehörigen Fenster 296 signifikante Werte (Tab. 7). Lediglich der *XP-EHH*-Vergleich DFV-GLW ist hier mit einem *eP*-Wert von 0,071 nicht signifikant.

Zusammenfassend ist die Signatur in Fenster 296 (~ 71,5 - 74,0 Mb) für eine anschließende Sequenzierung vielversprechender als jene in Fenster 284 (Abb. 23). Zum Einen sind hier acht, statt nur sechs *XP-EHH*-Vergleiche signifikant (Tab. 7), zum Anderen ist der Verlauf der einzelnen Kurven insgesamt kongruenter. Darüber hinaus liegen in der für die Sequenzierung ausgewählten Signatur gleich zwei potentielle kausale Gene anderer Studien, nämlich *PDGFRA* und *KIT*, während dies in Fenster 284 nur auf das Gen *KCNIP4* zutrifft.

Nachdem der Bereich der sequenziert werden sollte in etwa feststand, wurde nun in den Fleckvieh-Haplotypen nach einem langen und hochfrequenten Haplotypen gesucht, der in den anderen Rassen in dieser Form nicht anzutreffen war. Im Anschluss wurden vier repräsentative Fleckviehtiere ausgewählt, die die Fallgruppe bildeten. Als Kontrolle wählte man vier Tiere der Rasse RH. Dabei wurde darauf geachtet, dass diese sich möglichst stark von dem einheitlichen Haplotypen der Fallgruppe unterschieden und zugleich die Haplotypenverteilung des RH gut repräsentierten. Tabelle 9 zeigt einen Ausschnitt der Haplotypenmatrix, in welchem der für Fleckviehtiere typische Haplotyp in rot dargestellt wurde. Außerdem wurden in diesem Chromosomenabschnitt verschiedene Marker farblich gekennzeichnet. Dabei stellen die gelb unterlegten Marker die Grenzen des anschließend sequenzierten Bereichs dar. Darüber hinaus wurden Marker, welche sich jeweils in den zwei nächstgelegenen Genen distal (*KDR*, *SRD5A3*) sowie proximal (*PDGFRA*, *KIT*) des für das DFV typischen Haplotyps befanden, ebenfalls farbig (grau, blaugrün, blau, grün) gekennzeichnet.

Tabelle 9: Ausschnitt der Haplotypen-Matrix um den für das DFV typischen Chromosomenabschnitt.

Dargestellt ist ein Ausschnitt der Haplotypen-Matrix auf dem Rinderchromosom (BTA) 6. Dabei sind die Haplotypen sowohl nach der jeweiligen Zugehörigkeit (Gr.) zu Fall- (FALL) oder Kontrollgruppe (KONTROLLE) als auch nach Labornummern (Lab-Nr.) getrennt. Die aufgeführten Marker decken den Bereich von Position 71.189.705 bis 73.906.568 ab, wobei „0“ und „1“ jeweils für das „derived“ respektive „ancestral“ Allel steht. Der für das DFV typische Haplotyp wurde in roter Schrift abgebildet. Die gelb unterlegten Marker stellen den ersten und letzten Marker des Bereichs dar, der nachfolgend sequenziert wurde. Ebenfalls farbig unterlegt sind Marker, welche sich in den zwei nächstgelegenen Genen proximal (blau: PDGFRA; grün: KIT) sowie distal (grau: KDR; blaugrün: SRD5A3) des allen Haplotypen der Fallgruppe gemeinsamen Abschnitts befinden.

Gr.	Lab-Nr.	Teil der Haplotypen-Matrix von BTA6 (71.189.705- 73.906.568)
FALL (DFV)	020000	101100110011101111111011101010111001111111001111111111111111110
	020000	101100110011101111111011101010111001111111100111110111111111
	080000	10001111111101111111101110101011100111111100111111111111111110
	080000	10010111111101111111101110101011100111111100111111111111111110
	270000	10110011001110111111101110101011100111111100111111111111111110
	270000	1000111111111001111110111010101110011111110011111011111111111
	FV0239	1011001100111011111110001001111001111110111110101101101
	FV0239	1101111111111011111110111010101110011111110011111111111111110
KONTROLLE (RH)	RH106	1001111101111100111011110101011111100001011111111111100011
	RH106	011101111111111111111000100111100111101101100111111111110
	RH117	011101111111110111111000100111001110011100111111100011
	RH117	1001011101100101111110100111111110000011110110101111
	RH123	01110100101110101110111101010111111000000110110011100011
	RH123	0111011111111101111110001001110011100111001111111100011
	RH124	01110111111110110011111101001110011111100111110011101110011
	RH124	01001101101110111111100010111100110001011111110111101111

Tabelle 9 macht deutlich, dass im DFV in diesem Bereich nur ein Haplotyp vorherrschend ist (rot). Außerdem zeigt sie, dass drei der gewählten Tiere der Fallgruppe einen identischen 31 Marker umfassenden Haplotypen aufwiesen. Als viertes DFV (FV0239) wurde jedoch eines der wenigen Tiere gewählt, bei dem ein Chromosomen lediglich in einem deutlich kürzeren Teil (neun Marker) mit dem typischen Haplotypen übereinstimmte.

In der Kontrollgruppe hingegen sind nur vereinzelt unregelmäßige Teile dieses Haplotyps zu finden. Betrachtet man im speziellen die Marker, die im Bereich des *KIT*- respektive des *PDGFRA*-Gens (grün respektive blau unterlegt) liegen, so stimmen sieben und acht der insgesamt acht DFV-Haplotypen in diesen Allelen überein. In der Kontrollgruppe hingegen tragen nur drei beziehungsweise vier Haplotypen die im Fleckvieh häufige Markervariante. Ähnliches gilt für die Gene *KDR* und *SRD5A3*, welche bisher in keiner Selektionssignaturstudie auffällig waren. So tragen hier jeweils sieben DFV-Haplotypen das gleiche Allel, während dies für vier (*KDR*) respektive fünf (*SRD5A3*) RH-Chromosomen zutrifft.

Da die Genauigkeit von Sequenzierungen mit abnehmender Länge der zu sequenzierenden Region steigt, wurde diese auf den Bereich von Marker *BFGL-NGS-117708* (71.949.142) bis Marker *BTA-76807-no-rs* (73.469.085), beide in Tabelle 9 gelb unterlegt, eingeschränkt. Diese 2.088.236 Basen umfassende Region beinhaltet nicht nur die neun Marker, die allen DFV-Tieren der Fallgruppe gemein sind, sondern auch die insgesamt 31 Marker in denen immerhin sieben der acht Fall-Haplotypen übereinstimmen. Die Grenze wurde dort gesetzt, wo der gemeinsame Haplotyp nur noch in 50 Prozent der DFV-Tiere vorhanden war, so dass der Bereich im Ganzen 34 Marker umfasste (Tabelle 9).

In diesem Abschnitt des Genoms finden sich sechs Gene komplett wieder, sowie zusätzlich ein Teil eines Pseudogens am Randbereich. Tabelle 10 zeigt eine Übersicht über diese sieben Gene, inklusive ihrer jeweiligen Positionen.

Tabelle 10: Die sieben Gene im zu sequenzierenden Bereich auf BTA6

Aufgeführt sind sowohl Gensymbol, Genbeschreibung, physikalische Position in bp (Position) als auch Gentyp aller sieben Gene, die sich im Bereich 71.949.142 - 73.469.085 bp auf Rinderchromosom (BTA) 6 befinden, welcher im Anschluss sequenziert wurde.

Gensymbol	Genbeschreibung	Position	Gentyp
CHIC2	cysteine-rich hydrophobic domain 2	72.135.154-72.162.834	Protein-codierend
PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	72.298.965-72.343.408	Protein-codierend
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	72.779.297-72.826.422	Protein-codierend
KDR	kinase insert domain receptor (type III receptor tyrosine kinase)	73.148.811-73.193.668	Protein-codierend
SRD5A3	steroid 5 alpha-reductase 3	73.384.071-73.399.751	Protein-codierend
TMEM165	transmembrane protein 165	73.420.082-73.446.355	Protein-codierend
LOC540334	clock	73.455.785-73.482.485	Pseudogen

Die Reihenfolge der Marker basiert zwar auf der weniger aktuellen Version *Btau_4.0* des *Bos taurus* Genoms (<http://genome.ucsc.edu/cgi-bin/hgGateway>), jedoch stimmt sie in der kritischen Region mit der aktuelleren *Bos taurus* Version *UMD3.1* überein (http://www.cbcb.umd.edu/research/bos_taurus_assembly.shtml), so dass zu erwarten ist, dass die Signatur bei Verwendung des neueren Markersets ebenfalls entsprechend deutlich ausfallen würde.

Paired-End-Sequenzierung:

Die Paired-End-Sequenzierung des Bereichs 71.949.142 bis 73.469.085 auf BTA6 der acht ausgewählten Tiere wurde am Laboratory for Functional Genome Analysis (LAFUGA) mittels Genome Analyzer IIx (Illumina, San Diego, USA) durchgeführt (siehe 3.2.7.1.2) und die Sequenzabgleiche sowie die benötigten Auswertemodule, auf einem Webserver unter *GALAXY* zur weiteren Analyse zur Verfügung gestellt.

Die Betrachtung der Abgleiche mit dem *Integrative Genomics Viewer (IGV)* bestätigte eine ausgedehnte, weitestgehend mit der Referenzsequenz übereinstimmende Homozygotie im DFV. Besagte Referenzsequenz entspricht dabei einem Individuum der Rasse Hereford, das vom Bovine Genome Sequencing and Analysis Consortium vollständig sequenziert wurde (ELSIK *et al.*, 2009; <http://www.hgsc.bcm.tmc.edu/project-species-m-Bovine.hgsc?pageLocation=Bovine>). Abbildung 25 zeigt beispielhaft eine 16 Kilobasen umfassende Ansicht mittels *IGV*. Dargestellt ist der Übergang des allen DFV-Tieren gemeinsamen Haplotyps zu den Markern, welche im *KDR*-Gen liegen (siehe Tab. 9). Dabei werden in der dargestellten Vergrößerung Abschnitte, die homolog zur Referenzsequenz sind grau dargestellt. Das Vorhandensein heterozygoter SNPs sowie Homozygoter, die sich von der Referenzsequenz unterscheiden, ist hingegen durch eine den Basen entsprechende farbliche Kennzeichnung im Bereich der sogenannten „Coverage“-Kurve markiert. Die „Coverage“ stellt dabei ein Maß für die durchschnittliche Anzahl der sequenzierten Fragmente („reads“) dar, die sich laut Abgleich in diesem Bereich überlappen. Sie ist daher ein Maß der Qualität der jeweiligen Sequenzierungsergebnisse. Manche Genombereiche, insbesondere repetitive Sequenzen, erlauben jedoch keine sichere Zuordnung, so dass die „Coverage“ stark sinkt – unter Umständen sogar auf Null. In solchen Bereichen sind dann nur wenige oder keine „reads“ im Abgleich zu finden.

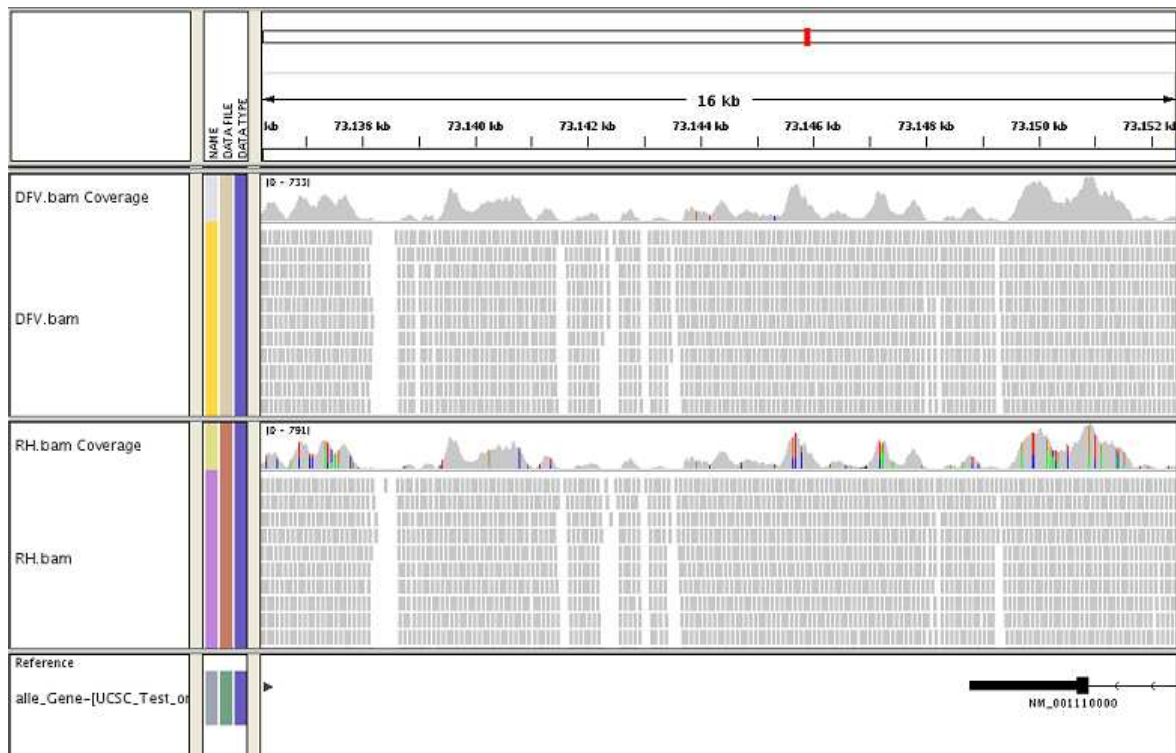


Abbildung 25: IGV-Ansicht des Sequenzabgleichs im Übergang des im DFV homozygoten Bereichs zum *KDR*-Gen.

Mittels „*Integrative Genomics Viewer*“ (*IGV*) dargestellt ist ein 16 kb umfassender Bereich an der Grenze des *KDR*-Gens (NM_001110000). Sowohl für die Fallgruppe (DFV) als auch die Kontrollgruppe (RH) ist die „*Coverage*“-Kurve, die ein Maß für die durchschnittliche Anzahl der sequenzierten DNA-Fragmente („*reads*“) darstellt, abgebildet. Ein Teil der „*reads*“ ist jeweils unterhalb dieser Kurven abgebildet. Übereinstimmungen mit der Referenzsequenz sind grau, Abweichungen dagegen farbige dargestellt. Die Farbverteilung entspricht dabei der Frequenz der Basen, wobei grün für Adenin, braun für Guanin, rot für Thymin und blau für Cytosin steht.

Dank der farblichen Kennzeichnung im Bereich der „*Coverage*“-Kurve ist klar ersichtlich, dass das DFV im Gegensatz zum RH in dem dargestellten Bereich (Abb. 25) eine ausgedehnte Haplotypenhomozygotie zeigt, was die Ergebnisse der *XP-EHH*-Berechnungen bestätigt.

Da eine alleinige Betrachtung der Sequenzabgleiche mittels *IGV* nicht ausreichend ist, um mögliche kausale Mutationen, welche dieser Selektionssignatur zugrunde liegen könnten, zu identifizieren, wurden die Daten durch die vom LAFUGA auf *GALAXY* zur Verfügung gestellten Module genauer ausgewertet. Diese Module machten es möglich durch verschiedene Filterschritte eine Vielzahl an Mutationen ausfindig zu machen. Dabei musste, um diese Filter korrekt zu setzen, folgendes beachtet werden:

Die *XP-EHH* unterscheidet nicht zwischen Homozygotie entsprechend oder abweichend der Referenzsequenz, sondern lediglich zwischen langen homozygoten Bereichen der Fall-Rasse und von Heterozygotie unterbrochenen

Bereichen der Kontroll-Rasse und *vice versa*. Das bedeutet, um einen Kurvenverlauf wie in Abbildung 24 dargestellt zu bekommen, bei dem die Selektion deutlich auf Seiten des Fleckviehs liegt, muss das DFV entweder entsprechend oder abweichend zur Referenzsequenz homozygot sein, während die Homozygotie im selben Bereich beim RH durch das Vorkommen von heterozygoten Allelen unterbrochen werden muss.

Um diesen Umstand gerecht zu werden, wurde der entsprechende Filter sowohl für die Ermittlung von SNPs als auch von INDELS (Insertionen und Deletionen) jeweils folgendermaßen gesetzt:

1. Filter: RH(30-70%) ungleich DFV(>30%) sowie
2. Filter: DFV(>70%) gleich RH(30-70%)

Dabei entsprechen die Angaben in Klammern der prozentualen Abweichung der „reads“ der jeweiligen gepoolten Rasse von der Referenzsequenz (*RefSeq*). Gibt es also an einer bestimmten Position für das Fleckvieh beispielsweise 100 „reads“ und weichen 40 davon von der *RefSeq* ab, dann entspricht das einer Frequenz von 40 %. Es wird angenommen, dass bei Frequenzen unter 30 % Homozygotie in Übereinstimmung mit der *RefSeq* sowie bei einer Frequenz größer 70 % abweichend von der *RefSeq* vorhanden ist. Folglich wird der Bereich von 30 bis 70 Prozent als heterozygot angesehen.

Der erste Filter gibt daher alle Mutationen zurück, in denen das RH heterozygot und das DFV gleichzeitig homozygot entsprechend der Referenzsequenz ist. Der zweite Filter gibt jene Varianten wieder, bei denen das DFV homozygot entgegengesetzt zur Referenzsequenz ist, während das RH wiederum heterozygot ist. Die etwas umständliche Formulierung dieser Filter beruht darauf, dass es mit den zur Verfügung gestellten Modulen nicht möglich ist, mit der Referenzsequenz übereinstimmende Basen auszulesen. Dies bringt allerdings den unumgänglichen Nachteil mit sich, dass durch den ersten Filter bei dem gepoolten DFV nicht nur Positionen berücksichtigt werden, deren Abweichung von der *RefSeq* kleiner ist als 30%. Es werden nämlich auch solche aufgeführt, für die es im DFV keine „reads“ gibt, sofern im RH „reads“ mit einer Frequenz von 30 bis 70 Prozent vorkommen. Zwar ist dieser Nachteil nicht unmittelbar zu kompensieren, jedoch würde bei einer Betrachtung dieses Bereichs mittels *IGV* das Fehlen von „reads“ im DFV sowie die entsprechend schlechte „Coverage“ auffallen, so dass diese Varianten von anschließenden Untersuchungen ausgeschlossen werden könnten.

Es ist anzunehmen, dass dieser Umstand, den der erste Filter birgt, auch ein Grund ist, weshalb die Anzahl der Varianten, die durch den ersten Filter gefunden wurden, sowohl für SNP- als auch für INDEL-Analysen deutlich über denen des zweiten Filters liegt, wie Tabelle 11 zu entnehmen ist.

Die Lage der so identifizierten Varianten bezüglich der Gene, sowie ihre Effekte auf die Transkription wurden dann mit dem online zugänglichen Programm „*Variant Effect Predictor*“ von „*Ensembl*“ bestimmt (<http://www.ensembl.org/tools.html>) und sind ebenfalls in Tabelle 11 dargestellt.

Tabelle 11: Ermittelte Varianten und ihre Auswirkungen

Dargestellt wurde die jeweilige Anzahl gefundener Mutationen, welche den jeweiligen Filterkriterien (Filter) entsprachen. Dabei wurde diese Auflistung nach der Art der Mutation (Mut) in Basenaustausch (SNP) sowie Insertion respektive Deletion (INDEL) unterteilt. Außerdem erfolgte eine Trennung entsprechend der durch den „*Variant Effect Predictor*“ (<http://www.ensembl.org/tools.html>) vorhergesagten Auswirkungen der einzelnen Mutationen in „*Intergenic*“ (Inter), „*Intronic*“ (Intro), „*Downstream*“ (Down), „*Upstream*“ (Up), „*Synonymous-Coding*“ (Syn), „*3prime UTR*“ (3'U), „*Splice Site*“ (Sp.S), „*Essential Splice Site*“ (E.S), „*Nonsynonymous-Coding*“ (Non) und „*Stop Gained*“ (St). Zusätzlich wurde die jeweilige Gesamtanzahl (Ges) der Varianten angegeben.

Mut	Filter	Auswirkung										
		Inter	Intro	Down	Up	Syn	3'U	Sp.S	E.S	Non	St	Ges
SNP	RH(30-70%)≠DFV>30%	290	93	27	14	4	12	1	3	2	1	447
	DFV>70%=RH(30-70%)	168	33	9	10	-	1	-	-	2	-	223
INDEL	RH(30-70%)≠DFV>30%	36	17	2	3	-	1	-	-	-	-	59
	DFV>70%=RH(30-70%)	15	4	-	-	-	-	-	-	-	-	19

Die genauen Bedeutungen der verschiedenen durch den „*Variant Effect Predictor*“ ermittelten Auswirkungen sind Tabelle 12 zu entnehmen.

Tabelle 12: Erklärung der durch den *Variant Effect Predictor* ermittelten Auswirkungen der Varianten

Es wird die genaue Bedeutung der durch den „*Variant Effect Predictor*“ (<http://www.ensembl.org/tools.html>) vorhergesagten Auswirkungen beschrieben, für welche auch die in Tabelle 11 verwendeten Abkürzungen (Abk.) aufgeführt sind.

Abk.	Auswirkung	Bedeutung
Inter	Intergenic	mehr als 5 kb distal oder proximal eines Transkripts
Intro	Intronic	im Intron liegend
Down	Downstream	innerhalb 5 kb distal des 3' Endes eines Transkripts
Up	Upstream	innerhalb 5 kb proximal des 5' Endes eines Transkripts
Syn	Synonymous SNP	in codierender Sequenz, keine Veränderung der Aminosäure (stille Mutation)
3'U	3prime UTR	in der 3' Untranslated Region (UTR) liegend
Sp.S	Splice Site	ein bis drei Basenpaare innerhalb eines Exons oder drei bis acht Basenpaare innerhalb eines Introns
E.S	Essential Splice Site	in den ersten zwei oder letzten zwei Basenpaaren eines Introns
Non	Non-synonymous SNP	in codierender Sequenz, Veränderung der Aminosäure der codierten Peptidsequenz
St	Stop gained	in codierender Sequenz, Zugewinn eines Stopp-Codons, Resultat z. B. verkürzte Peptidsequenz

Zwar kann nicht ausgeschlossen werden, dass auch Varianten, die in uncodierten Bereichen der Kategorien Intergenic, Intronic, Up- oder Downstream liegen, eine bedeutsame Wirkung auf eine etwaige Genfunktion besitzen, allerdings ist dies deutlich unwahrscheinlicher, so dass man sich ähnlich Owczarek-Lipska *et al.* (2011) aufgrund der Vielzahl der gefundenen Varianten im Folgenden dazu entschlossen hat zunächst diese vier Kategorien, sowie die ohnehin biologisch unbedeutsamen Stillen Mutationen (Syn) außer Acht zu lassen. Tabelle 13 zeigt eine nach Mutation und Filter geordnete Übersicht über die Positionen der 23 verbleibenden Varianten.

Tabelle 13: Die 23 Varianten im die Codierung betreffenden sowie im codierenden Bereich

Dargestellt wurden alle 23 Varianten, die in den regulierenden oder codierenden Bereichen gefunden wurden, sofern sie den jeweiligen Filterkriterien (Filter) entsprachen. Dabei wurde diese Auflistung nach der Art der Mutation (Mut), also in Basenaustausch (SNP) sowie Insertion respektive Deletion (INDEL), unterteilt. Für jede Variante ist die physikalische Position in bp, das betroffene Gen, die vom „Variant Effect Predictor“ (<http://www.ensembl.org/tools.html>) vorhergesagte Auswirkung und, sofern vorhanden, der entsprechende Aminosäureaustausch (AS-A.) angegeben. „N“ steht dabei für Asparagin, „K“ für Lysin, „E“ für Glutamat, „I“ für Isoleucin, „L“ für Leucin, „F“ für Phenylalanin, „H“ für Histidin, „Y“ für Tyrosin, „*“ für ein zusätzliches Stopp-Codon und „-“ für das Fehlen eines AS-A.

	Mut.	Filter	Position	Gen	Auswirkung	AS-A.	
INDEL		RH(30-70%) ≠ DFV>30%	73149227	KDR	3PRIME_UTR	-	
			72162485	CHIC2	ESSENTIAL_SPLICE_SITE	-	
SNP			72162486	CHIC2	ESSENTIAL_SPLICE_SITE	-	
			73148855	KDR	3PRIME_UTR	-	
			73148967	KDR	3PRIME_UTR	-	
			73149229	KDR	3PRIME_UTR	-	
			73149730	KDR	3PRIME_UTR	-	
			73149933	KDR	3PRIME_UTR	-	
			73149945	KDR	3PRIME_UTR	-	
			73150188	KDR	3PRIME_UTR	-	
		RH(30-70%) ≠ DFV>30%		73150255	KDR	3PRIME_UTR	-
				73150282	KDR	3PRIME_UTR	-
				73150337	KDR	3PRIME_UTR	-
				73150377	KDR	3PRIME_UTR	-
				73150555	KDR	3PRIME_UTR	-
				73150932	KDR	SPLICE_SITE	-
				73175972	KDR	NON_SYNONYMOUS_CODING	N/K
				73175983	KDR	STOP_GAINED	E/*
				73175995	KDR	NON_SYNONYMOUS_CODING	I/L
				73392405	SRD5A3	ESSENTIAL_SPLICE_SITE	-
				72136737	CHIC2	3PRIME_UTR	-
		DFV>70% = RH(30-70%)		73392306	SRD5A3	NON_SYNONYMOUS_CODING	F/L
			73392355	SRD5A3	NON_SYNONYMOUS_CODING	H/Y	

Hier wird klar ersichtlich, dass die Mutationen lediglich in drei der sechs Protein-codierenden Gene liegen, die sich in der sequenzierten Region befinden - nämlich in *CHIC2*, *KDR* und *SRD5A3*. Interessanterweise ist unter diesen drei Genen keines der bereits in anderen Studien deklarierten Kandidatengene (*PDGFRA*, *KIT*).

Mit Abstand die meisten Varianten befinden sich im *KDR*-Gen. Dabei sind bei 14 der insgesamt 17 Mutationen in diesem Gen wichtige nicht-codierende Bereiche (13mal 3'U, einmal Sp.S) betroffen, was allerdings eine Aussage über eine Veränderung der Genfunktion respektive des -produkts nicht zulässt. Bei den übrigen drei Varianten (zweimal Non, einmal St) hingegen, die sich in einem Areal befinden, das nur 24 Basen umfasst, erfolgt jeweils eine Änderung der codierten Aminosäure. So bewirkt der Basenwechsel von Adenin auf Guanin an Position 73.175.972 einen Wechsel von Asparagin zu Lysin, während der Ersatz von Cytosin durch Thymin an Position 73.175.983 dazu führt, dass anstelle der Aminosäure Glutamat ein verfrühtes Stopp-Codon gesetzt wird. Wenige Basen weiter (73.175.995) erfolgt dann ein Austausch von Thymin durch Cytosin, mit der Konsequenz, dass Isoleucin hier zu Leucin wird.

Die drei Varianten, die in *CHIC2* liegen, betreffen wiederum Bereiche, die keine Aussage über Auswirkungen auf Genfunktion oder -produkt zulassen (zweimal E.S, einmal 3'U). Dies gilt auch für eine Mutation in *SRD5A3* (73.392.405), während die zwei übrigen in diesem Gen wieder einen Aminosäurewechsel zur Folge haben. So wird aufgrund des Ersatzes von Cytosin durch Thymin an Position 73.392.306 aus Phenylalanin Lysin und an Position 73.392.355 aus Histidin Tyrosin, was in einem Wechsel von Cytosin auf Adenin begründet ist.

Abbildung 26 zeigt einen Ausschnitt von BTA6 (72.000.000-73.775.000), dem die Lage der Gene zueinander sowie die Positionen der Varianten zu entnehmen sind.

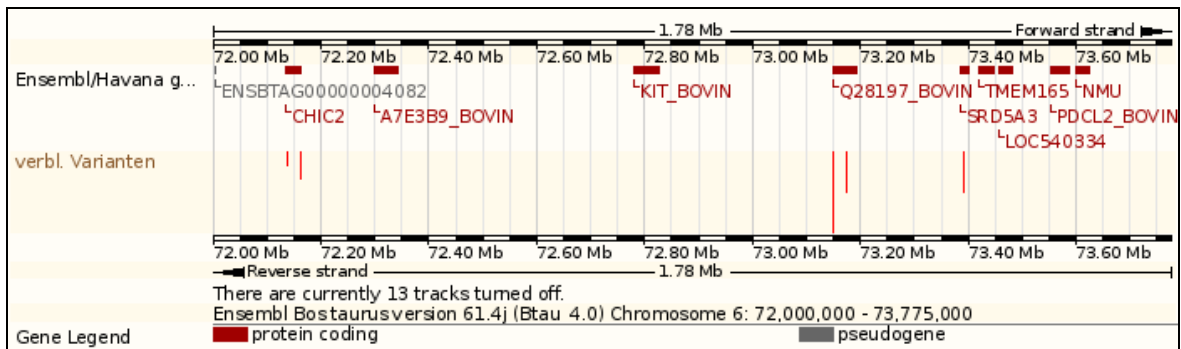


Abbildung 26: Ausschnitt von BTA6 mit Lage der Gene sowie der Varianten

Gezeigt ist ein mittels Genome-Browser von „Ensembl“ (<http://www.ensembl.org/index.html>) dargestellter 1,78 Mb umfassender Abschnitt des Rinderchromosoms (BTA) 6. Der Abbildung zu entnehmen sind sowohl die Position der hier liegenden Gene als auch die Lage der 23 Varianten (vertikale rote Linien, Länge abhängig von Anzahl der Varianten [1, 2, 3, 14]), welche codierende oder regulierende Bereiche betreffen. Das Gen *KDR* entspricht „Q28197_Bovine“ und *PDGFRA* „A7E3B9_Bovine“.

Die Abbildung verdeutlicht erneut, dass sich die 23 Varianten nur auf die Gene *CHIC2*, *KDR* und *SRD5A3* verteilen. Des Weiteren wird klar, dass sich innerhalb der Gene *CHIC2* und *KDR* jeweils zwei deutlich voneinander getrennte Regionen, welche von den Varianten betroffen sind, ausmachen lassen. Diese entsprechen bei *KDR* den codierenden und uncodierenden Bereichen, in denen die Varianten liegen.

Es galt nun die Frage zu klären, inwieweit diese Regionen mit dem Verlauf der *XP-EHH*-Kurve übereinstimmen. Betrachtet man dazu Abbildung 27, in welcher der Vergleich zwischen DFV und RH schwarz dargestellt ist, so stellt man fest, dass die Varianten im *KDR* mit dem höchsten Peak der Kurve zusammenfallen, was insbesondere für die codierenden Mutationen gilt. Auch bei *SRD5A3* stimmt die Lage der Varianten mit einem lokalen Peak überein. Lediglich die Mutationen in *CHIC2* stimmen mit keinem Maximum überein.

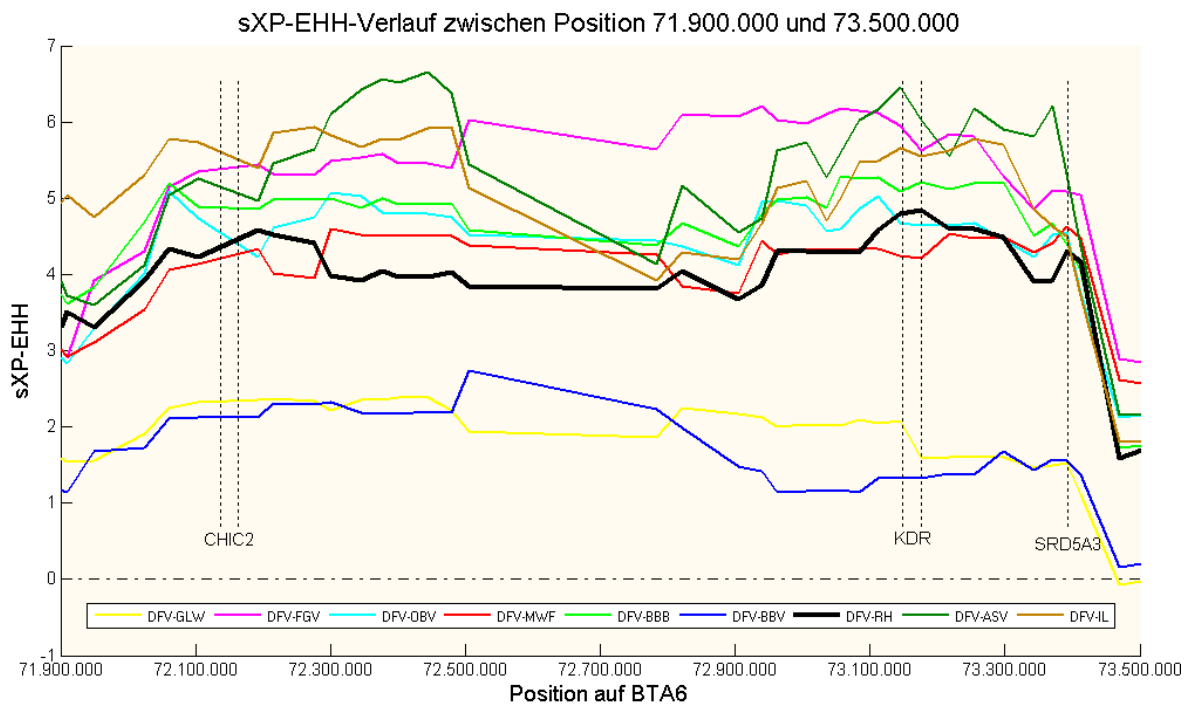


Abbildung 27: Lage der Varianten bezüglich dem Signal der positiven Selektion auf BTA6 im Bereich 71.900.000-73.500.000

Verlauf der standardisierten *XP-EHH*-Werte (*sXP-EHH*; y-Achse) in Abhängigkeit der physikalischen Position (bp) auf Rinderchromosom (BTA) 6 (x-Achse). Die Lage der Varianten ist mit vertikalen gepunkteten Linien gekennzeichnet und mit dem zugehörigen Gensymbol versehen. Der Vergleich von Fall- und Kontrollgruppe (DFV-RH), auf welchem das Auffinden der Varianten beruhte, ist durch eine breitere Linie gekennzeichnet.

Dies gemeinsam mit der Tatsache, dass sich der bei allen Fall-Tieren übereinstimmende Bereich, der in Tabelle 9 dargestellt wurde, zwischen den Genen *KIT* und *KDR* aufspannt, macht die Unterschiede von DFV zu RH im *KDR* zur mutmaßlichen Ursache der untersuchten Selektionssignatur. Insbesondere die drei Mutationen, die einen Aminosäurewechsel verursachen (Tab. 13), und deren Position mit dem Peak des *sXP-EHH*-Verlaufs des DFV-RH-Vergleichs (Abb. 27) übereinstimmt, sind als mögliche kausale Varianten zu betrachten.

Eine Betrachtung der Region um diese drei Varianten, die im codierenden Bereich des möglichen Kausalgens *KDR* gefunden wurden, ist in der folgenden Abbildung 28 dargestellt.

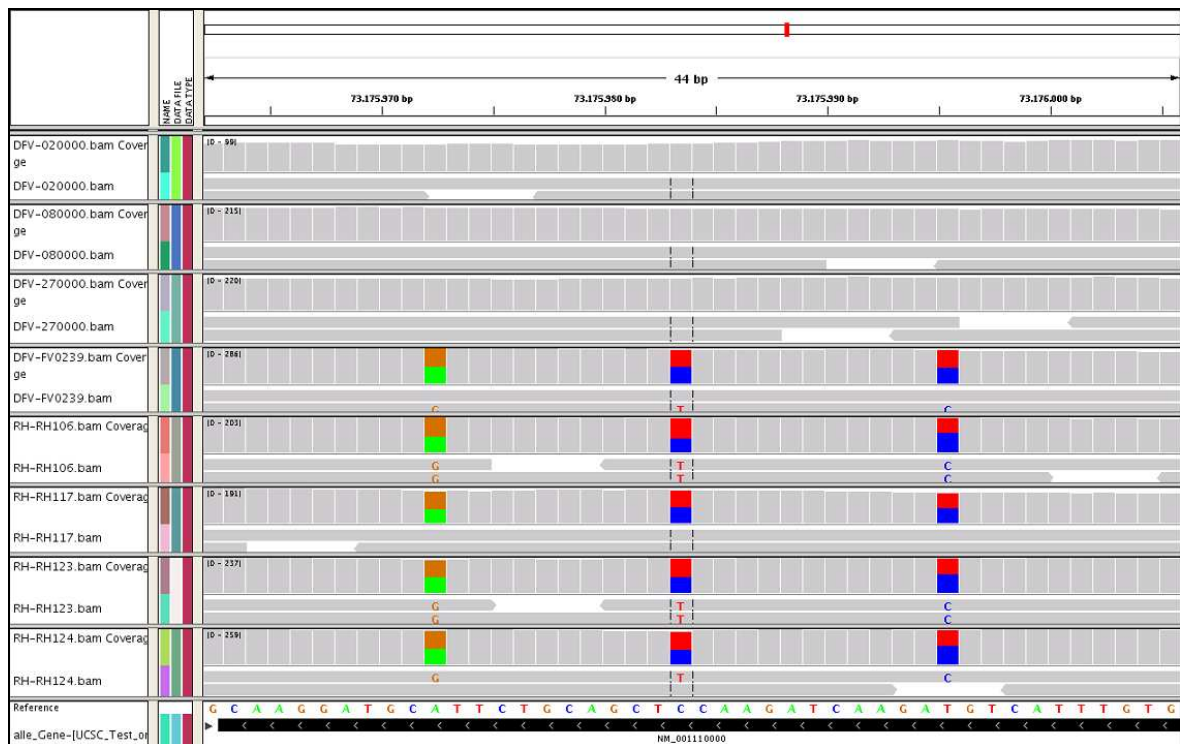


Abbildung 28: IGV-Ansicht der drei Varianten im codierenden Bereich des *KDR*-Gens

Mittels „*Integrative Genomics Viewer*“ (IGV) dargestellt ist ein 44 bp umfassender Bereich des Exons 13 des *KDR*-Gens (NM_001110000), in welchem die drei möglichen kausalen Mutationen liegen. Sowohl für die einzelnen Tiere der Fallgruppe (DFV) als auch die der Kontrollgruppe (RH) ist die „Coverage“, die ein Maß für die durchschnittliche Anzahl der sequenzierten DNA-Fragmente („reads“) darstellt, abgebildet. Unterhalb der „Coverage“ sind jeweils zwei Zeilen der „reads“ zu erkennen. Übereinstimmungen mit der am unteren Bildrand farbig dargestellten Referenzsequenz sind grau, Abweichungen dagegen farbig dargestellt. Die Farbverteilung (grün: Adenin, braun: Guanin, rot: Thymin, blau: Cytosin) der „Coverage“ lässt erkennen, dass alle Kontrolltiere sowie FV0239 heterozygot, die übrigen Tiere der Fallgruppe hingegen homozygot sind.

Für diese drei Varianten, welche sich alle in Exon 13 des insgesamt 30 Exons umfassenden Gens *KDR* befinden, ist jedes der vier RH heterozygot. Dies trifft ebenso auf das DFV mit der Labornummer FV0239 zu, welches bereits in der Haplotypen-Matrix in Tabelle 9 durch einen im Vergleich zu den anderen DFV-Tieren deutlich kürzeren Haplotypen auffiel.

DNA-Sequenzierung:

Für einen Teilbereich um diese drei möglichen kausalen Varianten erfolgte anschließend eine DNA-Sequenzierung. Durch diese wurde die genaue Basenabfolge des Teilbereichs ermittelt. Dies diente dazu die Ergebnisse der Paired-End-Sequenzierung zu validieren, da bekannt ist, dass die modernen

Hochdurchsatzverfahren der sogenannten „nächsten Generation“, denen auch die SBS-Verfahren von Illumina (San Diego, USA) angehören, weniger genau sind, als automatische Sequenzierungsmethoden, welche auf der Methode nach Sanger basieren (METZKER, 2010).

Als erster Schritt erfolgte die Erstellung und Auswahl der geeignetsten Primer mittels *Primer3 Input 0.4.0* und *Beacon DesignerTM Free Edition* (siehe 3.2.7.2.1). Die so ermittelten Primer flankierten den gewünschten Abschnitt und wurden KDR_F („forward“) sowie KDR_R („reverse“) benannt. Sie wiesen folgende Sequenzen auf:

- KDR_F: GCGATTTACCCACACCTGTC
- KDR_R: CTCCTGTCCTGAGCAAAGC

Im Anschluss an die Primererstellung wurde die eigentliche Sequenzierung mittels ABI 3100-Avant Genetic Analyzer (Applied Biosystems, California, USA), wie auch schon die Paired-End-Sequenzierung, vom LAFUGA durchgeführt. Die Ergebnisse wurden dann an den Lehrstuhl für Tierzucht und Allgemeine Landwirtschaftslehre zurück übermittelt.

Die DNA-Sequenzierung konnte das Vorhandensein der drei Varianten, welche aufgrund der Auswertung der Paired-End-Sequenzierung zu erwarten waren, verifizieren. Darüber hinaus konnte bestätigt werden, dass alle Tiere der Kontrollgruppe sowie das FV0239 für diese möglichen kausalen Varianten heterozygot waren.

In Abbildung 29 ist repräsentativ ein Ausschnitt des Elektropherogramms eines homozygoten Fall-Tieres (FV08000) sowie eines heterozygoten Kontroll-Tieres (RH123) gezeigt. Zur schnelleren Orientierung sind die heterozygoten Bereiche jeweils gelb unterlegt.

Die vom LAFUGA übermittelten Sequenzabgleiche aller acht Tiere der Fall- und Kontrollgruppe sind dem Anhang 5 zu entnehmen.

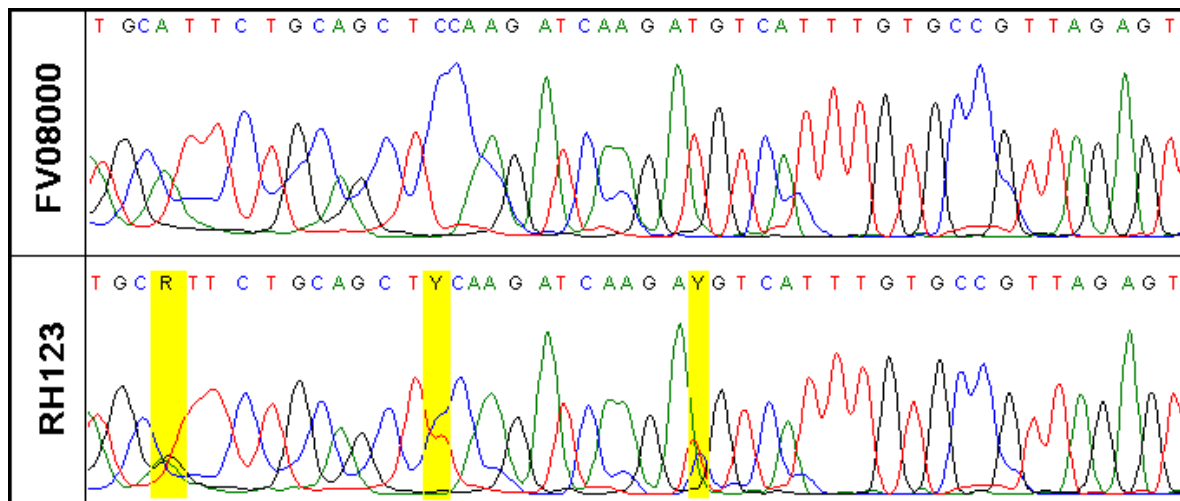


Abbildung 29: Elektropherogramm-Ausschnitt eines homozygoten Tieres der Fallgruppe sowie eines heterozygoten Tieres der Kontrollgruppe.

Dargestellt ist ein Elektropherogramm-Ausschnitt von FV08000, welches in dem gezeigten Bereich durchgehend homozygot ist, sowie von RH123, welches für die drei möglichen kausalen Varianten der Signatur um *KDR* heterozygot ist. Die Basenabfolge ist jeweils oberhalb der Kurven ablesbar. Dabei steht „A“ für Adenin, „C“ für Cytosin, „G“ für Guanin, „T“ für Thymin, „R“ für Purine (A oder G) und Y für Pyrimidine (T oder C).

Abbildung 29 macht deutlich, dass im DFV08000 (repräsentativ für die drei homozygoten Fall-Tiere) an den fraglichen Positionen nur die Basen Adenin, Cytosin und Thymin vorkommen. Im RH123 (repräsentativ für alle Kontrolltiere sowie FV0239) hingegen, kommen in den gelb markierten Bereichen auch die Basen Guanin, Thymin und Cytosin vor, so dass hier zwei verschiedene Sequenzvarianten zu finden sind. Dies bestätigt das in Abbildung 28 gezeigte Ergebnis der Paired-End-Sequenzierung.

Da das Vorhandensein der Varianten in dem möglichen kausalen Gen *KDR* somit verifiziert war, erfolgte anschließend eine Kontrolle der Richtigkeit der vom „*Variant Effect Predictor*“ vorhergesagten Auswirkungen dieser Basenaustausche. Dazu wurde dieser Genomabschnitt, von dem dank DNA-Sequenzierung nun die genaue Basenabfolge bekannt war, für die beiden verschiedenen Sequenzvarianten mittels „*ORF-Finder*“ (http://www.geneinfinity.org/sms/sms_orffinder.html) auf die dazugehörige Codierung geprüft. Überraschenderweise wichen die so gefundenen Aminosäureaustausche deutlich von denen des „*Variant Effect Predictor*“ ab. So gab es anstelle eines zusätzlichen Stopp-Codons und zweier Aminosäurewechsel lediglich zwei nicht mit den vorhergehenden Ergebnissen übereinstimmende Aminosäureaustausche. Tabelle 14 sind sowohl die ursprünglich gefundenen als auch die durch den „*ORF-Finder*“ ermittelten Auswirkungen auf die Codierung zu entnehmen.

Tabelle 14: Die Auswirkungen der drei näher untersuchten Varianten auf die Codierung.

Aufgeführt ist die physikalische Position (bp) der durch DNA-Sequenzierung bestätigten Varianten auf dem Rinderchromosom 6, der an dieser Position festgestellte Basenaustausch, sowie die Vorhersagen der Auswirkungen der Varianten auf die Codierung nach *Variant Effect Predictor* sowie *ORF-Finder*.

Position	Basenaustausch	Variant Effect Predictor	ORF-Finder
73175972	Adenin/Guanin	Asparagin/Lysin	Asparagin/Asparagin
73175983	Cytosin/Thymin	Glutamat/Stopp-Codon	Glutamat/Lysin
73175995	Thymin/Cytosin	Isoleucin/Leucin	Isoleucin/Valin

Da auf Position 73.175.972 kein Aminosäureaustausch stattfindet, scheidet diese Mutation als kausal aus. Bei einer Änderung von Isoleucin zu Valin, die beide aliphatische Seitenketten tragen, ist zu erwarten, dass die Auswirkungen auf das codierte Protein weniger deutlich sind, als bei einem Austausch der sauren Aminosäure Glutamat durch die Basische Lysin (STRYER, 1996).

Die gefundene Selektionssignatur im DFV lässt sich daher am wahrscheinlichsten auf den Polymorphismus an Position 73.175.983 auf BTA6 zurückführen, welcher sich in Exon 13 des Gens *KDR* befindet.

Die Erklärung für die falsche Vorhersage des „*Variant Effect Predictor*“ ist, dass dieser zwar durchaus an der richtigen Stelle des jeweiligen Codons das neue Nukleotid einsetzte, lag jedoch die Gensequenz, wie es auch hier der Fall war, auf dem Gegenstrang, wurde als neu einzusetzendes Nukleotid fälschlicherweise nicht die komplementäre Base genommen. Infolgedessen wurde eine fehlerhafte Sequenz erstellt und schließlich die jeweiligen Auswirkungen der einzelnen Mutationen falsch vorhergesagt. Ob das Problem letztlich beim „*Variant Effect Predictor*“ selbst oder bei einem Modul auf dem Webserver „*GALAXY*“ liegt, wird nun von Mitarbeitern des LAFUGA geklärt.

5 Diskussion

Seit der kommerziellen Verfügbarkeit einer großen Anzahl von SNP-Markern ist es möglich genomweit nach Selektionssignaturen zu suchen. Dabei ist das Auffinden dieser Signaturen wichtig, um Loci zu identifizieren, die an der Anpassung des Genoms an die verschiedensten Selektionsdrücke beteiligt sind, welche auf die einzelnen Phänotypen wirken.

In vorliegender Arbeit wurde diese genomweite Suche in zehn divergent selektierten Rinderrassen mittels zweier auf dem Auffinden von ausgedehnter Haplotypenhomozygotie basierender Methoden (*iHS* und *XP-EHH*) durchgeführt (VOIGHT *et al.*, 2006; SABETI *et al.*, 2007). Zur Auswertung wurde das Genom in 1.000 etwa gleich große Fenster unterteilt, von welchen in 94 mittels *iHS*, sowie in 90 durch *XP-EHH* Selektion nachgewiesen werden konnte. Für eine der so gefundenen Signaturen im DFV wurde dann eine Paired-End-Sequenzierung, sowie für einen Teilbereich anschließend eine DNA-Sequenzierung, durchgeführt und so das *KDR*-Gen als mögliches Kandidatengen identifiziert.

Die Festlegung des ancestralen Allels, die für die *iHS*-Berechnung unumgänglich ist, stellt einen potentiellen Störfaktor dar, der nicht unbeachtet bleiben darf. Idealerweise würde diese Festlegung anhand des direkten Vorfahrens, also des Auerochsen, erfolgen. Da dies jedoch nicht möglich ist, wurden diese bovinen ancestralen Allele in den verschiedenen Veröffentlichungen auf unterschiedlichste Weise definiert. Hayes *et al.* (2008) untersuchten den Ursprung von Selektionssignaturen auf BTA6 und legten hierfür das ancestrale Allel als jenes fest, das mit der Referenzsequenz, die sich von einer Hereford-Kuh ableitet, übereinstimmt. In einer späteren Arbeit hingegen nahmen Hayes *et al.* (2009) an, dass das Allel das ursprüngliche ist, welches bei Yak, Bison und Banteng (*Bison bison*, *Bos grunniens*, *Bos javanicus*) am häufigsten vorkam. Das Bovine HapMap Konsortium wiederum entschied sich für eine Kombination aus homozygoten Markern in Anoa und Wasserbüffel sowie einen Ansatz basierend auf Genomvergleichen zwischen Hund, Pferd und Mensch (GIBBS *et al.*, 2009).

Da das Hereford selbst eine stark durch Zucht beeinflusste Rasse darstellt, ist davon auszugehen, dass eine auf der Referenzsequenz basierende Zuteilung des Allelstatus wohl kaum den ursprünglichen Verhältnissen im Ur entspricht. Inwieweit eine Ableitung von anderen Rindergattungen, wie den oben genannten,

sinnvoll ist, ist schwer abzuschätzen. Es ist aber anzunehmen, dass hier Abweichungen auftreten, die bereits im Ur vorhanden waren. Dies sollte bei der Art der Zuteilung des Allelstatus anhand von unselektierten Rassen, wie sie in dieser Arbeit durchgeführt wurde, entsprechend nicht auftreten. Jedoch haben die etwa 10.000 Jahre, die sich Rinder bereits in menschlicher Obhut befinden, auch ohne gezielte Selektion in Form von Hochleistungszucht, ihre Spuren im Genom hinterlassen, so dass davon auszugehen ist, dass auch diese Zuteilung des ancestralen Allels nicht optimal ist. Darüber hinaus stellt diese beim Rind übliche uneinheitliche Allelstatuszuteilung gegebenenfalls ein Problem bezüglich der Vergleichbarkeit der verschiedenen Arbeiten dar.

Hinzukommt, dass die Annahme, das häufigste Allel sei immer das Älteste, erwiesenermaßen nicht immer zutrifft (WATTERSON & GUESS, 1977) und auch eine Rückmutation unter Umständen in einer inkorrekten Ableitung resultieren würde (FAY & WU, 2000). Dies hat zur Folge, dass auch hier zwangsläufig weitere leider unumgängliche Fehlerquellen vorhanden sind.

Da die Standardisierung der *iHS*-Werte, die vorgenommen wird um diese an das Alter der SNPs anzupassen, in Gruppen abhängig von der derived-Allelfrequenz durchgeführt wird, kann eine falsche Zuteilung der Allele zu ancestral und derived Auswirkungen auf den standardisierten *iHS*-Wert haben.

Dieses Unvermögen den Allelstatus mit absoluter Sicherheit zuteilen zu können, sowie die im Vergleich zur Humangenetik geringe SNP-Dichte limitieren diese Methode genauso, wie die Tatsache, dass der *iHS* nach der Fixation des selektierten Allels gelegentlich die umgebenden Bereiche identifiziert, nicht jedoch die eigentliche Selektion. Dies liegt daran, dass sich die Umgebung immer noch durch hohes LD auszeichnet, aber aufgrund der Fixation die Variation bei und um das selektierte Allel eliminiert wurde und daher durch den *iHS* nicht mehr nachweisbar ist (GIBBS *et al.*, 2009).

Ein Vergleich der Lage der signifikanten Fenster, die von *iHS* und *XP-EHH* erkannt wurden, wobei *XP-EHH* auf das Auffinden von Loci spezialisiert ist, die durch Selektion bereits fixiert oder nahe der Fixation sind, lässt die Vermutung zu, dass auch in dieser Studie einige vermeintliche durch *iHS* identifizierte Selektionssignaturen eher Ausläufer einer bereits sehr vorangeschrittenen Selektion darstellen, als tatsächlich selbst eine zu sein. So wäre dies unter anderem eine naheliegende Erklärung für folgende signifikante Fenster:

- RH: F737, $eP = 0,001$ → Nähe zu *XP-EHH*-Signatur F731-F735
- OBV: F370, $eP = 0,002$ → Nähe zu *XP-EHH*-Signatur F372
- OBV: F373, $eP = 0,008$ → Nähe zu *XP-EHH*-Signatur F372
- OBV: F931, $eP = 0,001$ → Nähe zu *XP-EHH*-Signatur F932/933
- FGV: F285, $eP = 0,007$ → Nähe zu *XP-EHH*-Signatur F281-F284

Ähnliches gilt für die Fenster 20, 21, 24, 25 und 26 im IL, wo sich signifikante *iHS*-Werte um eine eindeutige *XP-EHH*-Signatur in Fenster 22 anhäufen, allerdings mit dem Unterschied, dass hier auch Fenster 22 mit einem $eP = 0,002$ vom *iHS* erkannt wird. Diese Unsicherheit der durch *iHS* gefundenen Selektionssignaturen im Bezug auf die positionelle Übereinstimmung mit den eigentlich selektierten Loci, lässt den Schluss zu, dass eine alleinige Suche nach Selektionssignaturen mittels *iHS* nicht uneingeschränkt zu empfehlen ist.

Etwas überraschend war die Tatsache, dass die meisten *iHS*-Fenster solitär lagen und nur selten Cluster mit einer Länge von zwei (FGV, MWF, BBB, ASV) bis maximal drei Fenstern (RH, IL) auftraten. Zwar war die Anzahl signifikanter Fenster durch die Methodik auf genau zehn pro Rasse begrenzt, während der *XP-EHH* pro Rasse neun Vergleiche mit je zehn Fenstern, also insgesamt 90 Fenster, zugrunde lagen. Da bei *XP-EHH* aber für ein signifikantes Fenster mindestens drei Rassevergleiche übereinstimmen mussten, wären maximal 30 verschiedene signifikante Fenster möglich. Von diesem theoretischen Wert war jedoch auch die Rasse BBB, die mit insgesamt 14 die höchste Anzahl signifikanter Fenster aufwies, weit entfernt. Auch das GLW, das bei der *XP-EHH*-Berechnung die unter allen Rassen größte Fenstergruppe von sieben nebeneinander liegenden signifikanten Fenstern erreichte, lag mit zwölf Fenstern nur unweit über der bei *iHS* vorgegebenen Fensterzahl von zehn. Demzufolge sollte diese vorgegebene Fensterzahl nicht den einzigen Grund für die vergleichsweise geringe Anzahl sowie kleine Größe der gefundenen Fenstergruppen darstellen. Diese Annahme wird dadurch bestärkt, dass den durch *iHS* nachgewiesenen „*selective sweeps*“ prinzipiell positiv selektierte Allel zugrunde liegen, welche eine niedrigere Frequenz aufweisen, als die von *XP-EHH* erkannten. Das wiederum bedeutet, dass diese tendenziell jünger und somit von längeren Haplotypen umgeben sein müssten und daher eine ähnlich deutliche Gruppenbildung zeigen sollten. Die Erklärung hierfür ist vermutlich, dass in dem Frequenzbereich, den *iHS* erfasst, deutlich mehr und in etwa gleich starke Signaturen liegen. Hayes *et al.* (2008)

zeigten, dass der Anteil der Werte mit $|iHS| > 2$ in 20 Marker umfassenden Bereichen für QTL, die während der Periode der artifiziellen Selektion eine Frequenzänderung um 0,2 - 0,5 erfuhren, am höchsten war. Eine solche Frequenzveränderung sollten bei hinreichender Selektionsintensität, welche für Nutztier-Phänotypen von wirtschaftlichem Interesse vorausgesetzt wird, alle selektiven „sweeps“ erfahren, unabhängig davon, ob der Selektionsdruck ausreichend ist, um das bevorzugte Allel letztlich zu fixieren. Daher befinden sich im durch iHS nachweisbaren Frequenzbereich sowohl Loci, die unter moderater Selektion stehen, als auch solche, die zwar unter starker Selektion stehen, dies allerdings noch nicht lange respektive intensiv genug um sich bereits nahe der Fixation zu befinden. Dagegen werden durch $XP-EHH$ vor allem solche „sweeps“ effizient erfasst, deren Selektionsdruck und –dauer bereits ausreichend waren, um das bevorzugte Allel auf eine Frequenz über 0,8 zu bringen. In diesem Frequenzbereich nahe der Fixation ist nämlich die Effizienz des $XP-EHH$ -Tests („power“) der von iHS , bei welchem die „power“ bis zu einer Frequenz von etwa 0,75 steigt und dann rapide abnimmt, deutlich überlegen (VOIGHT *et al.*, 2006; SABETI *et al.*, 2007). Jedoch genau diese positive Korrelation zwischen iHS -Werten und moderat frequenten Allelen unter positiver Selektion, und das damit verbundene Potential Mutationen aufzudecken, die trotz geringerer Frequenz deutliche Effekte auf selektierte Merkmale haben können (HAYES *et al.*, 2008), stellt den Grund dar, aus dem eine Bestimmung der iHS -Werte trotz aller zu berücksichtigenden Probleme eine erstrebenswerte Ergänzung zur $XP-EHH$ darstellt.

Bei der Berechnung der $XP-EHH$ wird die Annahme gemacht, dass die Rekombinationsraten der zu vergleichenden Rassen identisch sind. Unter dieser Hypothese sind Unterschiede in der Länge der Haplotypenhomozygotie einzig auf Stärke und Dauer der zugrundeliegenden Selektion zurückzuführen. Sowohl Simianer *et al.* (1997) als auch Thomsen *et al.* (2001) zeigten jedoch, dass an manchen Loci zwischen einzelnen Rassen, Familien und sogar Individuen signifikante Unterschiede in der Rekombinationsrate nachweisbar sind. Zwar kamen beide Arbeiten zu dem Schluss, dass solche Unterschiede höchstwahrscheinlich nicht das gesamte Genom in gleicher Weise betreffen, dennoch sollten bereits für solche Unterschiede bekannte Bereiche bei der Interpretation der $XP-EHH$ -Ergebnisse berücksichtigt werden. Jedoch waren die

Übereinstimmungen der nachgewiesenen Selektionssignaturen mit den durch Simianer *et al.* (1997) und Thomsen *et al.* (2001) aufgezeigten Bereichen gering und beschränkten sich auf eine durch Thomsen *et al.* (2001) beschriebene Region auf BTA19, in welcher folgende zwei signifikante Fenster gefunden wurden:

- FGV: F770 mit drei von neun möglichen eP -Werten $\leq 0,01$
- BBV: F773 mit vier von neun möglichen eP -Werten $\leq 0,01$

Bei der Betrachtung dieser zwei Selektionssignaturen sollte dieser Umstand daher Beachtung finden. Dabei ist die insgesamt geringe Anzahl signifikanter Fenster in den bekannten Genomabschnitten mit unterschiedlichen Rekombinationsraten darauf zurückzuführen, dass mindestens drei der maximal möglichen neun Vergleiche pro Rasse einen eP -Wert $\leq 0,01$ aufweisen müssen, um als signifikant zu gelten. Diese Bedingung mindert damit den Effekt etwaiger nicht-selektionsbedingter Schwankungen einzelner Rassevergleiche und verdeutlicht den Vorteil der Kombination mehrerer Rassen.

Ein weiterer Punkt, der bei der Interpretation der *XP-EHH*-Ergebnisse beachtet werden muss, ist, dass der ursprünglich von Sabeti *et al.* (2007) eingeführte Algorithmus von J. Pickrell zur Anpassung an die im Vergleich zu Studien der Humangenetik geringere Markerdichte in dieser Arbeit verändert wurde. Daher erfolgte bei größeren Markerabständen im Gegensatz zur anfänglichen Version keine Korrektur (> 200 kb) respektive kein Abbruch (> 2.000 kb) mehr. Diese Angleichung an die Markerdichte machte eine lückenlose genomweite *XP-EHH*-Berechnung erst möglich, birgt jedoch gleichzeitig die Gefahr artifiziell erhöhter *XP-EHH*-Werte für Marker, die diese Lücken begrenzen. Müsste man davon ausgehen, dass sich Selektionssignaturen im Genom von Mensch und Rind in ihrer Ausprägung sehr ähnlich sind, könnte dies die Richtigkeit der Ergebnisse stark in Frage stellen. Spuren artifizieller Selektion, der Nutztiere unterliegen, sollten aber deutlich ausgeprägter sein als solche natürlicher Selektion auf fitnessbezogene Merkmale, wie sie in der humanen Population vorliegen. Daher kann angenommen werden, dass sich die jeweiligen Selektionssignaturen von Rind und Mensch deutlich unterscheiden und Ziele künstlicher Selektion in domestizierten Nutztierpopulationen einfacher aufzufinden sind (BISWAS & AKEY, 2006; QANBARI *et al.*, 2010a). Gestützt wird diese Annahme davon, dass die durchschnittliche Haplotypenblocklänge beim Rind etwa fünf- bis zehnmal so hoch ist, wie die des Menschen (KHATKAR *et al.*, 2007). Dies schließt das Vorkommen

artifiziell erhöhter *XP-EHH*-Werte bedingt durch hohe Markerabstände zwar nicht gänzlich aus, lässt aber vermuten, dass diese als Störfaktoren bei stark selektierten Populationen wie dem Rind weniger von Bedeutung sind.

Sabeti *et al.* (2007) stellten bereits bei der Einführung der *XP-EHH* fest, dass sowohl für Simulations- als auch Humangenomdaten in der Verteilung der Logratio-Werte immer eine geringe Asymmetrie in Richtung einer Population auszumachen war, ignorierten diese aber bei der Bestimmung der Signifikanzschwellen. Da in vorliegender Arbeit auf die Festlegung von für alle Vergleiche gemeinsamen Signifikanzgrenzen verzichtet wurde, und die Auswertung stattdessen durch empirische *P*-Werte erfolgte, ist diese Asymmetrie für die vorgestellten Ergebnisse ohne Belang. Dennoch wird an dieser Stelle darauf hingewiesen, dass je nach Rassevergleich teilweise sehr deutliche Abweichungen in Richtung einer der Rassen festzustellen war. Dabei war die Asymmetrie umso deutlicher, je mehr sich die Rassen in der effektiven Populationsgröße (N_e) und dem in der Regel damit zusammenhängenden Ausmaß an LD unterschieden. Bei Vergleichen ähnlicher Rassen wie IL und ASV, die beide ein sehr hohes N_e und sehr kleines LD aufweisen, war die Verteilung annähernd normal und symmetrisch um einen Mittelwert von Null (Abb. 30A). Bei Vergleichen einer dieser Rassen mit BBV, welches ein sehr kleines N_e und sehr großes LD auszeichnet, verschob sich die Verteilung stark auf die Seite von BBV, so dass sich die Ausläufer der Verteilung auch nach Standardisierung noch deutlich unterschieden (Abb. 30B).

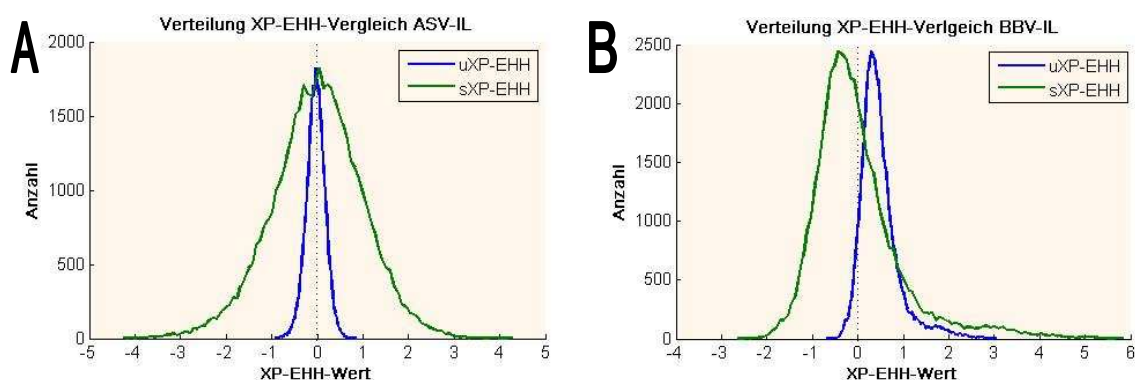


Abbildung 30: Verteilung der *XP-EHH*-Werte zweier Vergleiche von je zwei Rassen mit ähnlichen respektive stark unterschiedlichen N_e und LD.

Dargestellt ist die Anzahl an *XP-EHH*-Werten (y-Achse) in Abhängigkeit von deren Wert vor ($\mu XP-EHH$) und nach ($sXP-EHH$) Standardisierung (x-Achse) A) zeigt die Verteilung für den Vergleich der Rassen ASV und IL, welche sich in effektiver Populationsgröße (N_e) und Ausmaß an LD ähnlich sind. B) zeigt die Verteilung für den Vergleich von BBV und IL, welche sich in N_e und LD stark unterscheiden.

Dies lässt die Folgerung zu, dass die Bestimmung einer allgemeingültigen Signifikanzschwelle bei ähnlich inhomogenen Datenstrukturen ohne die Ergänzung eines oder mehrerer Korrekturfaktoren problematisch sein dürfte.

Wie zu erwarten war, zeigten die unselektierten Rassen IL und ASV mit vier respektive acht signifikanten Fenstern weniger Selektionssignaturen, als die selektierten Rassen (zehn bis 14). Dennoch war die Anzahl beim ASV, die näher an der der meisten selektierten Rassen lag als an der von IL, überraschend. Da in der Literatur daraufhingewiesen wird, dass in dieser Rasse eine Introgression durch Zebus stattgefunden hat (EDWARDS *et al.*, 2007), wurden die ASV-Tiere mittels „*maximum likelihood method*“ (WANG, 2003) genauer auf Subspezies „*Admixture*“ untersucht. Diese mit dem Programm *LEADMIX* (WANG, 2003) durchgeführte Analyse zeigte, dass die hier verwendeten ASV-Tiere einen *Bos-indicus*-Genanteil von 25,38 Prozent aufwiesen. Es wäre daher denkbar, dass die Ursache einiger dieser acht signifikanten Fenster, die für das ASV ermittelt wurden, im Auffinden Zebu-spezifischer Signaturen liegt.

Auffallend war außerdem der Umstand, dass BBB einerseits mit 14 signifikanten Fenstern die meisten Selektionssignaturen zeigte, andererseits jedoch mit maximal zwei nebeneinander liegenden signifikanten Fenstern die kleinsten Fenstergruppen abgesehen von den gänzlich gruppenlosen unselektierten Rassen hatte. Eine mögliche Erklärung wäre, dass mit der Ausprägung des Phänotyps Doppellender das Hauptzuchtziel bereits erreicht ist, und darüber hinaus vor allem – ähnlich den unselektierten Rassen – die Zucht auf Fitnesswerte erfolgt, was unzählige Loci genomweit in etwa gleich stark beeinflussen könnte. Ein anderer Grund für diese Ergebnisse könnte sein, dass sich der Fleischtyp erst in vergleichsweise junger Vergangenheit (1950) aus dem Doppelnutzungstyp des Weißblauen Belgiers entwickelte und diese Zuchtzielumstellung eventuell zu einem gleichzeitigen hohen Frequenzanstieg verhältnismäßig vieler Loci führte. Zwar erfolgte auch die Spezialisierung der Milchrassen (RH, BBV) etwa zur gleichen Zeit, jedoch unter Einkreuzung von bereits stark auf Milchleistung selektierten Tieren aus Nordamerika (siehe 2.1.2.3), wodurch sich deren Entwicklung von der des BBB unterscheidet.

Ein Punkt der sowohl *iHS*- als auch *XP-EHH*-Berechnungen beeinflussen könnte, ist die durch den verwendeten Illumina BovineSNP50 Genotyping BeadChip eingeführte Erhebungsverzerrung. Dieser entsteht aufgrund der Tatsache, dass die meisten auf diesem Chip befindlichen SNPs durch den Vergleich von Shotgun-Reads weniger Rassen, zu welchen auch das Holstein zählte, mit der auf der Rasse Hereford basierenden Referenzsequenz aufgedeckt wurden. Dies wiederum führt dazu, dass sich auf diesem Chip vermehrt SNPs befinden, die im Holstein informativer sind, was zur Folge hat, dass insbesondere in dieser Rasse die „*minor allele frequency*“ (MAF) verzerrt ist und es hier daher zu erhöhten MAFs kommt. Da für eine Korrektur dieser Erhebungsverzerrung die kombinierte Allelfrequenzverteilung über alle Rassen nötig wäre, diese aber nicht bekannt ist, schränkt das auch den Nutzen möglicher Simulationen ein, da die Auswirkungen einer solchen Erhebungsverzerrung ein wichtiges zu berücksichtigendes Merkmal wären (GIBBS *et al.*, 2009).

Da darüber hinaus für verlässliche Simulationen Wissen über demographische Parameter obligat, jedoch für die einzelnen Rinderrassen nicht ausreichend vorhanden ist, wäre eine hierauf basierende *P*-Wert-Bestimmung nicht verlässlich (PICKRELL *et al.*, 2009). Dies zusammen mit der Tatsache, dass bisher keine für *iHS*- oder *XP-EHH* geeignete Permutation bekannt ist, macht die Signifikanzbestimmung mittels empirischer *P*-Werte sinnvoll. Allerdings wird darauf hingewiesen, dass mit diesem Ansatz unter Umständen nur ein nichtrepräsentativer Anteil aller eigentlich selektierten Loci aufgefunden wird und insbesondere Selektion, die rezessive Loci und stehende anstelle neuer Variationen betrifft, unterrepräsentiert sein könnte (TESHIMA *et al.*, 2006).

Neben der oben erwähnten Erhebungsverzerrung der hier verwendeten ersten Version (v1) des Illumina Bovine BeadChips 50k, der in der mittlerweile verfügbaren neueren zweiten Version (v2) bereits geringer sein soll (http://www.illumina.com/Documents/products/datasheets/datasheet_bovine_snp50.pdf), stellt die gegenwärtige SNP-Dichte dieser einzigen bis vor kurzem verfügbaren SNP-Chips (v1: 54.001 SNPs, v2: 54.609 SNPs) einen limitierenden Faktor in der Auflösung dar. Dies gilt insbesondere bei der Auswertung von *iHS*, bei dem die lokale Anhäufung auffällig hoher Werte aussagekräftiger ist, als das für *XP-EHH* informativere Vorkommen einzelner maximaler Werte. Speziell für *iHS* ist daher, sofern eine ausreichende SNP-Anzahl pro Fenster erhalten werden will

(PICKRELL *et al.*, 2009), momentan noch die Verwendung relativ großer Fenster nötig. Um diese Fenster möglichst klein zu halten, werden sie in manchen Arbeiten (beispielsweise BARENDSE *et al.*, 2009) anhand einer bestimmten SNP-Anzahl festgelegt statt feste Größen zu bestimmen. Dies führt allerdings wiederum den Nachteil mit sich, dass sich diese Fenster je nach SNP-Dichte des jeweiligen Genombereichs in ihrer physikalischen Distanz unterscheiden, was ebenfalls Einfluss auf die erzielten Ergebnisse haben kann (PRASAD *et al.*, 2008). Barendse *et al.* (2009) schlossen aus ihren Studien, dass mindestens 150.000 gleichmäßig verteilte SNPs nötig sind, um die wirkliche Anzahl genomweiter Signale im Rind zu bestimmen. Da diese höheren Markerdichten seit kurzer Zeit mit dem BovineHD Genotyping BeadChip mit mehr als 777.000 SNPs verfügbar sind (http://www.illumina.com/Documents//products/datasheets/datasheet_bovineHD.pdf), werden diese Ansätze künftig noch besseren Einblick in die der Selektion unterliegenden biologischen Mechanismen bieten können (SIMIANER *et al.*, 2010).

Bei den insgesamt 170 verschiedenen, durch die *iHS*- respektive *XP-EHH*-Resultate bestimmten, signifikanten Fenstern, konnten Übereinstimmungen mit den Arbeiten von Qanbari *et al.* (2010a), Barendse *et al.* (2009), Flori *et al.* (2009), Gibbs *et al.* (2009), Hayes *et al.* (2009), MacEachern *et al.* (2009) und Stella *et al.* (2010) gefunden werden. Qanbari *et al.* (2010b) machten lediglich Angaben zur Anzahl der signifikanten Fenster jedoch keine zu ihrer genauen Position, was ein Abgleichen der hier gefundenen Ergebnisse mit diesen Resultaten unmöglich machte. Dass die Übereinstimmungen mit den oben genannten Arbeiten insgesamt eher moderat waren, war zu erwarten und liegt vermutlich an mehreren Faktoren. Die Gründe hierfür sind zum Einen bei der Verwendung sowohl verschiedener Methoden als auch unterschiedlicher Rassen, zum Anderen bei der Tatsache, dass die meisten Arbeiten nur über die signifikantesten Resultate berichten, zu suchen. Ferner ist die Rate falsch positiver Werte in genomweiten Studien wahrscheinlich relativ hoch (QANBARI *et al.*, 2010b), was zu einem großen Teil darin begründet ist, dass die Auswirkungen von Selektion und Gendrift sowie Migration nicht immer eindeutig zu unterscheiden sind (siehe 2.4). Des Weiteren kann die Verwendung von Fenstern zur Auswertung, deren Größe

subjektiv festgelegt wird, einen Einfluss auf die Endergebnisse zur Folge haben (QANBARI *et al.*, 2010b).

Für den potenziellen Nutzen der vorgestellten Methode bei der Identifikation von Selektionssignaturen spricht, neben einigen Überschneidungen mit den Ergebnissen anderer Studien, dass sich bei der Sequenzierung einer Signatur auf BTA6 die durch die *XP-EHH*-Ergebnisse zu vermutende lange Haplotypenhomozygotie im DFV im Vergleich zur Rasse RH bestätigte, und auch ein Kandidatengen (*KDR*) deklariert werden konnte.

Dabei wurden die auffälligsten Mutationen in Exon 13 dieses Gens gefunden. Bei diesen Varianten handelte es sich um zwei nichtsynonyme Mutationen an den Positionen 73.175.983 und 73.175.995. Letztere führt zu einem Austausch von Isoleucin durch Valin, welches beide Aminosäuren mit aliphatischen Seitenketten darstellen. Die Mutation an Position 73.175.983 hingegen führt zu einem Ersatz von Glutamat durch Lysin, also einer sauren Aminosäure durch eine Basische. Es wird daher davon ausgegangen, dass diese Mutation bedeutsamer und daher möglicherweise kausal für die gefundene Selektionssignatur ist. Für beide Varianten waren alle vier sequenzierten RH-Tiere sowie ein DFV heterozygot, während die übrigen drei DFV-Tiere hier homozygot waren und diese Mutationen nicht aufwiesen. Sollte eine Selektion auf die im DFV häufigere Variante des *KDR*-Gens tatsächlich Ursache der in dieser Rasse gefundenen Selektionssignatur sein, ist davon auszugehen, dass Tiere, die nicht Träger des entdeckten Alleles sind, einen Selektionsvorteil besitzen. Dies ist angesichts der Bedeutung dieses Gens durchaus vorstellbar, sofern sich mindestens eine dieser Mutationen tatsächlich auf die Genfunktion auswirkt.

KDR, der *kinase insert domain receptor*, gehört zur Genfamilie der *receptor tyrosin kinase (RTK)*, von welcher bekannt ist, dass ihre Mitglieder wichtig für die Differenzierung und das Wachstum vieler Zelltypen sind (ULLRICH & SCHLESSINGER, 1990). *KDR*, auch bekannt unter dem Namen *vascular endothelial growth factor receptor II (VEGFR-2)*, unterscheidet sich wie auch der von einem anderen Gen codierte *VEGFR-1* von dem Rest der Unterklasse III der *RTKs*, zu der sie gezählt werden, dadurch, dass sie sieben statt nur fünf immunoglobulin-ähnliche Wiederholungen in ihrer extrazellulären Domäne aufweisen. Dabei codiert das die genannten Mutationen tragende Exon 13 beim

Menschen für die sechste dieser sieben Wiederholungen (YIN *et al.*, 1998). Allerdings ist unklar, ob dies auf das Rind zu übertragen ist, da sich die Exonanzahl zwischen Mensch (28) und Rind (30) unterscheidet.

Es wird davon ausgegangen, dass *VEGFR-2* eine Schlüsselrolle in der Entwicklung von Gefäßen spielt, da seine Expression auf Endothelzellen und ihre Vorläufer beschränkt ist (SHALABY *et al.*, 1995). Neben der Angiogenese sind außerdem sowohl Gefäßdurchlässigkeit als auch Zellproliferation *VEGFR-2* vermittelt (TASAKI *et al.*, 2010). Der dazugehörige Ligand, der *vascular endothelial growth factor (VEGF)*, stellt ein endothelspezifisches Mitogen dar, das in verschiedenen Varianten sowohl von Zellen sezerniert wird als auch an Zellen gebunden vorkommt (BERISHA *et al.*, 2000). Da die Angiogenese ein Prozess ist, der hauptsächlich in der Embryonalphase vorkommt und in Adulten, abgesehen von einigen Ausnahmen wie der Wundheilung und pathologischen Prozessen wie dem Tumorwachstum, auf das weibliche Reproduktionssystem beschränkt ist (SCHAMS & BERISHA, 2002), sind mögliche Auswirkungen von Mutationen im *KDR*-Gen hier zu vermuten.

Shalaby *et al.* (1995) störten die Genfunktion durch Ersetzen des ersten Exons sowie des proximalen Teils des darauffolgenden Introns mit dem promoterlosen beta-Galaktosidase-Gen von *Escherichia coli*. Für dieses defekte Gen homozygote Mäuseembryos zeichneten sich durch das völlige Fehlen organisierter Blutgefäße aus und starben *in utero* zwischen Tag 8,5 und 9,5 *post coitum*. Zwar war keines der acht sequenzierten Tiere für die gefundenen Mutationen in Exon 13 homozygot, so dass die Möglichkeit, dass diese zu einem ähnlich frühen Absterben homozygoter Rinderembryos führen, gegenwärtig nicht gänzlich auszuschließen ist. Allerdings sollte ein solch gravierender Eingriff in das Reproduktionsgeschehen auch im RH zur Selektion des DFV-typischen Allels führen, so dass eine so deutliche Detektion einer Selektionssignatur im *XP-EHH*-Vergleich von DFV und RH nicht feststellbar sein sollte. Es wird daher davon ausgegangen, dass sich die beiden Allele durch weniger dramatische physiologische Auswirkungen unterscheiden. Den ersten Schritt nachfolgender Untersuchungen wird dennoch die Abklärung über das Vorkommen homozygoter Rinder darstellen. Sollten dann, wie vermutet, auch homozygote Tiere gefunden werden, wären weitere Studien nötig, die sich auf die Funktionen von *VEGFR-2* in adulten Tieren beziehen.

Ein Zusammenhang zwischen *VEGF* beziehungsweise *VEGFR-2* und dem Wachstum sowohl von dominantem Follikel als auch Corpus Luteum (CL) während des Zykluses weiblicher Rinder wurde bereits nachgewiesen (BERISHA *et al.*, 2000; SCHAMS & BERISHA, 2002). Die Ergebnisse beider Studien legen nahe, dass das vom Follikel- und Gelbkörpergewebe gebildete *VEGF*, das hauptsächlich der sekretorischen Form entspricht, parakrin als Chemoattraktor auf Endothelzellen wirkt. Darüber hinaus weist *VEGF* auf molekularer Basis eine 50.000-mal höhere Aktivität zur Gefäßpermeabilitätssteigerung auf als Histamin. Dies könnte für die Durchlässigkeit der Gefäße im dominanten Follikel verantwortlich sein, die eine wichtige Rolle beim Bezug von Gonadotropinen und LDL-Cholesterol („*low density lipoprotein*“) für die Biosynthese von Progesteron und Androgen spielt, was wiederum in einem gesteigerten Substratangebot für die Oestradiol-17 β -Produktion der Granulosazellen resultiert (BERISHA *et al.*, 2000). In einem ausgereiften CL befindet sich annähernd jede Parenchymzelle in Kontakt mit einer oder mehreren Kapillaren. Es ist daher nicht überraschen, dass bis zu 85 Prozent der Zellen, die während des CL-Wachstums proliferieren, Endothelzellen sind (SCHAMS & BERISHA, 2002) und eine ungenügende Gelbkörperfunktion bereits mit einer reduzierten Blutgefäßversorgung assoziiert wurde (REYNOLDS, 1986).

Dies in Kombination mit der Arbeit von Waltenberger *et al.* (1994), welche anhand von porcinen Aortenendothelzellen, die nach Transfektion entweder nur *VEGFR-1* oder *VEGFR-2* exprimierten, nachwies, dass nur Zellen mit *VEGFR-2* Expression in der Lage sind, in Richtung *VEGF* zu migrieren, bekräftigt einen möglichen Einfluss von Mutationen im *KDR*-Gen auf die Fruchtbarkeit.

Allerdings muss berücksichtigt werden, dass neben diesen zwei Mutationen eine Vielzahl anderer beobachtet werden konnte, wovon allerdings die meisten entweder zwischen den Genen oder in Introns lagen. Es wurde hier aber ähnlich der Arbeit von Owczarek-Lipska *et al.* (2011) argumentiert, dass die Wahrscheinlichkeit, dass eine solche Mutation kausal ist, im Vergleich zu Mutationen, die codierende respektive regulierende Bereiche betreffen, gering ist. Nach Ausschluss dieser Varianten blieben schließlich Mutationen in drei verschiedenen Genen (*CHIC2*, *KDR*, *SRD5A3*), wovon jene im *KDR*-Gen am besten mit dem Verlauf der *XP-EHH*-Kurve übereinstimmten und der in der Haplotypen-Matrix gefundenen auffälligen Region am nächsten waren. Dies

schließt die anderen beiden Gene als Ursache oder Mitverursacher der gefundenen Selektionssignatur jedoch nicht gänzlich aus, so dass diese Gene genauer analysiert werden sollten, falls in Folgeuntersuchungen keine physiologischen Auffälligkeiten zwischen den beiden Varianten des *KDR*-Gens feststellbar wären.

Könnte als Ursache für diese Selektionssignatur schließlich eine der gefunden Mutationen bestätigt werden, wäre ein weiterer interessanter Punkt, ob die im gleichen Bereich gefundenen Signaturen der *XP-EHH*-Vergleiche des DFV mit anderen Rassen als RH alle den gleichen Ursprung haben wie der DFV-RH-Vergleich, oder ob diesen Signaturen andere Mutationen zugrunde liegen.

Alles in allem stellen die aufgedeckten Selektionssignaturen eine Momentaufnahme basierend auf einer Auswahl gegenwärtig vorhandener Methoden dar, und sollten durch zukünftige Arbeiten, die im Zuge wissenschaftlichen Fortschritts weitere Erkenntnisse liefern werden, eine Verifizierung erfahren.

Die im DFV häufigere Genvariante des *KDR*, das bei genauerer Untersuchung einer auf BTA6 befindlichen Selektionssignatur des DFVs, als mögliches Kandidatengen deklariert wurde, sollte einen physiologischen Vorteil zu der im RH häufigen Variante mit sich bringen. Wie sehr sich die einzelnen Allele in ihren Auswirkungen auf die Genfunktion unterscheiden und inwieweit insbesondere die Mutation an Position 73.175.983, welche den Austausch einer sauren durch eine basische Aminosäure zur Folge hat, tatsächliche als Ursache dieser Signatur in Frage kommt, muss durch Folgestudien geklärt werden.

6 Zusammenfassung

Ziel dieser Arbeit war einerseits die umfassende Kartierung von Selektionssignaturen auf allen 29 Autosomen in zehn verschiedenen Rinderrassen, und andererseits die anschließende genauere Analyse einer dieser Signaturen, um das zugrundeliegende mögliche kausale Gen zu identifizieren.

Für die Suche nach Selektionssignaturen wurden zwei Methoden auf die mittels Illumina Bovine BeadChip 50K (Illumina, San Diego, USA) gewonnenen SNP-Daten (Single Nucleotide Polymorphism) angewandt, welche auf dem Auffinden ausgedehnter Haplotypen homozygotie basieren, namentlich *integrated Haplotype Score (iHS, VOIGHT et al., 2006)* und *Cross Population Extended Haplotype Homozygosity (XP-EHH, SABETI et al., 2007)*. Diese Methoden unterscheiden sich dadurch, dass die Erste innerhalb einer Rasse, die Zweite hingegen im Vergleich zweier Rassen Selektion aufdeckt.

Da die *XP-EHH* im Gegensatz zum *iHS* in der Analyse von Rindergenomdaten in keiner bisher veröffentlichten Studie Verwendung fand, erfolgte zunächst eine Validierung dieser Methode. Dabei wurde ihre Zuverlässigkeit ähnlich den Arbeiten von Qanbari *et al.* (2010a) und Stella *et al.* (2010) anhand des Auffindens der Selektionssignaturen bekannter einfach vererbter Merkmale (Doppellender, Rotfärbung, Hornlosigkeit) abgeschätzt.

Sowohl zur Auswertung von *iHS* als auch von *XP-EHH* wurde das Genom schließlich in 1.000 Fenster mit einer durchschnittlichen Größe von 2.544.053 Basen unterteilt, von denen 94 (*iHS*) respektive 90 (*XP-EHH*) verschiedene Fenster signifikante Selektionssignaturen beinhalteten.

Eine auf BTA6 befindliche Signatur der Rasse Deutsches Fleckvieh (DFV), welches für den süddeutschen Raum von besonderem wirtschaftlichen Interesse ist, wurde dann für weitere Untersuchungen ausgewählt. Vier DFV Tiere sowie als Vergleich weitere vier Tiere der Rasse Red Holstein (RH) wurden dann in der Region 71.949.142 bis 73.469.085 einer Paired-End-Sequenzierung unterzogen.

In drei (*CHIC2, KDR, SRD5A3*) der insgesamt sechs in diesem Bereich liegenden Protein-codierenden Gene konnten so in der codierenden Sequenz respektive in regulierenden Strukturen Abweichungen von der Referenzsequenz (*RefSeq*) nachgewiesen werden, deren Frequenz sich im DFV von der im RH unterschied.

Drei dieser Polymorphismen im *KDR*-Gen stimmten mit dem Peak des *XP-EHH*-Verlaufs überein. Für diese waren alle RH-Tiere sowie ein DFV-Tier heterozygot, während die übrigen DFV-Tiere Homozygotie in Übereinstimmung mit der *RefSeq* aufwiesen.

Zur Verifizierung der Ergebnisse der Paired-End-Sequenzierung erfolgte zusätzlich im Bereich um diese drei Mutationen eine DNA-Sequenzierung (nach SANGER *et al.* 1977). Diese bestätigte sowohl das Vorhandensein als auch die Genotypenverteilung der drei Varianten. Die jeweiligen Auswirkungen auf die Codierung waren dabei sehr unterschiedlich. So stellte ein Polymorphismus eine stille Mutation dar, während die beiden übrigen jeweils einen Aminosäureaustausch zur Folge hatten. Einer der beiden bedingte lediglich den Ersatz von Isoleucin durch Valin, beides Aminosäuren mit aliphatischen Seitenketten. Der Polymorphismus an der Position 73.175.983 führte hingegen zum Austausch einer sauren (Glutamat) durch eine basische (Lysin) Aminosäure. Diese Mutation ist als wahrscheinlichste kausale Ursache der untersuchten Selektionssignatur anzusehen.

Die mit der *RefSeq* übereinstimmende Genvariante, welche Tiere der Rasse DFV vornehmlich tragen, sollte gegenüber der häufig im RH gefundenen einen Selektionsvorteil aufweisen. Ein Zusammenhang der gefundenen Selektionssignatur mit der Fruchtbarkeit wäre denkbar, da das mögliche kausale Gen *KDR* für die Angiogenese unersetzlich ist, welche wiederum speziell in der Embryogenese (SHALABY *et al.*, 1995) und beim Wachstum von Gelbkörper und dominantem Follikel unentbehrlich ist (BERISHA *et al.*, 2000; SCHAMS & BERISHA, 2002).

Folgestudien werden zeigen, inwieweit diese Hypothese zutrifft, und weitere in dieser Arbeit gefundene Selektionssignaturen analysieren.

7 Summary

The aim of this thesis was a genome-wide detection of signatures of selection on all 29 autosomes of ten cattle breeds under divergent selection based on SNP (Single Nucleotide Polymorphism) data. One of those signatures was analysed in more detail in one breed to identify the putative gene causing it.

The detection of the signatures of selection was done with the Illumina Bovine BeadChip 50K (Illumina, San Diego, USA) using two approaches, both based on the finding of *extended haplotype homozygosity* (*EHH*, SABETI *et al.* 2002). While the *integrated Haplotype Score* (*iHS*, VOIGHT *et al.* 2006) searches for selection within one breed the *Cross Population Extended Haplotype Homozygosity* (*XP-EHH*, SABETI *et al.* 2007) compares two breeds.

In contrast to *iHS*, *XP-EHH* has never been used before for analysing bovine genome data, thus this method was validated at first. Similar to Qanbari *et al.* (2010a) and Stella *et al.* (2010) this was done by finding the signatures of selection of three well defined and simply inherited characteristics (polled, red-coat-colour, double-muscled).

For the evaluation of both, *iHS* and *XP-EHH*, the genome was split into 1.000 windows with an average size of 2.544.053 bases. Significant signatures of selection were finally found in 94 (*iHS*) and respectively 90 (*XP-EHH*) different windows.

One signature of selection of the German Fleckvieh (DFV), a breed of particular economic interest in the southern areas of Germany, was then chosen for detailed analysis. For this particular region on BTA6 (71.949.142 - 73.469.085), a paired-end sequencing was done for four animals of each of the two breeds German Fleckvieh and Red Holstein (RH).

Deviations from the reference sequence (*RefSeq*) were found with different frequencies in DFV and RH in some coding and regulating structures of three (*CHIC2*, *KDR*, *SRD5A3*) of the six protein-coding genes in this region.

Three of the polymorphisms found in the gene *KDR* matched with the peak of the *XP-EHH*-curve. Therefore, this is the most promising candidate gene for the analysed signature. All RH and one DFV animal were heterozygous for these polymorphisms whereas the three remaining DFV individuals were homozygous in congruence with the *RefSeq*.

To verify the results of the paired-end sequencing an additional sequencing with Sanger's method (Sanger *et al.* 1977) was done for the region surrounding those three most promising polymorphisms. Thereby the presence and genotype distribution of these variants was confirmed. Their specific impacts on the coding were different. One of the polymorphisms was a silent mutation, the remaining caused substitutions of the respective amino acid. One led to the replacement of Isoleucine by Valine, which are both aliphatic amino acids. The polymorphism on Position 73.175.983 however caused a substitution of an acidic (Glutamate) by an alkaline (Lysine) amino acid. This is therefore the most likely origin of the analysed signature of selection.

The allele in concordance to the *RefSeq*, which is found at high frequency in DFV, should imply a selective advantage compared to the other allele frequently found in RH. The putative gene *KDR* is essential for angiogenesis, which is *inter alia* especially important during embryogenesis (SHALABY *et al.*, 1995) and growth of the corpus luteum and the dominant follicle (BERISHA *et al.*, 2000; SCHAMS & BERISHA, 2002). Thus, this signature of selection can be seen in the context of fertility.

Follow-up studies will prove this hypothesis and analyse further signatures of selection found in this work.

8 Literaturverzeichnis

- Ahne, W., Liebich, H.-G., Stohrer, M., Wolf, E., (2000)** „Zoologie“, S. 298/299, Schattauer Verlagsgesellschaft mbH
- Akey, J.M., Zhang, G., Zhang, K., Jin, L., Shriver, M. D. (2002)** „Interrogating a high-density SNP map for signatures of natural selection“, *Genome Res* 12: 1805-1814
- Arber, W., Dussoix, D. (1962)** „Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda“, *J Mol Biol* 5: 18-36
- Arthur, P. F. (1995)** „Double muscling in cattle: a review“, *Australian Journal of Agricultural Research* 46: 1493-1515
- Barendse, W., Harrison, B.E., Bunch, R.J., Thomas, M.B., Turner, L.B. (2009)** „Genome wide signatures of positive selection: the comparison of independent samples and the identification of regions associated to traits“, *BMC Genomics* 10: 178
- Berisha, B., Schams, D., Kosmann, M., Amselgruber, W., Einspanier, R. (2000)** „Expression and localisation of vascular endothelial growth factor and basic fibroblast growth factor during the final growth of bovine ovarian follicles“, *J Endocrinol* 167: 371-382
- Biswas, S., Akey, J.M. (2006)** „Genomic insights into positive selection“, *Trends Genet* 22: 437-446
- Blott, S., Kim, J. J., Moiso, S., Schmidt-Kuntzel, A., Cornet, A., Berzi, P., Cambisano, N., Ford, C., Grisart, B., Johnson, D. *et al.* (2003)** „Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition.“, *Genetics* 163: 253-266.
- Bonin, A., Nicole, F., Pompanon, F., Miaud, C., Taberlet, P. (2007)** „Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation.“, *Conserv Biol* 21: 697-708
- Brade, W. (2010)** „Genombasierte Selektion: neue Ansätze in der Rinderzüchtung“, Landwirtschaftskammer Niedersachsen, <http://www.lwk-niedersachsen.de/index.cfm/portal/1/nav/1092/article/15546.html>
- Brem, G., Brenig, B., Müller, M., Springmann, K., Kräußlich, H. (1990)** „Genetische Vielfalt von Rinderrassen – Historische Entwicklung und moderne Möglichkeiten zur Konservierung“ S.40-49 Verlag Eugen Ulmer
- Brenneman, R.A., Davis, S.K., Sanders, J.O., Burns, B.M., Wheeler, T.C., Turner, J. W., Taylor, J. F. (1996)** „The polled locus maps to BTA1 in a *Bos indicus* x *Bos taurus* cross“, *J Hered* 87: 156-161
- Browning, S.R., Browning, B.L. (2007)** „Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering“, *Am J Hum Genet* 81: 1084-1097
- Browning, B.L., Browning, S.R. (2009)** „A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals“, *Am J Hum Genet* 84: 210-223

- Charlier, C., Coppieters, W., Farnir, F., Grobet, L., Leroy, P.L., Michaux, C., Mni, M., Schwers, A., Vanmanshoven, P., Hanset, R. et al. (1995)** „The mh gene causing double-muscling in cattle maps to bovine Chromosome 2”, *Mamm Genome* 6: 788-792
- Collins, F.S., Brooks, L.D., Chakravarti, A. (1998)** „A DNA polymorphism discovery resource for research on human genetic variation”, *Genome Research* 8: 1229-1231
- Department of Animal Science – Oklahoma State University (2001)** „Breeds of Livestock – Anatolian Black”, <http://www.ansi.okstate.edu/breeds/cattle/>
- Drögemüller, C., Wohlke, A., Leeb, T., Distl, O. (2005a)** „A 4 Mb high resolution BAC contig on bovine chromosome 1q12 and comparative analysis with human chromosome 21q22”, *Comp Funct Genomics* 6: 194-203
- Drögemüller, C., Wohlke, A., Mömke, S., Distl, O. (2005b)** „Fine mapping of the polled locus to a 1-Mb region on bovine Chromosome 1q12”, *Mammalian Genome* 16: 613-620
- Drossman, H., Luckey, J. A., Kostichka, A. J., D’Cunha, J., Smith, L. M. (1990)** „High-speed separations of DNA sequencing reactions by capillary electrophoresis.”, *Anal Chem* 62: 900-903
- Edwards, C.J., Baird, J.F., MacHugh, D.E. (2007)** „Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data”, *Anim Genet* 38: 520-524
- EFABIS – European Farm Animal Biodiversity Information System (2008)** „Breed Data Sheets - Busa/Cattle/Croatia”, <http://efabis.tzv.fal.de/>
- Elfrich, A., Roesicke, E. (2009)** „Rinderrassen”, aid infodienst Verbraucherschutz, Ernährung, Landwirtschaft e.V.
- Elsik, C.G., Tellam, R.L., Worley, K.C., Gibbs, R.A., Muzny, D.M. Weinstock, G.M., Adelson, D.L., Eichler, E.E., Elnitski, L., Guigo, R. et al. (2009)** „The genome sequence of taurine cattle: a window to ruminant biology and evolution”, *Science* 324: 522-528
- Falconer, D.S., Mackay, T.F.C. (1996)** „Introduction to Quantitative Genetics”, fourth edition, Pearson Education Limited Harlow
- FAO - Food And Agriculture Organization Of The United Nations (2000)** „WORLD WATCH LIST for domestic animal diversity”, 3rd Edition, FAO Rome, Editor: Scherf, B.D.
- FAO - Food And Agriculture Organization Of The United Nations (2010)** „STATUS AND TRENDS OF ANIMAL GENETIC RESOURCES”, FAO Rome, http://www.fao.org/ag/againfo/programmes/en/genetics/documents/ITWG_AnGR_6/CGRFA_WG_AnGR_6_10_inf_3.pdf
- Fay, J.C., Wu, C.I. (2000)** „Hitchhiking under positive Darwinian selection” *Genetics* 155: 1405-1413
- Felius, M. (1995)** „Cattle Breeds – An Encyclopedia“, Misset uitgeverij bv (NL)
- Finci, M. (1932)** „Die Rinderbestände in Bosnien und der Herzegowina“, Inaugural-Dissertation, Institut für Tierzüchtung und Haustiergenetik, Landwirtschaftliche Hochschule Berlin, Universitätsbuchdruckerei G. Neuenhahn, G.m.b.H., Jena, S: 25f
- Flori, L., Fritz, S., Jaffrezic, F., Boussaha, M., Gut, I. Heath, S., Foulley, J.L., Gautier, M. (2009)** „The genome response to artificial selection: a case study in dairy cattle”, *PLoS ONE* 4: e6595
- Fontanesi, L., Tazzoli, M., Russo, V., Beever, J. (2009)** „Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus”, *Anim Genet* 41: 295-303

- Frahm, K. (1982)** „Rinderrassen in den Ländern der Europäischen Gemeinschaft“, Ferdinand Enke Verlag Stuttgart
- GEH - Gesellschaft zur Erhaltung alter und gefährdeter Haustierrassen e.V. (2009)** „GEH - Rassekurzbeschreibungen Rinder - Original Braunvieh“, <http://www.g-e-h.de/geh-raku/rinder/rindbrau.htm>
- Geldermann, H. (2005a)** „Tier-Biotechnologie“, S. 198, Verlag Eugen Ulmer Stuttgart
- Geldermann, H. (2005b)** „Tier-Biotechnologie“, S. 165-171, Verlag Eugen Ulmer Stuttgart
- Georges, M., Drinkwater, R., King, T., Mishra, A., Moore, S.S., Nielsen, D., Sargeant, L.S., Sorensen, A., Steele, M.R., Zhao, X. et al. (1993)** „Micorsatellite mapping of a gene affecting horn development in *Bos taurus*“, *Nat Genet* 4: 206-210
- Gibbs, R.A., Taylor, J.F., Van Tassell, C.P., Barendse, W., Eversole, K.A., Gill, C.A., Green, R.D., Hamernik, D.L., Kappes, S.M., Lien, S. et al. (2009)** „Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds“, *Science* 324: 528-532
- Griffiths, A.J.F., Wessler, S.R., Lewontin, R.C., Carroll, S.B. (2008)** „Introduction to Genetic Analysis“, 9th Edition, S. 620-624, New York: W.H. Freeman and Company
- Grobet, L., Martin, L.J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Menissier, F., Massabanda, J. et al. (1997)** „A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle“, *Nat Genet* 17: 71-74
- Gunderson, K.L., Steemers, F.J., Lee, G., Mendoza, L.G., Chee, M.S. (2005)** „A genome-wide scalable SNP genotyping assay using microarray technology“, *Nat Genet* 37: 549-554
- Haller, M. (2000)** „Seltene Haus- & Nutztzierrassen“, S. 49, Leopold Stocker Verlag Graz
- Hampel, G. (1994)** „Fleischrinder- und Mutterkuhhaltung“, Verlag Eugen Ulmer Stuttgart
- Hayes, B.J., Lien, S., Nilsen, H., Olsen, H.G., Berg, P., MacEachern, S., Potter, S., Meuwissen, T.H.E. (2008)** „The origin of selection signatures on bovine chromosome 6“, *Animal Genetics* 39: 105-111
- Hayes, B.J., Chamberlain, A.J., MacEachern, S., Savin, K., McPartlan, H., MacLeod, I., Sethuraman, L., Goddard, M.E. (2009)** „A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle“, *Animal Genetics* 40: 176-184
- Hughes, A.L., Nei, M. (1988)** „Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection“, *Nature* 335: 167-170
- Ihaka, R. and Gentleman, R. (1996)** „R: A Language for Data Analysis and Graphics“, *Journal of Computational and Graphical Statistics* 5: 299-314
- Joerg, H., Fries, H.R., Meijerink, E., Stranzinger, G.F. (1996)** „Red coat color in Holstein cattle is associated with a deletion in the MSHR gene“, *Mamm Genome* 7: 317-318
- Kambadur, R., Sharma, M., Smith, T.P., Bass, J.J. (1997)** „Mutations in myostatin (GDF8) in double-muscling Belgian Blue and Piedmontese cattle“, *Genome Res* 7: 910-916

- Khatkar, M.S., Zenger, K.R., Hobbs, M., Hawken, R.J., Cavanagh, J.A., Barris, W., McClintock, A.E., McClintock, S, Thomson, P.C., Tier, B. et al. (2007)** „A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in hostein-friesian cattle”, *Genetics* 176: 763-772
- Klungland, H., Vage, D.I., Gomez-Raya, L., Adalsteinsson, S., Lien, S. (1995)** „The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination“, *Mamm Genome* 6: 636-639
- Kräußlich, H. (1994)** „Tierzüchtungslehre“, Verlag Eugen Ulmer Stuttgart
- Kräußlich, H., Brem, G. (1997)** „Tierzucht und Allgemeine Landwirtschaftslehre für Tiermediziner“, Ferdinand Enke Verlag Stuttgart
- Kreitman, M (2000)** „Methods to detect selection in populations with applications to the human”, *Annual Review of Genomics and Human Genetics* 1: 539-559
- Kugler, W. (2005)** „SAVE Focus - Special Edition Balkan”, Safeguard for Agricultural Varieties in Europe (Sicherung der landwirtschaftlichen Artenvielfalt Europa) http://www.save-foundation.net/Publications/SAVE_Focus_05_Balkan.pdf
- Landegren, U., Nilsson, M., Kwok, P.-Y. (1998)** „Reading Bits of Genetic Information: Methods for Single-Nucleotide Polymorphism Analysis“, *Genome Research* 8: 769-776
- Laurie, C.C., Dohney, K.F., Mirel, D.B., Pugh, E.W., Bierut, L.J., Bhangale, T., Boehm, F., Caporaso, N.E., Cornelis, M.C., Edenberg, H.J. et al. (2010)** „Quality control and quality assurance in genotypic data for genome-wide association studies”, *Genet Epidemiol* 34: 591-602
- Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M., Cunningham, P. (1994)** „Evidence for two independent domestications of cattle.”, *Proc Natl Acad Sci U S A* 91: 2757-2761
- Luckey, J. A., Drossman, H., Kostichka, A. J., Mead, D. A., D'Cunha, J., Norris, T. B., Smith, L. M. (1990)** „High speed DNA sequencing by capillary electrophoresis.”, *Nucleic Acids Res* 18: 4417-4421.
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., Taberlet, P. (2003)** „The power and promise of population genomics: from genotyping to genome typing.”, *Nat Rev Genet* 4: 981-994
- MacEachern, S., Hayes, B., McEwan, J., Goddard, M. (2009)** „An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle”, *BMC Genomics* 10: 181
- MacHugh, D. E., Loftus, R. T., Cunningham, P., Bradley, D. G. (1998)** „Genetic structure of seven European cattle breeds assessed using 20 microsatellite markers.”, *Anim Genet* 29: 333-340
- Mackrott, H. (1994)** „Milchviehhaltung” S. 9-17, Verlag Eugen Ulmer Stuttgart
- Mariasegaram, M., Reverter, A., Barris, W., Lehnert, S.A., Dalrymple, B., Prayaga, K. (2010)** „Transcription profiling provides insights into gene pathways involved in horn and scurs development in cattle”, *BMC Genomics* 11: 370
- Maxam, A.M., Gilbert W. (1977)** „A new method for sequencing DNA”, *Proc Natl Acad Sci U S A* 74: 560-564
- Maynard-Smith, J., Haigh, J. (1974)** „The hitch hiking effect of a favourable gene”, *Genet. Res.* 23: 23-35

- McPherron, A.C. , Lee, S.J. (1997)** „Double muscling in cattle due to mutations in the myostatin gene”, *Proc Natl Acad Sci U S A* 94: 12457-12461
- Medugorac, I., Medugorac, A., Russ, I., Veit-Kensch, C.E., Taberlet, P., Luntz, B., Mix, H.M., Förster, M. (2009)** „Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds with high effective population size”, *Molecular Ecology* 18: 3394-3410
- Metzker, M.L. (2010)** „Sequencing technologies - the next generation.”, *Nat Rev Genet* 11: 31-46
- Milligan, B. G. (2003)** „Maximum-likelihood estimation of relatedness.”, *Genetics* 163: 1153-1167
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. Erlich, H. (1986)** „Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction“, *Cold Spring Harb Symp Quant Biol* 51 Pt 1: 263-273
- Nielsen, R. (2005)** „Molecular signatures of natural selection”, *Annual Review of Genetics* 39: 197-218
- Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J. *et al.* (2005)** „A scan for positively selected genes in the genomes of humans and chimpanzees”, *PLoS Biol* 3: e170
- Ohta, T., Kimura, M. (1969)** „Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation”, *Genetics* 63: 229-238
- Owczarek-Lipska, M., Plattet, P., Zipperle, L., Drögemüller, C., Posthaus, H., Dolf, G., Braunschweig, M.H. (2011)** „A nonsense mutation in the optic atrophy 3 gene (OPA3) causes dilated cardiomyopathy in Red Holstein cattle”, *Genomics* 97: 51-57
- Pickrell, J.K., Coop, G., Novembre, J., Kudaravalli, S., Li, J.Z., Absher, D., Srinivasan, B.S., Barsh, G.S., Myers, R.M., Feldman, M.W., Pritchard, J.K. (2009)** „Signals of recent positive selection in a worldwide sample of human populations”, *Genome Res* 19: 826-837
- Prasad, A., Schnabel, R.D., McKay, S.D., Murdoch, B., Stothard, P., Kolbehdari, D., Wang, Z., Taylor, J.F., Moore, S.S. (2008)** „Linkage disequilibrium and signatures of selection on chromosomes 19 and 29 in beef and dairy cattle”, *Anim Genet* 39: 597-605
- Przeworski, M., (2002)** „The signature of positive selection at randomly chosen loci” *Genetics* 160: 1179-1189
- Przeworski, M., Coop G., Wall, J. D. (2005)** „The signature of positive selection on standing genetic variation.”, *Evolution* 59: 2312-2323
- 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., Sham, P.C. (2007)** „PLINK: a tool set for whole-genome association and population-based linkage analyses”, *Am J Hum Genet* 81: 559-575
- Qanbari, S., Pimentel, E.C., Tetens, J., Thaller, G., Lichtner, P., Sharifi, A.R., Simianer, H. (2010a)** „A genome-wide scan for signatures of recent selection in Holstein cattle“, *Anim Genet* 41: 377-389
- Qanbari, S., Gianloa, D., Thaller, G., Schenkel, F.S., Miller, S., Moore, S., Simianer, H. (2010b)** „Assessing Selection Footprints Based On Breed Differentiation – An Application To Whole Genome Cattle Data”, *Kongressband des 9th World Congress on Genetics Applied to Livestock Production (WCGALP)*
- Radman, M. (1999)** „Enzymes of evolutionary change.”, *Nature* 401: 866-867, 869

- Ramljak, J., Ivankovic, A., Veit-Kensch, C. E., Förster, M., Medugorac, I. (2011)** „Analysis of genetic and cultural conservation value of three indigenous Croatian cattle breeds in a local and global context.”, *J Anim Breed Genet* 128: 73-84
- Reinsch, N., Thomsen, H., Xu, N., Brink, M., Looft, C., Kalm, E., Brockmann, G.A., Grupe, S., Kuhn, C., Schwerin, M., et al. (1999)** „A QTL for the degree of spotting in cattle shows synteny with the KIT locus on chromosome 6”, *J Hered* 90: 629-634
- Reynolds, L.P. (1986)** „Utero-ovarian interactions during early pregnancy: role of conceptus-induced vasodilation”, *J Anim Sci* 62 Suppl 2:47-61
- Rosenberger, E. (2007)** „Max Obermayer, der Begründer der bayerischen Fleckviehzucht (I)”, *FleckviehWelt - Prüf- und Besamungsstation München Grub* 1/2007: 4-8
- Rouzaud, F., Martin, J., Gallet, P.F., Delourme, D., Goulemot-Leger, V., Amigues, Y., Menissier, F., Leveziel, H., Julien, R., Oulmouden, A. (2000)** „A first genotyping assay of French cattle breeds based on a new allele of the extension gene encoding the melanocortin-1 receptor (Mc1r)”, *Genet Sel Evol* 32: 511-520
- Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J. et al. (2002)** „Detecting recent positive selection in the human genome from haplotype structure”, *Nature* 419: 832-837
- Sabeti, P.C., Schaffner, S.F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T.S., Altshuler, D., Lander, E.S. (2006)** „Positive Natural Selection in the Human Lineage”, *Science* 312: 1614-1620
- Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R. et al. (2007)** „Genome-wide detection and characterization of positive selection in human populations”, *Nature* 449: 913-918
- Sambras, H.H. (2001)** „Farbatlas Nutztierassen“, Verlag Eugen Ulmer Stuttgart
- Sanger, F., Nicklen, S., Coulson, A.R. (1977)** „DNA sequencing with chain-terminating inhibitors”, *Proc Natl Acad Sci U S A* 74: 5463-5467
- Schams, D., Berisha, B. (2002)** „Angiogenic Factors (VEGF, FGF and IGF) in the Bovine Corpus Luteum”, *Journal of Reproduction and Development* 48: 233-242
- Schick, E. (2008)** „Original Braunvieh”, Rinderunion Baden-Württemberg e.V. http://www.rind-bw.de/braunvieh_original.htm
- Sendlinger, A., Irmer, J., Siedel, U. (2006)** „Grosses Handbuch Genetik Grundwissen und Gesetze“, S. 159/160 Compact Verlag München
- Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., Schuh, A.C. (1995)** „Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice”, *Nature* 376: 62-66
- Sibley, C.G., Ahlquist, J.E. (1984)** „The phylogeny of the hominoid primates, as indicated by DNA-DNA-hybridization”, *J Mol Evol* 20: 2-15
- Simianer, H., Szyda, J., Ramon, G., Lien, S. (1997)** „Evidence for individual and between-family variability of the recombination rate in cattle”, *Mamm Genome* 8: 830-835
- Simianer, H., Qanbari, S., Gianola, D. (2010)** „Detection of selection signatures within and between cattle populations”, *Kongressband des 9th World Congress on Genetics Applied to Livestock Production (WCGALP)*

- Slatkin, M. (1987)** „Gene flow and the geographic structure of natural populations”, *Science* 236: 787-792
- Smith, T.P., Lopez-Corrales, N.L., Kappes, S.M., Sonstegard, T.S. (1997)** „Myostatin maps to the interval containing the bovine mh locus”, *Mamm Genome* 8: 742-744
- Stella, A., Ajmone-Marsan, P., Lazzari, B., Boettcher, P. (2010)** „Identification of selection signatures in cattle breeds selected for dairy production”, *Genetics* 185: 1451-1461
- Storch, V., Welsch, U., Wink, M. (2007)** „Evolutionsbiologie“, S. 296-314, Springer-Verlag Heidelberg
- Strachan, T., Read, A.P. (2005)** „Molekulare Humangenetik”, ELSEVIER Spektrum Akademischer Verlag Heidelberg
- Stryer, L. (1996)** „Biochemie“, S. 18-23, Spektrum Akademischer Verlag Heidelberg
- Tariverdian, G., Buselmaier, W. (2004)** „Humangenetik“, S. 385-399 Springer-Verlag Heidelberg
- Tasaki, Y., Nishimura, R., Shibaya, M., Lee, H.Y., Acosta, T.J., Okuda, K. (2010)** „Expression of VEGF and its receptors in the bovine endometrium throughout the estrous cycle: effects of VEGF on prostaglandin production in endometrial cells”, *J Reprod Dev* 56: 223-229
- Teshima, K.M., Coop, G., Przeworski, M. (2006)** „How reliable are empirical genomic scans for selective sweeps?”, *Genome Res* 16: 702-712
- The Chimpanzee Sequencing and Analysis Consortium (2005)** „Initial sequence of the chimpanzee genome and comparison with the human genome”, *Nature* 437: 69-87
- Thomsen, H., Reinsch, N., Xu, N., Bennewitz, J., Looft, C., Grupe, S., Kuhn, C., Brockmann, G.A., Schwerin, M., Leyhe-Horn, B. et al. (2001)** „A whole genome scan for differences in recombination rates among three *Bos taurus* breeds“, *Mamm Genome* 12: 724-728
- Thrift, F.A. (2000)** „Simple example illustrating adjustment for merit of mates in a national cattle evaluation program”, *J Anim Sci* 78: 2475-2478
- Ullrich, A., Schlessinger, J. (1990)** „Signal transduction by receptors with tyrosine kinase activity”, *Cell* 61: 203-212
- Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K. (2006)** „A Map of Recent Positive Selection in the Human Genome”, *PLoS Biol* 4:e72
- Waltenberger, J., Claesson-Welsh, L., Siegbahn, A., Shibuya, M., Heldin, C.H. (1994)** „Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor”, *J Biol Chem* 269: 26988-26995
- Wang, J. (2003)** „Maximum-likelihood estimation of admixture proportions from genetic data”, *Genetics* 164: 747-765
- Watterson, G.A., Guess, H.A. (1977)** „Is the most frequent allele the oldest?“, *Theor Popul Biol* 11: 141-160
- Weiling, F. (1966)** „J.G. Mendels „Versuche über Pflanzen-Hybriden“ und ihre Würdigung in der Zeit bis zu ihrer Wiederentdeckung“, *TAG Theoretical and Applied Genetics* 36: 273-282
- Weir, B.S., Cockerham, C.C. (1984)** „Estimating F-Statistics for the Analysis of Population Structure”, *Evolution* 38: 1358-1370
- Weiss, K. M., Fullerton, S. M. (2000)** „Phenogenetic drift and the evolution of genotype-phenotype relationships.“, *Theor Popul Biol* 57: 187-195

-
- Weiss, K. M., Terwilliger, J. D. (2000)** „How many diseases does it take to map a gene with SNPs?” *Nat Genet* 26: 151-157
- Woolley, A. T., Mathies, R. A. (1995)** „Ultra-high-speed DNA sequencing using capillary electrophoresis chips.”, *Anal Chem* 67: 3676-3680
- Yauk, C.L., Bois, P.R., Jeffreys, A.J. (2003)** „High-resolution sperm typing of meiotic recombination in the mouse MHC E_{β} gene”, *EMBO J* 22: 1389-1397
- Yin, L.Y., Wu, Y., Ballinger, C.A., Patterson, C. (1998)** „Genomic structure of the human KDR/flk-1 gene”, *Mamm Genome* 9: 408-410
- Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P., del Cardayre, S. B. (2002)** „Genome shuffling leads to rapid phenotypic improvement in bacteria.” *Nature* 415: 644-646

9 Anhang

Anhang 1: Tierliste

Auflistung der 380 Tiere der Studie mit Angabe von Rasse und Labornummer (Labornr.)

<u>Rasse</u>	<u>Labornr.</u>	<u>Rasse</u>	<u>Labornr.</u>	<u>Rasse</u>	<u>Labornr.</u>
ASV	ASV001	BBB	BBB15	BBV	BV0135
ASV	ASV002	BBB	BBB16	BBV	BV0136
ASV	ASV003	BBB	BBB17	BBV	BV0137
ASV	ASV004	BBB	BBB18	BBV	BV0138
ASV	ASV005	BBB	BBB20	BBV	BV0141
ASV	ASV006	BBB	BBB23	BBV	BV0143
ASV	ASV007	BBB	BBB31	BBV	BV0144
ASV	ASV008	BBB	BBB32	BBV	BV0145
ASV	ASV009	BBB	BBB33	BBV	BV0146
ASV	ASV012	BBB	BBB35	BBV	BV0147
ASV	ASV013	BBB	BBB36	BBV	BV0150
ASV	ASV014	BBB	BBB39	BBV	BV0153
ASV	ASV016	BBB	BBB40	BBV	BV0155
ASV	ASV017	BBB	BBB44	BBV	BV0156
ASV	ASV018	BBB	BBB46	BBV	BV0157
ASV	ASV019	BBB	BBB47	BBV	BV0158
ASV	ASV020	BBB	BBB49	BBV	SM0496
ASV	ASV021	BBB	BBB51	BBV	W0226
ASV	ASV023	BBB	BBB52	BBV	W0297
ASV	ASV024	BBB	BBB54	BBV	W0362
ASV	ASV025	BBB	BBB57	BBV	W0630
ASV	ASV026	BBB	BBB66	BBV	W0631
ASV	ASV027	BBB	BBB67	BBV	WB0578
ASV	ASV029	BBB	BBB69	BBV	WB0781
ASV	ASV030	BBB	BBB70	BBV	WB1110
ASV	ASV031	BBB	BBB71	BBV	WB1127
ASV	ASV032	BBB	BBB75	GLW	GLW01
ASV	ASV033	BBV	0009139	GLW	GLW03
ASV	ASV034	BBV	0009727	GLW	GLW05
ASV	ASV035	BBV	0009729	GLW	GLW07
ASV	ASV036	BBV	0009793	GLW	GLW08
ASV	ASV037	BBV	0108900	GLW	GLW11
ASV	ASV038	BBV	0964857	GLW	GLW12
ASV	ASV039	BBV	0984342	GLW	GLW14
ASV	ASV040	BBV	BV0062	GLW	GLW15
ASV	ASV041	BBV	BV0125	GLW	GLW16
ASV	ASV042	BBV	BV0126	GLW	GLW17
ASV	ASV043	BBV	BV0127	GLW	GLW18
ASV	ASV044	BBV	BV0128	GLW	GLW19
ASV	ASV045	BBV	BV0129	GLW	GLW20
ASV	ASV046	BBV	BV0130	GLW	GLW21
ASV	ASV047	BBV	BV0131	GLW	GLW22
ASV	ASV050	BBV	BV0132	GLW	GLW23
BBB	BBB10	BBV	BV0134	GLW	GLW26

Rasse	Labornr.	Rasse	Labornr.	Rasse	Labornr.
GLW	GLW28	DFV	FV0249	FGV	FGV234
GLW	GLW29	DFV	FV0250	FGV	FGV244
GLW	GLW30	DFV	FV0251	FGV	FGV245
GLW	GLW31	DFV	FV0252	FGV	FGV246
GLW	GLW33	DFV	FV0273	FGV	FGV251
GLW	GLW37	DFV	FV0274	FGV	FGV252
GLW	GLW39	DFV	FV0275	IL	AT48
GLW	GLW40	DFV	FV0276	IL	AT60
GLW	GLW41	DFV	FV0277	IL	AT70
GLW	GLW42	DFV	FV0278	IL	AT76
GLW	GLW43	DFV	FV0279	IL	IL02
GLW	GLW47	DFV	R0026805	IL	IL03
DFV	010000	FGV	FGV004	IL	IL04
DFV	020000	FGV	FGV005	IL	IL06
DFV	050000	FGV	FGV007	IL	IL07
DFV	070000	FGV	FGV008	IL	IL08
DFV	080000	FGV	FGV009	IL	IL09
DFV	110000	FGV	FGV011	IL	IL10
DFV	130000	FGV	FGV012	IL	IL11
DFV	210000	FGV	FGV103	IL	IL12
DFV	250040	FGV	FGV105	IL	IL13
DFV	270000	FGV	FGV106	IL	IL14
DFV	300022	FGV	FGV107	IL	IL16
DFV	340001	FGV	FGV111	IL	IL17
DFV	360001	FGV	FGV115	IL	IL18
DFV	FV0126	FGV	FGV117	IL	IL19
DFV	FV0130	FGV	FGV119	IL	IL20
DFV	FV0133	FGV	FGV121	IL	IL21
DFV	FV0136	FGV	FGV122	IL	IL22
DFV	FV0137	FGV	FGV123	IL	IL23
DFV	FV0140	FGV	FGV126	IL	IL24
DFV	FV0169	FGV	FGV129	IL	IL25
DFV	FV0180	FGV	FGV135	IL	IL26
DFV	FV0191	FGV	FGV136	IL	IL27
DFV	FV0199	FGV	FGV169	IL	IL28
DFV	FV0223	FGV	FGV174	IL	IL29
DFV	FV0224	FGV	FGV180	IL	IL31
DFV	FV0225	FGV	FGV187	IL	IL32
DFV	FV0226	FGV	FGV188	IL	IL33
DFV	FV0230	FGV	FGV190	IL	IL34
DFV	FV0231	FGV	FGV195	IL	IL37
DFV	FV0232	FGV	FGV198	IL	IL38
DFV	FV0238	FGV	FGV204	IL	IL40
DFV	FV0239	FGV	FGV211	IL	IL41
DFV	FV0240	FGV	FGV214	IL	IL43
DFV	FV0241	FGV	FGV217	IL	IL44
DFV	FV0242	FGV	FGV220	IL	IL45
DFV	FV0243	FGV	FGV224	MWF	MWF001
DFV	FV0244	FGV	FGV230	MWF	MWF005
DFV	FV0247	FGV	FGV233	MWF	MWF011

Rasse	Labornr.	Rasse	Labornr.	Rasse	Labornr.
MWF	MWF012	RH	RH115	OBV	BV0148
MWF	MWF017	RH	RH116	OBV	OBV219
MWF	MWF020	RH	RH117	OBV	OBV215
MWF	MWF025	RH	RH118	OBV	OBV221
MWF	MWF028	RH	RH119	OBV	OBV222
MWF	MWF031	RH	RH120	OBV	OBV223
MWF	MWF034	RH	RH121	OBV	OBV225
MWF	MWF038	RH	RH122	OBV	OBV227
MWF	MWF039	RH	RH123	OBV	OBV229
MWF	MWF042	RH	RH124	OBV	OBV230
MWF	MWF055	RH	RH126	OBV	OBV231
MWF	MWF060	RH	RH127	OBV	OBV232
MWF	MWF063	RH	RH128	OBV	OBV233
MWF	MWF064	RH	RH129	OBV	OBV237
MWF	MWF074	RH	RH130	OBV	OBV239
MWF	MWF080	RH	RH131	OBV	OBV240
MWF	MWF084	RH	RH132	OBV	OBV242
MWF	MWF091	RH	RH134	OBV	OBV244
MWF	MWF093	RH	RH135	OBV	OBV245
MWF	MWF103	RH	RH136	OBV	OBV246
RH	RH101	RH	RH137	OBV	OBV247
RH	RH102	RH	RH138	OBV	OBV248
RH	RH103	RH	RH139	OBV	OBV249
RH	RH104	RH	RH140	OBV	OBV251
RH	RH105	RH	RH141	OBV	OBV252
RH	RH106	RH	RH142	OBV	OBV253
RH	RH107	RH	RH143	OBV	OBV254
RH	RH108	RH	RH144	OBV	OBV255
RH	RH109	RH	RH146	OBV	OBV257
RH	RH110	RH	RH147	OBV	OBV258
RH	RH112	RH	RH148	OBV	OBV259
RH	RH113	RH	RH149	OBV	OBV261
RH	RH114	RH	RH150		

Anhang 2: Fensterliste

Auflistung der 1.000 Fenster (F) mit Angabe von Chromosom (BTA) und Position der 1. Base (Startpos.) sowie der letzten Base (Endpos.)

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
1	1	483	2557020	51	1	127827383	130383920
2	1	2557021	5113558	52	1	130383921	132940458
3	1	5113559	7670096	53	1	132940459	135496996
4	1	7670097	10226634	54	1	135496997	138053534
5	1	10226635	12783172	55	1	138053535	140610072
6	1	12783173	15339710	56	1	140610073	143166610
7	1	15339711	17896248	57	1	143166611	145723148
8	1	17896249	20452786	58	1	145723149	148279686
9	1	20452787	23009324	59	1	148279687	150836224
10	1	23009325	25565862	60	1	150836225	153392762
11	1	25565863	28122400	61	1	153392763	155949300
12	1	28122401	30678938	62	1	155949301	158505838
13	1	30678939	33235476	63	1	158505839	161062329
14	1	33235477	35792014	64	2	4032	2561647
15	1	35792015	38348552	65	2	2561648	5119263
16	1	38348553	40905090	66	2	5119264	7676879
17	1	40905091	43461628	67	2	7676880	10234495
18	1	43461629	46018166	68	2	10234496	12792111
19	1	46018167	48574704	69	2	12792112	15349727
20	1	48574705	51131242	70	2	15349728	17907343
21	1	51131243	53687780	71	2	17907344	20464959
22	1	53687781	56244318	72	2	20464960	23022575
23	1	56244319	58800856	73	2	23022576	25580191
24	1	58800857	61357394	74	2	25580192	28137807
25	1	61357395	63913932	75	2	28137808	30695423
26	1	63913933	66470470	76	2	30695424	33253039
27	1	66470471	69027008	77	2	33253040	35810655
28	1	69027009	71583546	78	2	35810656	38368271
29	1	71583547	74140084	79	2	38368272	40925887
30	1	74140085	76696622	80	2	40925888	43483503
31	1	76696623	79253160	81	2	43483504	46041119
32	1	79253161	81809698	82	2	46041120	48598735
33	1	81809699	84366236	83	2	48598736	51156351
34	1	84366237	86922774	84	2	51156352	53713967
35	1	86922775	89479312	85	2	53713968	56271583
36	1	89479313	92035850	86	2	56271584	58829199
37	1	92035851	94592388	87	2	58829200	61386815
38	1	94592389	97148926	88	2	61386816	63944431
39	1	97148927	99705464	89	2	63944432	66502047
40	1	99705465	102262002	90	2	66502048	69059663
41	1	102262003	104818540	91	2	69059664	71617279
42	1	104818541	107375078	92	2	71617280	74174895
43	1	107375079	109931616	93	2	74174896	76732511
44	1	109931617	112488154	94	2	76732512	79290127
45	1	112488155	115044692	95	2	79290128	81847743
46	1	115044693	117601230	96	2	81847744	84405359
47	1	117601231	120157768	97	2	84405360	86962975
48	1	120157769	122714306	98	2	86962976	89520591
49	1	122714307	125270844	99	2	89520592	92078207
50	1	125270845	127827382	100	2	92078208	94635823

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
101	2	94635824	97193439	151	3	81861854	84420009
102	2	97193440	99751055	152	3	84420010	86978165
103	2	99751056	102308671	153	3	86978166	89536321
104	2	102308672	104866287	154	3	89536322	92094477
105	2	104866288	107423903	155	3	92094478	94652633
106	2	107423904	109981519	156	3	94652634	97210789
107	2	109981520	112539135	157	3	97210790	99768945
108	2	112539136	115096751	158	3	99768946	102327101
109	2	115096752	117654367	159	3	102327102	104885257
110	2	117654368	120211983	160	3	104885258	107443413
111	2	120211984	122769599	161	3	107443414	110001569
112	2	122769600	125327215	162	3	110001570	112559725
113	2	125327216	127884831	163	3	112559726	115117881
114	2	127884832	130442447	164	3	115117882	117676037
115	2	130442448	133000063	165	3	117676038	120234193
116	2	133000064	135557679	166	3	120234194	122792349
117	2	135557680	138115295	167	3	122792350	125350505
118	2	138115296	140672838	168	3	125350506	127908629
119	3	862	2559017	169	4	1621	2534759
120	3	2559018	5117173	170	4	2534760	5067898
121	3	5117174	7675329	171	4	5067899	7601037
122	3	7675330	10233485	172	4	7601038	10134176
123	3	10233486	12791641	173	4	10134177	12667315
124	3	12791642	15349797	174	4	12667316	15200454
125	3	15349798	17907953	175	4	15200455	17733593
126	3	17907954	20466109	176	4	17733594	20266732
127	3	20466110	23024265	177	4	20266733	22799871
128	3	23024266	25582421	178	4	22799872	25333010
129	3	25582422	28140577	179	4	25333011	27866149
130	3	28140578	30698733	180	4	27866150	30399288
131	3	30698734	33256889	181	4	30399289	32932427
132	3	33256890	35815045	182	4	32932428	35465566
133	3	35815046	38373201	183	4	35465567	37998705
134	3	38373202	40931357	184	4	37998706	40531844
135	3	40931358	43489513	185	4	40531845	43064983
136	3	43489514	46047669	186	4	43064984	45598122
137	3	46047670	48605825	187	4	45598123	48131261
138	3	48605826	51163981	188	4	48131262	50664400
139	3	51163982	53722137	189	4	50664401	53197539
140	3	53722138	56280293	190	4	53197540	55730678
141	3	56280294	58838449	191	4	55730679	58263817
142	3	58838450	61396605	192	4	58263818	60796956
143	3	61396606	63954761	193	4	60796957	63330095
144	3	63954762	66512917	194	4	63330096	65863234
145	3	66512918	69071073	195	4	65863235	68396373
146	3	69071074	71629229	196	4	68396374	70929512
147	3	71629230	74187385	197	4	70929513	73462651
148	3	74187386	76745541	198	4	73462652	75995790
149	3	76745542	79303697	199	4	75995791	78528929
150	3	79303698	81861853	200	4	78528930	81062068

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
201	4	81062069	83595207	251	5	82999179	85514297
202	4	83595208	86128346	252	5	85514298	88029416
203	4	86128347	88661485	253	5	88029417	90544535
204	4	88661486	91194624	254	5	90544536	93059654
205	4	91194625	93727763	255	5	93059655	95574773
206	4	93727764	96260902	256	5	95574774	98089892
207	4	96260903	98794041	257	5	98089893	100605011
208	4	98794042	101327180	258	5	100605012	103120130
209	4	101327181	103860319	259	5	103120131	105635249
210	4	103860320	106393458	260	5	105635250	108150368
211	4	106393459	108926597	261	5	108150369	110665487
212	4	108926598	111459736	262	5	110665488	113180606
213	4	111459737	113992875	263	5	113180607	115695725
214	4	113992876	116526014	264	5	115695726	118210844
215	4	116526015	119059153	265	5	118210845	120725963
216	4	119059154	121592292	266	5	120725964	123241082
217	4	121592293	124125394	267	5	123241083	125756132
218	5	252	2515370	268	6	1887	2554834
219	5	2515371	5030489	269	6	2554835	5107782
220	5	5030490	7545608	270	6	5107783	7660730
221	5	7545609	10060727	271	6	7660731	10213678
222	5	10060728	12575846	272	6	10213679	12766626
223	5	12575847	15090965	273	6	12766627	15319574
224	5	15090966	17606084	274	6	15319575	17872522
225	5	17606085	20121203	275	6	17872523	20425470
226	5	20121204	22636322	276	6	20425471	22978418
227	5	22636323	25151441	277	6	22978419	25531366
228	5	25151442	27666560	278	6	25531367	28084314
229	5	27666561	30181679	279	6	28084315	30637262
230	5	30181680	32696798	280	6	30637263	33190210
231	5	32696799	35211917	281	6	33190211	35743158
232	5	35211918	37727036	282	6	35743159	38296106
233	5	37727037	40242155	283	6	38296107	40849054
234	5	40242156	42757274	284	6	40849055	43402002
235	5	42757275	45272393	285	6	43402003	45954950
236	5	45272394	47787512	286	6	45954951	48507898
237	5	47787513	50302631	287	6	48507899	51060846
238	5	50302632	52817750	288	6	51060847	53613794
239	5	52817751	55332869	289	6	53613795	56166742
240	5	55332870	57847988	290	6	56166743	58719690
241	5	57847989	60363107	291	6	58719691	61272638
242	5	60363108	62878226	292	6	61272639	63825586
243	5	62878227	65393345	293	6	63825587	66378534
244	5	65393346	67908464	294	6	66378535	68931482
245	5	67908465	70423583	295	6	68931483	71484430
246	5	70423584	72938702	296	6	71484431	74037378
247	5	72938703	75453821	297	6	74037379	76590326
248	5	75453822	77968940	298	6	76590327	79143274
249	5	77968941	80484059	299	6	79143275	81696222
250	5	80484060	82999178	300	6	81696223	84249170

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
301	6	84249171	86802118	351	7	89142992	91689799
302	6	86802119	89355066	352	7	91689800	94236607
303	6	89355067	91908014	353	7	94236608	96783415
304	6	91908015	94460962	354	7	96783416	99330223
305	6	94460963	97013910	355	7	99330224	101877031
306	6	97013911	99566858	356	7	101877032	104423839
307	6	99566859	102119806	357	7	104423840	106970647
308	6	102119807	104672754	358	7	106970648	109517455
309	6	104672755	107225702	359	7	109517456	112064213
310	6	107225703	109778650	360	8	491	2542623
311	6	109778651	112331598	361	8	2542624	5084756
312	6	112331599	114884546	362	8	5084757	7626889
313	6	114884547	117437494	363	8	7626890	10169022
314	6	117437495	119990442	364	8	10169023	12711155
315	6	119990443	122543360	365	8	12711156	15253288
316	7	4712	2551519	366	8	15253289	17795421
317	7	2551520	5098327	367	8	17795422	20337554
318	7	5098328	7645135	368	8	20337555	22879687
319	7	7645136	10191943	369	8	22879688	25421820
320	7	10191944	12738751	370	8	25421821	27963953
321	7	12738752	15285559	371	8	27963954	30506086
322	7	15285560	17832367	372	8	30506087	33048219
323	7	17832368	20379175	373	8	33048220	35590352
324	7	20379176	22925983	374	8	35590353	38132485
325	7	22925984	25472791	375	8	38132486	40674618
326	7	25472792	28019599	376	8	40674619	43216751
327	7	28019600	30566407	377	8	43216752	45758884
328	7	30566408	33113215	378	8	45758885	48301017
329	7	33113216	35660023	379	8	48301018	50843150
330	7	35660024	38206831	380	8	50843151	53385283
331	7	38206832	40753639	381	8	53385284	55927416
332	7	40753640	43300447	382	8	55927417	58469549
333	7	43300448	45847255	383	8	58469550	61011682
334	7	45847256	48394063	384	8	61011683	63553815
335	7	48394064	50940871	385	8	63553816	66095948
336	7	50940872	53487679	386	8	66095949	68638081
337	7	53487680	56034487	387	8	68638082	71180214
338	7	56034488	58581295	388	8	71180215	73722347
339	7	58581296	61128103	389	8	73722348	76264480
340	7	61128104	63674911	390	8	76264481	78806613
341	7	63674912	66221719	391	8	78806614	81348746
342	7	66221720	68768527	392	8	81348747	83890879
343	7	68768528	71315335	393	8	83890880	86433012
344	7	71315336	73862143	394	8	86433013	88975145
345	7	73862144	76408951	395	8	88975146	91517278
346	7	76408952	78955759	396	8	91517279	94059411
347	7	78955760	81502567	397	8	94059412	96601544
348	7	81502568	84049375	398	8	96601545	99143677
349	7	84049376	86596183	399	8	99143678	101685810
350	7	86596184	89142991	400	8	101685811	104227943

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
401	8	104227944	106770076	451	10	5058213	7586847
402	8	106770077	109312209	452	10	7586848	10115482
403	8	109312210	111854342	453	10	10115483	12644117
404	8	111854343	114396475	454	10	12644118	15172752
405	8	114396476	116938581	455	10	15172753	17701387
406	9	671	2513964	456	10	17701388	20230022
407	9	2513965	5027258	457	10	20230023	22758657
408	9	5027259	7540552	458	10	22758658	25287292
409	9	7540553	10053846	459	10	25287293	27815927
410	9	10053847	12567140	460	10	27815928	30344562
411	9	12567141	15080434	461	10	30344563	32873197
412	9	15080435	17593728	462	10	32873198	35401832
413	9	17593729	20107022	463	10	35401833	37930467
414	9	20107023	22620316	464	10	37930468	40459102
415	9	22620317	25133610	465	10	40459103	42987737
416	9	25133611	27646904	466	10	42987738	45516372
417	9	27646905	30160198	467	10	45516373	48045007
418	9	30160199	32673492	468	10	48045008	50573642
419	9	32673493	35186786	469	10	50573643	53102277
420	9	35186787	37700080	470	10	53102278	55630912
421	9	37700081	40213374	471	10	55630913	58159547
422	9	40213375	42726668	472	10	58159548	60688182
423	9	42726669	45239962	473	10	60688183	63216817
424	9	45239963	47753256	474	10	63216818	65745452
425	9	47753257	50266550	475	10	65745453	68274087
426	9	50266551	52779844	476	10	68274088	70802722
427	9	52779845	55293138	477	10	70802723	73331357
428	9	55293139	57806432	478	10	73331358	75859992
429	9	57806433	60319726	479	10	75859993	78388627
430	9	60319727	62833020	480	10	78388628	80917262
431	9	62833021	65346314	481	10	80917263	83445897
432	9	65346315	67859608	482	10	83445898	85974532
433	9	67859609	70372902	483	10	85974533	88503167
434	9	70372903	72886196	484	10	88503168	91031802
435	9	72886197	75399490	485	10	91031803	93560437
436	9	75399491	77912784	486	10	93560438	96089072
437	9	77912785	80426078	487	10	96089073	98617707
438	9	80426079	82939372	488	10	98617708	101146342
439	9	82939373	85452666	489	10	101146343	103674977
440	9	85452667	87965960	490	10	103674978	106203563
441	9	87965961	90479254	491	11	496	2562617
442	9	90479255	92992548	492	11	2562618	5124739
443	9	92992549	95505842	493	11	5124740	7686861
444	9	95505843	98019136	494	11	7686862	10248983
445	9	98019137	100532430	495	11	10248984	12811105
446	9	100532431	103045724	496	11	12811106	15373227
447	9	103045725	105559018	497	11	15373228	17935349
448	9	105559019	108072259	498	11	17935350	20497471
449	10	943	2529577	499	11	20497472	23059593
450	10	2529578	5058212	500	11	23059594	25621715

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
501	11	25621716	28183837	551	12	42638966	45147113
502	11	28183838	30745959	552	12	45147114	47655261
503	11	30745960	33308081	553	12	47655262	50163409
504	11	33308082	35870203	554	12	50163410	52671557
505	11	35870204	38432325	555	12	52671558	55179705
506	11	38432326	40994447	556	12	55179706	57687853
507	11	40994448	43556569	557	12	57687854	60196001
508	11	43556570	46118691	558	12	60196002	62704149
509	11	46118692	48680813	559	12	62704150	65212297
510	11	48680814	51242935	560	12	65212298	67720445
511	11	51242936	53805057	561	12	67720446	70228593
512	11	53805058	56367179	562	12	70228594	72736741
513	11	56367180	58929301	563	12	72736742	75244889
514	11	58929302	61491423	564	12	75244890	77753037
515	11	61491424	64053545	565	12	77753038	80261185
516	11	64053546	66615667	566	12	80261186	82769333
517	11	66615668	69177789	567	12	82769334	85277438
518	11	69177790	71739911	568	13	6040	2561741
519	11	71739912	74302033	569	13	2561742	5117443
520	11	74302034	76864155	570	13	5117444	7673145
521	11	76864156	79426277	571	13	7673146	10228847
522	11	79426278	81988399	572	13	10228848	12784549
523	11	81988400	84550521	573	13	12784550	15340251
524	11	84550522	87112643	574	13	15340252	17895953
525	11	87112644	89674765	575	13	17895954	20451655
526	11	89674766	92236887	576	13	20451656	23007357
527	11	92236888	94799009	577	13	23007358	25563059
528	11	94799010	97361131	578	13	25563060	28118761
529	11	97361132	99923253	579	13	28118762	30674463
530	11	99923254	102485375	580	13	30674464	33230165
531	11	102485376	105047497	581	13	33230166	35785867
532	11	105047498	107609619	582	13	35785868	38341569
533	11	107609620	110171705	583	13	38341570	40897271
534	12	450	2508597	584	13	40897272	43452973
535	12	2508598	5016745	585	13	43452974	46008675
536	12	5016746	7524893	586	13	46008676	48564377
537	12	7524894	10033041	587	13	48564378	51120079
538	12	10033042	12541189	588	13	51120080	53675781
539	12	12541190	15049337	589	13	53675782	56231483
540	12	15049338	17557485	590	13	56231484	58787185
541	12	17557486	20065633	591	13	58787186	61342887
542	12	20065634	22573781	592	13	61342888	63898589
543	12	22573782	25081929	593	13	63898590	66454291
544	12	25081930	27590077	594	13	66454292	69009993
545	12	27590078	30098225	595	13	69009994	71565695
546	12	30098226	32606373	596	13	71565696	74121397
547	12	32606374	35114521	597	13	74121398	76677099
548	12	35114522	37622669	598	13	76677100	79232801
549	12	37622670	40130817	599	13	79232802	81788503
550	12	40130818	42638965	600	13	81788504	84344187

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
601	14	1182	2542518	651	15	46125656	48687938
602	14	2542519	5083855	652	15	48687939	51250221
603	14	5083856	7625192	653	15	51250222	53812504
604	14	7625193	10166529	654	15	53812505	56374787
605	14	10166530	12707866	655	15	56374788	58937070
606	14	12707867	15249203	656	15	58937071	61499353
607	14	15249204	17790540	657	15	61499354	64061636
608	14	17790541	20331877	658	15	64061637	66623919
609	14	20331878	22873214	659	15	66623920	69186202
610	14	22873215	25414551	660	15	69186203	71748485
611	14	25414552	27955888	661	15	71748486	74310768
612	14	27955889	30497225	662	15	74310769	76873051
613	14	30497226	33038562	663	15	76873052	79435334
614	14	33038563	35579899	664	15	79435335	81997617
615	14	35579900	38121236	665	15	81997618	84559855
616	14	38121237	40662573	666	16	2240	2514922
617	14	40662574	43203910	667	16	2514923	5027605
618	14	43203911	45745247	668	16	5027606	7540288
619	14	45745248	48286584	669	16	7540289	10052971
620	14	48286585	50827921	670	16	10052972	12565654
621	14	50827922	53369258	671	16	12565655	15078337
622	14	53369259	55910595	672	16	15078338	17591020
623	14	55910596	58451932	673	16	17591021	20103703
624	14	58451933	60993269	674	16	20103704	22616386
625	14	60993270	63534606	675	16	22616387	25129069
626	14	63534607	66075943	676	16	25129070	27641752
627	14	66075944	68617280	677	16	27641753	30154435
628	14	68617281	71158617	678	16	30154436	32667118
629	14	71158618	73699954	679	16	32667119	35179801
630	14	73699955	76241291	680	16	35179802	37692484
631	14	76241292	78782628	681	16	37692485	40205167
632	14	78782629	81323942	682	16	40205168	42717850
633	15	4562	2566844	683	16	42717851	45230533
634	15	2566845	5129127	684	16	45230534	47743216
635	15	5129128	7691410	685	16	47743217	50255899
636	15	7691411	10253693	686	16	50255900	52768582
637	15	10253694	12815976	687	16	52768583	55281265
638	15	12815977	15378259	688	16	55281266	57793948
639	15	15378260	17940542	689	16	57793949	60306631
640	15	17940543	20502825	690	16	60306632	62819314
641	15	20502826	23065108	691	16	62819315	65331997
642	15	23065109	25627391	692	16	65331998	67844680
643	15	25627392	28189674	693	16	67844681	70357363
644	15	28189675	30751957	694	16	70357364	72870046
645	15	30751958	33314240	695	16	72870047	75382729
646	15	33314241	35876523	696	16	75382730	77895388
647	15	35876524	38438806	697	17	7913	2556124
648	15	38438807	41001089	698	17	2556125	5104336
649	15	41001090	43563372	699	17	5104337	7652548
650	15	43563373	46125655	700	17	7652549	10200760

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
701	17	10200761	12748972	751	18	61030901	63573763
702	17	12748973	15297184	752	18	63573764	66116595
703	17	15297185	17845396	753	19	993	2509184
704	17	17845397	20393608	754	19	2509185	5017376
705	17	20393609	22941820	755	19	5017377	7525568
706	17	22941821	25490032	756	19	7525569	10033760
707	17	25490033	28038244	757	19	10033761	12541952
708	17	28038245	30586456	758	19	12541953	15050144
709	17	30586457	33134668	759	19	15050145	17558336
710	17	33134669	35682880	760	19	17558337	20066528
711	17	35682881	38231092	761	19	20066529	22574720
712	17	38231093	40779304	762	19	22574721	25082912
713	17	40779305	43327516	763	19	25082913	27591104
714	17	43327517	45875728	764	19	27591105	30099296
715	17	45875729	48423940	765	19	30099297	32607488
716	17	48423941	50972152	766	19	32607489	35115680
717	17	50972153	53520364	767	19	35115681	37623872
718	17	53520365	56068576	768	19	37623873	40132064
719	17	56068577	58616788	769	19	40132065	42640256
720	17	58616789	61165000	770	19	42640257	45148448
721	17	61165001	63713212	771	19	45148449	47656640
722	17	63713213	66261424	772	19	47656641	50164832
723	17	66261425	68809636	773	19	50164833	52673024
724	17	68809637	71357848	774	19	52673025	55181216
725	17	71357849	73906060	775	19	55181217	57689408
726	17	73906061	76454249	776	19	57689409	60197600
727	18	2189	2545051	777	19	60197601	62705792
728	18	2545052	5087914	778	19	62705793	65213966
729	18	5087915	7630777	779	20	642	2524136
730	18	7630778	10173640	780	20	2524137	5047631
731	18	10173641	12716503	781	20	5047632	7571126
732	18	12716504	15259366	782	20	7571127	10094621
733	18	15259367	17802229	783	20	10094622	12618116
734	18	17802230	20345092	784	20	12618117	15141611
735	18	20345093	22887955	785	20	15141612	17665106
736	18	22887956	25430818	786	20	17665107	20188601
737	18	25430819	27973681	787	20	20188602	22712096
738	18	27973682	30516544	788	20	22712097	25235591
739	18	30516545	33059407	789	20	25235592	27759086
740	18	33059408	35602270	790	20	27759087	30282581
741	18	35602271	38145133	791	20	30282582	32806076
742	18	38145134	40687996	792	20	32806077	35329571
743	18	40687997	43230859	793	20	35329572	37853066
744	18	43230860	45773722	794	20	37853067	40376561
745	18	45773723	48316585	795	20	40376562	42900056
746	18	48316586	50859448	796	20	42900057	45423551
747	18	50859449	53402311	797	20	45423552	47947046
748	18	53402312	55945174	798	20	47947047	50470541
749	18	55945175	58488037	799	20	50470542	52994036
750	18	58488038	61030900	800	20	52994037	55517531

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
801	20	55517532	58041026	851	22	38642846	41218685
802	20	58041027	60564521	852	22	41218686	43794525
803	20	60564522	63088016	853	22	43794526	46370365
804	20	63088017	65611511	854	22	46370366	48946205
805	20	65611512	68135006	855	22	48946206	51522045
806	20	68135007	70658501	856	22	51522046	54097885
807	20	70658502	73181996	857	22	54097886	56673725
808	20	73181997	75705448	858	22	56673726	59249565
809	21	2711	2564511	859	22	59249566	61825382
810	21	2564512	5126312	860	23	9738	2548773
811	21	5126313	7688113	861	23	2548774	5087809
812	21	7688114	10249914	862	23	5087810	7626845
813	21	10249915	12811715	863	23	7626846	10165881
814	21	12811716	15373516	864	23	10165882	12704917
815	21	15373517	17935317	865	23	12704918	15243953
816	21	17935318	20497118	866	23	15243954	17782989
817	21	20497119	23058919	867	23	17782990	20322025
818	21	23058920	25620720	868	23	20322026	22861061
819	21	25620721	28182521	869	23	22861062	25400097
820	21	28182522	30744322	870	23	25400098	27939133
821	21	30744323	33306123	871	23	27939134	30478169
822	21	33306124	35867924	872	23	30478170	33017205
823	21	35867925	38429725	873	23	33017206	35556241
824	21	38429726	40991526	874	23	35556242	38095277
825	21	40991527	43553327	875	23	38095278	40634313
826	21	43553328	46115128	876	23	40634314	43173349
827	21	46115129	48676929	877	23	43173350	45712385
828	21	48676930	51238730	878	23	45712386	48251421
829	21	51238731	53800531	879	23	48251422	50790457
830	21	53800532	56362332	880	23	50790458	53329482
831	21	56362333	58924133	881	24	2261	2500072
832	21	58924134	61485934	882	24	2500073	4997884
833	21	61485935	64047735	883	24	4997885	7495696
834	21	64047736	66609536	884	24	7495697	9993508
835	21	66609537	69171298	885	24	9993509	12491320
836	22	5246	2581085	886	24	12491321	14989132
837	22	2581086	5156925	887	24	14989133	17486944
838	22	5156926	7732765	888	24	17486945	19984756
839	22	7732766	10308605	889	24	19984757	22482568
840	22	10308606	12884445	890	24	22482569	24980380
841	22	12884446	15460285	891	24	24980381	27478192
842	22	15460286	18036125	892	24	27478193	29976004
843	22	18036126	20611965	893	24	29976005	32473816
844	22	20611966	23187805	894	24	32473817	34971628
845	22	23187806	25763645	895	24	34971629	37469440
846	22	25763646	28339485	896	24	37469441	39967252
847	22	28339486	30915325	897	24	39967253	42465064
848	22	30915326	33491165	898	24	42465065	44962876
849	22	33491166	36067005	899	24	44962877	47460688
850	22	36067006	38642845	900	24	47460689	49958500

F	BTA	Startpos.	Endpos.	F	BTA	Startpos.	Endpos.
901	24	49958501	52456312	951	27	17965877	20529246
902	24	52456313	54954124	952	27	20529247	23092616
903	24	54954125	57451936	953	27	23092617	25655986
904	24	57451937	59949748	954	27	25655987	28219356
905	24	59949749	62447560	955	27	28219357	30782726
906	24	62447561	64945342	956	27	30782727	33346096
907	25	3230	2592541	957	27	33346097	35909466
908	25	2592542	5181853	958	27	35909467	38472836
909	25	5181854	7771165	959	27	38472837	41036206
910	25	7771166	10360477	960	27	41036207	43599576
911	25	10360478	12949789	961	27	43599577	46162946
912	25	12949790	15539101	962	27	46162947	48726297
913	25	15539102	18128413	963	28	6821	2563162
914	25	18128414	20717725	964	28	2563163	5119504
915	25	20717726	23307037	965	28	5119505	7675846
916	25	23307038	25896349	966	28	7675847	10232188
917	25	25896350	28485661	967	28	10232189	12788530
918	25	28485662	31074973	968	28	12788531	15344872
919	25	31074974	33664285	969	28	15344873	17901214
920	25	33664286	36253597	970	28	17901215	20457556
921	25	36253598	38842909	971	28	20457557	23013898
922	25	38842910	41432221	972	28	23013899	25570240
923	25	41432222	44021516	973	28	25570241	28126582
924	26	6050	2592052	974	28	28126583	30682924
925	26	2592053	5178055	975	28	30682925	33239266
926	26	5178056	7764058	976	28	33239267	35795608
927	26	7764059	10350061	977	28	35795609	38351950
928	26	10350062	12936064	978	28	38351951	40908292
929	26	12936065	15522067	979	28	40908293	43464634
930	26	15522068	18108070	980	28	43464635	46020951
931	26	18108071	20694073	981	29	912	2599834
932	26	20694074	23280076	982	29	2599835	5198757
933	26	23280077	25866079	983	29	5198758	7797680
934	26	25866080	28452082	984	29	7797681	10396603
935	26	28452083	31038085	985	29	10396604	12995526
936	26	31038086	33624088	986	29	12995527	15594449
937	26	33624089	36210091	987	29	15594450	18193372
938	26	36210092	38796094	988	29	18193373	20792295
939	26	38796095	41382097	989	29	20792296	23391218
940	26	41382098	43968100	990	29	23391219	25990141
941	26	43968101	46554103	991	29	25990142	28589064
942	26	46554104	49140106	992	29	28589065	31187987
943	26	49140107	51726098	993	29	31187988	33786910
944	27	22287	2585656	994	29	33786911	36385833
945	27	2585657	5149026	995	29	36385834	38984756
946	27	5149027	7712396	996	29	38984757	41583679
947	27	7712397	10275766	997	29	41583680	44182602
948	27	10275767	12839136	998	29	44182603	46781525
949	27	12839137	15402506	999	29	46781526	49380448
950	27	15402507	17965876	1000	29	49380449	51979343

Anhang 3: signifikante *iHS*-Fenster ($eP \leq 0,01$)

Nach Rassen getrennte Auflistung aller signifikanten *iHS*-Fenster mit Fensternummern und zugehörigem *eP*-Wert in Klammern.

DFV	FGV	OBV	MWF	RH
187(0,009)	285(0,007)	22(0,004)	27(0,008)	116(0,007)
248(0,008)	370(0,009)	185(0,005)	342(0,005)	227(0,008)
294(0,001)	395(0,004)	224(0,007)	343(0,003)	591(0,009)
448(0,007)	484(0,001)	370(0,002)	491(0,001)	601(0,002)
463(0,003)	485(0,002)	373(0,008)	493(0,010)	602(0,006)
602(0,002)	580(0,005)	405(0,009)	662(0,007)	603(0,004)
736(0,005)	599(0,010)	500(0,003)	725(0,002)	605(0,005)
827(0,006)	603(0,008)	552(0,006)	745(0,006)	737(0,001)
848(0,004)	604(0,003)	786(0,010)	746(0,009)	762(0,003)
941(0,010)	606(0,006)	931(0,001)	995(0,004)	984(0,010)

BBV	BBB	GLW	ASV	IL
62(0,010)	2(0,004)	108(0,003)	34(0,007)	20(0,009)
165(0,004)	283(0,005)	371(0,010)	92(0,003)	21(0,001)
482(0,003)	316(0,001)	533(0,004)	229(0,009)	22(0,002)
492(0,007)	516(0,009)	602(0,009)	250(0,005)	24(0,003)
553(0,008)	695(0,008)	605(0,005)	251(0,002)	25(0,004)
678(0,002)	699(0,003)	677(0,007)	369(0,006)	26(0,007)
729(0,006)	826(0,006)	777(0,008)	466(0,004)	42(0,008)
775(0,005)	928(0,007)	863(0,002)	548(0,008)	281(0,010)
857(0,001)	975(0,010)	950(0,006)	665(0,010)	526(0,005)
876(0,009)	976(0,002)	955(0,001)	682(0,001)	731(0,006)

Anhang 4: signifikante *XP-EHH*-Fenster

Als signifikant gilt jedes Fenster, in dem mindestens drei der maximal neun möglichen *XP-EHH*-Vergleiche einer Rasse einen *eP*-Wert $\leq 0,01$ aufweisen. Sie sind getrennt nach den zehn verschiedenen Rassen im Folgenden aufgelistet.

Anhang 4. 1: Illyrisches Rind

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit IL mit Angabe von Chromosom (BTA) und Fensternummer (F)

BTA	F	<i>XP-EHH</i> der Rasse IL mit der jeweiligen Vergleichsrasse								
		DFV	FGV	OBV	MWF	RH	BBV	BBB	GLW	ASV
1	22	0,001	0,007	0,002			0,002	0,010		0,001
1	32	0,002		0,001			0,010			
15	655	0,003	0,003					0,001		
22	836	0,008	0,002			0,006				

Anhang 4. 2: Anatolisches SchwarzviehSignifikante *XP-EHH*-Fenster aller Rassevergleiche mit ASV mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse ASV mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	FGV	OBV	MWF	RH	BBV	BBB	GLW	IL
3	153			0,010		0,005		0,009		
3	168	0,002				0,001			0,006	0,001
6	283					0,002		0,002	0,009	0,008
10	486			0,009		0,009			0,002	
16	676	0,001	0,002	0,001	0,001				0,003	0,002
20	780	0,007	0,001				0,001	0,008		
22	859		0,005	0,003	0,003					
25	920	0,010		0,005			0,009			

Anhang 4. 3: Deutsches FleckviehSignifikante *XP-EHH*-Fenster aller Rassevergleiche mit DFV mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse DFV mit der jeweiligen Vergleichsrasse								
BTA	F	OBV	MWF	FGV	RH	BBV	GLW	BBB	ASV	IL
3	119	0,007				0,002	0,006			
5	244	0,004	0,005	0,002	0,008		0,004	0,006		0,008
5	245	0,003	0,007		0,010			0,007	0,010	0,009
6	282				0,001		0,002	0,003	0,005	0,003
6	283				0,002		0,001	0,008	0,002	0,002
6	284	0,002			0,003		0,003	0,004	0,001	0,001
6	294		0,002		0,005			0,002	0,007	0,007
6	295							0,010	0,006	0,004
6	296	0,001	0,001	0,001	0,004	0,005		0,001	0,003	0,005
22	848	0,008		0,005	0,009		0,008	0,009		
28	965		0,006	0,004		0,008				

Anhang 4. 4: Original BraunviehSignifikante *XP-EHH*-Fenster aller Rassevergleiche mit OBV mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse OBV mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	MWF	FGV	RH	BBV	GLW	BBB	ASV	IL
2	100		0,003	0,006		0,001	0,006	0,004		0,010
2	101		0,008	0,007	0,010	0,004	0,010			
5	232	0,007	0,010						0,009	
6	282				0,003		0,005	0,003	0,007	0,008
6	283				0,005		0,004		0,005	0,007
8	372	0,010			0,007	0,005			0,010	
11	516	0,004	0,006	0,003	0,006				0,004	0,003
11	517	0,002	0,002	0,002	0,002		0,001	0,002	0,002	0,002
11	518	0,001	0,001	0,001	0,001	0,002	0,002	0,001	0,001	0,001
11	519	0,003	0,007	0,004	0,004	0,010	0,009	0,005	0,003	0,005
26	932	0,006	0,004	0,005			0,007	0,006	0,006	0,004
26	933	0,005	0,005	0,010			0,003	0,007	0,008	0,006

Anhang 4. 5: Fränkisches Gelbvieh

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit FGV mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse FGV mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	MWF	OBV	RH	BBV	BBB	GLW	ASV	IL
3	119			0,002		0,001		0,002		
6	281	0,006		0,001	0,004		0,002	0,005	0,001	0,003
6	282				0,002		0,001	0,004	0,004	0,004
6	283				0,001			0,003	0,002	0,001
6	284			0,008	0,003			0,007	0,003	0,002
6	295						0,007		0,010	0,006
10	484	0,008	0,004			0,002	0,005			0,008
10	485	0,005	0,005		0,010					
14	601	0,001	0,001			0,005	0,003	0,010		
19	770			0,007	0,007		0,006			

Anhang 4. 6: Murnau-Werdenfelser

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit MWF mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse MWF mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	FGV	SOBV	RH	BBV	BBB	GLW	ASV	IL
3	128					0,002			0,006	0,004
6	282				0,003		0,001	0,002	0,003	0,003
6	283				0,001		0,003	0,001	0,002	0,001
6	284			0,001	0,002		0,002	0,003	0,001	0,002
11	491			0,005			0,009	0,005		
13	593	0,002	0,010		0,008		0,004		0,005	
15	658	0,010		0,010			0,010			
16	669			0,006					0,009	0,006
20	799		0,001			0,005		0,008		
20	800		0,008			0,001		0,007		0,005

Anhang 4. 7: Braunvieh (neue Zuchtrichtung)

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit BBV mit Angabe von Chromosom (BTA) und Fensternummer (F)

		<i>XP-EHH</i> der Rasse BBV mit der jeweiligen Vergleichsrasse								
BTA	F	DFV	MWF	FGV	OBV	RH	GLW	BBB	ASV	IL
4	170			0,010		0,007		0,006		
4	171			0,005	0,004	0,001	0,005	0,004		0,008
4	172	0,009	0,009	0,003	0,001	0,006	0,001	0,002		0,009
4	173			0,009	0,009	0,008	0,002	0,007		
5	227	0,002	0,010				0,004			
6	303	0,008	0,006					0,001	0,009	0,003
6	304	0,004	0,002	0,006	0,007	0,009		0,003	0,001	0,001
6	305	0,006	0,003	0,004	0,003			0,010	0,003	0,005
6	306	0,001		0,002	0,006				0,007	0,002
6	307	0,003		0,001	0,008	0,010			0,010	0,010
19	773	0,005		0,007		0,003				0,006

Anhang 4. 8: Red Holstein

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit RH mit Angabe von Chromosom (BTA) und Fensternummer (F)

BTA	F	<i>XP-EHH</i> der Rasse RH mit der jeweiligen Vergleichsrasse								
		DFV	MWF	FGV	SOBV	BBV	GLW	BBB	ASV	IL
9	442	0,008		0,010		0,002		0,006		0,009
9	443			0,007		0,010		0,007	0,010	0,010
18	731	0,004	0,005	0,005					0,004	0,008
18	732	0,003	0,001	0,003		0,003		0,003	0,003	0,003
18	733	0,001	0,002	0,001	0,002	0,005	0,001	0,001	0,001	0,001
18	734	0,002	0,010	0,002	0,004		0,002	0,005	0,002	0,002
18	735				0,007		0,009	0,002	0,009	
20	793		0,006	0,009	0,001	0,004	0,003		0,005	0,004
26	932	0,005	0,003				0,007	0,010	0,007	0,006
26	933	0,006	0,004				0,004			0,007
29	998	0,010	0,009		0,005				0,006	

Anhang 4. 9: Galloway

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit GLW mit Angabe von Chromosom (BTA) und Fensternummer (F)

BTA	F	<i>XP-EHH</i> der Rasse GLW mit der jeweiligen Vergleichsrasse								
		DFV	MWF	FGV	SOBV	RH	BBV	BBB	ASV	IL
1	58	0,007			0,007	0,010	0,008		0,008	
2	108		0,009	0,008				0,008		0,009
9	423	0,008		0,006	0,006	0,006			0,006	0,006
9	424	0,005	0,004	0,004	0,004	0,005	0,005		0,003	0,004
9	425	0,002	0,003	0,005	0,003	0,004	0,004	0,005	0,004	0,003
9	426	0,001	0,002	0,002	0,001	0,002	0,002	0,004	0,002	0,001
9	427	0,004	0,005	0,003	0,005	0,003	0,003	0,007	0,001	0,002
9	428	0,009		0,007						0,008
9	429			0,009		0,008	0,009			
10	455			0,010	0,008	0,007				
18	745						0,007	0,002	0,009	0,007
21	809	0,003	0,001	0,001	0,002	0,001	0,001	0,001	0,005	0,005

Anhang 4. 10: Weißblauer Belgier

Signifikante *XP-EHH*-Fenster aller Rassevergleiche mit BBB mit Angabe von Chromosom (BTA) und Fensternummer (F)

<i>XP-EHH</i> der Rasse BBB mit der jeweiligen Vergleichsrasse										
BTA	F	DFV	MWF	FGV	SOBV	RH	BBV	GLW	ASV	IL
2	66	0,008	0,005	0,009	0,006	0,002	0,008		0,007	0,005
3	148	0,005			0,010		0,006		0,009	
4	197	0,007	0,002	0,010		0,006	0,009		0,006	0,006
4	198	0,010	0,010	0,008		0,007				
5	226			0,003		0,008		0,004	0,005	0,007
7	331				0,005	0,005			0,008	
7	332				0,001			0,010	0,002	0,002
8	378	0,001		0,001	0,003	0,001	0,002	0,005	0,001	0,001
8	379	0,004				0,003			0,010	
8	382	0,002		0,002	0,002	0,004	0,001	0,003	0,003	0,004
8	383	0,006	0,008	0,004	0,004		0,004			0,008
11	503		0,009	0,006	0,008					
14	610		0,006				0,010			0,009
16	690	0,003	0,001	0,005			0,003		0,004	0,003

Anhang 5: Alignment des Teilbereichs des *KDR*-Gens der acht sequenzierten Tiere

```

FV0800_KDR-F -----ATGCCACC
FV2700_KDR-F -----ATGCCACC
FV0200_KDR-F -----ATGCCACC
RH124_KDR-F -----ATGCCACC
RH106_KDR-F -----ATGCCACC
FV0239_KDR-F -----ATGCCACC
RH123_KDR-F -----ATGCCACC
RH117_KDR-F -----ATGCCACC
FV0800_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
FV2700_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
FV0200_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
RH117_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
FV0239_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
RH106_KDR-R -GCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
RH124_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
RH123_KDR-R TGCGATTTACCCACACCTGTCTGCAAGAACTTGGATGCTCTTTGGAAAATGAATGCCACC
*****

FV0800_KDR-F ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCATCCTTGCAG
FV2700_KDR-F ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCATCCTTGCAG
FV0200_KDR-F ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCATCCTTGCAG
RH124_KDR-F ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCTTGCAG
RH106_KDR-F ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCTTGCAG
FV0239_KDR-F ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCTTGCAG
RH123_KDR-F ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCTTGCAG
RH117_KDR-F ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCTTGCAG
FV0800_KDR-R ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCAT-----
FV2700_KDR-R ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCAT-----
FV0200_KDR-R ATGAACTCTAACGGCACAATGACATCTTGATCTTGGAGCTGCAGAATGCATCCT-----
RH117_KDR-R ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAG-----
FV0239_KDR-R ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATC-----
RH106_KDR-R ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCAT-----
RH124_KDR-R ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATC-----
RH123_KDR-R ATGAACTCTAACGGCACAATGACRTCTTGATCTTGRAGCTGCAGAAAYGCATCCT-----
*****

FV0800_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
FV2700_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
FV0200_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
RH124_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
RH106_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
FV0239_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
RH123_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
RH117_KDR-F GACCAAGGAGACTATGTCTGCTTTGCTCAGGACAGGAAGA
FV0800_KDR-R -----
FV2700_KDR-R -----
FV0200_KDR-R -----
RH117_KDR-R -----
FV0239_KDR-R -----
RH106_KDR-R -----
RH124_KDR-R -----
RH123_KDR-R -----

```

Danksagung

Herrn Univ.-Prof. Dr. Dr. habil. Martin Förster danke ich herzlichst für die Überlassung des interessanten Themas sowie die Durchsicht der Arbeit.

Herrn Dr. Ivica Međugorac danke ich für die geduldige Einführung und freundliche Unterstützung, die konstruktiven Ratschläge, die fachliche Hilfe sowie die kritische Durchsicht des Manuskripts.

Allen wissenschaftlichen Assistenten (Dr. Jan Maxa, Dr. Jennifer Schär) und wissenschaftlichen Mitarbeitern (Ina Vogl, Janey Heine, Ashraf Awad, Markus Neuditschko, Verena Schneider, Monika Höpfl), welche ich in den vergangenen zwei Jahren mehr oder weniger lange meine Kolleginnen und Kollegen nennen durfte, danke ich für die guten fachlichen Gespräche, die lustige Zeit und die gelegentlich nötigen aufbauenden Worte. *Dies gilt insbesondere für dich Verena, da du immer ein offenes Ohr für mich hattest und immer an mich geglaubt hast.*

Dem technischen Personal (Tatiana Schröter, Renate Damian, Gisela Andorfer, Heike Zierahn, Uschi Werner, Martin Dinkel) danke ich für die stets gewährte Hilfe bei Laborfragen und das nette Arbeitsklima.

Bei Frau Dr. Claudia Veit-Kensch bedanke ich mich für die Durchsicht des Summaries.

Den folgenden Institutionen danke ich für die gute Kooperation respektive die Bereitstellung von Probenmaterial:

- Laboratory for Functional Genome Analysis (LAFUGA) der LMU München, insbesondere Herrn Dr. Stefan Krebs
- Tierzuchtforschung e.V. München
- Der bayerischen Genreserve, LfL, Grub sowie allen Züchtern, Zuchtverbänden und Besamungsstationen, die zum Zustandekommen der umfangreichen Probensammlung beigetragen haben.

Besonderer Dank gilt darüber hinaus der Deutschen Forschungsgemeinschaft (DFG), die durch ihre finanzielle Unterstützung diese Untersuchungen mit ermöglichte (Projektnummer: ME 3404/2-1).

Meiner Großtante Thea Mayr danke ich für die finanzielle Unterstützung sowohl während des Studiums als auch in der Zeit der Doktorarbeit.

Meiner Freundin Nicole Knak danke ich für die seelische Unterstützung und den schier endlosen Optimismus.

Der größte Dank jedoch gilt meiner Familie, insbesondere meinen Eltern Waltraud und Alois Rothammer, ohne deren Motivation und Zuspruch, das in mich gesetzte Vertrauen und nicht zuletzt die finanzielle Unterstützung, weder mein Tiermedizinstudium noch die Fertigstellung dieser Arbeit möglich gewesen wären.
DANKESCHÖN!