

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Tierzucht

**Die Analyse von genomweiten Re-Sequenzierungsdaten zur
Identifizierung kausaler DNA-Variation beim Rind**

Sandra Kipp

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für
Ernährung, Landnutzung und Umwelt der Technischen Universität München zur
Erlangung des akademischen Grades eines

Doktors der Agrarwissenschaften (Dr. agr.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. H. Luksch

Prüfer der Dissertation: 1. Univ.-Prof. Dr. H. R. Fries
2. Univ.-Prof. Dr. T. Meitinger

Die Dissertation wurde am 19.01.2015 bei der Technischen Universität München
eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für
Ernährung, Landnutzung und Umwelt am 24.06.2015 angenommen.

Inhaltsverzeichnis

Abbildungsverzeichnis.....	V
Tabellenverzeichnis.....	VII
Abkürzungen.....	VIII
1 Einleitung.....	1
2 Literaturübersicht.....	3
2.1 Identifizierung kausaler Varianten.....	3
2.1.1 Phänotypbasierter Ansatz.....	4
2.1.2 Genotypbasierter Ansatz.....	5
2.1.2.1 Missing homozygotes.....	5
2.1.2.2 Vorzeitige Stoppmutationen.....	5
2.2 Nutzen genomweiter Re-Sequenzierungsdaten.....	7
2.3 Merkmal „Totgeburten“ beim Rind.....	8
3 Tiere, Material und Methoden.....	10
3.1 Genomweite Re-Sequenzierung mit NGS.....	10
3.1.1 Identifizierung von Schlüsseltieren im Deutschen Fleckvieh.....	10
3.1.2 Auswahl weiterer Tiere.....	10
3.1.3 Quantifizierung und Qualitätskontrolle der DNA-Proben.....	11
3.1.4 Generierung der Sequenzdaten.....	11
3.1.5 Identifizierung von Varianten.....	12
3.1.6 Evaluierung der Varianten.....	12
3.1.7 Funktionelle Annotation NGS-basierter Varianten.....	13
3.2 Genomweite Assoziationsstudien.....	14
3.2.1 Genotypdaten und Imputation.....	14
3.2.2 Phänotyp „direkte Totgeburten bei Kühen“.....	15
3.2.3 Durchführung der genomweiten Assoziationsstudie.....	16
3.2.4 Abschätzung der von den Polymorphismen erklärten Varianz	16
3.2.5 Berechnung des Kopplungsungleichgewichts.....	17
3.3 In-silico-Analysen.....	17
3.3.1 Auswahl und Annotation der Kandidatengene.....	17
3.3.2 Charakterisierung der Kandidatengene.....	17
3.3.3 Haplotypenanalyse der Stoppmutationen.....	18
3.3.4 Überprüfung von Mapping-Fehlern.....	18
3.4 PCR.....	18
3.4.1 Primerdesign.....	18
3.4.2 Standard-PCR.....	19
3.4.3 PCR-Optimierung.....	19
3.5 Sanger-Re-Sequenzierung.....	20
3.5.1 Aufreinigung der PCR.....	20
3.5.2 Sequenzierreaktion, Aufreinigung und Kapillargelelektrophorese.....	20

3.5.3	Analyse der <i>Sanger-Re-Sequenzierung</i>	21
4	Ergebnisse	22
4.1	Abschätzung der genomischen Variation.....	22
4.1.1	Auswahl der Schlüsseltiere im Deutschen Fleckvieh für die genomweite Re-Sequenzierung	22
4.1.2	Allgemeine Daten der genomweiten Re-Sequenzierung.....	23
4.1.3	Identifizierung NGS-basierter Varianten.....	24
4.1.4	Evaluierung NGS-basierter Varianten.....	26
4.1.5	Funktionelle Annotation NGS-basierter Varianten.....	27
4.1.6	Identifizierung bekannter kausaler Varianten.....	29
4.2	Verifizierung vorzeitiger Stopp-Codons.....	33
4.2.1	Potentielle Auswirkungen vorzeitiger Stopp-Codons auf die Proteinsequenz	35
4.2.2	Allelfrequenzverteilung vorzeitiger Stopp-Codons.....	38
4.2.3	Hinweise vorzeitiger Stopp-Codons auf den Phänotyp.....	41
4.2.3.1	Stoppmutation in <i>VCAMI</i>	46
4.2.3.2	Stoppmutation in <i>FREMI</i>	48
4.2.3.3	Stoppmutation in <i>STRC</i>	50
4.2.3.4	Stoppmutation in <i>TAS2R38</i>	52
4.2.4	Vorzeitige Stopp-Codons im Deutschen Fleckvieh.....	53
4.2.4.1	Vorzeitige Stopp-Codons ohne homozygote Genotypen im Deutschen Fleckvieh.....	54
4.2.4.2	Vorzeitige Stopp-Codons im Deutschen Fleckvieh mit Abweichung vom Hardy-Weinberg-Gleichgewicht.....	55
4.2.4.2.1	Stoppmutation in <i>PNLIP</i>	56
4.2.4.2.2	Stoppmutation in <i>QRFPR</i>	58
4.2.4.2.3	Stoppmutation in <i>VTCNI</i>	61
4.3	Sequenzbasierte genomweite Assoziationsstudie für das Merkmal direkte Totgeburten bei Holstein-Kühen.....	64
4.3.1.1	QTL auf Chromosom 5 (56,8 bis 57,9 Mb).....	65
4.3.1.2	QTL auf Chromosom 5 (105,7 bis 106,8 Mb).....	67
4.3.1.3	QTL auf Chromosom 7.....	70
4.3.1.4	QTL auf Chromosom 17.....	72
4.3.1.5	QTL auf Chromosom 18 (14,5 bis 15,6 Mb).....	76
4.3.1.6	QTL auf Chromosom 18 (59 bis 60,1 Mb).....	79
4.3.1.7	QTL auf Chromosom 29.....	82
5	Diskussion	85
5.1	Sichtung der genomischen Variation.....	85
5.2	Identifizierung kausaler Mutationen.....	87
5.2.1	Vorzeitiger Abbruch der Translation.....	88
5.2.1.1	Validierung der Auswirkungen der Stoppmutationen.....	94
5.2.2	Kausalität der identifizierten QTL für das Merkmal direkte Totgeburten bei Kühen... ..	96
5.2.3	Vergleich der identifizierten QTL mit anderen Studien.....	98
5.3	Herausforderungen bei der Identifizierung kausaler Varianten.....	102
5.3.1	Dichte der Markerkarte.....	102

5.3.2	Qualität der Phänotypinformationen.....	105
5.3.3	Effekt der Polymorphismen.....	106
5.3.4	Qualität der Imputation.....	107
5.3.5	Falsch-positive Varianten.....	110
5.3.6	Qualität des UMD3.1-Assemblys.....	110
5.4	Anwendung in genomischer Selektion.....	112
6	Zusammenfassung.....	114
7	Summary.....	116
8	Literaturverzeichnis.....	118
9	Anhang.....	150

Abbildungsverzeichnis

Abbildung 3.1.1: Berechnung für Übereinstimmung, NRS- und NRD-Rate der Genotypen.....	13
Abbildung 3.2.1: Verteilung des Merkmals direkte Totgeburten bei Kühen	15
Abbildung 4.1.1: Anteil der erklärten genetischen Variation durch die Re-Sequenzierung von 107 BFV-Tieren.....	22
Abbildung 4.1.2: Durchschnittliche Abdeckung der re-sequenzierten 214 Tiere.....	23
Abbildung 4.1.3: Verteilung der Genotypqualität vor und nach der Imputation mit <i>Beagle</i>	24
Abbildung 4.1.4: Anzahl entdeckter Varianten in Abhängigkeit zur Genomabdeckung.....	25
Abbildung 4.1.5: Übereinstimmung der Genotypen je Chromosom.....	26
Abbildung 4.1.6: Übereinstimmung, NRD und NRS zwischen NGS-basierten Genotypen und HD- Genotypen.....	27
Abbildung 4.1.7: Funktionelle Annotation der Varianten.....	28
Abbildung 4.2.1: Multiple Varianten innerhalb eines Codons.....	34
Abbildung 4.2.2: Nonsense-vermittelter mRNA-Abbau in Säugetieren.....	36
Abbildung 4.2.3: Anteil der Verkürzung der kodierenden Sequenz durch vorzeitige Stopp-Codons	37
Abbildung 4.2.4: Verteilung der Allelfrequenzen vorzeitiger Stoppmutationen im Vergleich zu nicht- synonymen Aminosäureaustauschen.....	38
Abbildung 4.2.5: Verteilung der Allelfrequenzen vorzeitiger Stoppmutationen zwischen den Rassen	39
Abbildung 4.2.6: Verteilung der Allelfrequenzen von vorzeitigen Stoppmutationen mit unterschiedlichen Genotypen.....	41
Abbildung 4.2.7: Genstruktur von <i>VCAM1</i>	47
Abbildung 4.2.8: Transmembranstruktur von <i>VCAM1</i>	48
Abbildung 4.2.9: Genstruktur von <i>FREM1</i>	49
Abbildung 4.2.10: Proteinstruktur von <i>FREM1</i>	50
Abbildung 4.2.11: Genstruktur von <i>STRC</i>	51
Abbildung 4.2.12: Genstruktur von <i>TAS2R38</i>	52
Abbildung 4.2.13: Transmembranstruktur von <i>TAS2R38</i>	53
Abbildung 4.2.14: Genstruktur von <i>PNLIP</i>	57
Abbildung 4.2.15: Peptidstruktur von <i>PNLIP</i>	58
Abbildung 4.2.16: Genstruktur von <i>QRFPR</i>	59
Abbildung 4.2.17: Transmembranstruktur von <i>QRFPR</i>	60
Abbildung 4.2.18: Genstruktur von <i>VTCN1</i>	61
Abbildung 4.2.19: Transmembranstruktur von <i>VTCN1</i>	62
Abbildung 4.3.1: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen.....	64
Abbildung 4.3.2: Region-Plot auf Chromosom 5 für direkte Totgeburten bei Holsteinkühen (56,8- 57,9Mb).....	66
Abbildung 4.3.3: Region-Plot auf Chromosom 5 für direkte Totgeburten bei Holsteinkühen (105,7- 106,8Mb).....	67
Abbildung 4.3.4: Region-Plot auf Chromosom 7 für direkte Totgeburten bei Holsteinkühen.....	70
Abbildung 4.3.5: Region-Plot auf Chromosom 17 für direkte Totgeburten bei Holsteinkühen.....	73

Abbildung 4.3.6: Region-Plot auf Chromosom 18 für direkte Totgeburten bei Holsteinkühen (14,5-15,6Mb).....	76
Abbildung 4.3.7: Region-Plot auf Chromosom 18 für direkte Totgeburten bei Holsteinkühen (59-60,1Mb).....	79
Abbildung 4.3.8: Region-Plot auf Chromosom 29 für direkte Totgeburten bei Holsteinkühen.....	82
Abbildung 5.2.1: Benötigte Tieranzahl zur Überprüfung der Stoppmutationen.....	95
Abbildung 5.2.2: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen auf Chromosom 14.....	100
Abbildung 5.2.3: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Färsen.....	101
Abbildung 5.3.1: Region-Plot auf Chromosom 5 für den direkten Kalbeverlauf bei Kühen.....	104
Abbildung 5.3.2: Ergebnisse der Prinzipalkomponentenanalyse innerhalb der Holsteintiere.....	109
Abbildung 9.1: Bestimmung der Grenze zwischen pseudoautosomaler und nicht-pseudoautosomaler Region des bovinen X-Chromosoms.....	163
Abbildung 9.2: Protein-BLAST-Ergebnis für LOC787812.....	179
Abbildung 9.3: Manhattan-Plot der Ergebnisse der HD-basierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen.....	180

Tabellenverzeichnis

Tabelle 4.1: Identifizierte OMIA-Mutationen in den Rassen.....	29
Tabelle 4.2: Genotypen der Stoppmutation in <i>SHBG</i>	31
Tabelle 4.3: Identifizierte QTN in den Rassen.....	33
Tabelle 4.4: Anzahl homozygoter und nicht-homozygoter Genotypen.....	40
Tabelle 4.5: Stoppmutationen in Genen mit bekanntem Phänotyp.....	42
Tabelle 4.6: Übersicht der Stoppmutationen mit fehlenden Homozygoten im Deutschen Fleckvieh	55
Tabelle 4.7: Auflistung der signifikantesten Polymorphismen für das Merkmal direkte Totgeburten bei Holsteinkühen.....	65
Tabelle 4.8: Signifikante Polymorphismen auf Chromosom 5 (56,8 bis 57,9 Mb).....	66
Tabelle 4.9: Polymorphismen im exonischen Bereich von <i>FGF23</i>	69
Tabelle 4.10: Polymorphismen im exonischen Bereich von <i>ARRDC3</i>	72
Tabelle 4.11: Signifikante Polymorphismen auf Chromosom 17 (72,7 bis 73,8 Mb).....	75
Tabelle 4.12: Polymorphismen im exonischen Bereich von <i>VPS35</i>	77
Tabelle 4.13: Polymorphismen im exonischen Bereich von <i>LOC787812</i>	81
Tabelle 4.14: Polymorphismen im exonischen Bereich von <i>H19</i>	83
Tabelle 9.1: Ausgewählte Tiere zur genomweiten Re-Sequenzierung.....	150
Tabelle 9.2: Verteilung der Readlängen je Tier [in %].....	154
Tabelle 9.3: Mapping Statistik.....	159
Tabelle 9.4: Anzahl Varianten je Chromosom.....	164
Tabelle 9.5: Übersicht der Allelaustausche.....	164
Tabelle 9.6: Übersicht der verifizierten vorzeitigen Stopp-Codons.....	165
Tabelle 9.7: Übersicht über die bei der Sanger-Re-Sequenzierung verwendeten Primer.....	167
Tabelle 9.8: Identifizierte Varianten im <i>FGF23</i> -Gen.....	169
Tabelle 9.9: Identifizierte Varianten im <i>ARRDC3</i> -Gen.....	171
Tabelle 9.10: Identifizierte Varianten im <i>VPS35</i> -Gen.....	173
Tabelle 9.11: Identifizierte Varianten im <i>LOC787812</i> -Gen.....	177

Abkürzungen

54K	BovineSNP50 BeadChip
777K	BovineHDBeadChip
A (Allel)	Adenin
A (Aminosäure)	Alanin
ABCG2	ATP-binding cassette sub-family G member 2 kodierendes Gen
AF	Allelfrequenz
ALOX15	Arachidonate 15-Lipoxygenase kodierendes Gen
ALT	Alternativallel
ANG	Angus
APAF1	Apoptotic protease activating factor 1 kodierendes Gen
ARRDC3	Arrestin domain-containing 3 kodierendes Gen
ARRDC3	Arrestin domain-containing 3
AS	Aminosäure
ATP	Adenosintriphosphat
AYR	Ayrshire
BAQ	base alignment quality
BBL	Belgian Blue
BBO	Beef Booster
BCO2	Beta-carotene oxygenase 2 kodierendes Gen
BDKRB2	Bradykinin Rezeptor B2 kodierendes Gen
beob	beobachtet
BFV	Deutsches Fleckvieh
BLAST	basic local alignment search tool
BLUP	Best-Linear-Unbiased-Prediction
BOSTAUV1R430	Vomeronal 1 Rezeptor bosTauV1R430 kodierendes Gen
bp	Basenpaar
BSW	Brown Swiss
BTNL2	Butyrophilin-like 2 kodierendes Gen
BWA	Burrows-Wheeler-Alignment Tool
C (Allel)	Cytosin
C (Aminosäure)	Cystein
CASP12	Caspase 12 kodierendes Gen
CBX3	Chromobox homolog 3 kodierendes Gen
CHA	Charolais
CHCHD10	Coiled-coil-helic-coiled-coil-helix domain containing 10 kodierendes Gen
Chr	Chromosom
CNDP1	Carnosine Deipeptidase 1 kodierendes Gen
DDX58	DEAD (Asp-Glu-Ala-Asp) Box Polypeptid 58 kodierendes Gen
DERL3	Derlin 3 kodierendes Gen
DGAT1	Diaglycerol O Transferase
DHV	Deutscher Holstein Verband
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DNASE1	Desoxyribonuclease 1 kodierendes Gen
dNTP	Desoxy-Nukleotidtriphosphat
dSBc	direct stillbirth cow
DSG2	Desmoglein 2 kodierendes Gen
DYTN	Dystrotelin kodierendes Gen
EDTA	Ethylendiamintetraacetat
ENOSF1	Enolase Superfamily Member 1 kodierendes Gen
ERAP1	Endoplasmatic reticulum aminopeptidase 1 kodierendes Gen
erw	erwartet

EST	expressed sequence tags
F	Phenylalanin
<i>FGF23</i>	Fibroblast Growth Factor kodierendes Gen
FGF23	Fibroblast Growth Factor
<i>FREM1</i>	FRAS1 related extracellular matrix 1 kodierendes Gen
FREM1	FRAS1 related extracellular matrix 1
g	Gramm
G (Allel)	Guanin
GCO	Guelph Composite
<i>GHR</i>	Growth Hormone Receptor
<i>GIP</i>	Gastric inhibitory polypeptide kodierendes Gen
<i>GLRA3</i>	Glycin receptor, alpha 3 subunit kodierendes Gen
GLUT11	Glukosetransporter 11
GPI	Glykosylphosphatidylinositol
GPR103	G-Protein gekoppelter Rezeptor 103
<i>GPR128</i>	G-Protein gekoppelter Rezeptor 128 kodierendes Gen
<i>GPR75</i>	G-Protein gekoppelter Rezeptor 75 kodierendes Gen
<i>GSTM1</i>	Glutathione S-Transferase 1 kodierendes Gen
GWAS	Genomweite Assoziationsstudie
<i>GYLTL1B</i>	Glycosyltransferase-like 1B kodierendes Gen
h	Stunden
h2	Heritabilität
HD	High-density
HER	Hereford
HH1	Holstein Haplotyp 1
HH2	Holstein Haplotyp 2
HH3	Holstein Haplotyp 3
HOL	Holstein, Farbrichtung schwarzbunt
HWG	Hardy-Weinberg-Gleichgewicht
<i>IFI44</i>	Interferon-induced protein 44 kodierendes Gen
Ig	Immunglobulin
IGF2	Insulin-ähnlicher Wachstumsfaktor 2
<i>IGF2</i>	Insulin-ähnlicher Wachstumsfaktor 2 kodierendes Gen
IGV	integrative genome viewer
<i>IL19</i>	Interleukin 19 kodierendes Gen
InDel	Insertionen und Deletionen
<i>INMT</i>	Indolethylamine N-Methyltransferase kodierendes Gen
<i>ITGB4</i>	Integrin-beta-4
<i>IVL</i>	Involucrin kodierendes Gen
JER	Jersey
kb	Kilobasen
<i>KCNB1</i>	Potassium voltage gated channel, Shab-related subfamily, member 1 kodierendes Gen
<i>KDSR</i>	3-Ketodihydrosphingosine kodierendes Gen
KRAB	Krüppel-assoziierte Box
<i>KRT74</i>	Keratin 74 kodierendes Gen
L	Leuzin
LD	Linkage Disequilibrium
LIM	Limousin
LoF	loss of function
<i>LRP2BP</i>	LRP2 bindendes Protein kodierendes Gen
<i>LYSB</i>	Lysozym B kodierendes Gen
MAF	minor allele frequency
<i>MANBA</i>	Mannosidose, beta A, lysosomal kodierendes Gen
MAX	Maximum

Mb	Megabasen
MC1R	Melanocortinrezeptor kodierendes Gen
mg	Milligramm
mgCl2	Magnesiumchlorid
MGI	Mouse Genome Informatics
MH1	Montbéliarde Haplotyp 1
MIF	Macrophage migration inhibitory factor kodierendes Gen
min	Minute
min	Minimum
MITF	Microphthalmia-associated transcription factor kodierendes Gen
mM	Millimol
MMP11	Matrix Metalloproteinase 11 kodierendes Gen
MMP11	Matrix Metalloproteinase 11
MOCS1	Molybdenum cofactor synthesis 1 kodierendes Gen
Mrd	Milliarden
mRNA	messenger RNA
N	Asparagin
NCAPG	Non-SMC condensin I complex, subunit G kodierendes Gen
NCBI	National Center for Biotechnology Information
Ne	effektive Populationsgröße
NFC	Nordic Finncattle
ng	Nanogramm
NGS	Next Generation Sequencing
NMD	Nonsense-mediated mRNA decay
non PAR	Nicht-pseudoautosomale Region
non-syn	nicht-synonymer Austausch
NPL	N-Acetylneuraminat pyruvat Lyase kodierendes Gen
NPT2	Natrium/Phosphat Transporter II
NPY5R	Neuropeptid Y Rezeptor Y5 kodierendes Gen
NRD	Non-Reference-Discrepancy
NRS	Non-Reference-Sensitivity
OMIA	Online Mendelian Inheritance in Animals
OMIM	Online Mendelian Inheritance in Man
P	Prolin
PAR	pseudoautosomale Region des X-Chromosoms
PCR	Polymerasekettenreaktion
PIE	Pietmonteser
PLA2G3	Phospholipase A2, Group III kodierendes Gen
PLAT	Polycystin-1 Lipoxygenase Alpha-Toxin
PNLIP	Pankreas Lipase kodierendes Gen
PNLIP	Pankreas Lipase
POLI	Polymerase iota kodierendes Gen
PRKDC	Protein Kinase, DNA aktiviert, katalytisches Polypeptid kodierendes Gen
PRKG2	Protein Kinase, cGMP-abhängig, Typ II kodierendes Gen
PRLR	Prolaktin Rezeptor kodierendes Gen
PROKR1	Prokinectin Rezeptor 1 kodierendes Gen
Q	Glutamin
QRFP	Pyroglutamylated Rfamidpeptidrezeptor kodierendes Gen
QTL	quantitative trait loci
QTN	quantitative trait nucleotide
QUAL	SNP-Qualität
R	Arginin
RASGRP2	RAS, Guanin freisetzendes Protein 2 kodierendes Gen
RDC	Red Dairy Cattle

RED	Holstein, Farbrichtung rotbunt
REF	Referenzallel
RNA	Ribonukleinsäure
s	Sekunde
s	Serin
SAM	Sequence Alignment Map Format
SHBG	Sex hormone binding globulin kodierendes Gen
SIM	Simmental
SLC2A1	Glukosetransporter 1 kodierendes Gen
SLC2A11	Glukosetransporter 11 kodierendes Gen
SLC35A3	Solute carrier 35, member A3 kodierendes Gen
SMA	spinale muskuläre Atrophie
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 kodierendes Gen
SMC2	structural maintenance of chromosomes 2 kodierendes Gen
SNP	single nucleotide polymorphism
STRC	Stereozilin kodierendes Gen
STRC	Stereozilin
SULT6B1	Sulfotransferase family, cytosolic, 6B. Member 1 kodierendes Gen
syn	synonymer Austausch
T (Allel)	Thymin
T (Aminosäure)	Threonin
Taq	Thermus aquaticus
TAS2R38	Taste receptor, type 2, member 38 kodierendes Gen
TAS2R38	Taste receptor, type 2, member 38
TBE	Tris-Borat-EDTA-Puffer
Ti	Transition
TIGAR	TP53-induced glycolysis and apoptosis regulator kodierendes Gen
TMED2	Transmembran emp24 domain trafficking protein 2 kodierendes Gen
TMEM95	Transmembranprotein 95
Tris	Tris(hydroxymethyl)aminoethan
Tv	Transversion
U	Unit
UTR	untranslatierter Bereich
V	Volt
VCAM1	Vascular cell adhesion molecule 1 kodierendes Gen
VCAM1	Vascular cell adhesion molecule 1
VPREB3	Pre-B lymphocyte 3 kodierendes Gen
VPS35	Vacuolar protein sorting 35 homolog kodierendes Gen
VPS35	Vacuolar protein sorting 35 homolog
VTCN1	V-Set domain containing T cell activation inhibitor 1 kodierendes Gen
VTCN1	V-Set domain containing T cell activation inhibitor 1
W	Tryptophan
WDR8	WD repeat containing, antisense to TP73 kodierendes Gen
X	Stopp-Codon
Y	Tyrosin
ZFP	Zinkfingerprotein
ZNF70	Zinkfingerprotein 79 kodierendes Gen
ZNF845	Zinkfingerprotein 845
α	Allelsubstitutionseffekt
μl	Mikroliter
μM	Mikrogramm

1 Einleitung

Die Verfügbarkeit moderner Methoden zur Genotypisierung genomweiter Polymorphismen bei Nutztieren trägt nicht nur zur genomischen Leistungsvorhersage in Form von Zuchtwerten, sondern auch zur Identifizierung kausaler Regionen bei. Basierend auf dem Kopplungsungleichgewicht zwischen den verwendeten Markern und der kausalen Variante können solche Regionen detektiert werden, ohne die dem Merkmal zugrunde liegende Mutation zu kennen. Für eine optimale Nutzung dieser genomischen Informationen in Zuchtprogrammen ist jedoch das Wissen über die ursächliche Variante notwendig. Die Kenntnis über alle variablen Positionen im Genom kann den Vorgang zur Identifizierung der kausalen Mutation mit Hilfe vorhandener Genomsequenzen beschleunigen.

Für eine vollständige und zuverlässige Aufdeckung der kompletten genomischen Variation ist die Verfügbarkeit populations- und genomweiter Sequenzinformation erforderlich. Die Re-Sequenzierung von Genomen der gesamten Population ist aufgrund der hohen Kosten bisher jedoch noch nicht realisierbar. Eine kosteneffiziente Methodik zur Abfrage der populationsweiten Varianten ist die Re-Sequenzierung von sogenannten Schlüsseltieren, die einen Großteil der genetischen Diversität einer Population erklären. Diese können sich bei Populationen mit geringer effektiver Populationsgröße auf wenige Tiere beschränken. Sobald die Sequenzen ausgewählter Schlüsseltiere vorliegen, kann die Genomsequenz über vorhandene Genotypdaten eines SNP-Chips mittels Imputation für die Nachkommen abgeleitet werden. Eine geringe bis mittlere Abdeckung des re-sequenzierten Genoms der Schlüsseltiere ist für eine genaue Abfrage der Varianten und den entsprechenden Genotypen ausreichend. Auch seltene Varianten, die bei komplexen Merkmalen eine bedeutende Rolle spielen, können abgefragt werden.

Innerhalb der genomischen Variation kommen auch solche Mutationen vor, die zu einer frühzeitigen Beendigung der Translation eines Gens führen und somit ein verkürztes Protein entstehen lassen. Einige dieser Varianten können einen unbedeutenden oder auch einen vorteilhaften Effekt haben, sodass diese Mutationen eine höhere Allelfrequenz aufweisen oder womöglich in der Population fixiert sind. Andere hingegen führen zu einer enormen Beeinträchtigung der Funktion des verkürzten Proteins. Die Proteinfunktion ist möglicherweise vollständig verloren. Diese als vorzeitige Stoppmutationen bezeichneten Varianten sind in vielen Fällen ursächlich für Erkrankungen, die nach Mendelschen Regeln vererbt werden. Aber auch bei komplexen

Merkmalsausprägungen nehmen sie eine bedeutende Rolle ein. Aufgrund des negativen Selektionsdrucks sind diese schädlichen Mutationen in einer Population selten vorzufinden. Das Fehlen von Mutationen im homozygoten Zustand deutet auf eine letale Wirkung dieser Variante hin.

Für das komplexe Merkmal Totgeburten sind bereits verschiedene solcher Varianten beim Rind bekannt. Die Totgeburtenrate ist in der Rinderzucht ein ökonomisch wichtiges Merkmal, da der Verlust eines Kalbes mit hohen Kosten verbunden ist. Neben zahlreichen Umweltbedingungen wird die Totgeburtenrate von genetischen Faktoren beeinflusst, deren Aufklärung zu einer Verbesserung der Überlebensrate führen kann. Oft werden Totgeburten durch eng korrelierte Merkmale begründet, wie z.B. ein erhöhtes Geburtsgewicht des Kalbes und damit einhergehende Schwierigkeiten bei der Kalbung. Im Genom des Rindes existieren jedoch auch Regionen, die das Auftreten von Totgeburten direkt beeinflussen. Neben autosomal-rezessiv vererbten Defekten wird das komplexe Merkmal Totgeburten von weiteren Genorten bestimmt.

Das Ziel dieser Dissertation war es, auf Basis genomweiter Re-Sequenzierungsdaten verschiedener Rinderpopulationen kausale Mutationen für verschiedene Phänotypausprägungen zu identifizieren. Dabei wurden potentielle Stoppmutationen systematisch analysiert sowie eine sequenzbasierte genomweite Assoziationsstudie für das Merkmal Totgeburtenrate bei der Rasse Holstein durchgeführt.

2 Literaturübersicht

2.1 Identifizierung kausaler Varianten

Die Identifizierung von kausalen Varianten ist notwendig, um diese Information optimal in einem Zuchtprogramm nutzen zu können (MEUWISSEN und GODDARD 2010; EICHLER *et al.* 2010). Die Kenntnis über den Trägerstatus eines Tieres bzgl. eines Gendefekts kann die Anpaarung von heterozygoten Tieren verhindern und somit die Geburt von Kälbern mit Erbdefekten vermeiden. Außerdem kann durch die Implementierung der kausalen Mutationen bei der genomischen Leistungsvorhersage eine höhere Genauigkeit und ein größerer Zuchtfortschritt erreicht werden (MEUWISSEN und GODDARD 2010).

In den vergangenen Jahren wurden in zahlreichen Studien beim Rind kausale Regionen oder auch Mutationen für phänotypische Merkmale identifiziert (z.B. WINTER *et al.* 2002; PAUSCH *et al.* 2012, 2014; ONLINE MENDELIAN INHERITANCE IN ANIMALIS, 2014). Die Kartierung dieser Regionen begann in Rinderpopulationen zunächst mit Hilfe des „Granddaughter Designs“ (WELLER *et al.* 1990). Die dafür verwendeten Mikrosatelliten wiesen jedoch zu große Abstände zueinander auf, sodass der ursächliche Bereich nur bedingt eingegrenzt werden konnte (VISSCHER und GODDARD 2004). Die Verfügbarkeit der bovinen Referenzsequenz (THE BOVINE HAPMAP CONSORTIUM 2009) ermöglichte die systematische Erfassung von Sequenzpolymorphismen beim Rind, insbesondere von Einzelnukleotidaustauschen (*Single Nucleotide Polymorphism*, SNP). Mit Hilfe der Positionsinformation eines Polymorphismus können dank verschiedenster Methoden die kausalen Regionen präzise lokalisiert werden. Der rasante Technologiefortschritt zur Abfrage von SNP-Genotypen (ECK *et al.* 2009; MATUKUMALLI *et al.* 2009) sowie die Erstellung und kontinuierliche Verbesserung der Genomannotation (ZIMIN *et al.* 2009) revolutionierte die Entdeckung kausaler Mutationen (NICHOLAS und HOBBS 2014). Einen bedeutenden Einfluss auf die erfolgreiche Identifizierung der kausalen Regionen hat die Anzahl der untersuchten Individuen sowie die Dichte der verwendeten Marker (RISCH 2000; MATUKUMALLI *et al.* 2009). Zur Erreichung einer höheren Markerdichte für möglichst viele Individuen können Datensätze mit unterschiedlichen Markerdichten kombiniert und fehlende Genotypen mit hoher Genauigkeit imputiert werden (PAUSCH *et al.* 2013). Die höchste Effizienz wird erreicht, wenn die gesamte genomische Variation in die Analyse integriert

ist. Je nach genetischer Architektur und phänotypischer Ausprägung eines Merkmals werden unterschiedliche Ansätze zur Identifizierung der kausalen Variante verwendet.

2.1.1 Phänotypbasierter Ansatz

Bei komplexen, quantitativen Merkmalen wird die phänotypische Ausprägung von wenigen Genorten mit großen Effekten und vielen Genorten mit infinitesimal kleinen Effekten bestimmt (GODDARD und HAYES 2009b). Diese dem Phänotyp zugrunde liegenden Genorte werden als Quantitative Trait Loci (QTL) bezeichnet (GELDERMANN 1975). Die Aufdeckung dieser QTL wurde insbesondere durch die Anwendung genomweiter Assoziationsstudien (GWAS) mittels hochdichter Markerkarten und der Verfügbarkeit von Phänotypinformationen vieler Individuen ermöglicht. Für die Identifizierung eines QTL ist es notwendig, dass der mit dem Merkmal assoziierte Marker mit dem kausalen Quantitative Trait Nukleotid (QTN) mindestens im Kopplungsungleichgewicht (Linkage Disequilibrium, LD) liegt (HAYES *et al.* 2010). Das Ausmaß des LD ist somit entscheidend für die Detektion der kausalen Region (KHATKAR *et al.* 2008). In Populationen mit einer geringen effektiven Populationsgröße (N_e) liegen größere LD-Blöcke vor, sodass für die Aufdeckung des QTL eine geringere Anzahl an Markern benötigt wird. Eine weitere Voraussetzung für die erfolgreiche Detektion von QTL sind möglichst genaue, hoch heritable Phänotypen. In Rinderpopulationen werden die meisten Phänotypen der Besamungsbullen basierend auf deren Nachkommen erhoben. Die nachkommenbasierten Zuchtwerte bieten sich aufgrund ihrer hohen Sicherheit für GWAS an, auch wenn die Heritabilität für manche Merkmale niedrig ist (PAUSCH *et al.* 2011). Phänotypen mit niedriger Sicherheit könnten zu falsch-positiven Resultaten führen und sollten daher nicht für die Anwendung in GWAS genutzt werden (EKINE *et al.* 2014). Die Individuenanzahl ist ein zusätzlicher beeinflussender Faktor für eine erfolgreiche Detektion von QTL. Um auch QTL mit geringem Beitrag an der genetischen Varianz zu identifizieren, ist je nach Heritabilität des Merkmals die Stichprobengröße zu erhöhen (GODDARD und HAYES 2009b). Genetische Mechanismen, wie z.B. Epistasie und Pleiotropie erschweren zusätzlich die Identifizierung solcher QTL (NADEAU 2001). Der effizienteste Ansatz ist die Implementierung aller genomischer Varianten und deren Genotypen, erhoben an einer möglichst hohen Tieranzahl. Somit ist gewährleistet, dass die kausale Variante im Datensatz enthalten ist.

2.1.2 Genotypbasierter Ansatz

Beim genotypbasierten Ansatz sind zunächst keine Phänotypdaten vorhanden. Die Analyse beruht anfangs ausschließlich auf Genotypinformationen. Ein potentieller Phänotyp kann anschließend abgeleitet werden.

2.1.2.1 Missing homozygotes

In den letzten Jahren konnten rezessive Erbdefekte und die zugrundeliegende kausale Mutation mit Hilfe der Homozygotiekartierung aufgedeckt werden (CHARLIER *et al.* 2008; JUNG 2013). Zur Durchführung dieser Analyse sind Phänotypdaten von betroffenen Tieren (Falltiere) und Genotypinformationen dieser Falltiere sowie zusätzlicher Kontrolltiere notwendig. Nach Anwendung der Daten in Fall-Kontroll-Studien können homozygote Regionen kartiert und nach potentiell kausalen Polymorphismen überprüft werden.

Mutationen, die zum embryonalen oder fetalen Fröhrtod führen, sind mit dieser Methodik schwierig zu detektieren. Im Gegensatz zu genetischen Defekten, die peri- oder postnatal auftreten, ist die Erfassung embryonal letaler Defekte aufgrund der fehlenden Phänotypen erschwert. Der embryonale Fröhrtod bleibt oft unerkannt und wird somit häufig Fruchtbarkeitsproblemen zugeordnet (VANRADEN and MILLER 2006). Die Verfügbarkeit zahlreicher und hochdichter Genotypen schaffte die Voraussetzung für die Identifizierung kausaler Mutationen mit letalem Ausgang (VANRADEN *et al.* 2011b; ADAMS 2012; SONSTEGARD *et al.* 2013). Mit Hilfe genomischer Daten können Haplotypen, die relativ häufig, aber ohne homozygote Genotypen in der Population auftreten, lokalisiert werden. Basierend auf dem Hardy-Weinberg-Gleichgewicht (HWG) kann bei einer ausreichenden Menge an Genotypen die erwartete Anzahl homozygoter Individuen berechnet werden. Haplotypen, die bei einer relativ hohen Frequenz ohne homozygote Genotypen in Erscheinung treten, stellen letal wirkende, rezessive Kandidaten dar. Für diese Vorgehensweise benötigt man keine Phänotypinformationen eines vom dem Defekt betroffenen Individuums. Die Methodik basiert allein auf Genotypinformation, ohne der Nutzung phänotypischer Daten.

2.1.2.2 Vorzeitige Stoppmutationen

Polymorphismen, die ein vorzeitiges Stopp-Codon entstehen lassen, führen aufgrund des verfröhrteten Translationsabbruchs zu einem verkürzten Protein und beeinträchtigen somit dessen

Funktion. Möglicherweise kann es auch zu einem vollständigen Proteinverlust kommen. Jedes Individuum trägt unabhängig von seinem Gesundheitszustand Stoppmutationen, sowohl im heterozygoten als auch im homozygoten Zustand (YNGVADOTTIR *et al.* 2009; MACARTHUR *et al.* 2012). Die Beeinträchtigung oder der Verlust der Proteinfunktion kann erhebliche Schäden für den Organismus mit sich bringen. Beim Menschen sind 15 bis 30% der Mendelschen Erbkrankheiten durch einen vorzeitigen Abbruch der Translation bedingt (MORT *et al.* 2008). Auch beim Rind sind Stoppmutationen für genetische Defekte verantwortlich. So wird die bovine männliche Subfertilität beim Fleckvieh durch ein vorzeitiges Stopp-Codon im Gen *TMEM95* verursacht (PAUSCH *et al.* 2014). Ein weiterer Polymorphismus im Gen *PRKG2*, das beim American Angus für den Zwergwuchs verantwortlich ist, verkürzt das kodierte Protein um 85 Aminosäuren mit einem Verlust der funktionellen Kinasedomäne (KOLTES *et al.* 2009). LEIPPRANDT *et al.* (1999) zeigten, dass der durch eine Stoppmutation verursachte Funktionsverlust der Beta-Mannosidase, welche vom Gen *MANBA* kodiert wird, die lysosomale Speicherkrankheit Beta-Mannosidosis auslöst.

Aufgrund der gesundheitsschädigenden Wirkung und der damit zusammenhängenden negativen Selektion der Stoppmutationen treten diese Varianten bei einer niedrigen Allelfrequenz auf, insbesondere wenn sie kausal für eine rezessiv vererbte Erkrankung sind. Stoppmutationen können aber auch nur geringe oder möglicherweise keinerlei Konsequenzen auf einen Phänotyp haben. Die Allelfrequenz dieser harmlosen Mutationen kann sich durch zufällige Effekte in beide Richtungen entwickeln (YNGVADOTTIR *et al.* 2009). Die von diesen Varianten betroffenen Proteine haben meist eine untergeordnete Funktion, die womöglich von anderen Proteinen übernommen werden kann (GU *et al.* 2003; MACARTHUR *et al.* 2012). Beeinflusst eine Stoppmutation ein komplexes Merkmal, könnte der Polymorphismus lediglich einen kleinen Beitrag zur Phänotypausprägung leisten. Der Verlust einer Proteinfunktion kann jedoch auch vorteilhafte Auswirkungen haben, sodass das Allel durch positive Selektion fixiert wird (OLSON 1999; MACARTHUR und TYLER-SMITH 2010). Beim Menschen sind einige solcher positiv wirkenden Varianten bekannt. Ein Beispiel ist das von einer Stoppmutation betroffene Gen *CASP12*, das für das Protein Caspase12 kodiert. Eine Inaktivierung dieses Gens führt zur Resistenz des menschlichen Organismus gegenüber Sepsis (XUE *et al.* 2006). Das gezielte Ausschalten von Genen, dem sogenannten Knockout, ist somit nicht zwangsläufig mit negativen Resultaten verbunden, sondern kann sich auch positiv auf den Organismus auswirken und dementsprechend bei der Therapie bestimmter Erkrankungen nützlich sein.

2.2 Nutzen genomweiter Re-Sequenzierungsdaten

Für eine optimale Nutzung der Informationen über genetische Defekte und QTL sowie deren Anwendung in Zuchtprogrammen ist die Kenntnis der kausalen Variante erforderlich (MEUWISSEN und GODDARD 2010; EICHLER *et al.* 2010). Da die Markerdichte bislang nicht groß genug war, um kausale Regionen mit geringem Effekt identifizieren zu können, war die präzise Aufklärung quantitativer Merkmale wenig erfolgreich (SNELLING *et al.* 2010; PETERS *et al.* 2012). Bisher wurde mittels hochdichter SNP-Arrays und GWAS sowie Homozygotiekartierung meist lediglich die kausale Region identifiziert (VANRADEN *et al.* 2011b; PAUSCH *et al.* 2011; WANG *et al.* 2012). Nur mit Hilfe der Information des Kopplungsungleichgewichts zwischen den Markern und der tatsächlichen kausalen Mutation konnte die kausale Region populationspezifisch detektiert werden (GODDARD und HAYES 2009b). Die Übertragbarkeit und Überprüfung in andere Populationen war aufgrund unterschiedlicher LD-Strukturen und ohne Kenntnis der kausalen Variante nicht möglich. Die für das Merkmal ursächlichen DNA-Varianten waren nicht notwendigerweise auf den bisher verfügbaren Chips enthalten und konnten nur über aufwendige Post-GWAS Studien identifiziert werden. Insbesondere seltene Mutationen, die bei komplexen Merkmalen von großer Bedeutung sind (BODMER und BONILLA 2008; SCHORK *et al.* 2009), bleiben bei der Array-Genotypisierung unberücksichtigt.

Die Einführung von Next-Generation-Sequencing (NGS)-Technologien ermöglicht die Re-Sequenzierung kompletter Genome und die vollständige Aufdeckung der genomischen Variation (ECK *et al.* 2009; JANSEN *et al.* 2013; DAETWYLER *et al.* 2014). Seit der Veröffentlichung des ersten diploiden Genoms wurden bereits zahlreiche Genome und Exome re-sequenziert (NG *et al.* 2008; STOTHARD *et al.* 2011; QIU *et al.* 2012). Es werden nicht nur SNP, sondern auch Insertionen, Deletionen und strukturelle Varianten identifiziert (CONRAD *et al.* 2010; STOTHARD *et al.* 2011; MILLS *et al.* 2011). Der größte Nutzen steckt zweifellos im Einschluss der kausalen Variante in den Genomsequenzen. Bei der Identifizierung kausaler Regionen oder Varianten wird die LD-Information zwischen Marker und kausaler Mutation nicht mehr benötigt. Es besteht nun die Möglichkeit, die Varianten direkt auf Assoziation mit dem vorgegebenen Merkmal zu testen (DRUET *et al.* 2014). Ein weiterer Vorteil ist die Abfrage seltener Varianten, die bisher aufgrund der populationsübergreifenden Nutzung auf SNP-Chips wenig integriert werden konnten (ZHAN *et al.* 2011).

Trotz der stetig sinkenden Anwendungskosten von NGS-Technologien (WETTERSTRAND 2014) ist eine populationsweite Genomresequenzierung aufgrund der weiterhin hohen Kosten- und Arbeitsintensität (DAVEY *et al.* 2011) jedoch noch nicht realisierbar. In vielen Rinderrassen führt

die intensive Bullenselektion einzelner Tiere in Kombination mit künstlicher Besamung zu einem Absinken ihrer N_e , wodurch sich die genetische Diversität dieser Populationen auf wenige sogenannte Schlüsseltiere beschränkt (GODDARD und HAYES 2009a). Die Re-Sequenzierung dieser Schlüsseltiere ist eine kosteneffiziente Lösung zur Aufdeckung der genomischen Variation der gesamten Population (JANSEN *et al.* 2013; DAETWYLER *et al.* 2014). Falls zusätzlich für die Nachkommen hochdichte Genotypinformationen eines SNP-Chips vorliegen, kann deren Sequenz anhand der Genominformation der Gründertiere anschließend mittels Imputation abgeleitet werden (PAUSCH *et al.* 2013). Diese kostengünstig erschlossene populations- und genomweite Sequenzinformation für eine große Zahl von Individuen gewährleistet die Identifikation der kausalen Mutation und erspart somit die teure Sequenzierung weiterer Tiere (HOLM *et al.* 2011; ZEGGINI 2011). Sie kann sowohl für die Vorhersage genomisch geschätzter Zuchtwerte als auch für die Identifizierung kausaler DNA-Varianten herangezogen werden. Genomisch getestete Bullen können noch effizienter und kostengünstiger selektiert werden (SCHAEFFER 2006; BOICHARD *et al.* 2012). Des Weiteren ist eine Überprüfung der ursächlichen Polymorphismen in anderen Rassen unabhängig von der populationspezifischen LD-Struktur durchführbar.

Die populationsweite Imputation der sequenzbasierten Genotypen ist jedoch mit einer Fehlerrate verbunden (DRUET *et al.* 2014). Falsch-positive Ergebnisse können entstehen. Seltene Varianten, deren Genotyp unzulänglich vorhergesagt wurde, können eine geringe Genauigkeit aufweisen oder möglicherweise gar nicht imputiert werden. Die Anzahl der Individuen sowie die Verwandtschaftsstruktur innerhalb der Individuen spielen zusätzlich eine entscheidende Rolle bei der Imputationsgenauigkeit (HUANG *et al.* 2009; DRUET *et al.* 2010).

2.3 Merkmal „Totgeburten“ beim Rind

Das komplexe Merkmal „Totgeburten“ ist ein funktionell wichtiges Merkmal beim Rind mit enormem Einfluss auf die Gesundheit und auf die Produktivität eines Tieres (LÓPEZ DE MATURANA *et al.* 2007). In der Vergangenheit war eine stetige Erhöhung der Totgeburtenrate bei verschiedenen Rinderpopulationen zu beobachten (MEYER *et al.* 2001; HANSEN *et al.* 2004; COLE *et al.* 2007; GHAVI HOSSEIN-ZADEH *et al.* 2008). Insbesondere die Holsteinpopulation weist hohe Totgeburtenraten auf (HEINS *et al.* 2006). Im Vergleich zu anderen Rassen oder Kreuzungstieren sind in dieser Population mitunter die höchsten Totgeburtenraten zu verzeichnen. Aufgrund der hohen resultierenden Kosten, die sowohl durch das fehlende Kalb selbst als auch durch anschlie-

bende Behandlungen von einhergehenden Fruchtbarkeitsproblemen entstehen, steht die Verbesserung dieses Merkmals in hohem Interesse der Landwirte und Zuchtorganisationen.

Die Totgeburtenrate wird in der Zuchtwertschätzung als direktes binäres Merkmal erhoben. In dieses „Alles-oder-Nichts“ (0/1) Merkmal fließen sowohl die Daten totgeborener Kälber ein, als auch die Daten von Kälbern, die innerhalb von 48 Stunden nach der Geburt verenden. Zusammen mit dem genetisch eng korrelierten, subjektiv erhobenen Merkmal Kalbeverlauf wird die Totgeburtenrate im Zuchtwert Kalbeverlauf berücksichtigt.

In den letzten Jahren wurden vermehrt autosomal rezessive Erbdefekte mit letalem Ausgang beim Rind identifiziert, die ursächlich für Totgeburten beim Rind sind (u.a. DRÖGEMÜLLER *et al.* 2010; BUTTKAMP *et al.* 2011; ADAMS 2012; CHARLIER *et al.* 2012; SONSTEGARD *et al.* 2013; FRITZ *et al.* 2013). Die identifizierten kausalen Mutationen führen bereits perinatal zum Tod des Tieres, da die Kälber aufgrund des Defekts nicht überlebensfähig sind. Jedoch konnten Studien zeigen, dass die Totgeburten nicht allein auf diese Genorte zurückzuführen sind (SAHANA *et al.* 2011; COLE *et al.* 2014). Weitere Analysen berichten über QTL, die mit verschiedenen Kalbeverlaufsmerkmalen assoziiert sind und mit der Totgeburtenrate korrelieren (SEIDENSPINNER *et al.* 2009; PAUSCH *et al.* 2011; SAHANA *et al.* 2011; COLE *et al.* 2014). Etwa 50% der Totgeburten sind Folge einer schwierigen Kalbung (MEYER *et al.* 2000). Ein kausaler QTN konnte bisher noch nicht identifiziert werden.

Die Heritabilität des Merkmals ist gering ($h^2 \sim 0,05$) (STEINBOCK *et al.* 2003; SEIDENSPINNER *et al.* 2009). Das Merkmal Totgeburt wird sowohl von verschiedenen genetischen Faktoren - maternale und direkte Effekte – als auch von Umweltbedingungen und Managementfaktoren beeinflusst. Es ist eng korreliert mit dem direkten Effekt des Geburtsgewichts des Kalbes sowie wie mit dem maternalen Einfluss der Trächtigkeitsdauer und dem Geburtsverlauf (MEJERING 1984; STEINBOCK *et al.* 2003; JOHANSON *et al.* 2011; EAGLEN *et al.* 2012). Da nicht alle totgeborenen Kälber aufgrund einer Schweregeburt mit schwergewichtigen Kälbern oder gestörter Wehenfunktion verenden, ist der schwere Geburtsverlauf nicht die einzige Ursache für Totgeburten (PHILIPSSON 1996; HANSEN *et al.* 2004). Die Inzidenz von Totgeburten liegt auch bei leichten Geburtsverläufen, bei denen keine externe Hilfe benötigt wird, vor. Auch die Anzahl der Kalbungen bzw. die Laktationsanzahl der gebärenden Kuh wirkt sich auf die Totgeburtenrate aus (BERGER *et al.* 1992).

3 Tiere, Material und Methoden

3.1 Genomweite Re-Sequenzierung mit NGS

3.1.1 Identifizierung von Schlüsseltieren im Deutschen Fleckvieh

Tiere, die einen Großteil der genetischen Diversität einer Population erklären, werden als Schlüsseltiere bezeichnet. Für die Identifizierung von Schlüsseltieren der Deutschen Fleckviehpopulation (BFV) waren Pedigreeinformationen von 3645 Tieren verfügbar, die die aktuelle Fleckviehpopulation repräsentieren (PAUSCH *et al.* 2012). Die Tiere wurden zwischen 1966 und 2009 geboren. Die Schlüsseltiere dieser Subpopulation wurden iterativ nach GODDARD und HAYES (2009a) mit dem Modell $p_m = A_m^{-1} * c_m$ identifiziert (A = Zähler der Verwandtschaftsmatrix von m selektierten Tieren mit Pedigreeinformationen seit 1920, c = Vektor der durchschnittlichen Verwandtschaft der selektierten Tiere zur gesamten Population, p = Vektor mit dem Anteil der genetischen Diversität, der von den m selektierten Tieren erfasst wird).

3.1.2 Auswahl weiterer Tiere

Neben den identifizierten Schlüsseltieren des deutschen Fleckvieh wurden zehn zusätzliche BFV-Tiere re-sequenziert. Zum Einen wurde ein TRIO (ESMERALDA VANSTEIN und VALERO) analysiert. Die genomische Sequenz VANSTEINs stand aus ECK *et al.* (2009) zur Verfügung (vgl. Anhang Tabelle 9.1). Zum Anderen wurden vier BFV-Tiere zur genomweiten Re-Sequenzierung für Erbfehlerstudien ausgewählt. Die drei übrigen BFV-Individuen wurden aufgrund ihrer hohen Besamungszahlen in der BFV-Population re-sequenziert. Von der Rasse Simmental (SIM) wurden 15 Bullen genomweit re-sequenziert. Im Rahmen weiterer Projekte in Kooperation mit dem Deutschen Holstein Verband (DHV) und der Qualitas AG standen Blut- und Spermaproben von 50 Holstein Tieren (30 Farbrichtung schwarzbunt (HOL), 20 Farbrichtung rotbunt (RED)) und 33 Brown Swiss Bullen (BSW) für die genomweite Re-Sequenzierung zur Verfügung. Außerdem wurde DNA von lehrstuhlinternen Projekten von drei Nordic Finncattle (NFC) und einem weiteren BSW Tier (9741) für die Analyse mit NGS bereitgestellt. Insgesamt wurden somit 214 Individuen aus sechs verschiedenen Rinderrassen genomweit re-sequenziert (vgl. Anhang Tabelle 9.1).

Darüber hinaus konnte auf die Information des dritten Runs des 1000-Bullen-Genom-Projekts (DAETWYLER *et al.* 2014) zurückgegriffen werden. Dieser beinhaltete Daten von 16 Rinderrassen bestehend aus 429 Tieren.

3.1.3 Quantifizierung und Qualitätskontrolle der DNA-Proben

Aus vorangegangenen Arbeiten waren die DNA-Proben der Tiere bereits aufbereitet. Die DNA-Lösungen wurden spektralphotometrisch über die Absorption bei einer Wellenlänge von 260 nm mit dem *NanoDrop 1000* Spektrophotometer (ND-1000, Nanodrop Technologies, Wilmington, DE, USA) quantifiziert. Zur Überprüfung auf Proteinrückstände und zur Kontrolle auf Verunreinigung aus der Extraktion wurden die Parameter A260/280 sowie A260/230 herangezogen. Die 214 Proben, die für die genomweite Re-Sequenzierung vorhergesehen waren, wurden mit dem Qubit®2.0 Fluorometer (Life Technologies, Darmstadt, Deutschland) quantifiziert.

Die Überprüfung der Qualität der DNA erfolgte mittels Gelelektrophorese (0,8%iges Agarosegel; 0,001% Ethidiumbromid, 0,5x Tris-Borat-Ethylendiamintetraacetat (TBE)-Puffer, 50 V Spannung, 1 h Laufzeit) nach Zugabe von *MassRuler DNA Loading Dye* (R0621, Thermo Fisher Scientific (Fermentas), St. Leon-Rot, Deutschland). Diese wurde anschließend unter UV-Licht (Wellenlänge 254 nm) betrachtet. DNA-Proben, die verunreinigt oder degradiert waren, wurden ausgeschlossen.

3.1.4 Generierung der Sequenzdaten

Paired-End Bibliotheken („Libraries“) wurden mit dem *True DNA Sample Prep Kit* (Illumina inc., San Diego, CA, USA) angefertigt und auf den Geräten *GenomeAnalyzerIIx* (*GAIIx*), *HiSeq2000* und *HiSeq2500* (Illumina inc., San Diego, CA, USA) re-sequenziert. Die Read-Längen betragen 76 bis 101 bp (vgl. Anhang Tabelle 9.2) bzw. 36 bp beim zusätzlichen Bullen VANSTEIN (ECK *et al.* 2009). Die resultierenden Reads wurden mit der Software *Illumina BaseCaller* verarbeitet und in *fastq*-Dateien (COCK *et al.* 2010) formatiert. Das Alignment der Reads an die bovine Referenzsequenz UMD3.1 (University of Maryland) (ZIMIN *et al.* 2009) erfolgte mit den Standardeinstellungen des *Burrows-Wheeler Alignment Tool* (BWA, version 0.6.1-r104) (LI and DURBIN 2009). Die resultierenden SAM (Sequence Alignment Map Format) - Dateien wurden für jedes Individuum mit *SAMtools* (version 0.1.18) (LI *et al.* 2009a) in die binäre Version konvertiert (*BAM*). PCR-Duplikate wurden mit der Option *MarkDuplicates* von *PICARD* (<http://picard.sourceforge.net>) identifiziert und markiert.

Die genomweiten Sequenzen wurden mittels des „*integrative genomic viewer*“ (IGV) (THORVALDSDÓTTIR *et al.* 2013) visualisiert.

3.1.5 Identifizierung von Varianten

Innerhalb der 214 genomweit re-sequenzierten Individuen wurden Einzelnukleotidaustausche (*single nucleotide polymorphisms*, SNP) sowie kurze Insertionen und Deletionen (InDels) mit dem in *SAMtools* (version 0.1.18) implementierten multi-sample SNP-calling Ansatz (mpileup) und *BCFtools* identifiziert (LI *et al.* 2009a). Die Genotypen der Polymorphismen wurden mit *SAMtools* (LI *et al.* 2009a) abgeleitet. Zur Überprüfung der an InDels angrenzenden SNP wurde die Option der *extended base alignment quality* (BAQ) angewandt. Das X-Chromosom wurde in die rekombinante pseudoautosomale Region (PAR) und in die nicht-pseudoautosomale Region (non PAR) unterteilt. Die PAR wurde bei der Entdeckung der Varianten als ein autosomales Chromosom berücksichtigt (vgl. Anhang Abbildung 9.1). PCR-Duplikate und Reads mit einer Abdeckung größer als 4576 Reads je Position (ergibt sich aus Anzahl Tiere (n=214) x 2 x durchschnittliche Genomabdeckung (=10,69 Reads je Position)) wurden bei der Identifizierung der Polymorphismen nicht berücksichtigt. Die Ableitung der Haplotypen sowie die Imputation fehlender Genotypen wurde mit der Software *Beagle* (BROWNING und BROWNING 2009) durchgeführt.

3.1.6 Evaluierung der Varianten

Für 104 der 214 genomweit re-sequenzierten Tiere waren Genotypinformationen nach der Genotypisierung mit dem Illumina BovineHD BeadChip (777K) (Illumina, San Diego, USA) vorhanden. Diese array-basierten Genotypen wurden mit den abgeleiteten Genotypen der Sequenz verglichen. Als Parameter für die Verlässlichkeit der Genotypen wurden die von DEPRISTO *et al.* (2011) vorgeschlagenen Kenngrößen Non-Reference-Sensitivity (NRS), Non-Reference-Discrepancy (NRD) und die Übereinstimmung der Genotypen berechnet (Abbildung 3.1.1).

		Array		
		AA	AB	BB
Sequenz	AA	a	b	c
	AB	d	e	f
	BB	g	h	i
	--	k	l	m

$$\text{Übereinstimmung} = \frac{a+e+i}{a+b+c+d+e+f+g+h+i}$$

$$\text{NRS} = \frac{e+f+h+i}{b+c+e+f+h+i+l+m}$$

$$\text{NRD} = \frac{b+c+d+f+g+h}{b+c+d+e+f+g+h+i}$$

Abbildung 3.1.1: Berechnung für Übereinstimmung, NRS- und NRD-Rate der Genotypen

Um die Genauigkeit der sequenz-basierten Genotypen zu ermitteln, wurden diese mit den Array-basierten Genotypen verglichen. Die Formeln für die Berechnung der Übereinstimmung, NRS- und NRD-Rate der Genotypen entsprechen denen von DEPRISTO *et al.* (2011).

3.1.7 Funktionelle Annotation NGS-basierter Varianten

Die funktionellen Effekte der ungefilterten Varianten wurden basierend auf dem UMD3.1-Assembly des bovinen Genoms (ZIMIN *et al.* 2009) mit 20.188 Transkripten vorhergesagt. Genmodelle mit einem oder mehreren Stopp-Codons innerhalb des offenen Leserahmens (in-frame Mutationen) wurden nicht berücksichtigt und begrenzt die Annotation der Varianten auf 18.900 proteinkodierende Transkripte, wobei je Gen nur ein Transkript verwendet wurde. 77 der verwendeten Transkripte waren manuell annotiert (Kapitel 3.3.1). Die Bestimmung der Positionen und die Analyse der Varianten im Promotor (1.000 bp aufwärts vom Transkriptionsstart), im 5' untranslatierten Bereich (untranslated region, UTR), in aminosäurekodierenden Regionen, an Spleißstellen und im 3'UTR wurden mit einer in MySQL aufgesetzten Datenbank und mit in Python erstellten Skripten durchgeführt. Die Effekte der nicht-synonymen Austausch auf die Proteinsequenz wurden mit *Polyphen2-2 v.2.2.2* (ADZHUBEI *et al.* 2010) und *SIFT* (NG und HENIKOFF 2003) vorhergesagt.

Die Varianten wurden außerdem auf bereits bekannte kausale Mutationen für nach Mendel vererbte Merkmale im bovinen Genom untersucht. Dazu wurden die Positionen der in dieser Arbeit

identifizierten Varianten mit den Mutationen der OMIA-Datenbank (ONLINE MENDELIAN INHERITANCE IN ANIMALS 2014) verglichen. Des Weiteren wurde überprüft, ob Varianten an Positionen kausaler Mutationen komplexer Merkmale beim Rind in den Genomsequenzen segregieren.

3.2 Genomweite Assoziationsstudien

3.2.1 Genotypdaten und Imputation

Von der Holstein Population wurden Genotypdaten vom Illumina BovineSNP50 BeadChip (54K) (MATUKUMALLI *et al.* 2009) sowie vom hochdichten Illumina BovineHD BeadChip (777K) von Vereinigte Informationssysteme Tierhaltung w.V., Verden (vit) zur Verfügung gestellt. Für 8840 Tiere waren 54K-Genotypen (Genotypisierungsrate 95,07%) und für 345 Individuen 777K-Daten (Genotypisierungsrate 99,26%) vorhanden. Innerhalb der 345 Tiere mit 777K-Information lagen für 106 Tiere auch 54K-Daten vor. Zunächst wurden die Haplotypen mit *Beagle* (BROWNING und BROWNING 2009) abgeleitet und die Daten anschließend mittels Imputation mit der Software *Minimac* (HOWIE *et al.* 2012) und Standard-Qualitätsanforderungen aufbereitet (vgl. PAUSCH *et al.* 2013). Für 9013 Individuen lagen somit Genotypen an 588.145 Positionen (588K) der Autosomen vor. Durch einen weiteren Schritt wurden die sequenzbasierten Varianten, die in den 115 HOL- und sieben RED-Tieren des 1000-Bullen-Genom-Projekts (DAETWYLER *et al.* 2014) segregieren, *via* der hochdichten 588K-Genotypen mit *Minimac* (HOWIE *et al.* 2012) imputiert. Die Varianten setzen sich aus SNP und kurzen InDels zusammen. Somit standen für die sequenzbasierten genomweiten Analysen 17.290.678 variable Positionen für 9013 Tiere zur Verfügung. Auf Basis des UMD3.1-Assemblys (ZIMIN *et al.* 2009) wurden die chromosomalen Positionen der einzelnen Varianten angegeben.

Für die Qualitätskontrolle wurde *PLINK* (v1.90a, <https://www.cog-genomics.org/plink2>) (PURCELL *et al.* 2007; YANG *et al.* 2011) verwendet. Polymorphismen mit einer Frequenz des selteneren Allels (minor allele frequency, MAF) unter 0,5% und Varianten, die nicht in mindestens 85% der Tiere ermittelt wurden, wurden verworfen. Des Weiteren wurden variable Positionen, die nicht im Hardy-Weinberg-Gleichgewicht lagen (P-Wert < 0,00001) nicht berücksichtigt. Außerdem wurden diejenigen Tiere ausgeschlossen, bei denen mehr als 5% der Genotypen fehlten.

3.2.2 Phänotyp „direkte Totgeburten bei Kühen“

Phänotypinformationen des Merkmals direkte Totgeburten bei Kühen (dSBc) waren in Form konventioneller, nach der Best-Linear-Unbiased-Prediction (BLUP) Methode geschätzter Relativzuchtwerte für 8870 Tiere von der vit, Verden vorhanden (Stand Januar 2014). Für die 8870 Tiere wurde der pedigreebasierte Verwandtschaftskoeffizient mit dem genom-basierten Verwandtschaftskoeffizienten der genomischen Verwandtschaftsmatrix (GRM) verglichen. Individuen mit einer extremen Abweichung in ihren Koeffizienten wurden ohne einen Phänotyp ausgewiesen (n=13), sodass der Phänotyp des Merkmals direkte Totgeburten bei Kühen für 8857 Tiere zur Verfügung stand. Die Analysen der deskriptiven Statistik wurden mit der Open-Source-Software R (www.r-project.org) durchgeführt. Die Ergebnisse sind in Abbildung 3.2.1 dargestellt.

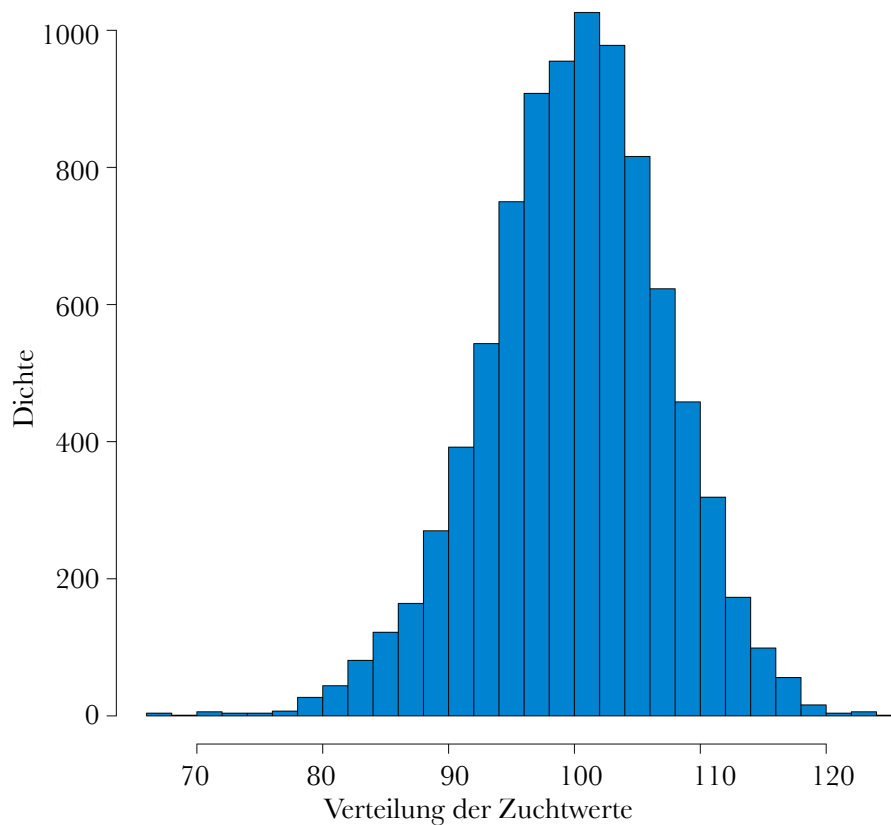


Abbildung 3.2.1: Verteilung des Merkmals direkte Totgeburten bei Kühen

Für das Merkmal direkte Totgeburten bei Kühen lagen 8857 Phänotypen vor. Das Minimum beträgt 66,53, der Mittelwert 100,01, der Median 100,27 und das Maximum 125,32. Die Standardabweichung der Phänotypverteilung beträgt 7,19.

3.2.3 Durchführung der genomweiten Assoziationsstudie

Die genomweite Assoziationsstudie wurde mit Hilfe des gemischten linearen Modells aus *EMMAX* (KANG *et al.* 2010) durchgeführt, um den Effekt der Populationsstratifikation zu berücksichtigen. Das entsprechende Modell $y = Xb + u + \varepsilon$ wurde verwendet, in welchem y den Vektor der geschätzten Zuchtwerte des jeweiligen untersuchten Merkmals, b den fixen Effekt des Polymorphismus und X die Design-Matrix der Polymorphismengenotypen darstellt. u ist der zufällige polygene Effekt mit $(0, \sigma_a^2 G)$, wobei σ_a^2 die additive genetische Varianz und G die GRM darstellt, und ε ist die nicht erbliche Komponente der zufälligen Variation. Die GRM wurde entsprechend VANRADEN (2008) für die 9013 HOL-Tiere basierend auf den 588K-Genotypen (Kapitel 3.2.1) aufgestellt.

Die Ergebnisse der Assoziationsstudie wurden nach Überschreitung der genomweiten und chromosomalen Bonferroni-korrigierten Signifikanzgrenze als signifikant auf genomweiter und chromosomenweiter Ebene angesehen.

3.2.4 Abschätzung der von den Polymorphismen erklärten Varianz

Die phänotypische Varianz, die durch die genomweiten autosomalen Polymorphismen erklärt wird, wurde mit *GCTA* (version 1.24) (YANG *et al.* 2011) bestimmt. Dazu wurden zunächst die Dateien der Imputation mit *MACH* (LI *et al.* 2009b) (mach.dose und mach.info) mit *GCTA* (YANG *et al.* 2011) zu binären *PLINK*-Dateien formatiert. Für jedes bovine Chromosom wurde auf Basis der genomweiten sequenzbasierten Polymorphismen mit *PLINK* (v1.90a, <https://www.cog-genomics.org/plink2>) (YANG *et al.* 2011) die genomische Verwandtschaftsmatrix erstellt. Die chromosomalen Verwandtschaftsmatrizen wurden mit *GCTA* (YANG *et al.* 2011) zu einer genomweiten Verwandtschaftsmatrix zusammengefügt. Die Informationen der genomweiten Verwandtschaftsmatrix diente zur Schätzung des Anteils der genetischen Varianz an der phänotypischen Varianz.

Der Anteil der Varianz, die von einem identifizierten QTL erklärt wird, zur phänotypischen Varianz wurde durch $2pq\alpha^2 / \sigma_a^2$ erklärt, wobei p und q die Frequenzen der beiden Allele des signifikantesten Markers des QTL darstellen, α beschreibt den Allelsubstitutionseffekt des signifikantesten Markers, der aus der Varianz-Komponenten basierten GWAS mit *EMMAX* (KANG *et al.* 2010) stammt und σ_a^2 ist die additiv-genetische Varianz, die mit Hilfe von REML innerhalb von *EMMAX* berechnet wurde. Der Allelsubstitutionseffekt einiger QTN war laut *EMMAX* größer als

die gesamte additiv-genetische Varianz. Dieser Effekt war somit überschätzt. Diese QTN blieben in weiteren Analysen unberücksichtigt.

3.2.5 Berechnung des Kopplungsungleichgewichts

Das Kopplungsungleichgewicht (Linkage Disequilibrium, LD), angegeben als r^2 , wurde mit *PLINK* (v1.90a, <https://www.cog-genomics.org/plink2>) (YANG *et al.* 2011) bestimmt und farblich im Manhattan-Plot dargestellt. Dabei wurde das LD zwischen dem Top-Marker eines QTL zu seinen umgebenden Varianten bestimmt. Die Basis bildeten die für die GWAS verwendeten Genotypen der Imputation.

3.3 In-silico-Analysen

3.3.1 Auswahl und Annotation der Kandidatengene

Beruhend auf den Ergebnissen der funktionellen Annotation NGS-basierter Varianten wurden die Gene ausgewählt, in denen ein vorzeitiges Stopp-Codon vorhergesagt wurde. Außerdem wurden positionelle und mit dem Phänotyp funktionell zusammenhängende Kandidatengene selektiert, welche aus den genomweiten Assoziationsstudien resultierten. Das UMD3.1-Assembly (ZIMIN *et al.* 2009) diente dabei als Grundlage für die Position der Gene.

Die Genstruktur der ausgewählten Gene wurde mit der Software *GENOMETHREADER* (GREMME *et al.* 2005) unter Verwendung von Expressed-Sequenz-Tags (ESTs) der Spezies Mensch, Maus, Ratte, Schwein, Hund und Rind, basierend auf dem bovinen Referenzgenom UMD3.1 (ZIMIN *et al.* 2009) und dem bovinen Genindex (Dana-Farber Cancer Institute, 12.0), vorhergesagt. Die mit *MegaBLAST Bos taurus* (<http://blast.cbcb.umd.edu/megablast.html>) lokalisierten Referenzsequenzen der Genregionen wurden in diesem Prozess implementiert. Die daraus resultierende *.xml*-Datei wurde mit dem *Apollo Annotation and Curation Tool* (v.6.1) (LEWIS *et al.* 2002) bearbeitet.

3.3.2 Charakterisierung der Kandidatengene

Die Konservierung der Kandidatengene erfolgte über das Multiple-Sequenz-Alignment mit Hilfe der Software *ClustalW2* (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) (Larkin *et al.* 2007). Die dafür verwendeten Sequenzen verschiedener Spezies stammen aus den Datenbanken des Natio-

nal Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.gov/>) und wurden mit der manuell annotierten Version der bovinen Sequenz (Kapitel 3.3.1) verglichen.

3.3.3 Haplotypenanalyse der Stoppmutationen

Zur Identifizierung der den Kandidatenmutationen zugrunde liegenden Haplotypen standen für die Rasse BFV Genotypinformationen für 10.363 Tiere und 614.677 SNP (614K) zur Verfügung, die für vorangegangene Analysen mittels Imputation aufbereitet wurden (vgl. PAUSCH *et al.* (2013)). Die aus der Genotypisierung mit dem Illumina BovineSNP50 BeadChip (MATUKUMALLI *et al.* 2009) resultierenden Genotypen wurden mit denen des Illumina BovineHD BeadChip kombiniert und fehlende Informationen mit *Beagle* (BROWNING und BROWNING 2009) und *Minimac* (HOWIE *et al.* 2012) imputiert. Die Haplotypen wurden unter Verwendung der Software *Beagle* abgeleitet.

3.3.4 Überprüfung von Mapping-Fehlern

Sequenzregionen mit Mutationen, die in den Genomsequenzen mit Hilfe von *SAMtools* (version 0.1.18) (LI *et al.* 2009a) vorhergesagt wurden, jedoch nicht über die *Sanger-Re-sequenzierung* (Kapitel 3.5) ermittelt werden konnten, wurden mit Hilfe von *BLAST* (Basic Local Alignment Search Tool) (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) (ALTSCHUL *et al.* 1990) in der bovinen Referenzsequenz auf eine eindeutige Zuordnung im Genom untersucht. Dabei wurde die Datenbank „*Reference genomics sequences (refseq_genomic)*“ in *Bos taurus* (taxid: 9913) verwendet.

3.4 PCR

3.4.1 Primerdesign

Für die Amplifikation mittels PCR wurden Primer verwendet, die mit der Software *Primer3* (<http://frodo.wi.mit.edu/primer3/>) (ROZEN und SKALETSKY 1999) auf Basis des UMD3.1-Assemblies (ZIMIN *et al.* 2009) designed wurden. Die Primer sollten eine Länge von 20 Nukleotiden und einen GC-Gehalt von etwa 50% bei einer Schmelztemperatur zwischen 57 und 63°C (Optimum 60°C) aufweisen. Um das Risiko einer Rückfaltung der Primer oder einer Primerdimerbildung möglichst gering zu halten, wurde auf eine geringe Komplementarität der Primer geachtet. Die amplifizierten Produkte sollten eine Länge zwischen 350 und 800 Basen haben. Die verwendeten Primer sind im Anhang in Tabelle 9.7 aufgeführt.

3.4.2 Standard-PCR

Die Primer wurden von der Firma Sigma-Aldrich Chemie GmbH (Taufkirchen, Deutschland) synthetisiert. Nach der Eluierung der lyophilisierten Primer in Tris-Puffer (10 mM Tris, pH 8,0) wurde eine Konzentration der Stammlösung der Oligonukleotide auf 100 μ M erreicht. Für die Anwendung der Primer in der PCR wurde durch Verdünnung der Stammlösung mit Tris-Puffer eine Arbeitslösung (5 μ M) hergestellt. Für alle PCR-Reaktionen wurde ein 20 μ l Ansatz gewählt. Dieser enthielt 10 ng genomische DNA, 1 x *Coral Load PCR Puffer* mit 15 mM MgCl₂ (Qiagen, Hilden, Deutschland), 0,2 mM dNTP-Mix (K039.1, Carl Roth, Karlsruhe, Deutschland), 0,25 μ M des Forward- und des Reverse-Primers und 1 U Taq-Polymerase (201203, Qiagen, Hilden, Deutschland). Unter Verwendung des *TProfessional Basic Gradient*-Thermocycler (070-601, Analytik Jena (Biometra), Jena, Deutschland) wurde die PCR amplifiziert. Die angewandten Temperaturabläufe umfassten die initiale Denaturierung bei 94°C für 4 Minuten, die zyklische Wiederholung der Schritte Denaturierung bei 95°C für 40s, Annealing bei 60°C für 1 min und Elongation bei 72°C für 1 min, sowie einen finalen Elongationsschritt bei 72°C für 4 min. In der Standard-einstellung wurden 30 Zyklen implementiert.

Die PCR wurde mit einem 1,5%igem Agarosegel (0,001% Ethidiumbromid, 0,5x TBE-Puffer, 100 V Spannung, 30 min Laufzeit) überprüft. Der *GeneRuler 100bp-DNA Ladder* (SM0241, Thermo Fisher Scientific (Fermentas), St. Leon-Rot, Deutschland) diente als Längenstandard.

3.4.3 PCR-Optimierung

Falls die Primer unter den Standard-PCR-Bedingungen nicht optimal funktionierten, wurde für diese Oligonukleotide die Annealingtemperatur optimiert. Dies erfolgte durch die Anwendung eines Temperaturgradienten mit zwölf Temperaturschritten (54,0°C, 54,3°C, 55,0°C, 56,0°C, 57,2°C, 58,4°C, 59,6°C, 60,8°C, 62,0°C, 63,0°C, 63,7°C und 64,0°C). Außerdem wurde die Zyklenzahl bei schwachen PCR-Produkten auf 35 erhöht. Zusätzlich wurde dem Mastermix *Q-Lösung* (finale Konzentration 1x) (Qiagen, Hilden, Deutschland) oder Dimethylsulfoxid (DMSO) (finale Konzentration 5%) (D9170, Sigma-Aldrich, Taufkirchen, Deutschland) hinzugegeben. Bei hohen GC-Gehalten in den zu amplifizierenden Fragmenten erfolgte die Amplifizierung mit der *AccuPrime GC-Rich* DNA-Polymerase (12337016, Life Technologies, Darmstadt, Deutschland).

3.5 Sanger-Re-Sequenzierung

Die mittels PCR amplifizierten Fragmente wurden nach der Kettenabbruchmethode nach SANGER *et al.* (1977) re-sequenziert. Dabei wurden die *Big Dye Terminator v1.1 Cycle Sequencing Reagenzien* (4337450, Life Technologies, Darmstadt, Deutschland) und das *ABI Prism 3130xl Genetic Analyzer* (Life Technologies, Darmstadt, Deutschland) verwendet.

3.5.1 Aufreinigung der PCR

Um einzelsträngige DNA-Moleküle, wie z.B. überschüssige Primer, aus dem PCR-Produkt zu entfernen, wurde 2 U Exonuklease 1 (ENO0581, Thermo Fisher Scientific (Fermentas), St. Leon-Rot, Deutschland) verwendet. Nach Zugabe von 0,4 U alkaline Phosphatase *FastAP* (EF0654, Thermo Fisher Scientific (Fermentas), St. Leon-Rot, Deutschland) wurde der Aufreinigungsansatz bei 37°C inkubiert. Die anschließende Deaktivierung der Enzyme erfolgte für 15 min bei einer Temperatur von 80°C. Anschließend wurde das aufgereinigte PCR-Produkt bei -20°C gelagert.

3.5.2 Sequenzierreaktion, Aufreinigung und Kapillargelelektrophorese

Die Re-Sequenzierung der Fragmente wurden mit dem *Big Dye Terminator v1.1 Cycle Sequencing Kit* (4337450, Life Technologies, Darmstadt, Deutschland) durchgeführt. Der 10 µl Ansatz der Sequenzierreaktion bestand aus 0,25 µM Primer, 1x *Big Dye Terminator*, 1x *Big Dye Sequenzierpuffer*, 1x Dilutionpuffer (5 mM MgCl₂, 0,2 mM Tris-HCl, pH 8,0) und dem aufgereinigtem PCR-Produkt. Es erfolgte eine initiale Denaturierung bei 95°C für 20 s, und eine in 25 Zyklen ablaufende Denaturierung bei 95°C für 12 s, Annealing bei 51°C für 8 s und Elongation bei 60°C für 4 min.

Die Aufreinigung der Sequenzierreaktion wurde mit der Sephadex-Gelfiltrationsmethode durchgeführt. Dabei wurden 25 mg *Sephadex G-50* (G5050-100G, Sigma-Aldrich, Taufkirchen, Deutschland) in *MultiScreen HV* Platten (MAHVN4510, Merck Millipore, Darmstadt, Deutschland) vorgelegt und mit Wasser aufgefüllt. Nach anschließender Inkubation bei Raumtemperatur für 2 h wurde das überschüssige Wasser in der Platte abzentrifugiert (960 g, 5 min). Nach der Verdünnung der Sequenzierreaktion mit 0,1 mM EDTA wurde das Gemisch auf die Sephadex-Platte pipettiert und anschließend mittels Zentrifugation (960g, 5 min) aufgereinigt.

Mittels Kapillargelelektrophorese und der Verwendung des Polymers *3130 POP-7* (4352759, Life Technologies, Darmstadt, Deutschland) sowie einer 36 cm *3130xl & 3100*-Kapillare (4315931,

Life Technologies, Darmstadt, Deutschland) wurden die Fragmente im *ABI Prism 3130xl Genetic Analyzer* (Life Technologies, Darmstadt, Deutschland) aufgetrennt. Die Daten wurden mit der *Run 3130xl Data Collection Software* (v.3.0) erfasst.

3.5.3 Analyse der Sanger-Re-Sequenzierung

Die *AB Sequencing Analysis Software* (v.5.2) wurde für die qualitative Beurteilung, sowie für das Calling der Basen verwendet. Die Entdeckung von Polymorphismen in den Sequenzen erfolgte mit der Phred/Phrap/Polyphred/Consed-Software (NICKERSON *et al.* 1997; EWING und GREEN 1998; EWING *et al.* 1998; GORDON *et al.* 1998).

4 Ergebnisse

4.1 Abschätzung der genomischen Variation

4.1.1 Auswahl der Schlüsseltiere im Deutschen Fleckvieh für die genomweite Re-Sequenzierung

Die Tiere, die für die genomweite Re-Sequenzierung vorgesehen waren, repräsentieren 87,4% der genomischen Variation der Deutschen Fleckviehpopulation. Jedoch war es nicht möglich für jedes dieser Schlüsseltiere DNA zu erhalten, da einige Individuen schon vor Jahren gestorben waren und kein DNA-fähiges Material mehr zur Verfügung stand. Die schließlich zur Re-Sequenzierung ausgewählten 107 BFV-Tiere tragen 78,1% zur genetischen Diversität bei, wobei die zwölf einflussreichsten Tiere bereits 49,2% repräsentieren (Abbildung 4.1.1).

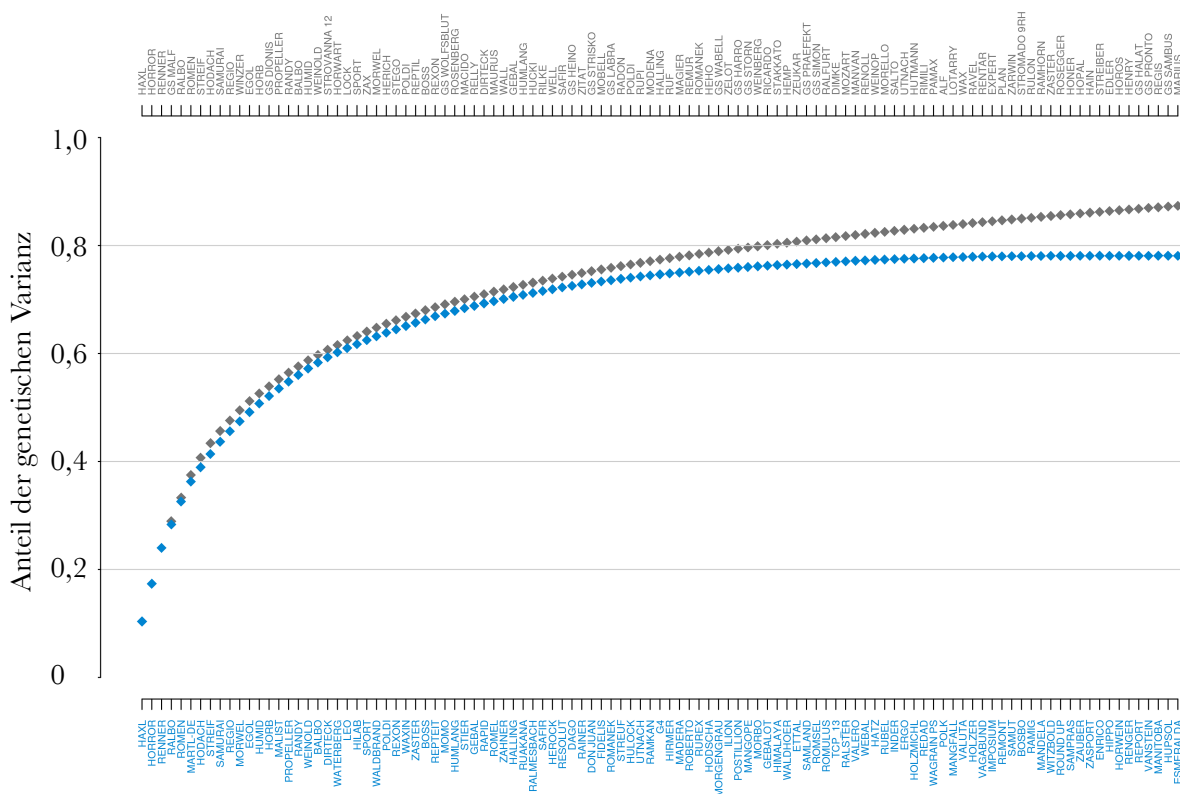


Abbildung 4.1.1: Anteil der erklärten genetischen Variation durch die Re-Sequenzierung von 107 BFV-Tieren

Die 107 für die Re-Sequenzierung ausgewählten Tiere des Deutschen Fleckvieh (BFV) sind nach der Höhe ihres genetischen Beitrags zur Population sortiert. In grau sind die optimal ausgewählten Tiere dargestellt, in blau die aktuell selektierten Individuen. Für fünf der insgesamt 112 re-sequenzierten Tiere lagen keine Pedigreeinformationen vor.

Diese 107 BFV-Tiere gelten nicht alle als Schlüssel-tiere. Innerhalb dieser 107 BFV-Tiere existiert das TRIO (VANSTEIN (ECK *et al.* 2009), ESMERALDA und VALERO), sowie zwei der insgesamt vier Tiere aus Erbfehlerstudien (58953 und 65657), die keine Gründertiere der BFV-Population sind (vgl. Anhang Tabelle 9.1). Aufgrund fehlender Pedigreeinformationen konnte für die übrigen beiden Individuen aus Erbfehlerstudien sowie für drei zusätzliche Tiere der Anteil an der genetischen Diversität der BFV-Population nicht berechnet werden. Diese wurden für den Zweck der Aufdeckung der Erbfehler als auch aufgrund ihrer hohen Besamungszahlen in der BFV-Population (65651, 9735, 9740) für die genomweite Re-Sequenzierung selektiert. Alles in allem wurden 112 BFV-Tiere genomweit re-sequenziert.

4.1.2 Allgemeine Daten der genomweiten Re-Sequenzierung

Insgesamt wurden 68,85 Mrd. paired-end Reads mit einer Länge von 36, 76, 96, 100 und 101 Basen für die 214 Tiere auf dem *GenomeAnalyzerIIx* (*GAIIx*), *HiSeq2000* und *HiSeq2500* (Illumina inc., San Diego, CA, USA) generiert. 86,6 bis 98,6% (durchschnittlich 97,3%) der Reads konnten auf die Scaffolds der 30 Chromosomen (Autosomen 1-29, X-Chromosom) der bovinen Referenzsequenz (ZIMIN *et al.* 2009) mit einer Gesamtlänge von 2,66 Mrd. Basen kartiert werden. Nach Ausschluss von PCR-Duplikaten in Höhe von 3,6 bis 33,8% (durchschnittlich 9,15%) konnten netto 88,4% der generierten Reads kartiert werden (MIN: 63,7%, MAX: 94,2%).

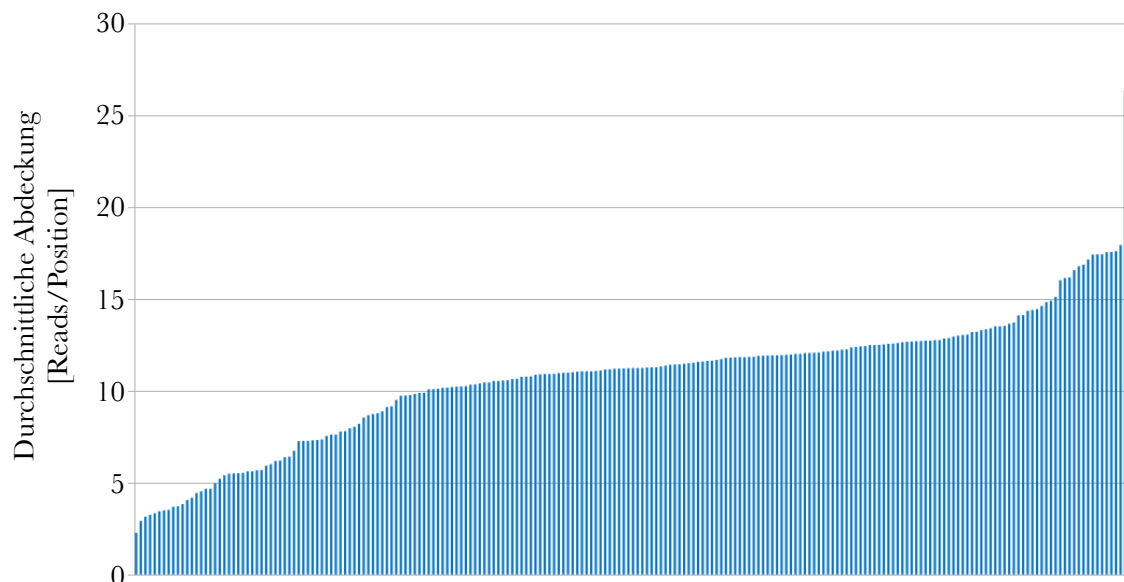


Abbildung 4.1.2: Durchschnittliche Abdeckung der re-sequenzierten 214 Tiere

Die durchschnittliche Abdeckung des Genoms wird in Reads je Position angegeben. Sie ergibt sich aus der Anzahl der erfolgreich kartierten Basen im Verhältnis zur Genomlänge des Rindes (vgl. Anhang Tabelle 9.3). Jeder Balken repräsentiert die Abdeckung des Genoms eines einzelnen Tieres.

Nach Berücksichtigung der Readlängen je Tier (vgl. Anhang Tabelle 9.2) lag die Abdeckung zwischen 2,26 und 26,35 Reads je Position (Durchschnitt über alle Tiere: 10,69 Reads je Position) (Abbildung 4.1.2 und vgl. Anhang Tabelle 9.3).

Die Genomsequenzen wurden für den vierten Run des 1000-Bullen-Genom-Projekts (DAETWYLER *et al.* 2014) zur Verfügung gestellt.

4.1.3 Identifizierung NGS-basierter Varianten

Mit dem multi-sample calling Ansatz (LI 2011) konnten für jedes der 214 re-sequenzierten Tiere Genotypen an 22.937.922 Positionen ermittelt werden, wobei für jede Position mindestens ein Individuum von der UMD3.1-Referenzsequenz (ZIMIN *et al.* 2009) abweichen muss. An Positionen mit fehlender oder unzureichender Abdeckung wurde für das entsprechende Tier der Genotyp mit der höchsten Wahrscheinlichkeit basierend auf der gegebenen Allelfrequenz an diesem Ort bestimmt.

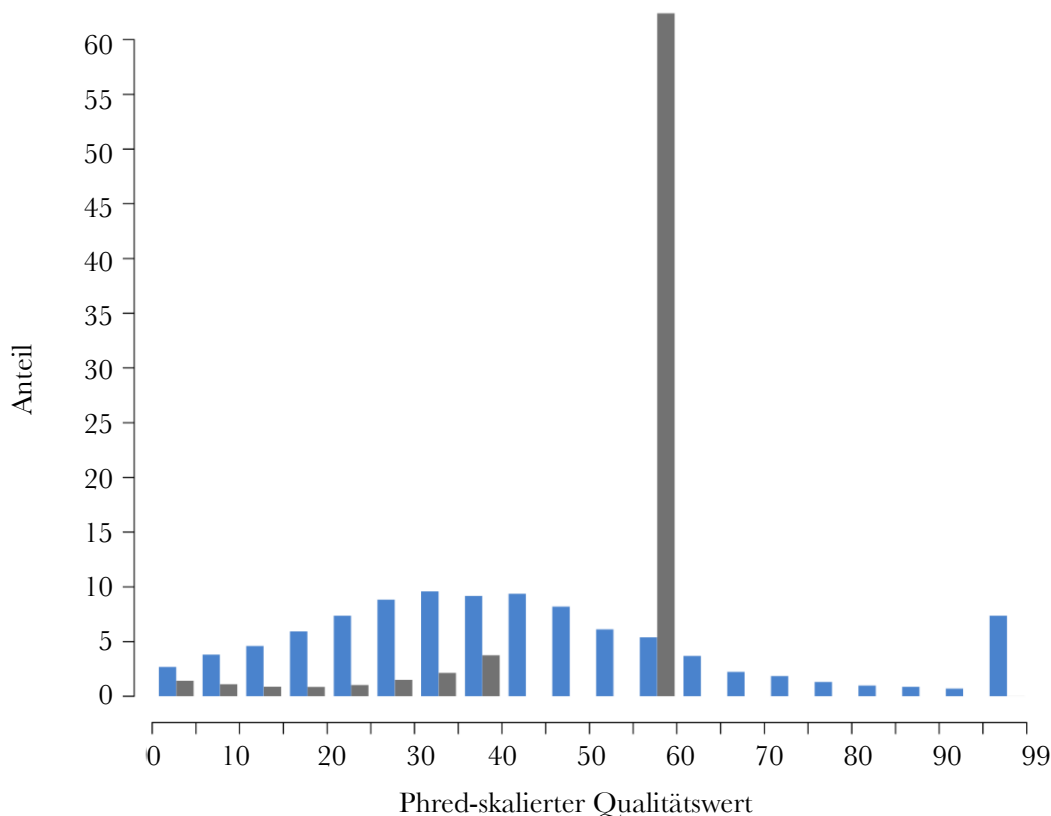


Abbildung 4.1.3: Verteilung der Genotypqualität vor und nach der Imputation mit Beagle

Die Qualität eines jeden Genotypen wird durch die Phred-skalierte Wahrscheinlichkeit bestimmt. Die Beagle-Imputation erreicht einen maximalen Qualitätswert von 60. Werte in Höhe von 40 oder 50 wurden mit Beagle nicht erreicht. Die Werte vor der Beagle-Imputation sind blau, nach der Imputation grau dargestellt.

Das Phasing und Imputieren der Haplotypen mit *Beagle* (BROWNING und BROWNING 2009) erhöhte den Anteil an Genotypen mit einem Phred-skalierten Qualitätswert > 10 von 93,51% vor der Imputation auf 97,47% nach der Imputation (Abbildung 4.1.3).

Der Großteil der Varianten wurde durch biallelische SNP (90,71%) erklärt (vgl. Anhang Tabelle 9.4). Bei 8,51% der identifizierten Polymorphismen handelte es sich um Insertionen/Deletionen (InDels) (45,52% Insertionen, 54,48% Deletionen). Die verbleibenden 0,78% waren triallelische Polymorphismen. Die Rate von Transitionen zu Transversionen (Ti/Tv) betrug 2,2 (vgl. Anhang Tabelle 9.5). Die Anzahl biallelischer Varianten je Tier lag zwischen 5,64 Mio. und 6,06 Mio. nach der Imputation mit *Beagle*, unabhängig von der Abdeckung des Genoms des Tieres (Abbildung 4.1.4). Das Tier mit der geringsten Genomabdeckung wies nicht die wenigsten Varianten auf (5,8 Mio. Varianten bei einer Abdeckung von 2,26 Reads je Position), jedoch zeigte dieser Bulle den höchsten Wert für die Varianten, die nicht abgedeckt sind (5487 Positionen). Über alle 214 Tiere betrug der Anteil biallelischer SNP ohne Abdeckung im Verhältnis zu den entdeckten Polymorphismen des einzelnen Tieres 0,008% (n=496) bis 0,94% (n=5487) (durchschnittlich 0,03 Prozent).

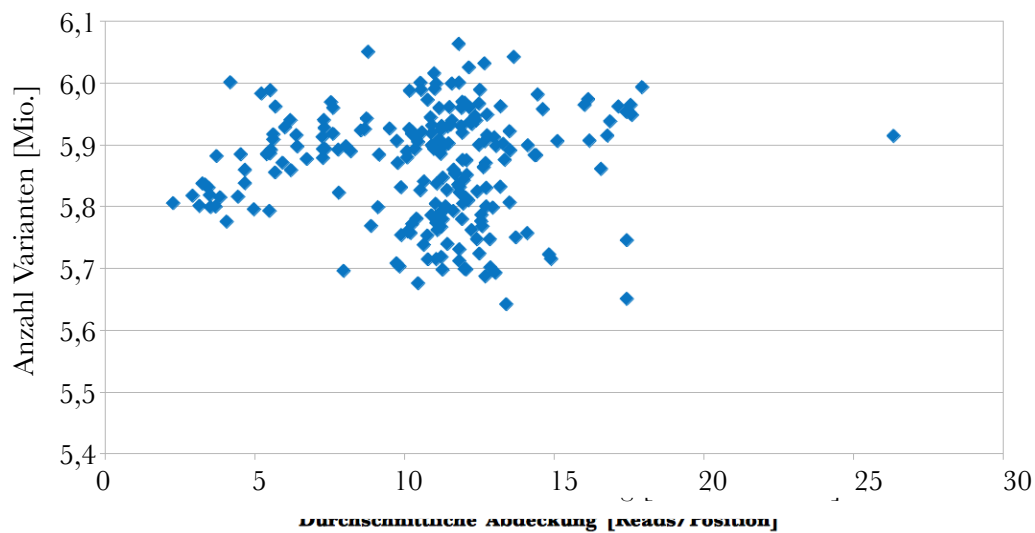


Abbildung 4.1.4: Anzahl entdeckter Varianten in Abhängigkeit zur Genomabdeckung

Die Anzahl variabler Positionen je Individuum ergibt sich aus den Positionen, deren Genotyp nach der Imputation mit *Beagle* (BROWNING und BROWNING 2009) vom homozygoten Genotyp der UMD3.1-Referenzsequenz (ZIMLIN et al. 2009) abweicht.

4.1.4 Evaluierung NGS-basierter Varianten

Zur Bestimmung der Genauigkeit der Genotypen, die aus der genomweiten Re-Sequenzierung resultierten, wurden diese mit den array-basierten Genotypen verglichen. Für 104 Individuen der 214 genomweit re-sequenzierten Tiere lagen Genotypen an durchschnittlich 38.761 Positionen auf Chromosom 1 vor, welches die Ergebnisse der autosomalen Chromosomen repräsentiert (Abbildung 4.1.5).

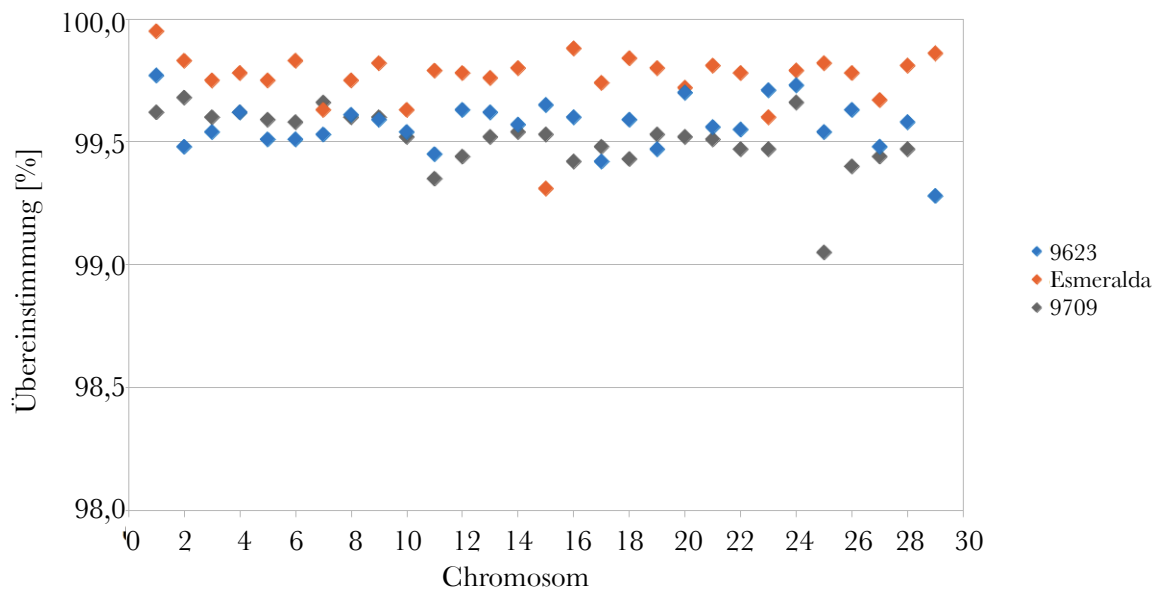


Abbildung 4.1.5: Übereinstimmung der Genotypen je Chromosom

Die Genotypen der genomweiten Re-Sequenzierung wurden mit den Genotypen für jedes Chromosom an den Positionen vom HD-Chip verglichen. Die Ergebnisse werden für drei Individuen dargestellt. Die blauen Rauten repräsentieren ein Tier mit einer mittleren Abdeckung (9623: 10,16 Reads je Position), die orangefarbenen Rauten das Tier mit der höchsten Abdeckung (Esmeralda: 26,35 Reads je Position) und die grauen Rauten das Tier mit der niedrigsten Abdeckung des Genoms (9709: 2,26 Reads je Position).

Die berechneten Werte für die NRS- und NRD-Rate standen mit der Abdeckung des Tiergenoms, mit der Anwendung der *Beagle*-Imputation sowie mit der Zuverlässigkeit der Genotypen, angegeben als Phred-skalierten Qualitätswert, im engen Zusammenhang (Abbildung 4.1.6). Bei einer niedrigen Abdeckung ($< 7x$) war der Einfluss der *Beagle*-Imputation besonders stark. Das Tier mit der geringsten Abdeckung (2,26 Reads je Position) wies vor der Imputation mit 24,98% die insgesamt höchste NRD-Rate auf, welche nach der Imputation nur noch 0,53% betrug. Die höchsten NRS-Werte bei vergleichbar geringen NRD-Werten wurden mit der *Beagle*-Imputation und nach Vernachlässigung der Genotypen mit einem Phred-Wert < 10 festgestellt. Nach der

Beagle-Imputation und nach Ausschluss der Genotypen mit geringer Qualität betrug die NRD-Rate maximal 3,82%, die Übereinstimmung der Genotypen mindestens 97,21% und die NRS-Rate mindestens 99,63%.

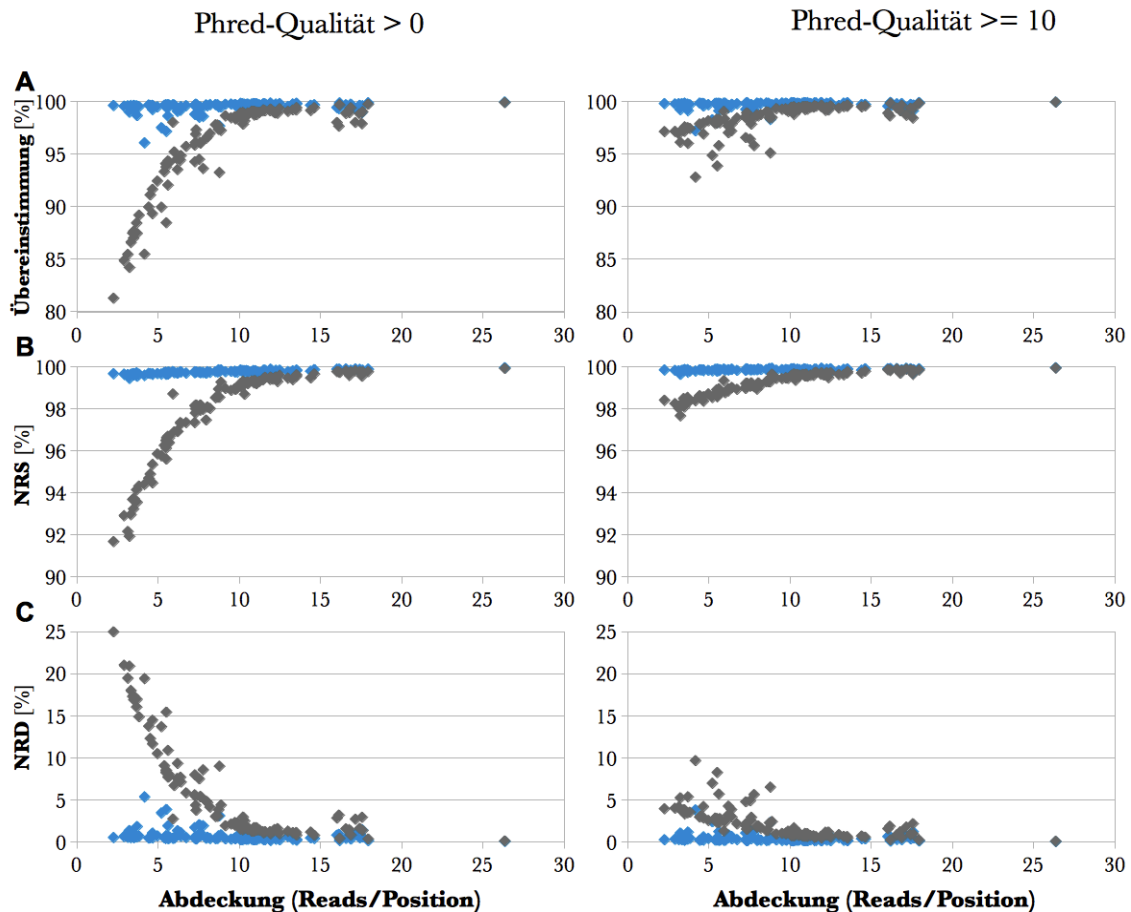


Abbildung 4.1.6: Übereinstimmung, NRD und NRS zwischen NGS-basierten Genotypen und HD-Genotypen

Die Genotypen, die aus der genomweiten Re-Sequenzierung resultieren, wurden mit den Genotypen an den Positionen vom HD-Chip verglichen. In grau sind die Werte vor der Beagle-Imputation dargestellt, in blau sind die Daten nach der Imputation mit Beagle gekennzeichnet. Die Ergebnisse wurden in Bezug zur durchschnittlichen Abdeckung des Genoms (in Reads je Position) jedes einzelnen Individuums aufgetragen. **A** Übereinstimmung der Genotypen **B** NRS-Rate der Genotypen **C** NRD-Rate der Genotypen

4.1.5 Funktionelle Annotation NGS-basierter Varianten

Bei der funktionellen Annotation der genombasierten, entdeckten Mutationen wurde für jedes Transkript eine Promotorlänge von 1000 bp festgelegt. Die Ergebnisse der Annotation wurden in Varianten im Promotorbereich, im Bereich von Spleißstellen und in exonischer Region unterteilt.

Die exonische Sequenz setzt sich aus dem 3'- und 5'-UTR sowie aus der kodierenden Sequenz zusammen.

Von den insgesamt 20.188 verfügbaren Transkripten war für 18.900 Transkripte die kodierende Sequenz mit einer Gesamtlänge von 28.461.287 bp verfügbar. In 128 Genen konnte weder im Promotor, noch im exonischen Bereich oder an den Spleißstellen eine Mutation identifiziert werden. In den verbleibenden 18.772 Transkripten wurden 358.391 Varianten annotiert. In der nicht-kodierenden Region wurden 219.043 Varianten identifiziert, während 139.348 Mutationen der kodierenden Sequenz des Genoms zugeordnet wurden. Es konnten 72.816 synonyme (synonymous, syn) (52,25%) und 66.532 nicht-synonyme (non-synonymous, non-syn) (47,75%) in der kodierenden Sequenz identifiziert werden. Von den nicht-synonymen Austauschen wurden 1038 Varianten als ein vorzeitiges Stopp-Codon vorhergesagt und führen damit zu einem vorzeitigem Abbruch der Translation (s. Kapitel 4.2). Außerdem existierten 937 Varianten, die eine Verschiebung des Leserahmens der Aminosäuresequenz (frame shift) und damit einen vorzeitigen Translationsstopp verursachten (Abbildung 4.1.7).

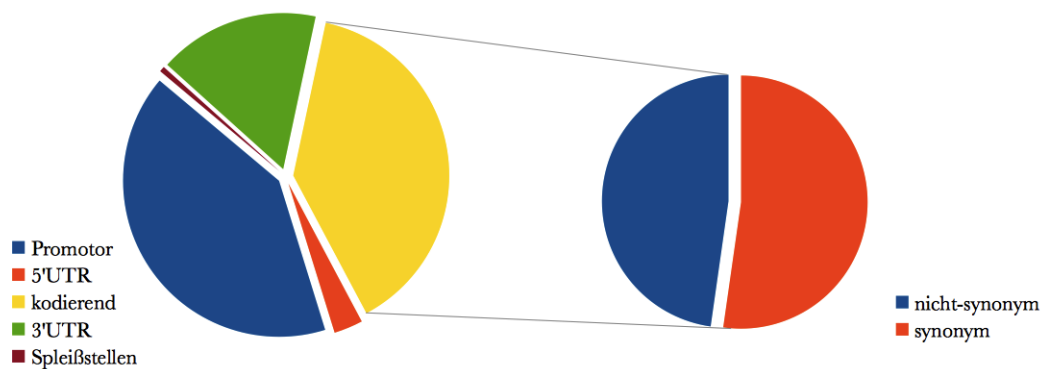


Abbildung 4.1.7: Funktionelle Annotation der Varianten

Die sequenzbasierten Varianten wurden mit in Python geschriebenen Skripten und den auf Basis der UMD3.1-Referenzsequenz (ZIMIN *et al.* 2009) enthaltenen Transkripten funktionell annotiert.

Bei der Vorhersage der potentiellen Effekte der nicht-synonymen Austausche mit *Polyphen2-2 v.2.2.2* (ADZHUBEI *et al.* 2010) wurden 12,66% als „possibly damaging“, 15,29% als „probably damaging“ und 61,76% als „benign“ eingestuft. Für die verbleibenden Varianten (10,29%) konnte kein Effekt prognostiziert werden. *SIFT* (NG und HENIKOFF 2003) charakterisierte 9,08%

der nicht-synonymen Austausch als „deleterious“; als „tolerated“ wurden 66,13% eingestuft. 4,49% Aminosäureaustausche konnten nicht bewertet werden.

4.1.6 Identifizierung bekannter kausaler Varianten

Von den insgesamt 93 Einträgen der OMIA-Datenbank (ONLINE MENDELIAN INHERITANCE IN ANIMALS 2014) sind 85 Gene bekannt, in denen beim Rind eine Mutation für einen vererbten Defekt oder Phänotyp als kausal postuliert wird. In den Genomsequenzen der 214 Tiere konnten 762 Mutationen im kodierenden Bereich von 71 Genen der OMIA-Datenbank identifiziert werden. 315 dieser Varianten führten zu einem Aminosäureaustausch, wovon fünf ein vorzeitiges Stopp-Codon verursachten.

Für die Identifizierung bereits bekannter kausaler Varianten der OMIA-Datenbank konnten die UMD3.1-Positionen für 83 der in der Datenbank verfügbaren Mutationen abgeleitet werden. Elf bereits bekannte kausale Varianten segregieren in mindestens einer der in dieser vorliegenden Arbeit analysierten Rassen (s. Tabelle 4.1).

Tabelle 4.1: Identifizierte OMIA-Mutationen in den Rassen

Gen	Strang	OMIA	Chr	Position	REF	ALT	Allelfrequenz					
							NFC	SIM	HOL	BSW	BFV	RED
<i>APAF1</i>	+	000001	5	63150400	C	T	0,000	0,000	0,000	0,000	0,000	0,025
<i>BCO2</i>	+	001079	15	22877552	G	A	0,000	0,000	0,017	0,000	0,000	0,000
<i>KDSR</i>	-	000939	24	62138835	G	A	0,000	0,000	0,000	0,015	0,000	0,000
<i>MC1R</i>	+	001199	18	14757924	G	-	0,500	0,967	0,033	0,015	0,978	0,950
<i>MITF</i>	-	000214	22	31769189	A	T	0,000	0,067	0,000	0,647	0,129	0,050
<i>MOCS1</i>	-	001541	23	13833868	CA	-	0,000	0,000	0,000	0,000	0,009	0,000
<i>RASGRP2</i>	-	001003	29	43599204	T	C	0,000	0,033	0,000	0,000	0,076	0,000
<i>SHBG</i>	+	001827	19	27956790	C	T	0,000	0,033	0,000	0,000	0,027	0,000
<i>SLC35A3</i>	-	001340	3	43412427	G	T	0,000	0,000	0,017	0,000	0,000	0,000
<i>SMC2</i>	+	001824	8	95410507	T	C	0,000	0,000	0,050	0,000	0,000	0,000
<i>TMEM95</i>	+	001902	19	27689622	C	A	0,000	0,000	0,000	0,000	0,098	0,000

Die Spalte OMIA enthält die Information zur dazugehörigen Identifikationsnummer; Chr entspricht dem bovinen Chromosom, auf dem die Mutation lokalisiert ist; REF steht für das Allel in der Referenzsequenz; ALT beschreibt das alternative Allel der Mutation. Die Rassen sind mit ihren Abkürzungen dargestellt (NFC = Nordic Fincattle, SIM = Simmental, HOL = Holstein, Farbrichtung schwarzbunt, BSW = Brown Swiss, BFV = Fleckvieh, RED = Holstein, Farbrichtung rotbunt).

Das Gen *APAF1* enthält eine Variante, die mit dem Haplotyp HH1 vererbt wird und zum frühzeitigen Abort führt (OMIA 000001-9913; Chr5: 63.150.400) (VANRADEN *et al.* 2011b; ADAMS

2012; FRITZ *et al.* 2013). Das vorzeitige Stopp-Codon wurde in der Sequenz eines einzelnen RED-Tieres identifiziert. Der SNP im Gen *BCO2*, der den Anteil des sogenannten gelben Fettes in der Milch beeinflusst, konnte in einem HOL-Tier (MAGNUM) heterozygot detektiert werden (OMIA 001079-9913; Chr15: 22.877.552; c.306G>A) (BERRY *et al.* 2009). Das resultierende vorzeitige Stopp-Codon (p.W102X) wirkt sich auf den Gehalt des β -Karotins in der Milch aus. Die spinale muskuläre Atrophie (SMA) beim Rind (OMIA 000939-9913; Chr24: 62.138.835) ist auf einen nicht-synonymen Austausch im Gen *KDSR* (p.A59T) zurückzuführen (MEDUGORAC *et al.* 2003; KREBS *et al.* 2007; VANRADEN *et al.* 2011b). Aus dem vorliegenden Datensatz geht ein BSW-Tier hervor, das Anlageträger für diese Mutation ist. In den anderen Rassen ist die Mutation nicht zu beobachten. Die Basendeletion in *MC1R* (OMIA 001199-9913; Chr18: 14.757.924, c.310delG), die die Produktion von Eumelanin und Phäomelanin und somit die Aufhellung der Fellfarbe bestimmt, lag mit einer Allelfrequenz von 68,2% in den 214 Tieren vor. Das rezessive Allel *e* führt im homozygoten Zustand zur typischen roten bzw. gelben Fellfarbe beim Rind und wird daher als Rotfaktor bezeichnet (KLUNGLAND *et al.* 1995). Diese Mutation wurde in allen sechs Rassen gefunden, wobei der homozygote Genotyp für das alternative Allel nur bei den Populationen beobachtet wurde, in denen die rote Fellfarbe charakteristisch für die entsprechende Rasse ist (SIM, BFV und RED). In diesen Populationen wurde wie erwartet kein Tier mit dem homozygoten Genotyp für das Referenzallel gesichtet. Nur in den Rassen HOL und BSW existierte der REF-homozygote Genotyp, der für die dunkle Fellpigmentierung dieser Tiere verantwortlich ist. Eine weitere, für die Farbvererbung relevante Mutation ist ein intronischer SNP im Gen *MITF*. Er wird für die Ausprägung weißer Flecken im Fell als kausal beschrieben (OMIA 000214-9913; Chr22: 31.769.189) (FONTANESI *et al.* 2012). Die gescheckten Rassen (NFC, SIM, HOL, BFV, RED) wiesen eine signifikant höhere Allelfrequenz für das Referenzallel A auf, während die einfarbige Rasse BSW einen erhöhten Anteil des alternativen T-Allel zeigt (AF = 0,647). Ein BFV-Tier wurde mit dem ALT-homozygoten und fünf BSW-Tiere mit dem REF-homozygoten Genotyp beobachtet. Die für die Arachnomelie verantwortliche 2bp-Deletion im Gen *MOCSI* konnte in zwei der genomweit re-sequenzierten BFV-Tiere heterozygot beobachtet werden (OMIA 001541-9913; Chr23: 13.833.86; c.1224_1225delCA) (BUTKAMP *et al.* 2011). Kein Tier einer anderen Rasse zeigte das mutierte Allel. Da der Defekt der Spinnengliedrigkeit in anderen Rassen von Mutationen in abweichenden Genen ausgelöst wird (DRÖGEMÜLLER *et al.* 2010; BUTKAMP *et al.* 2011), ist es unwahrscheinlich, dass die auf dem bovinen Chromosom 23 lokalisierte Variante in anderen Rinderpopulationen außer der BFV- und SIM-Population beobachtet werden

kann. Ein nicht-synonymer Aminosäureaustausch (p.L234P), der für die Thrombopathie in SIM-Tieren als kausal postuliert wird (OMIA 001001-9913; Chr29:43.599.204; c.701C>T) (BOUDREAU *et al.* 2007), wurde in SIM- und BFV-Tieren identifiziert. Bestätigt werden konnte dies durch ein Tier, das im Rahmen eines Thrombopathie-Projekts am Lehrstuhl für Tierzucht der Technischen Universität München genomweit resequenziert wurde und den homozygoten Genotyp für diese Variante trägt. Jedoch kann nicht mit Sicherheit gesagt werden, dass diese Variante die tatsächlich kausale Mutation für die Blutgerinnungsstörung ist (JUNG 2013). FRITZ *et al.* (2013) identifizierten in der Rasse Montbéliarde einen Haplotypen (MH1), der nie im homozygoten Zustand vorzufinden war und mit Fruchtbarkeitsstörungen und Embryosterblichkeit assoziiert wurde. In dem Haplotyp ist das Gen *SHBG* lokalisiert, das die von FRITZ *et al.* (2013) beschriebene kausale Variante trägt. Die Variante führt zu einem vorzeitigen Stopp-Codon und somit zu einem verkürzten Protein (OMIA 001827-9913; Chr19: 27.956.790; c.154C>T; p.Q52X). Auf Basis der vorliegenden 214 Genomsequenzen konnte diese potentiell kausale Mutation und somit auch MH1 in der SIM- und BFV-Population beobachtet werden. In einem BFV-Tier (MANGFALL; ID 9770) wurde der homozygote Genotyp TT entdeckt, was der Theorie des frühzeitigen Aborts widersprechen würde. Die Position der Mutation in diesem Individuum war gering abgedeckt (6-fach) und wies eine sehr niedrige Genotypqualität (5) auf. Der NGS-basierte homozygote Genotyp in diesem Tier an dieser Position wurde mit Hilfe der *Sanger-Re-Sequenzierung* widerlegt. Die populationsweite Genotypisierung an dieser Position mit einem TaqMan®-Assay identifizierte jedoch ein homozygotes Tier. Der Genotyp wurde zusätzlich mittels *Sanger-Re-Sequenzierung* bestätigt. Aufgrund des Auftretens von mindestens einem homozygoten Tier ist diese Mutation nicht die kausale Variante für die Embryosterblichkeit. Die von FRITZ *et al.* (2013) vorgeschlagene Theorie konnte damit in der BFV-Population widerlegt werden (s. Tabelle 4.2).

Tabelle 4.2: Genotypen der Stoppmutation in SHBG

	Genotypen		
	CC	CT	TT
Anzahl	1724	63	1

Die Genotypen für die Punktmutation, die zu einem vorzeitigem Stopp-Codon im Gen SHBG führt, wurde mittels TaqMan® -Typisierung für insgesamt 1788 zufällig ausgewählte BFV Tiere ermittelt.

Der Holstein Haplotyp 3 (HH3) ist ähnlich wie HH1 und HH2 mit einem frühzeitigem Abort assoziiert (OMIA 001824-9913; Chr8: 95.410.507) (VANRADEN *et al.* 2011b; FRITZ *et al.* 2013).

McCLURE *et al.* (2014) identifizierten die dafür zugrunde liegende Mutation im Gen *SMC2* (p.F1135S). Die kausale Variante konnte in dieser Studie ausschließlich in der HOL-Rasse in drei Tieren heterozygot identifiziert werden. Die für die komplexe Verformung der Wirbelsäule bei HOL kausale Mutation in *SLC35A3* (OMIA 001340-9913; Chr3: 43.412.427) (THOMSEN *et al.* 2006) konnte in einem in dieser Studie genomweit re-sequenzierten HOL-Tier heterozygot beobachtet werden. Kein Individuum einer anderen Rasse wies diese Variante auf. Eine weitere kausale Mutation im Gen *TMEM95*, die zur männlichen Subfertilität beim Rind führt, trugen zwei Individuen homozygot (OMIA 001902-9913; Chr19: 27689622; c.483C>A; p.C161X) (PAUSCH *et al.* 2014). Diese Variante wurde mittels 43 genomweit re-sequenzierter BFV-Tiere, die im vorliegend analysierten Datensatz der 214 Tiere enthalten sind, identifiziert.

Nicht nur Polymorphismen für autosomal rezessiv vererbte Defekte oder Merkmale konnten in den vorliegenden Genomsequenzen detektiert werden, sondern auch Allelaustausche, die mit quantitativen Merkmalen assoziiert sind (Tabelle 4.3). Von den insgesamt fünf bekannten Varianten beeinflussen vier die Gehalte an Inhaltsstoffen und die Menge der Milch. COHEN-ZINDER *et al.* (2005) konnten im Gen *ABCG2* eine A>C-Substitution identifizieren, die zu einem Aminosäureaustausch an der 581. Position im Protein führt (p.Y581S). Dabei wirkt sich das C-Allel negativ auf die Milchleistung aus. Dieser SNP segregiert innerhalb der 214 Tiere nur innerhalb der RED-Population in einem Tier (ROCHESTER) heterozygot. Zwei gekoppelte Mutationen im kodierenden Bereich des *DGATI*-Gens (c.694_695 GC>AA) bestimmen den Fettgehalt der Milch sowie die Milchmenge (GRISART *et al.* 2002; WINTER *et al.* 2002). Tiere mit den beiden A-Allelen tragen anstelle der Aminosäure Alanin die Aminosäure Lysin an der 232. Position im Protein, welches zur extremen Erhöhung der Milchleistung und des Fettgehaltes in der Milch führt. Diese Mutationen wurde in den Rassen SIM, HOL, BFV und RED mit einer Allelfrequenz von weniger als 40% beobachtet, wobei nur in BFV, HOL und RED homozygote Tiere vorlagen. Dagegen segregiert der Polymorphismus in *GHR*, der eine ähnliche Wirkung wie die Varianten in *DGATI* aufweist (BLOTT *et al.* 2003; VIITALA *et al.* 2006), in allen der untersuchten Rassen mit einer Allelfrequenz von mindesten 6%. Das aufgrund des T>A-Allelaustauschs entstehende Tyrosin (p.F279Y) bewirkt einen Anstieg der Milchmenge bei gleichzeitig sinkendem Gehalt an Inhaltsstoffen. Von den 214 genomweit re-sequenzierten Tieren tragen vier Individuen aus drei Rassen (SIM, HOL und RED) den homozygoten Genotyp. Ein weiterer nicht-synonymer Austausch im Gen *PRLR* steuert außerdem die Menge an Protein und Fett in der Milch (VIITALA *et al.*

2006). Die gekoppelten SNP verursachen die Substitution der Aminosäuren Serin zu Asparagin (p.S18N). Das die Protein- und Fettmenge erhöhende Asparagin-Allel liegt in allen Rassen in allen drei Genotypvariationen vor. Eine hohe Frequenz ($AF > 0,8$) für das auf den Inhaltsstoffgehalt positiv wirkende Allel war in den Rassen NFC, HOL und RED zu erkennen.

Tabelle 4.3: Identifizierte QTN in den Rassen

Gen	Merkmal	Chr	Position	REF	ALT	Allelfrequenz					
						NFC	SIM	HOL	BSW	BFV	RED
<i>ABCG2</i>	Milchmenge und -zusammensetzung	6	38.027.010	A	C	0,000	0,000	0,000	0,000	0,000	0,025
<i>DGAT1</i>	Fettgehalt, Milchmenge	14	1.802.265	G	A	0,000	0,100	0,400	0,000	0,067	0,400
			1.802.266	C	A	0,000	0,100	0,400	0,000	0,067	0,400
<i>GHR</i>	Milchmenge und -zusammensetzung	20	31.909.478	A	T	0,167	0,200	0,150	0,103	0,063	0,175
<i>NCAPG</i>	fetales Wachstum	6	38.777.311	T	G	0,167	0,600	0,383	1,000	0,786	0,300
<i>PRLR</i>	Protein- und Fettmenge	20	39.115.344	G	A	0,833	0,367	0,883	0,176	0,232	0,800
			39.115.345	G	A	0,833	0,367	0,883	0,176	0,232	0,800

Die Gene mit bekannten QTN sind alphabetisch sortiert. Die Spalte Chr entspricht dem bovinen Chromosom, auf dem die Mutation lokalisiert ist; REF steht für das Allel in der Referenzsequenz; ALT beschreibt das alternative Allel der Mutation. Die Rassen sind mit ihren Abkürzungen dargestellt (NFC = Nordic Fincattle, SIM = Simmental, HOL = Holstein, Farbrichtung schwarzbunt, BSW = Brown Swiss, BFV = Fleckvieh, RED = Holstein, Farbrichtung rotbunt).

Der für das fetale Wachstum kausale QTN in *NCAPG* (EBERLEIN *et al.* 2009), der eine Veränderung von Isoleucin zu Methionin herbeiführt, segregiert in den re-sequenzierten 214 Tieren. Das wachstumsfördernde Methionin-Allel liegt bei der BSW-Population ausschließlich homozygot vor ($AF = 1$), während es bei den anderen vier Rassen mindestens als heterozygoter Genotyp in Erscheinung tritt. Die hohe Allelfrequenz in BFV und SIM ($AF > 0,6$) steht den niedrigeren Allelfrequenzen in HOL und RED gegenüber ($AF = 0,3$ bis $0,4$).

4.2 Verifizierung vorzeitiger Stopp-Codons

Mit Hilfe der funktionellen Charakterisierung der aus der genomweiten Re-Sequenzierung vorliegenden Varianten konnten 1038 vorzeitige Stopp-Codons vorhergesagt werden. Nachdem diese hinsichtlich der höchsten Variantenqualität (QUAL=999) als Parameter für die Zuverlässigkeit gefiltert wurden, verblieben 642 Stoppmutationen. 424 Varianten wurden in Genen lokalisiert, die bisher noch nicht beschrieben (z.B. „LOCs“) oder die dem Hauptkompatibilitätskomplex

eingegliedert wurden. Diese Mutationen wurden für weitere Analysen ausgenommen. Die verbleibenden 215 Stopp-Codons wurden in 197 verschiedenen Genen ermittelt. Um Stopp-Codons, die aufgrund von Annotationsfehlern entstanden sind, auszuschließen, wurden diese Gene manuell reannotiert (*GENOMETHREADER* und *Apollo Annotation and Curation Tool* (v.6.1)). 65 vorzeitige Stoppmutationen wurden so aufgrund einer fehlerhaften Annotation der kodierenden Sequenz der Gene entfernt.

Die Überprüfung der verbleibenden Mutationen erfolgte mittels Re-Sequenzierung nach *Sanger*. Neun Varianten konnten aufgrund uneindeutiger Amplifikationsprodukte nicht weiter untersucht werden.

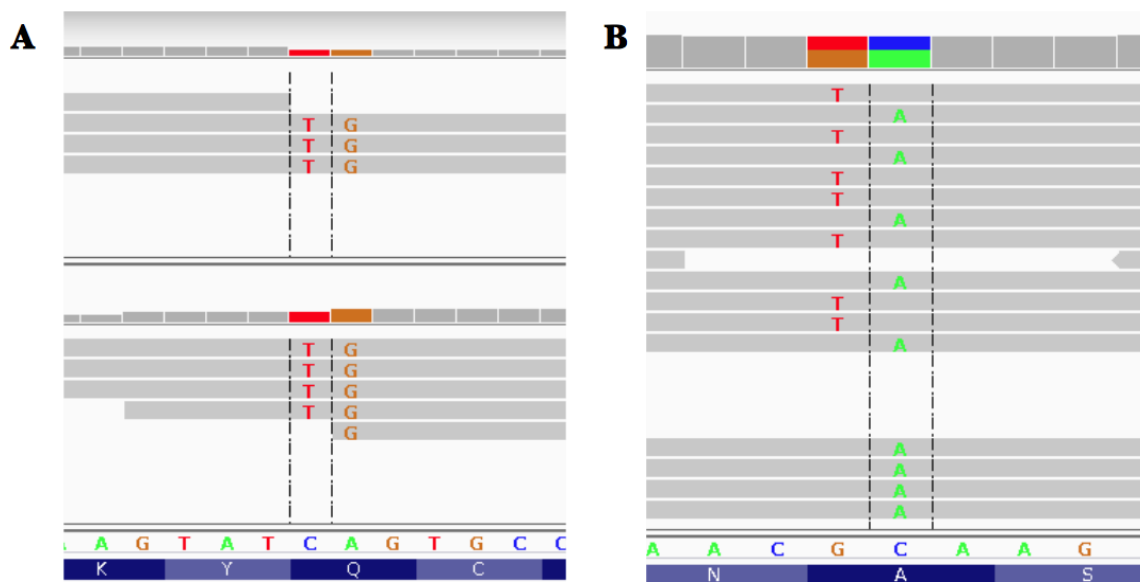


Abbildung 4.2.1: Multiple Varianten innerhalb eines Codons

Wird ein vorzeitiges Stopp-Codon auf Basis eines einzelnen SNP vorhergesagt, obwohl auf dem gleichen Haplotyp eine weitere Mutation in demselben Codon vorliegt, entsteht anstelle des Stopp-Codons ein Aminosäureaustausch ($CAG \rightarrow TGG$)(A). Beruht die Vorhersage eines vorzeitiges Stopp-Codons auf der vollständigen Kopplung zweier SNP innerhalb eines Codons, obwohl diese nicht gekoppelt sind, entstehen anstelle der Stopp-Codons zwei Aminosäureaustausche ($GCA \rightarrow TCA$ und GAA)(B).

Beim Auftreten mehrerer SNP innerhalb eines Codons können diese auf unterschiedlichen Haplotypen lokalisiert sein oder im vollständigen Kopplungsungleichgewicht vorliegen. Der funktionelle Effekt der einzelnen SNP kann sich von dem der gekoppelten SNP unterscheiden. Alle hoch qualitativen vorzeitigen Stopp-Codons wurden mit Hilfe der Visualisierung im IGV (THORVALDSDÓTTIR *et al.* 2013) auf multiple Varianten untersucht. Diese Analyse identifizierte 26

vorhergesagte Stopp-Codons, die auf multiplen Varianten beruhen. Unter diesen 26 vorzeitigen Stopp-Codons waren solche, die auf Basis eines einzelnen SNP vorhergesagt wurden, jedoch aufgrund eines mittels *Sanger-Re-Sequenzierung* identifizierten gekoppelten SNP in demselben Codon zu einem Aminosäureaustausch führten. Die Stoppmutation wurde somit aufgelöst (Abbildung 4.2.1(A)). Bei anderen vorhergesagten vorzeitigen Stopp-Codons mit multiplen Varianten wurde ein vollständiges Kopplungsungleichgewicht zwischen den umliegenden Mutationen vorausgesetzt. Falls die Varianten auf unterschiedlichen Haplotypen lokalisiert wurden und somit kein vollständiges Kopplungsungleichgewicht bestand, entstanden anstelle des vorzeitigen Stopp-Codons zwei Aminosäureaustausche (Abbildung 4.2.1(B)). Diese falsch vorhergesagten Stoppmutationen wurden verworfen.

Bei der Auswertung der Re-Sequenzierung konnten 15 Mutationen nicht detektiert werden. Mittels lokaler BLAST-Suche (ALTSCHUL *et al.* 1990) gegen das UMD3.1-Assembly des Rindes wurden diese Reads mit höherer Übereinstimmung an anderen Stellen des Rindergenoms lokalisiert. Die Mutationen wurden durch fehlerhaftes Alignment der Reads hervorgerufen und stellen keinen echten SNP dar. Diese aufgrund von Alignmentfehlern entstandenen Mutationen wurden für weitere Analysen nicht berücksichtigt.

Nach Anwendung gezielter Filter verblieben von den 1038 vorzeitigen Stopp-Codons 100 hoch qualitative Stoppmutationen in 94 verschiedenen Genen, von denen 91 Varianten auch im 1000-Bullen-Genom-Projekt identifiziert wurden (vgl. Anhang Tabelle 9.6). Die meisten Gene beinhalteten eine Stoppmutation; sechs Gene wiesen zwei Stoppmutationen auf. Dieser Datensatz enthält zuverlässige Varianten, die ein vorzeitiges Stopp-Codon einleiten und somit zu einer Veränderung oder zu einem Verlust der Proteinfunktion führen können (Loss of function = LoF). Weitere Mutationen, die die Translation eines Gens vorzeitig abbrechen und die Funktion des Proteins beeinträchtigen können, wie z.B. nicht-synonyme Austausche, wurden nicht berücksichtigt. Die Analyse konzentriert sich ausschließlich auf Mutationen, die zum potentiellen Knockout eines Gens führen können.

4.2.1 Potentielle Auswirkungen vorzeitiger Stopp-Codons auf die Proteinsequenz

Um das potentielle Ausmaß der LoF-Varianten auf das Genprodukt zu bestimmen, wurde für jede Mutation überprüft, ob sie bei Rind, Mensch und Maus alle Isoformen betreffen. SNP, die nur einzelne Isoformen beeinflussen, könnten gewebsspezifische Effekte haben. Falls alle

Isoformen beeinträchtigt sind, können die Punktmutationen zu einer schädlichen Veränderung (gain of function) oder zum vollständigen Verlust (LoF) der Proteinfunktion führen, insbesondere wenn das Stopp-Codon in den Transkripten einen Nonsense-vermittelten mRNA-Abbau (nonsense-mediated mRNA decay, NMD) auslöst. Dabei wird die mRNA durch Exonukleasen abgebaut, um die Expression eines verkürzten Proteins zu verhindern (SCHWEINGRUBER *et al.* 2013). Der Mechanismus des NMD findet nicht statt, wenn das vorzeitige Stopp-Codon in einem Bereich von 50 bis 55 Basenpaaren vor der nächsten Exon-Exon-Bindung lokalisiert ist (NAGY und MAQUAT 1998). Eine weitere Voraussetzung für eine Verhinderung des NMD ist, dass die prä-mRNA nur aus einem Exon besteht und die Stoppmutation sich nicht im letzten Exon befindet (NAGY und MAQUAT 1998) (Abbildung 4.2.2).

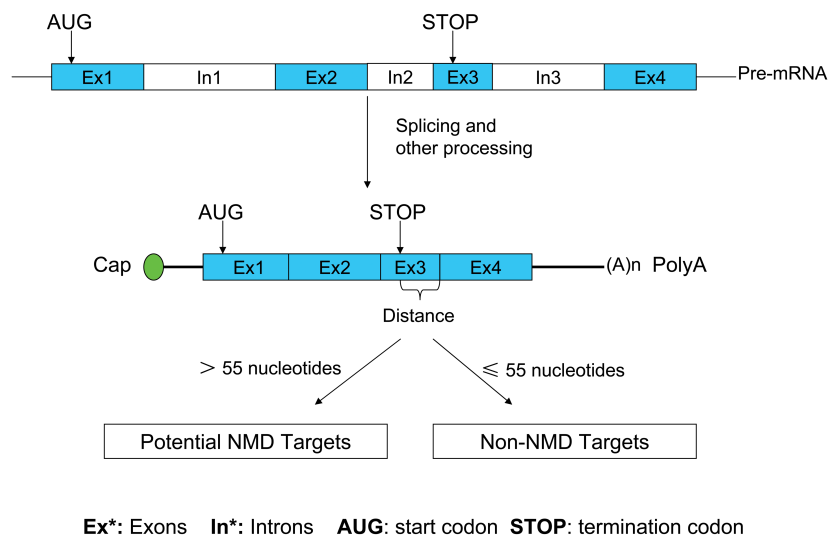


Abbildung 4.2.2: Nonsense-vermittelter mRNA-Abbau in Säugetieren

Gene in Säugetieren werden vom Genom abgelesen und lassen eine prä-mRNA entstehen, die sowohl Exons (Ex) als auch Introns (In) inkludiert. Durch das Spleißen werden Introns entfernt; Exon-Exon-Grenzen werden durch Proteinkomplexe markiert. Falls das vorzeitige Stopp-Codon (STOP) > 50-55 bp in 5'-3'-Richtung vor der nächsten Spleißstelle des Exon-Junction-Komplex lokalisiert ist, initiiert diese Mutation einen Nonsense-vermittelten mRNA-Abbau (NMD) (NAGY und MAQUAT 1998). Angelehnt an (ZHANG *et al.* 2009)

Von den 100 untersuchten LoF-Varianten beeinflussen 63 alle alternativen untersuchten Transkripte des Gens. Basierend auf den oben genannten Regelfällen lösen 62 der 100 hoch qualitativen LoF-Varianten keinen NMD aus. Für die anderen 38 Stoppmutationen treffen die Voraussetzungen für einen voraussichtlichen NMD zu. Von besonderem Interesse sind 30 vorzeitige

Stopp-Codons, die alle Isoformen des Gens berühren und gleichzeitig die Voraussetzungen für einen NMD erfüllen. Neun dieser 30 LoF-Varianten wurden weder innerhalb der 214 Individuen noch in den Daten des 1000-Bullen-Genom-Projekts im homozygoten Zustand beobachtet. Die Re-Sequenzierung konnte für drei Stoppmutationen (*GIP*, *PNLIP* und *VTCNI*) homozygote Individuen aufdecken.

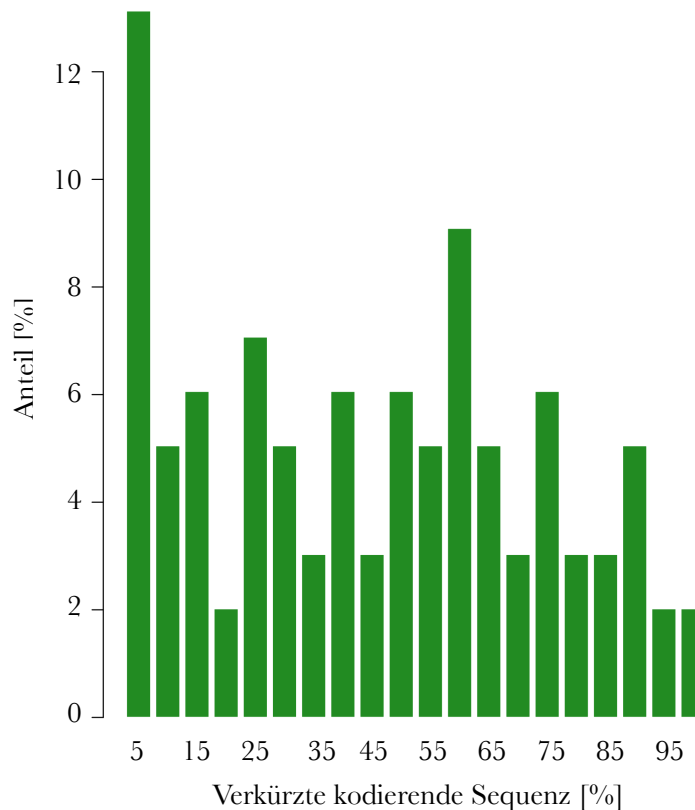


Abbildung 4.2.3: Anteil der Verkürzung der kodierenden Sequenz durch vorzeitige Stopp-Codons

Bei der Analyse der Stoppmutationen wurde je Gen die Isoform mit der längsten, kodierten Aminosäuresequenz gewählt. Die Größe der verkürzten Proteinsequenz durch das jeweilige vorzeitige Stopp-Codon wurde ins Verhältnis zur Gesamtlänge der Isoform gesetzt.

Bei der Analyse des Anteils der durch die vorzeitigen Stopp-Codons verkürzten kodierenden Sequenz wird deutlich, dass 26 der Stoppmutationen in den letzten 15% der Proteinsequenz lokalisiert sind. Allein 15 LoF-Varianten verkürzen die Sequenz um nur 5%. Etwa die gleiche Anzahl an Stoppmutationen (n=12) dagegen befindet sich am Anfang der kodierenden Sequenz und würde somit zum Verlust von mehr als 80% der kodierenden Sequenz führen (Abbildung 4.2.3). Im Durchschnitt lag die Verkürzung der Aminosäuresequenz durch die Stoppmutationen

bei 42,05%. Berücksichtigt man nur die LoF-Varianten, die voraussichtlich nicht zu einem NMD und somit zu keinem vollständigen Verlust, sondern nur zu einer Verkürzung des Proteins führen, werden durchschnittlich 39,15% des Peptids abgeschnitten (MIN: 0,00%, MAX: 95,86%).

4.2.2 Allelfrequenzverteilung vorzeitiger Stopp-Codons

Mehr als 70% der hoch qualitativen SNP, die ein vorzeitiges Stopp-Codon verursachen, weisen innerhalb der 214 re-sequenzierten Tiere eine niedrige Allelfrequenz auf ($AF < 0,05$) (Abbildung 4.2.4). Je höher die Allelfrequenz, umso geringer ist der Anteil der Stoppmutationen in der entsprechenden Kategorie. Die höchste erreichte Allelfrequenz beträgt 0,7. Nicht-synonyme Aminosäureaustausche, die als „deleterious“ oder „probably damaging“ vorhergesagt wurden (Kapitel 4.1.5), zeigen einen ähnlichen Verlauf.

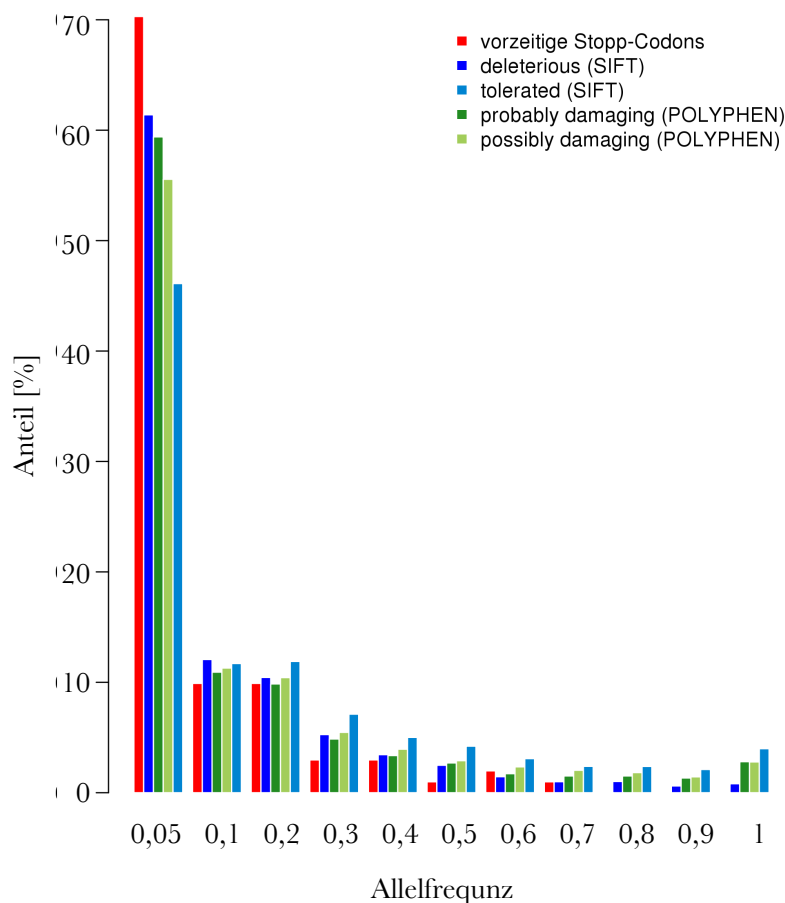


Abbildung 4.2.4: Verteilung der Allelfrequenzen vorzeitiger Stoppmutationen im Vergleich zu nicht-synonymen Aminosäureaustauschen

Die Grundlage der Allelfrequenzen sind die 214 genomweit re-sequenzierten Tiere. Die nicht-synonymen Austausche werden nach der vorhergesagten Auswirkung von SIFT und Polyphen differenziert.

Der Anteil aller non-syn Austausch für seltene Allele beträgt mehr als 50%, wovon etwa 36% als „deleterious“ bzw. etwas 18% als „probably damaging“ eingestuft wurden. Einzelne schädliche vorhergesagte non-syn Mutationen treten in allen Tieren homozygot auf (AF=1). In diesem Allelfrequenzbereich existieren keine vorzeitigen Stopp-Codons.

Beim Rassevergleich innerhalb der 214 Tiere fällt auf, dass nicht jede Stoppmutation in jeder Rasse zu erkennen ist. In NFC mit nur drei Tieren lassen sich die wenigsten LoF-Varianten beobachten (n=23), während in BFV die meisten Stoppmutationen lokalisiert wurden (n=70) (BSW: 53; HOL: 47; RED: 44; SIM: 43). In nur einer Rasse wurden 35 Mutationen entdeckt, wohingegen 13 Varianten in allen sechs untersuchten Populationen erscheinen. Keines dieser 13 vorzeitigen Stopp-Codons zeigt eine Abwesenheit von homozygoten Genotypen an dieser Position (Abbildung 4.2.5).

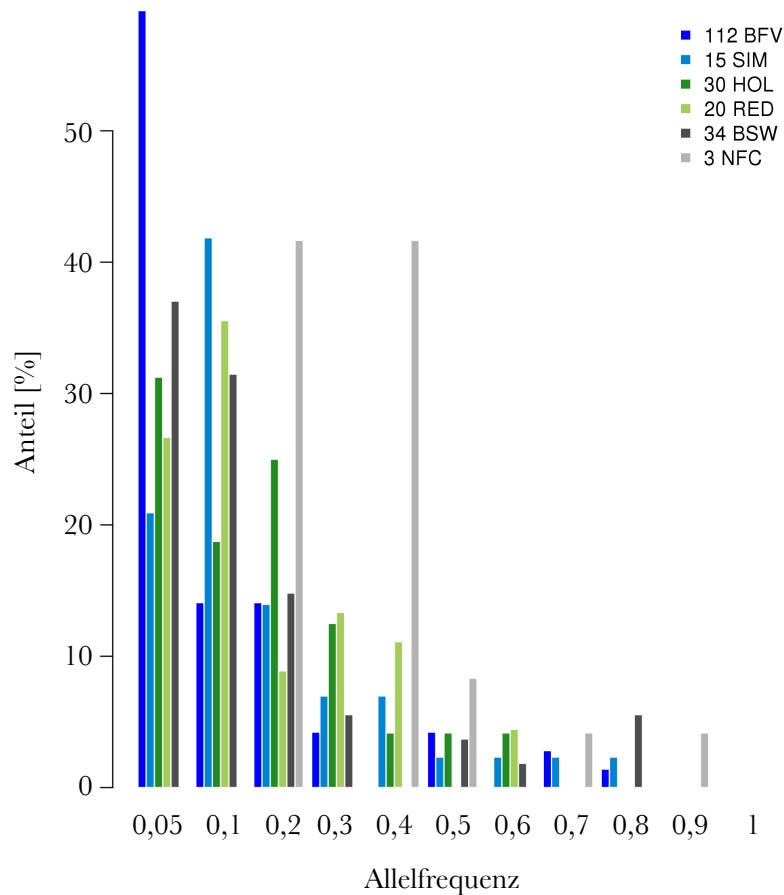


Abbildung 4.2.5: Verteilung der Allelfrequenzen vorzeitiger Stoppmutationen zwischen den Rassen

Die Allelfrequenzen beziehen sich auf die einzelnen Rassen. BFV = Deutsches Fleckvieh; SIM = Simmental; HOL = Holstein, Farbrichtung schwarzbunt; RED = Holstein, Farbrichtung rotbunt; BSW = Brown Suisse; NFC = Nordic Finncattle

Den größten Unterschied zwischen den Allelfrequenzen der einzelnen Stoppmutationen weist *PROKR1* (Chr 11: 66.998.234 bp) auf. *PROKR1* liegt nur in der Rasse BSW vor bei einer Allelfrequenz von 0,794. Von den 13 Varianten beschreibt die Mutation in *LISB* den größten Unterschied in der Allelfrequenz von 0,633 (MIN: 0,167 (HOL); MAX: 0,8 (SIM)).

Die Verteilung der Allelfrequenzen in den einzelnen Rassen verläuft ähnlich wie bei der Betrachtung über alle Rassen. Jedoch zeigen Populationen, in denen wenige Tiere vertreten sind, Abweichungen hinsichtlich des Verlaufs. Je kleiner die Anzahl der Individuen einer Population ist, umso größer sind die Deviationen (Abbildung 4.2.5).

Innerhalb der 214 Genomsequenzen konnte für 38 Stoppmutationen ein homozygoter Genotyp beobachtet werden, von denen 34 auch homozygot im 1000-Bullen-Genom-Projekt erschienen (Tabelle 4.4).

Tabelle 4.4: Anzahl homozygoter und nicht-homozygoter Genotypen

		214 Tiere		
		Homozygote	Keine Homozygote	
1000-Bullen-Genom-Projekt	Homozygote	34	15	49
	Keine homozygote	3	39	42
	Nicht detektiert	1	8	9
		38	62	100

Die 100 vorzeitigen Stopp-Codons, die in dieser Arbeit auf Basis von 214 bovinen Genomsequenzen detektiert wurden, wurden über die UMD3.1-Position im 1000-Bullen-Genom-Projekt überprüft. Die Stoppmutationen wurden in den beiden Datensätzen auf die Präsenz von homozygoten Genotypen kontrolliert.

Bei der Hälfte der Varianten war ein relativ häufiges Auftreten des mutierten Allels zu erkennen (AF > 0,1) mit einer maximalen Allelfrequenz von ~70% (Abbildung 4.2.6). Dahingegen wurden 39 Stopp-Codons weder innerhalb der 214 Tiere noch im 1000-Bullen-Genom-Projekt im homozygoten Zustand entdeckt. Für 37 dieser Mutationen lag die Allelfrequenz unter 5%, während ein SNP eine Allelfrequenz von mehr als 10% aufzeigte (*BOSTAUVIR430*). Jedes Tier trägt im Durchschnitt 12 der 100 hoch qualitativen vorzeitigen Stoppmutationen (MIN: 4, MAX: 20), von denen durchschnittlich zwei im homozygoten Zustand vorkommen (MIN: 0, MAX: 7).

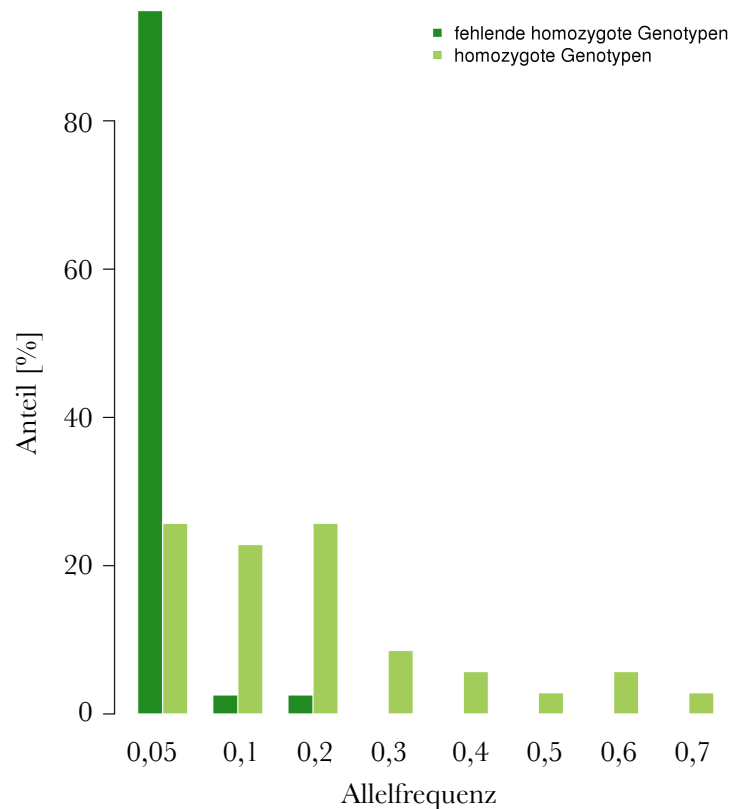


Abbildung 4.2.6: Verteilung der Allelfrequenzen von vorzeitigen Stoppmutationen mit unterschiedlichen Genotypen

Die Allelfrequenzen beziehen sich auf die 214 re-sequenzierten Tiere. Hellgrüne Balken repräsentieren Stoppmutationen, die im homozygoten Zustand erscheinen ($n=36$), dunkelgrüne Balken stellen Stoppmutationen dar, die nicht homozygot zu beobachten waren ($n=49$).

4.2.3 Hinweise vorzeitiger Stopp-Codons auf den Phänotyp

Um einen möglichen Phänotyp für die gefilterten und bestätigten LoF-Varianten ($n=100$) für das Rind abzuleiten, wurden die Gene in den Datenbanken OMIM (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) 2014) und MGI (Blake *et al.* 2014) auf assoziierte Phänotypen hin überprüft. Für 27 Gene ist ein Phänotyp aufgrund von Knockout-Studien an Mäusen beschrieben. In der OMIM-Datenbank sind acht der Gene enthalten, in denen sich Mutationen befinden, die beim Menschen mit Merkmalen assoziiert oder als kausal für vererbte Defekte beschrieben sind (Tabelle 4.5).

Ergebnisse

Tabelle 4.5: Stoppmutationen in Genen mit bekanntem Phänotyp

Chr	Position	REF	ALT	Gen	Knockout	OMIM	Homo	AF	Rasse
19	27333935	C	T	<i>ALOX15</i>	myoproliferatives Syndrom (MIDDLETON <i>et al.</i> 2006; TAYLOR <i>et al.</i> 2012)	-	ja	0,0400	SIM, BSW, BFV, GCO
21	62743832	C	T	<i>BDKRB2</i>	Schutz vor Thrombose (SHARIAT-MADAR <i>et al.</i> 2006)	-	ja	0,0260	NFC, BFV, BBO, LIM, SIM
23	25620856	G	T	<i>BTNL2</i>	-	erhöhte Anfälligkeit für Sarkoidose (VALENTONYTE <i>et al.</i> 2005; RYBICKI <i>et al.</i> 2005)	ja	0,0230	BSW, BFV, CHA, BBL, LIM, JER, RDC, SIM, AYR
1	264831	T	A	<i>CBX3</i>	Infertilität (BROWN <i>et al.</i> 2010; TAKADA <i>et al.</i> 2011)	-	ja	0,1140	NFC, SIM, HOL, BSW, BFV, RED, BBO, LIM, JER, ANG, RDC, AYR
8	11434686	C	T	<i>DDX58</i>	Anfälligkeit ggü. Virusinfektionen (KATO <i>et al.</i> 2006), Kolitis-ähnlicher Phänotyp (WANG <i>et al.</i> 2007)	-	ja	0,0630	SIM, HOL, BSW, BFV, BBO, CHA, LIM, JER, ANG
25	2993664	C	G	<i>DNASE1</i>	Symptome eines systemischen Lupus erythematodes (NAPIREI <i>et al.</i> 2000)	Systemischer Lupus erythematodes (YASUTOMO <i>et al.</i> 2001; SHIN <i>et al.</i> 2004)	ja	0,0300	BSW, BFV, PIE, SIM
24	25915403	A	C	<i>DSG2</i>	embryonale Sterblichkeit (ESHKIND <i>et al.</i> 2002)	Arrhythmogene rechtsventrikuläre Kardiomyopathie (PILICHOU <i>et al.</i> 2006; AWAD <i>et al.</i> 2006), dilatative Kardiomyopathie (POSCH <i>et al.</i> 2008)	ja	0,0420	SIM, HOL, BFV, RED, ANG
7	98612967	G	A	<i>ERAPI</i>	reduzierte Immunantwort ggü. Viren (YORK <i>et al.</i> 2006; FIRAT <i>et al.</i> 2007)	-	nein	0,0093	BSW, JER, SIM
8	29433977	C	T	<i>FREMI</i>	Fraser Syndrom und Frühtod (SMYTH <i>et al.</i> 2004; KIZOZUMI <i>et al.</i> 2012)	Anomalien der Nase mit renalen und anorektalen Missbildungen (AL-GAZALI <i>et al.</i> 2002; ALAZAMI <i>et al.</i> 2009); Manitoba okulotrichoanales Syndrom (SLAVOTINEK <i>et al.</i> 2011); Trigonozephalie (VISSERS <i>et al.</i> 2011)	nein	0,0070	HOL, RED
19	38204969	C	T	<i>GIP</i>	verminderte Fettsucht (MIYAWAKI <i>et al.</i> 2002)	-	ja	0,0093	BFV
8	6857405	G	A	<i>GLRA3</i>	reduzierte Schmerzsensibilisierung (HARVEY <i>et al.</i> 2004)	-	nein	0,0120	BSW

Ergebnisse

Chr	Position	REF	ALT	Gen	Knockout	OMIM	Homo	AF	Rasse
3	33798941	A	T	<i>GSTM1</i>	Autismus (YOCHUM <i>et al.</i> 2010)	-	ja	0,1000	NFC, SIM, HOL, BSW, BFV, CHA, LIM, ANG, RDC, AYR
16	4484297	C	T	<i>IL19</i>	Anfälligkeit ggü. Kolitis (AZUMA <i>et al.</i> 2010), epidermale Hyperplasie (CHAN <i>et al.</i> 2006)	-	ja	0,1290	NFC; SIM, HOL, BSW, BFV, RED, BBO, CHA, LIM, JER, ANG, AYR
3	17883354	T	A	<i>IVL</i>	reduzierte Proteolyse (SEVILLA <i>et al.</i> 2007)	-	ja	0,0070	BFV, RED, BBO, SIM
13	78195467	G	C	<i>KCNB1</i>	Erhöhte Insulinsekretion (JACOBSON <i>et al.</i> 2007)	-	ja	0,0120	SIM, BFV, PIE
5	27490704	C	T	<i>KRT74</i>	-	Hypotrichose (WASIF <i>et al.</i> 2011); gekräuseltes Haar (SHIMOMURA <i>et al.</i> 2010; WASIF <i>et al.</i> 2011)	ja	0,0140	BSW, BFV, BBL, LIM, ANG, RDC, SIM
16	65368596	G	T	<i>NPL</i>	Immunologische Anomalien (BLAKE <i>et al.</i> 2014)	-	nein	0,0140	SIM, BSW, BFV
6	2386214	G	A	<i>NPY5R</i>	reduzierte Futteraufnahme (KANATANI <i>et al.</i> 2000); Gewichtszunahme (ISHIHARA <i>et al.</i> 2006)	-	ja	0,0350	BSW, BFV, SIM
17	72089872	C	T	<i>PLA2G3</i>	reduzierte männliche Fruchtbarkeit (SAITO <i>et al.</i> 2010); kutane Anaphylaxie (TAKETOMI <i>et al.</i> 2013)	-	ja	0,0190	HOL, RED, HER, LIM, ANG, AYR
26	37302490	C	G	<i>PNLIP</i>	verzögerte Absorption von Nahrungsfetten (HUGGINS <i>et al.</i> 2003); geringere Gewichtszunahme (GILHAM <i>et al.</i> 2007)	Exokrine Pankreasinsuffizienz (FIGARELLA <i>et al.</i> 1980)	ja	0,0280	BFV, RED
24	54170070	C	G	<i>POLI</i>	erhöhte Anfälligkeit für Hautkrebs (OHKUMO <i>et al.</i> 2006)	-	nein	0,0070	HOL
14	21121932	G	A	<i>PRKDC</i>	Immundefizienz (GAO <i>et al.</i> 1998); Abnormalitäten im Verdauungstrakt (KURIMASA <i>et al.</i> 1999)	-	nein	0,0023	SIM
3	117646636	C	T	<i>PRLH</i>	Polyphagie und Fettleibigkeit (TAKAYANAGI <i>et al.</i> 2008)	-	ja	0,1540	SIM, HOL, BSW, BFV, RED, BBO, CHA, BBL, LIM, PIE, GCO, JER, ANG, RDC, AYR

Ergebnisse

Chr	Position	REF	ALT	Gen	Knockout	OMIM	Homo	AF	Rasse
11	66998234	C	A	<i>PROKR1</i>	Anomalien des kardiovaskulären Systems (URAYAMA <i>et al.</i> 2008); gestörte Schmerzempfindung (NEGRI <i>et al.</i> 2006)	-	ja	0,1260	BSW
6	4152699	C	T	<i>QRFPR</i>	Kyphose (BARIBAULT <i>et al.</i> 2006)	-	ja	0,3290	NFC, SIM, HOL, BSW, BFV, RED, BBO, CHA, BBL, LIM, PIE, JER, ANG, RDC, AYR
21	55891444	G	A	<i>STRC</i>	gestörtes Hörvermögen (VERPY <i>et al.</i> 2008)	Taubheit (VERPY <i>et al.</i> 2001)	ja	0,0120	HOL, RED
4	106103710	G	T	<i>TAS2R38</i>	-	bittere Geschmacksempfindung (KIM <i>et al.</i> 2003; WOODING <i>et al.</i> 2004)	nein	0,0140	NFC, BSW, BBO, CHA, LIM, GCO, ANG
1	70697630	C	T	<i>TMED2</i>	embryonaler Fröhntod, embryonale Wachstumsverzögerung (JEROME-MAJEWSKA <i>et al.</i> 2010)	-	ja	0,0400	NFC, SIM, HOL, BSW, BFV, RED, BBO, BBL, LIM, ANG
3	42748263	A	C	<i>VCAMI</i>	Embryonaler Fröhntod (KWEE <i>et al.</i> 1995b; a; GURTNER <i>et al.</i> 1995)	-	nein	0,0047	BSW
3	26175832	C	T	<i>VTCNI</i>	geringere Anfälligkeit ggü. Infektionen (ZHU <i>et al.</i> 2009)	-	ja	0,0280	SIM, BFV, PIE

Die Gene, die ein vorzeitiges Stopp-Codon inkludierten, wurden systematisch in den OMIM- und MGI-Datenbanken analysiert. Gene, die mit einem Phänotyp beschrieben sind, sind in der Tabelle alphabetisch sortiert aufgeführt. Chr= bovines Chromosom; bp = Position auf dem Chromosom; REF = Referenzallele; ALT = Alternatives Allel; Homo = Detektion homozygoter Tiere; AF = Allelfrequenz des alternativen Allels in den 214 genomweit-resequenzierten Tieren; Rasse = Rassen, in denen die Mutation segregiert.

Die beschriebenen Phänotypen dieser Gene variieren von der Beeinflussung des Haarkleids (z.B. *KRT74*) über die Beeinträchtigung des Stoffwechsels (z.B. *GIP*, *KCNB1*, *NPY5R*, *PNLIP*, *PRLH*) und Veränderungen bzw. Störungen der Immunantwort (*DDX58*, *ERAP1*, *VTCNI*) bis hin zu Fehlbildungen und Anomalien (*PRKDC*, *PROKR1* und *QRFPR*) sowie zum embryonalen Tod (*DSG2*, *TMED2*, *VCAMI*).

Die höchste Allelfrequenz liegt für eine Stoppmutation im Gen *QRFPR* vor (AF=0,3290). Die Mutation segregiert mit Ausnahme von GCO und HER in allen Rassen mit einer Allelfrequenz von mindestens sechs Prozent bei Anwesenheit von homozygoten Tieren. Diese Stoppmutation wird in Kapitel 4.2.4.2.2 genauer analysiert.

Bezogen auf alle untersuchten Tiere beträgt die geringste Allelfrequenz einer Stoppmutation 0,0023 und es wurde kein homozygoten Tier entdeckt. Diese Variante segregiert nur in der SIM-Population (AF=0,033), wo es in nur einem einzigen Tier heterozygot erscheint. Im 1000-Bullen-Genom-Projekt wurde die Mutation nicht detektiert. Die LoF-Variante ist in dem Gen *PRKDC* lokalisiert, das für die DNA-abhängige katalytische Untereinheit der Proteinkinase kodiert. Mäuse aus Knockout-Studien, bei denen dieses Gen ausgeschaltet wurde, weisen einerseits eine kombinierte Immundefizienz (GAO *et al.* 1998) und andererseits Anomalien im Verdauungstrakt auf (KURIMASA *et al.* 1999). Die Mutation verändert ein Arginin zu einem Stopp-Codon (c.3550C>T; p.R1184X) und führt zu einer Verkürzung der Peptidsequenz um 2940 Aminosäuren (71,3%). Aufgrund des extrem seltenen Auftretens dieser Mutation kann keine Aussage über dessen Wirkung getroffen werden.

Mutationen, die kausal für autosomal rezessive Letalfaktoren sind, treten homozygot nicht in Erscheinung. Tiere mit dem homozygoten Genotyp einer solchen Mutation sind nicht überlebensfähig. Um mögliche rezessive letale Effekte zu identifizieren, wurden die Genotypverteilung und die Anwesenheit von homozygoten Tieren an den hoch qualitativen Stoppmutationen mit Phänotypinformationen betrachtet. Basierend auf den Genomsequenzen und der *Sanger-Re-Sequenzierung* konnte für acht LoF-Varianten, die sich in Genen mit beschriebenem Phänotyp befinden, kein homozygoter Genotyp detektiert werden (*ERAPI*, *FREMI*, *GLRA3*, *NPL*, *POLI*, *PRKDC*, *TAS2R38* und *VCAMI*) (Tabelle 4.5). Um einen potenziellen schädlichen Effekt dieser Mutationen zu identifizieren, sollten die Allelfrequenzen der Polymorphismen berücksichtigt werden. Die Allelfrequenzen dieser vorzeitigen Stopp-Codons mit fehlenden Homozygoten reichen über alle Rassen gesehen von 0,23 bis maximal 1,40%. In den einzelnen Populationen waren die Allelfrequenzen für diese Mutationen mit fehlenden Homozygoten so gering, dass bei gegebener Anzahl an total re-sequenzierten Individuen kein homozygoten Tier erwartet wurde. Allein aufgrund der Allelfrequenzen konnte somit nicht auf einen letalen Effekt geschlossen werden. Interessanterweise lösen fünf dieser Stoppmutationen mit fehlenden Homozygoten möglicherweise einen NMD aus (*ERAPI*, *FREMI*, *GLRA3*, *POLI* und *VCAMI*). Studien konnten zeigen, dass Mutationen in zwei dieser Gene (*FREMI* und *VCAMI*) mit einem frühzeitigen Tod assoziiert sind (GURTNER *et al.* 1995; KWEE *et al.* 1995a; SMYTH *et al.* 2004; KIYOZUMI *et al.* 2012). Die Stoppmutationen in diesen beiden Genen werden im weiteren Verlauf genauer analysiert (Kapitel 4.2.3.1 und 4.2.3.2).

Neben den beiden Genen *FREMI* und *VCAM1*, deren Knockout sich letal auswirken kann, löst der Verlust der Proteine von zwei weiteren Genen (*DSG2* und *TMED2*) einen vergleichbaren Phänotyp aus. Das Ausschalten dieser Gene bei Mäusen führt zum embryonalen Tod oder Frühtod (ESHKIND *et al.* 2002; JEROME-MAJEWSKA *et al.* 2010). Die Vermutung ist, dass vorzeitige Stopp-Codons in diesen Genen, insbesondere im homozygoten Zustand, beim Rind zu einem ähnlichen Phänotyp führen können und sich letal auswirken. Es konnte jedoch nachgewiesen werden, dass die Stoppmutationen in den Genen *DSG2* und *TMED2* nicht den letalen Phänotyp auslösen. Die Polymorphismen dieser Gene lagen im homozygoten Zustand für das alternative Allel vor und bestätigen damit, dass Tiere mit einer doppelten Kopie des vorzeitigen Stopp-Codons überlebensfähig sind.

4.2.3.1 Stoppmutation in *VCAM1*

VCAM1, kodierend für das vaskuläre Adhäsionsmolekül 1, auch bekannt als CD106, ist ein transmembranes Glykoprotein. Als Mitglied der Immunoglobulin(Ig)-Gen-Unterfamilie ist *VCAM1* an zahlreichen Entzündungsprozessen beteiligt. Während einer Entzündungsreaktion haftet es an den Integrinen, die in der Membran von Leukozyten und Thrombozyten verankert sind (OSBORN *et al.* 1989; ELICES *et al.* 1990). Außerdem ist *VCAM1* in die Reifung von T- und B-Zellen involviert. Die Expression des Gens findet in einer Vielzahl von vaskulären Zellen statt und wird insbesondere durch inflammatorische Zytokine gesteigert (CARLOS *et al.* 1990; CONSTANS und CONRI 2006). Bei Mäusen wurde *VCAM1*-mRNA zusätzlich in folliculären dendritischen Zellen und in Knochenmark-Zellen entdeckt (FREEDMAN *et al.* 1990; MIYAKE *et al.* 1991).

Neben seiner Transmembrandomäne besitzt *VCAM1* sieben homologe Ig-ähnliche Domänen (OSBORN *et al.* 1989; POLTE *et al.* 1991) sowie zwei Bindestellen für α_4 Integrin in der ersten und vierten Domäne der extrazellulären Region (OSBORN *et al.* 1994; VONDERHEIDE *et al.* 1994).

Der Verlust der Funktion von *VCAM1* ist bei der Maus mit dem embryonalen Frühtod assoziiert. GURTNER *et al.* (1995) erzeugten *VCAM1*-defiziente Mäuse, die aufgrund einer gestörten Verbindung zwischen dem Embryo und der Plazenta in der mittleren Trächtigkeit starben. Weitere Knockout-Studien bei Mäusen konnten eine Beeinträchtigung der Überlebensfähigkeit bestätigen (KWEE *et al.* 1995a; b). Einige dieser nicht lebensfähigen Tiere wiesen Defekte am Herzen auf.

Im bovinen Genom ist das Gen *VCAM1* auf dem negativen Strang des dritten Chromosoms lokalisiert (42.743.604 - 42.766.061 bp). Das ~22,4 kb große Gen wird in neun Exons unterteilt (Abbildung 4.2.7), die ein 739 Aminosäuren langes Transkript kodieren.

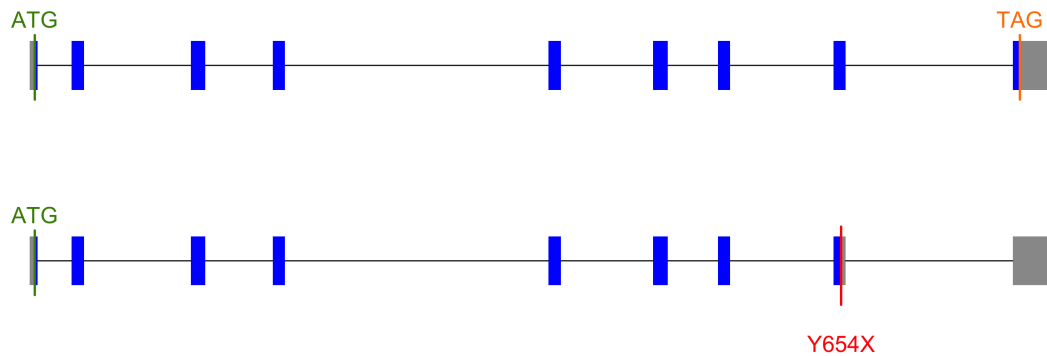


Abbildung 4.2.7: Genstruktur von *VCAM1*

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.Y654X) ist rot gekennzeichnet.

In den vorliegenden Daten konnte eine Stoppmutation in *VCAM1* identifiziert werden, die nur in BSW-Rindern segregiert ($AF(214) = 0,029$; $AF(1000) = 0,012$). Zwei BSW-Tiere tragen die Mutation heterozygot, ALT-homozygote Tiere wurden nicht beobachtet. Bei der geringen Allelfrequenz und Tieranzahl ($n(214) = 34$; $n(1000) = 43$) wurde keine signifikante Abweichung vom HWG festgestellt ($P\text{-Wert} = 0,8597$). Die Stoppmutation könnte basierend auf den Regeln für den NMD (Kapitel 4.2.1) zu einem kompletten Verlust des Proteins führen. Falls es nicht zum enzymatischen Abbau der mRNA kommt, entsteht ein um 11,5% verkürztes Protein mit einem Verlust des letzten und einem Teil des vorletzten Exons (c.1962T>G; p.Y654X) (Abbildung 4.2.7). Diese Region ist über die Spezies hinweg hoch konserviert. Die Transmembrandomäne wird möglicherweise aufgrund des vorzeitigen Stopp-Codons nicht translatiert (Abbildung 4.2.8). Es bleibt jedoch offen, inwieweit die Stoppmutation die Funktion des Gens beeinflusst, da nicht jede Isoform des Gens von der Variante beeinträchtigt wird.

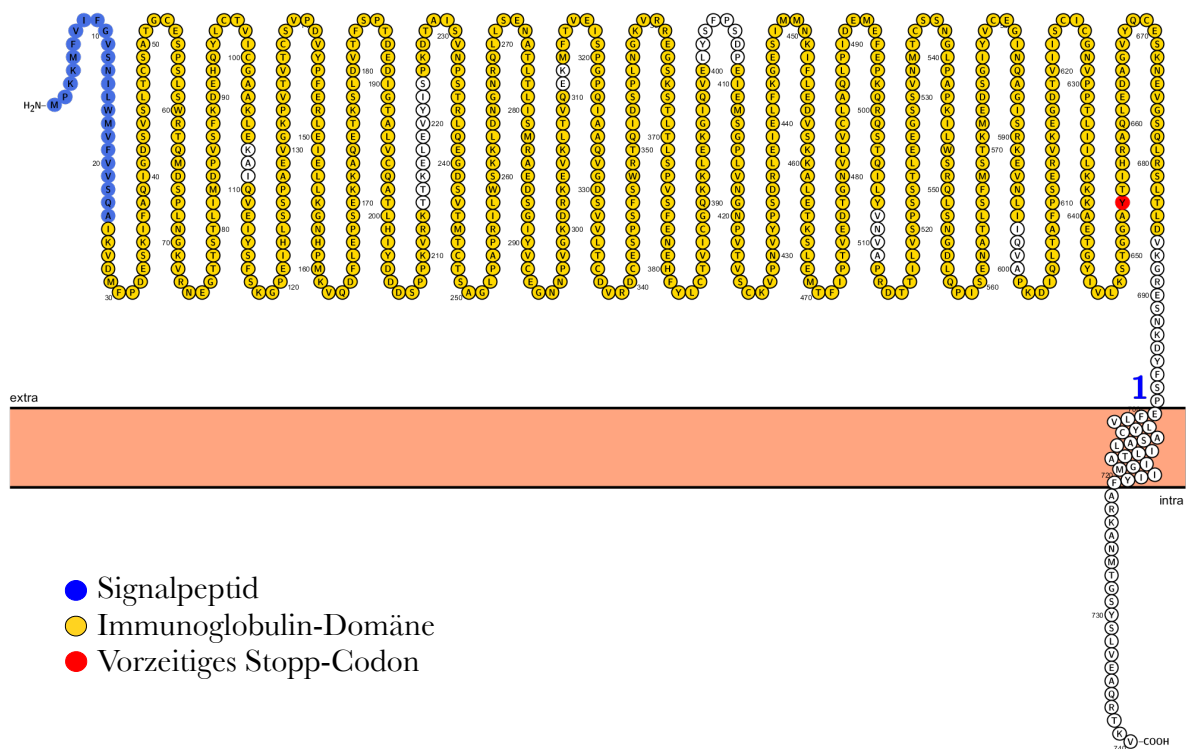


Abbildung 4.2.8: Transmembranstruktur von VCAM1

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS *et al.* 2013) die Positionen der Transmembrandomäne, der Immunoglobulin-(Ig)-Domänen und des Signalpeptids visualisiert. Das vorzeitige Stopp-Codon (p.Y654X) ist rot gekennzeichnet.

4.2.3.2 Stoppmutation in *FREM1*

Eine weitere Stoppmutation, die Hinweise auf einen letalen Effekt gibt, ist im Gen *FREM1* lokalisiert. Das Gen kodiert das FRAS1-verwandte extrazelluläre Matrixprotein 1. Dieses Protein wird während der Embryonalentwicklung bei der Formung der Epidermiszellen exprimiert, insbesondere in der nasalen Region, an den Augenlidern, bei der Verbindung der Schädelplatten sowie in der Niere (SMYTH *et al.* 2004; ALAZAMI *et al.* 2009). Es wird für die Adhäsion der Basalmembran benötigt, die für die Stabilisierung der Epithelschicht verantwortlich ist (SMYTH *et al.* 2004).

FREM1 besitzt eine C-Typ-Lektin- (AS 2061-2175) und eine Calx-beta-Domäne (AS 1732-1831) (SMYTH *et al.* 2004). C-Typ-Lektin-Domänen agieren bei der Verbindung zwischen einzelnen Zellen, bei der Immunantwort zu Pathogenen sowie bei der Apoptose (DRICKAMER 1999; CAMBI and FIGDOR 2009). Das Calx-beta-Motiv ist für die Ausscheidung von Kalzium und den Austausch mit Natrium in den Zellen zuständig (SCHWARZ und BENZER 1997). Dieser Austausch ist besonders

wichtig bei elektrisch erregbaren Zellen, wie z.B. Muskel- und Nervenzellen (DiPOLO und BEAUGÉ 2006). *FREM1* gehört außerdem zu den sogenannten 12-CSPG Proteinen, die zwölf Kollagen bindende CSPG-repeats aufweisen (STAUB *et al.* 2002; SMYTH *et al.* 2004).

FREM1-Knockout-Studien mit Mäusen zeigten, dass diese Tiere an dem autosomal rezessiv vererbten Fraser-Syndrom leiden und früher sterben (SMYTH *et al.* 2004; KIYOZUMI *et al.* 2012). Beim Menschen bewirken Mutationen im Gen *FREM1* einen ähnlichen Phänotyp mit Anomalien des Kopfes, des Gesichts, der Nieren und des Enddarms (AL-GAZALI *et al.* 2002; ALAZAMI *et al.* 2009; SLAVOTINEK *et al.* 2011; VISSERS *et al.* 2011). Diese Erscheinungsbilder sind bekannt als Manitoba-okulo-tricho-anales Syndrom (OMIM 248450), die Trigonecephaly (OMIM 514485) und die zweigeteilte Nase mit anorektalen Anomalien (OMIM 608980).

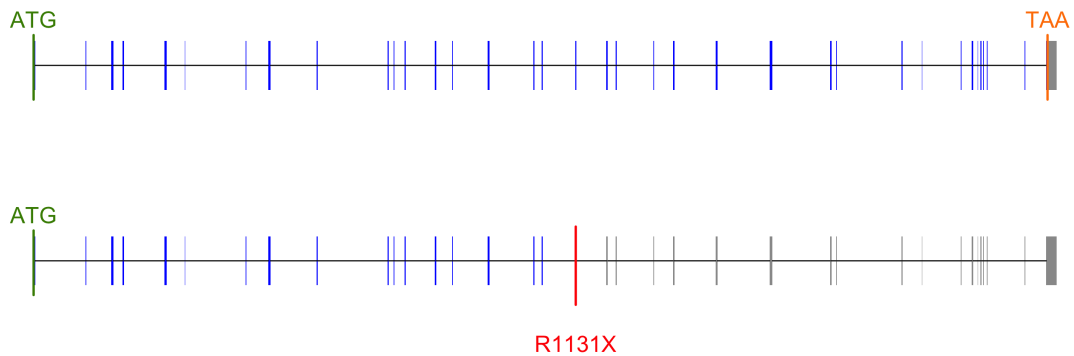


Abbildung 4.2.9: Genstruktur von *FREM1*

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.R1131X) ist rot gekennzeichnet.

Das bovine *FREM1* befindet sich auf dem positiven Strang von Chromosom 8. Die 36 Exons erstrecken sich über ~152 kb (29.353.304 - 29.505.352 bp), die ein 2180 Aminosäuren langes Protein kodieren (Abbildung 4.2.9). In den Rassen HOL und RED wurde ein vorzeitiges Stoppcodon identifiziert. Die Mutation lag in keinem Tier homozygot vor. Bei den gegebenen Allelfrequenzen in den Populationen (HOL: 2,6%; RED: 7,1%) wurde kein ALT-homozygotes Tier erwartet. Die Mutation befindet sich im 18. von insgesamt 36 Exons (Abbildung 4.2.9) und ist ein potentieller Kandidat, um einen NMD herbeizuführen. Falls dies nicht der Fall ist, wäre die Aminosäuresequenz um etwa die Hälfte reduziert (48,12%), wobei sich die LoF-Variante nur auf einzelne Isoformen des Proteins von Mensch, Maus und Rind beschränkt. Mit Ausnahme des

betroffenen Codons (c.3391C>T; p.R1131X) ist die umgebende Region hoch konserviert. Das vorzeitige Stopp-Codon wird die Translation der Calx-beta- und der C-Typ-Lektin-Domänen verhindern (Abbildung 4.2.10).

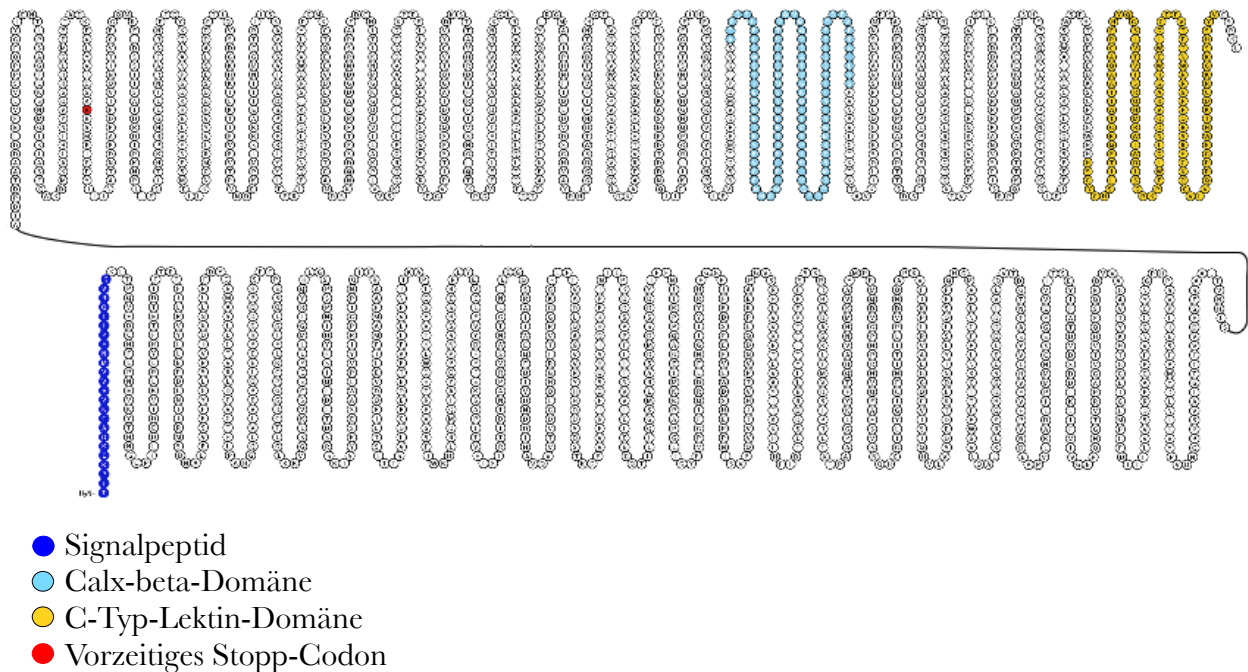


Abbildung 4.2.10: Proteinstruktur von FREM1

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS *et al.* 2013) die Positionen der Domänen und des Signalpeptids visualisiert. Das vorzeitige Stopp-Codon (p.R1131X) ist rot gekennzeichnet.

Vorzeitige Stopp-Codons müssen nicht notwendigerweise einen letalen Einfluss auf oder einen selektiven Nachteil für das Individuum haben. Sie können sich ebenso schädlich auf ein Protein auswirken, das für das Überleben eines Tieres eine untergeordnete Rolle spielt. Dabei sind *STRC* und *TAS2R38* phänotypisch interessante Gene (Tabelle 4.5).

4.2.3.3 Stoppmutation in STRC

Das Protein STRC gehört zur Stereozilin-Familie. Dieses extrazelluläre Protein wird hauptsächlich in den Sinneshaarzellen im Ohr (Stereozilien), aber auch im Hoden und in geringen Mengen im Auge exprimiert (VERPY *et al.* 2001, 2011). Im Ohr ist Stereozilin über einen Glykosylphosphatidylinositol-Anker (GPI) mit der Membran der sensorischen und nicht-sensorischen Zellen verbunden (VERPY *et al.* 2001). SATHYANARAYANA *et al.* (2009) vermuten, dass Stereozilin eine

Superhelix-Struktur besitzt, die für die Interaktion mit extrazellulären Glykoproteinen verantwortlich ist.

Mutationen im Gen *STRC* werden bei Mensch und Maus mit einem autosomal rezessiv vererbten Phänotyp assoziiert. *STRC* defiziente Mäuse weisen ein beeinträchtigtes Hörvermögen auf (VERPY *et al.* 2008) und auch beim Menschen wurden Polymorphismen im Gen *STRC* als kausal für die Taubheit beschrieben (OMIM 603720) (VERPY *et al.* 2001; AVIDAN *et al.* 2003; ZHANG *et al.* 2007).

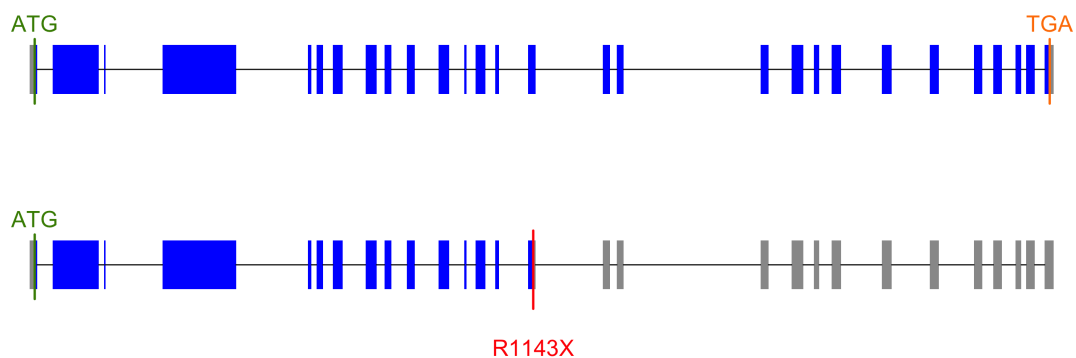


Abbildung 4.2.11: Genstruktur von *STRC*

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.R1143X) ist rot gekennzeichnet.

Das bovine *STRC* konnte auf dem negativen Strang des 21. Chromosoms identifiziert werden (55.882.606 - 55.899.996 bp). Die 28 kodierenden Exons ergeben ein 1754 Aminosäuren langes Stereozilin. In den Genomsequenzen der 214 Tiere konnte ein vorzeitiges Stopp-Codon in diesem Gen bestätigt werden. Die Mutation trat ausschließlich bei HOL- und RED-Individuen auf, von denen ein Tier den ALT-homozygoten Genotyp trug. Die Allelfrequenzen im 1000-Bullen-Genom-Projekt lagen bei 0,122% (HOL) und 0,071% (RED). Die LoF-Variante befindet sich im 15. Exon im 1143. Codon (c.3427C>T; p.R1143X) (Abbildung 4.2.11). Dieser Bereich ist konserviert, jedoch nicht in allen untersuchten Isoformen vorhanden. Vermutlich löst die Mutation keinen Abbau der mRNA aus, jedoch würde das Protein durch die Mutation um 35,2% verkürzt sein. Dieser Verlust eines großen Teils der Aminosäuresequenz könnte das Protein in seiner Funktion stark beeinträchtigen.

4.2.3.4 Stoppmutation in *TAS2R38*

Eine weitere Mutation, die beim Rind von Bedeutung sein könnte, befindet sich im Gen *TAS2R38*, das für den Bitter-Geschmack-Rezeptor TAS2R38 kodiert. Neben 25 weiteren funktionellen Genen gehört es der Familie der Geschmacksrezeptoren an (ADLER *et al.* 2000; CONTE *et al.* 2002), wobei TAS2R38 als einziger Rezeptor zur Wahrnehmung des bitteren Geschmacks gilt (BACHMANOV und BEAUCHAMP 2007). Der G-Protein gekoppelte Rezeptor ist ein Transmembranprotein mit sieben Transmembrandomänen (ADLER *et al.* 2000). Er befindet sich neben weiteren Geschmacksrezeptoren in den Geschmackszellen der Zunge (CONTE *et al.* 2002). Nicht-synonyme Aminosäureaustausche im Gen *TAS2R38* ermöglichen dem Menschen den bitteren Geschmack in Nahrungsmitteln wahrzunehmen (KIM *et al.* 2003; WOODING *et al.* 2004). So weisen Menschen, die den ALT-homozygoten Genotyp tragen, die höchste Empfindlichkeit gegenüber bitterem Geschmack auf.

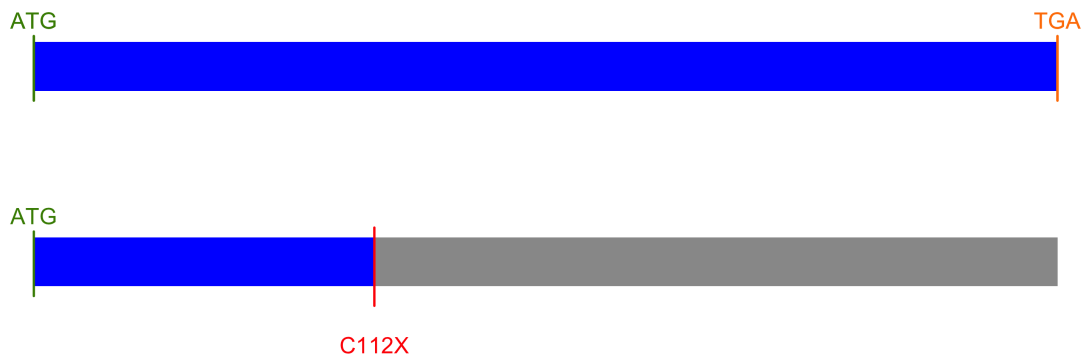


Abbildung 4.2.12: Genstruktur von *TAS2R38*

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.C112X) ist rot gekennzeichnet.

TAS2R38 wurde auf dem 4. Chromosom des bovinen Genoms identifiziert. Das auf dem negativen Strang lokalisierte Gen mit einer Länge von ~1kb kodiert 335 Aminosäuren. Beim Rind wurde in sieben Rassen ein vorzeitiges Stopp-Codon, bedingt durch einen SNP im einzigen Exons des Gens identifiziert. Bei einer Frequenz von 1,4% in 214 re-sequenzierten Tieren wurde der ALT-homozygote Genotyp in keiner Population beobachtet. Bei den beobachteten Allelfrequenzen in den einzelnen Rassen (MIN: 3,7%; MAX: 11,1%) und gegebener Individuenanzahl wird in keiner Rasse ein ALT-homozygotes Rind erwartet. Zwar wird die LoF-Variante

(c.336C>A.; p.C112X) vermutlich keinen Abbau der mRNA induzieren, jedoch verursacht sie eine Verkürzung der Peptidsequenz um 66,56% (Abbildung 4.2.12).

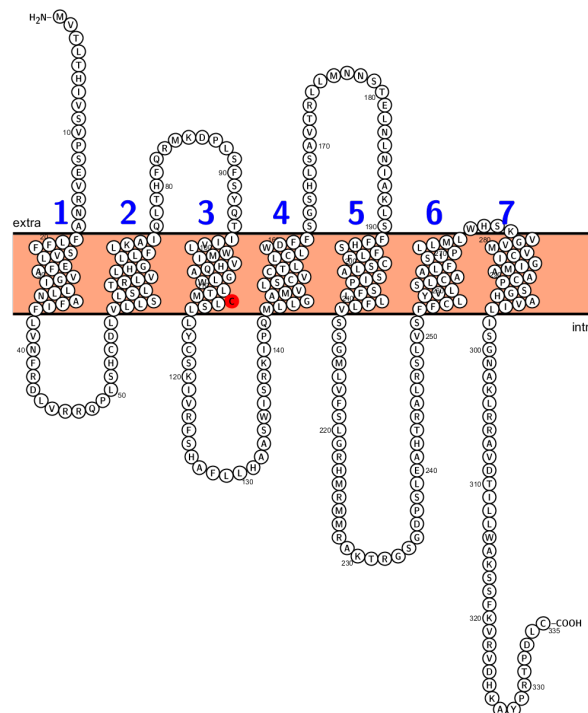


Abbildung 4.2.13: Transmembranstruktur von TAS2R38

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS et al. 2013) die Positionen der Transmembrandomänen visualisiert. Das vorzeitige Stopp-Codon (p.C112X) ist rot gekennzeichnet.

Das vorzeitige Stopp-Codon ist in der dritten Transmembrandomäne lokalisiert (Abbildung 4.2.13) und würde zu einem Verlust der vier nachfolgenden Domänen führen. Alle untersuchten Isoformen in Rind, Mensch und Maus sind von dieser Mutation betroffen. Das Protein ist stark konserviert.

4.2.4 Vorzeitige Stopp-Codons im Deutschen Fleckvieh

In den 112 genomweit re-sequenzierten Tieren der BFV-Population segregieren 70 der 100 hochqualitativen Stoppmutationen. Neben der Sanger-Re-Sequenzierung wurden die Varianten zur Validierung in der gesamten BFV-Population in einem imputierten Datensatz überprüft. Der Datensatz besteht aus 10.363 Tieren der Rasse BFV. Für jedes der 10.363 Individuen stehen Genotypinformationen für die sequenzbasierten Mutationen aus der Genomsequenzierung der

214 Tiere zur Verfügung (Kapitel 4.1). Die sequenzbasierten Genotypen wurden *via* der Genotypinformationen des bovinen HD-BeadChips von Illumina imputiert (nach PAUSCH *et al.* 2013). Des Weiteren wurde die HD-Information (614K) nach möglichen Stoppmutation tragenden Haplotypen überprüft.

Alle 70 LoF-Varianten, die in den Sequenzen der 112 BFV-Tiere identifiziert wurden, lagen im sequenz-imputierten Datensatz vor. Neun dieser 70 Mutationen segregieren nicht im imputierten Datensatz, da nur der REF-homozygote Genotyp detektiert wurde. Die Allelfrequenzen dieser neun Varianten waren in den 112 genomweit re-sequenzierten BFV-Tieren relativ niedrig (AF < 5%). Eine Ausnahme ist die Mutation im Gen *BOSTAUVIR430*, die eine Frequenz von 20,1% aufwies. Fünf dieser neun Mutationen kamen in nur je einem BFV-Tier heterozygot vor (AF = 0,4%). Die verbleibenden 61 Varianten wurden für weitere Analysen in BFV herangezogen.

4.2.4.1 Vorzeitige Stopp-Codons ohne homozygote Genotypen im Deutschen Fleckvieh

Um in der Rasse BFV mögliche autosomal rezessiv vererbte Stoppmutationen mit letalem Effekt aufzudecken, wurden zunächst Mutationen untersucht, die keinen homozygoten Genotyp in den imputierten Daten aufwiesen. Für zwölf Varianten der 61 Stoppmutationen konnte in den 10.363 BFV-Tieren kein homozygoter Genotyp gefunden werden. Jedoch wurden für drei dieser LoF-Varianten homozygote Tiere in anderen Rassen der 214 Tiere oder im 1000-Bullen-Genom-Projekt beobachtet (*DSG2*, *SHBG* und *SULT6B1*). Eine weitere Mutation im Gen *WDR8* konnte mit Hilfe der *Sanger*-Re-Sequenzierung homozygot entdeckt werden. Somit können diese vier Varianten als mögliche kausale Mutationen für einen letalen Effekt ausgeschlossen werden. Die verbleibenden acht Stoppmutationen wurden in keinem der vorhandenen Datensätze homozygot erfasst (Tabelle 4.6). Die Allelfrequenzen dieser Mutationen mit fehlenden Homozygoten betragen innerhalb der 10.363 Tiere 0,006% (*GYLTL1B*) bis 1,25% (*GPR128*), wobei die Mutationen mit der geringsten Allelfrequenz im imputierten Datensatz nicht die niedrigste Frequenz bei den 112 genomweit re-sequenzierten BFV-Tieren aufwiesen. Die Stoppmutation in *GYLTL1B* hatte bei der Genomsequenzierung in den 112 BFV-Rindern eine Frequenz von 0,4%, nach der Imputation sank sie auf 0,006%. Dagegen stieg die innerhalb der 112 BFV-Rinder niedrigste Allelfrequenz der Variante in *ENOSF1* (0,4%) auf 1,15% im imputierten Datensatz an.

Tabelle 4.6: Übersicht der Stoppmutationen mit fehlenden Homozygoten im Deutschen Fleckvieh

Chr	Position	Gen	AF (NGS)	AF (imputiert)	P-Wert (HWG)	AF (Haplo)
24	42.48.577	<i>CNDP1</i>	0,004	0,0011	0,91386	-
2	95.390.113	<i>DYTN</i>	0,004	0,0028	0,778913	-
24	35.956.168	<i>ENOSF1</i>	0,004	0,0115	0,235	-
1	45.468.030	<i>GPR128</i>	0,009	0,0125	0,197671	0,01
15	76.831.914	<i>GYLTL1B</i>	0,009	0,00006	0,952973	-
3	66.286.574	<i>IFI44</i>	0,013	0,0057	0,559965	-
27	14.698.133	<i>LRP2BP</i>	0,018	0,0069	0,476325	0,0066
16	65.368.596	<i>NPL</i>	0,009	0,0053	0,590439	-

Die Stoppmutationen sind alphabetisch nach den Namen der Gene sortiert. AF (NGS) = Allelfrequenz in BFV basierend auf den Genomsequenzen in 112 Tieren; AF (imputiert) = Allelfrequenz in BFV basierend auf den Genotypen in 10.363 Tieren; P-Wert (HWG) = Irrtumswahrscheinlichkeit, dass die Genotypverteilung in den 10.363 BFV Tieren vom HWG abweicht; AF (Haplo) = Status und Frequenz des identifizierten Haplotypen.

Für keine dieser LoF-Varianten mit fehlenden Homozygoten konnte eine signifikante Abweichung vom HWG festgestellt werden (P-Wert = 0,19 - 0,95).

Mit Hilfe von 614K-Genotypinformationen für die 10.363 BFV-Tiere gelang es den zu Grunde liegenden Haplotypen für zwei vorzeitige Stoppmutationen mit fehlenden Homozygoten zu identifizieren (*GPR128* und *LRP2BP*). Der für die in *GPR128* lokalisierte Stoppmutation zugrunde liegende potentielle Haplotyp besteht aus 80 SNP mit einer Länge von etwa 362 kb (Chromosom1: 45.110.815 - 45.472.868 Mb). Bei einer Haplotypenfrequenz von etwa 1% wurden 174 heterozygote und kein homozygoter Haplotypenträger beobachtet (P-Wert = 0,389). Der Haplotyp, der das vorzeitige Stopp-Codon des Gens *LRP2BP* enthält, weist eine Länge von 13,8 Mb mit 3384 SNP auf (Chromosom 27: 10.331.457 - 24.087.408 bp). 69 BFV-Tiere tragen diesen Haplotypen heterozygot. Homozygote Tiere wurden bei einer Frequenz des Haplotypen von 0,66% nicht entdeckt. Die Anzahl erwarteter homozygoter Genotypen weicht demnach nicht signifikant von der Anzahl der beobachteten homozygoten Genotypen ab (P-Wert = 0,72). Basierend auf der Information der zwei identifizierten Haplotypen, die jeweils eine geringe Allelfrequenz aufwiesen ($AF \leq 1\%$), wurde keine signifikante Abweichung vom HWG festgestellt.

4.2.4.2 Vorzeitige Stopp-Codons im Deutschen Fleckvieh mit Abweichung vom Hardy-Weinberg-Gleichgewicht

Bei der Analyse der übrigen 53 Stoppmutationen mit homozygotem Genotyp lag die Allelfrequenz zwischen 0,63% und 47,6% im imputierten Datensatz. Die Irrtumswahrscheinlichkeit, die angibt, dass die Genotypverteilung der Stoppmutationen signifikant vom HWG abweicht, erreichte für die Mutation im Gen *PNLIP* das Minimum (0,00246). Eine weniger signifikante

Abweichung vom HWG (P-Wert < 0,05) zeigten fünf weitere Stoppmutationen (*GPR75*, *INMT*, *LYSB*, *QRFPR*, *VTCN1*). Die relative Häufigkeit des mutierten Allels der Varianten lag zwischen 4,54% (*PNLIP*) und 35,24% (*LSYB* und *INMT*).

PNLIP gehört wie *QRFPR* und *VTCN1* zu den Genen, die bereits in Knockout-Studien bei Mäusen genauer untersucht wurden bzw. beim Menschen mit einem Phänotyp assoziiert sind (Tabelle 4.5). Diese drei Gene werden im Folgenden intensiver analysiert.

4.2.4.2.1 Stoppmutation in *PNLIP*

Das Gen *PNLIP* kodiert für die Pankreaslipase (PNLIP), die im Dünndarm für die Hydrolyse von langkettigen Fettsäuren zuständig ist und die Absorption von Fetten ermöglicht (LOWE *et al.* 1989; DAVIS *et al.* 1991; LOWE 2002). Die Fettsäuren werden anschließend direkt für den Energiehaushalt und Lipidspeicher verwendet. Das ausschließlich im Pankreas exprimierte Enzym ist somit für die Fettverdauung unentbehrlich. Eine weitere Funktion dieser Lipase ist die Hydrolyse von Retinylestern zu Fettsäuren und Retinol, das im Stoffwechsel zu Vitamin A synthetisiert wird (VAN BENNEKUM *et al.* 2000).

Beim Menschen wurden Mutationen im Gen *PNLIP* mit einer angeborenen Pankreasdefizienz, die auch als exokrine Pankreasinsuffizienz bekannt ist, beschrieben (OMIM 614338) (FIGARELLA *et al.* 1980). Betroffene Individuen haben einen ölartigen, hellen und schmierigen Stuhl (Fettstuhl) und verlieren trotz erhöhter Nahrungsaufnahme an Körpergewicht (SHELDON 1964; REY *et al.* 1966). In Mäusen, bei denen die Aktivität der Pankreaslipase ausgeschaltet wurde, konnte eine verzögerte Aufnahme von Nahrungsfetten festgestellt werden (HUGGINS *et al.* 2003). Ihre Gewichtszunahme war im Vergleich zu Mäusen mit aktiver Pankreaslipase reduziert (GILHAM *et al.* 2007).

Die Pankreaslipase kann strukturell in zwei Regionen eingeteilt werden. Im N-terminalen Bereich besitzt PNLIP die typische Lipase-Domäne mit katalytischer Serin-Histidin-Asparagin-Triade. Beim Menschen wird diese Domäne durch eine Schleife bedeckt (Deckel-Domäne), die eine Substratbindung an die Lipase verhindert (WINKLER *et al.* 1990). Die Anwesenheit von Mizellen bewirkt eine Bewegung dieses Deckels, sodass die katalytische Domäne für die Substrate geöffnet wird (VAN TILBEURGH *et al.* 1993). Die C-terminale Region enthält eine Polycystin-1-Lipoxygenase-Alpha-Toxin-Domäne (PLAT), die für die Bindung von PNLIP an den Kofaktor Colipase ausschlaggebend ist (BATEMAN und SANDFORD 1999; CHAHINIAN *et al.* 2000; LOWE 2002). Erst durch die

Erstellung des Pankreas-Lipase-Colipase-Komplexes kann die Lipase ihre Funktion im Rahmen der Hydrolyse von Fettsäuren ausüben.

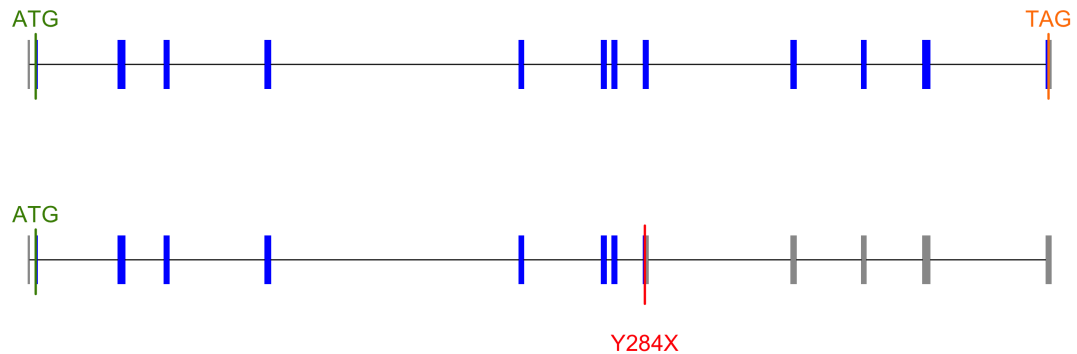


Abbildung 4.2.14: Genstruktur von PNLIP

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.Y284X) ist rot gekennzeichnet.

PNLIP befindet sich im Rindergenom auf dem 26. Autosom mit einer Länge von ~ 20,1 kb (37.290.364 - 37.310.481 bp). Das Gen ist auf dem positiven Strang lokalisiert. Es besteht mit einem vollständig untranslatiertem Exon im 5'-Bereich aus insgesamt 13 Exons, die für 466 Aminosäuren kodieren.

In der BFV-Population konnte mit Hilfe der genomweiten Re-Sequenzierung ein vorzeitiges Stoppcodon im neunten Exon von *PNLIP* identifiziert werden (c.852C>G; p.Y284X) (Abbildung 4.2.14). Die Frequenz des G-Allels betrug sowohl innerhalb der 112 NGS-Tiere als auch im imputierten Datensatz etwa 5%. Bei der gegebenen Allelfrequenz sollten innerhalb der 10.363 BFV-Tiere 21 mit homozygotem Genotyp für die Mutation erscheinen, jedoch wurden nur acht reinerbige Tiere entdeckt (P-Wert = 0,00246). Die LoF-Variante wurde in einem Haplotypen von 9,12 Mb Länge lokalisiert (Chromosom 26: 33.228.016 - 42.356.782 bp) mit einer Frequenz von 3,08%. Ähnlich wie bei den genombasierten Genotypen wurden bei der Betrachtung des Haplotypen weniger homozygote Tiere beobachtet (n=2) als erwartet (n=10) (P-Wert = 0,0102).

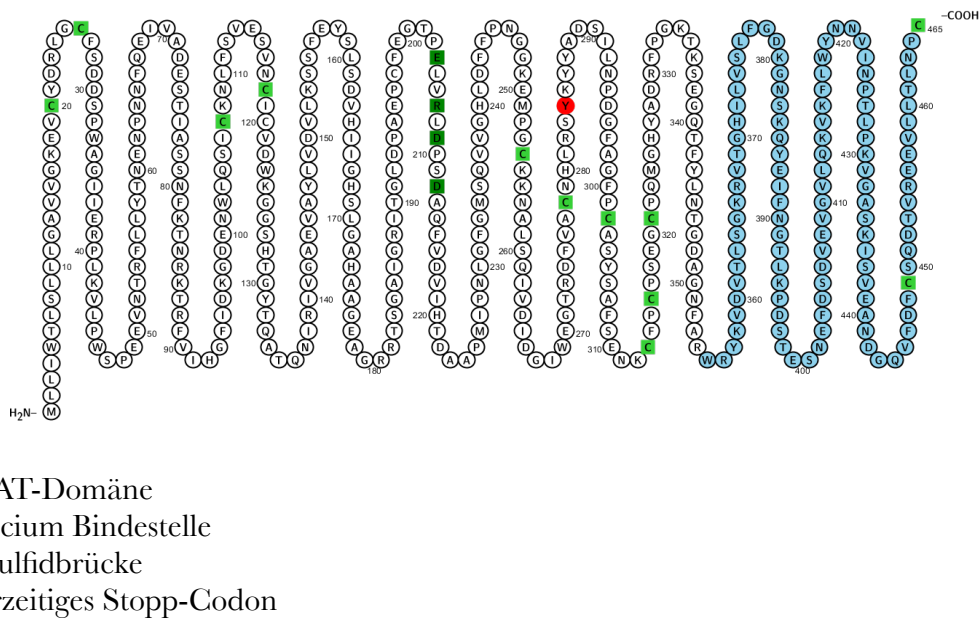


Abbildung 4.2.15: Peptidstruktur von PNLIP

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS *et al.* 2013) die Positionen der Transmembrandomänen visualisiert. Das vorzeitige Stopp-Codon (p.Y284X) ist rot gekennzeichnet.

Die Variante befindet sich in einer stark konservierten Region, wobei das Protein insgesamt über einen hohen Konservierungsgrad verfügt. Beim Vergleich der Isoformen ist deutlich zu erkennen, dass jede der untersuchten Proteinsequenzen von der Stoppmutation beeinträchtigt wird. Das vorzeitige Stopp-Codon an der 284. Aminosäure wird möglicherweise den NMD einleiten. Falls der Abbau der mRNA nicht eintritt, kann die Stoppmutation durch die Verkürzung der Proteinsequenz um 38,93% einen Funktionsverlust des Proteins entstehen lassen. Die funktionelle PLAT-Domäne wird nicht translatiert (Abbildung 4.2.15).

4.2.4.2.2 Stoppmutation in *QRFPR*

Der Pyroglutamylierte-RFamid-Peptid-Rezeptor (orexigenischer Neuropeptid-QRFP-Rezeptor, GPR103) wird vom Gen *QRFPR* kodiert. Er gehört zur Familie der G-Protein gekoppelten Rezeptoren (GPR). GPR nehmen eine zentrale Rolle bei der Signaltransduktion innerhalb und zwischen einzelnen Zellen ein (PIERCE *et al.* 2002; KROEZE *et al.* 2003). Sie stellen sowohl für Signale als auch für kleine Moleküle einen Rezeptor dar, sodass sie in nahezu allen physiologischen Prozessen im Körper involviert sind (BARIBAULT *et al.* 2006). Ein Ligand des GPR103 ist

das 26 Aminosäuren langes RF-Amid, das die Nahrungsaufnahme stimuliert (JIANG *et al.* 2003). QRFP, ein 43 Aminosäuren langes RF-Amid, das an GPR103 bindet, reguliert bei Ratten die Aldosteronkonzentration im Plasma (FUKUSUMI *et al.* 2003). Der Metabolismus wird dabei gesteigert (TAKAYASU *et al.* 2006). Beim Menschen wird GPR103 vorwiegend im Gehirn, insbesondere in der Großhirnrinde, der Hypophyse, dem Thalamus, dem Hypothalamus, im basalen Vorderhirn, im Mittelhirn und im Pons exprimiert (LEE *et al.* 2001). Aber auch im Herzen, in der Niere, in der Retina, im Hoden und in verschiedenen Zelllinien von Osteoblasten wurde GPR103-mRNA entdeckt (JIANG *et al.* 2003; BARIBAULT *et al.* 2006).

BARIBAULT *et al.* (2006) erzeugten GPR103 defiziente Mäuse, die eine geringe Knochendichte (Osteopenie) aufwiesen. Bei den weiblichen Tieren wurde zusätzlich eine abnormale Wirbelsäulenkrümmung (Kyphose) registriert.

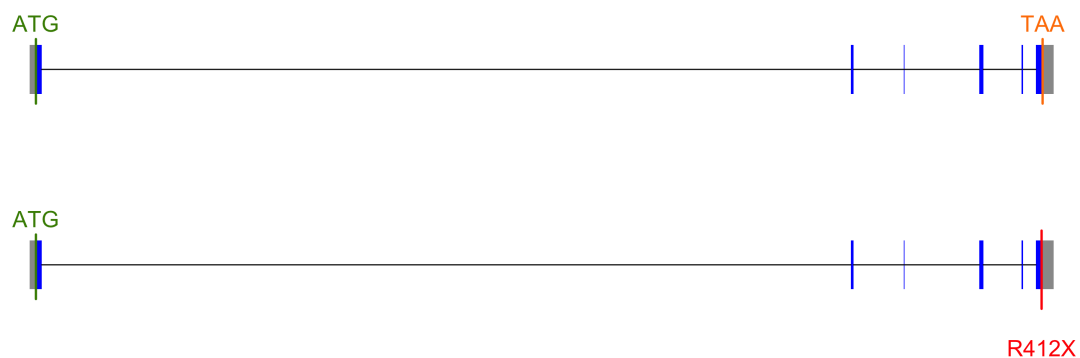


Abbildung 4.2.16: Genstruktur von QRFP

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (p.R412X) ist rot gekennzeichnet.

Das bovine *QRFP* ist auf dem positiven Strang des sechsten Chromosoms positioniert (40.93.248 - 41.53.398 bp). Das Protein besteht aus 432 Aminosäuren. Durch einen C>T-Polymorphismus im letzten der insgesamt sechs Exons (c.1234C>T; p.R412X) entsteht ein vorzeitiges Stopp-Codon (Abbildung 4.2.16). Die Allelfrequenz dieser Mutation innerhalb der 112 BFV-Genomsequenzen entsprach der des imputierten Datensatzes der 10.363 Individuen (AF = 29,34%). Sowohl innerhalb der 112 BFV-Tiere, als auch im 1000-Bullen-Genom-Projekt und im imputierten Datensatz lagen homozygote Genotypen für diese Variante vor. 847 Tiere wurden im ALT-homozygoten Zustand beobachtet, 892 wurden nach dem HWG erwartet (P-Wert = 0,0338).

Der Haplotyp, auf dem diese Mutation lokalisiert ist, konnte auf Basis der 614K-Informationen nicht identifiziert werden.

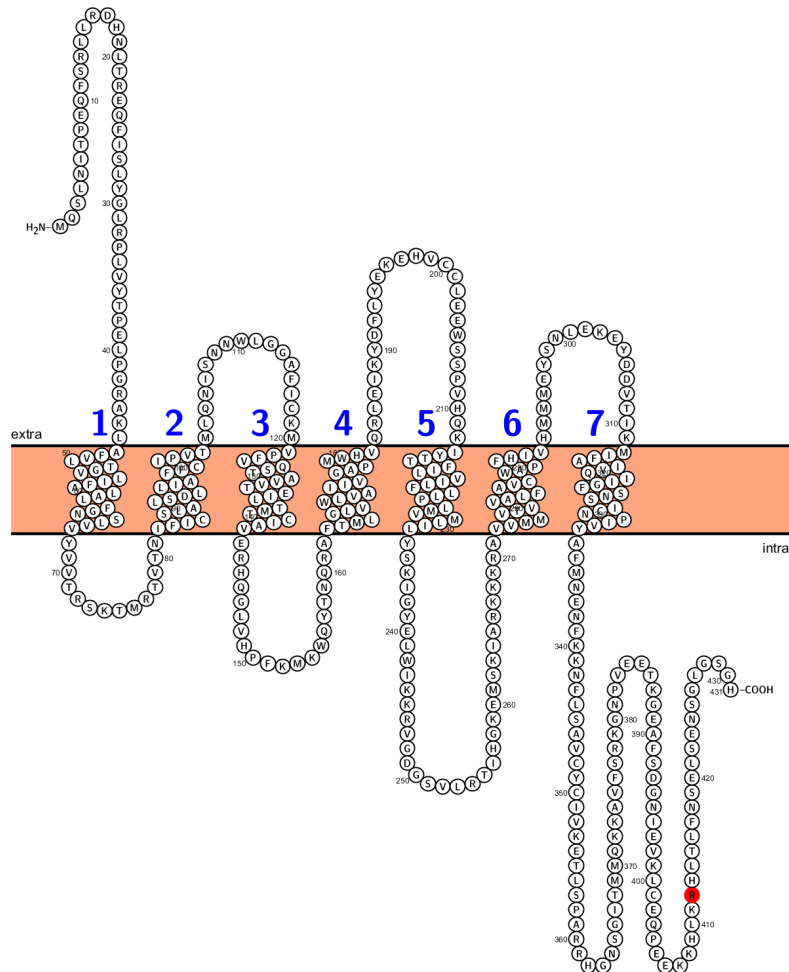


Abbildung 4.2.17: Transmembranstruktur von QRFPR

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS et al. 2013) die Positionen der Transmembrandomänen visualisiert. Das vorzeitige Stopp-Codon (p.R412X) ist rot gekennzeichnet.

Charakteristisch für G-Protein-gekoppelte Rezeptoren sind ihre sieben Transmembrandomänen. Das N-terminale Ende ist extrazellulär lokalisiert, während die C-terminale Region in der Zelle endet (Abbildung 4.2.17). Durch die Veränderung der Aminosäure Arginin, die alle Isoformen von GPR103 betrifft, wird das entstehende Protein um 4,63% verkürzt. Ein Teil des hoch konservierten intrazellulären Bereichs geht somit verloren. Es ist zu vermuten, dass der Rezeptor auch durch diesen nur geringen Verlust der Aminosäuresequenz in seiner Funktion gestört ist.

4.2.4.2.3 Stoppmutation in *VTCN1*

Mitglieder der B7 kostimulatorischen Proteinfamilie haften an der Oberfläche von Antigen-präsentierenden Zellen (KRYCZEK *et al.* 2006). Dort interagieren sie mit den Liganden, die an die spezifischen Rezeptoren der Zellen mit Antigenpräsentation binden. Das Transmembranprotein VTCN1 (V-set domain-containing T-cell activation inhibitor 1), ein Protein der B7-Familie, schwächt die durch T-Zellen bedingte erworbene Immunantwort (SICA *et al.* 2003; CARRENO und COLLINS 2003) sowie den angeborenen Abwehrmechanismus (SICA *et al.* 2003; ZHU *et al.* 2009). Außerdem hat VTCN1 ohne Beeinflussung des Immunsystems einen direkten Einfluss auf die Tumorentwicklung (QIAN *et al.* 2013). Polymorphismen im Gen *VTCN1* sind beim Menschen mit rheumatoider Arthritis assoziiert (DAHA *et al.* 2012; ALBERS *et al.* 2014). Bei VTCN1-defizienten Mäusen wurden keine gesundheitlichen Nachteile festgestellt, im Gegenteil sind diese Mäuse resistenter gegenüber Infektionen (ZHU *et al.* 2009).

In normalen Geweben und Zellen wurde *VTCN1*-mRNA kaum entdeckt (SICA *et al.* 2003; ZANG *et al.* 2003). Die Expression von *VTCN1* steigt mit fortschreitendem Krankheitsverlauf (FLIES und CHEN 2007) und das Gen wird vorwiegend entweder in der Membran oder im Zytoplasma von T-Zellen, B-Zellen, Monozyten und dendritischen Zellen exprimiert (SICA *et al.* 2003). Insbesondere bei der Entwicklung verschiedenster Tumorarten steigt die Expression von *VTCN1* (SALCEDA *et al.* 2005; QIAN *et al.* 2011).

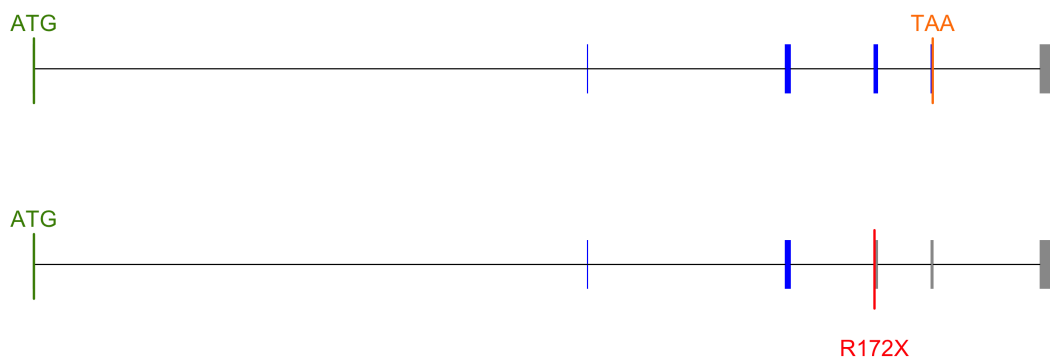


Abbildung 4.2.18: Genstruktur von *VTCN1*

Kodierende Sequenzbereiche sind blau, untranslatierte Bereiche grau markiert. Die vorzeitige Stoppmutation (*p.R172X*) ist rot gekennzeichnet.

Im Rindengenom konnte das Gen *VTCN1* auf Chromosom 3 lokalisiert werden (26.126.115 - 26.186.582 bp). Mit einer Transkriptlänge von ~60 kb und sechs Exons kodiert es ein Protein

bestehend aus 282 Aminosäuren (Abbildung 4.2.18). Der beim Rind detektierte C>T-Polymorphismus im vierten der insgesamt fünf kodierenden Exons verändert die Aminosäure Arginin zu einem vorzeitigem Stopp-Codon (c.700C>T; p.R172X). Beim BFV betrug die Allelfrequenz dieser Mutation sowohl innerhalb der 112 genomweit re-sequenzierten als auch in den 10.363 Tieren ~5%. Homozygote Tiere konnten nur im imputierten Datensatz identifiziert und mit Hilfe der *Sanger-Re-Sequenzierung* bestätigt werden, jedoch entsprach die Anzahl der beobachteten (n = 16) ALT-homozygoten Tiere nicht der erwarteten Anzahl (n = 26; P-Wert = 0,0367). Die Mutation segregiert auch in den Rassen SIM und PIE.

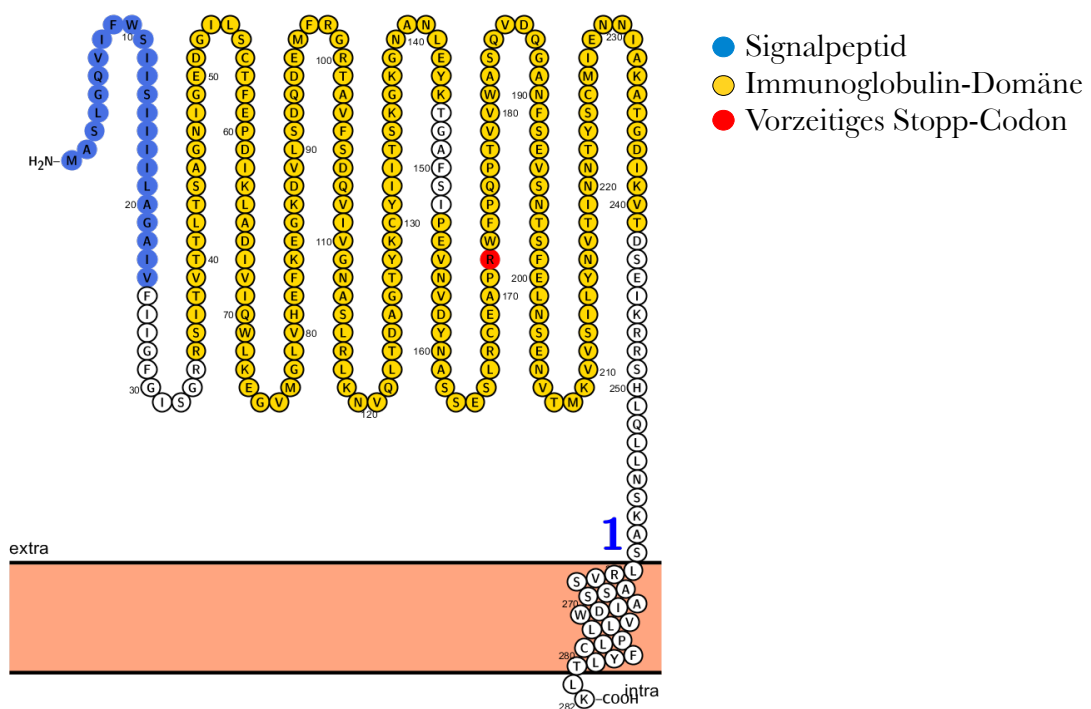


Abbildung 4.2.19: Transmembranstruktur von VTCN1

Basierend auf den Informationen der UniProtKB-Datenbank wurden mit Protter (OMASITS *et al.* 2013) die Positionen der Transmembran- und die Immunoglobulindomänen visualisiert. Das vorzeitige Stopp-Codon (p.R172X) ist rot gekennzeichnet.

VTCN1 besitzt neben seiner hydrophoben Transmembrandomäne ein Signalpeptid und eine sehr kurze interzelluläre Domäne (SICA *et al.* 2003). In der extrazellulären, C-terminalen Region sind die Ig-Domänen lokalisiert (CHEN 2004). Das vorzeitige Stopp-Codon würde die Peptidsequenz von VTCN1 um 110 Aminosäuren (39%) verkürzen. Dabei würden die hoch konservierten Regionen mit der zweiten Ig-Domäne (AS 153-241) und der Transmembrandomäne (AS 260-280) verloren gehen (Abbildung 4.2.19). Alle untersuchten Isoformen des Proteins sind

von dem Stopp-Codon betroffen. Nach den Regeln des NMD wird die mRNA aufgrund der vorzeitigen Stoppmutation abgebaut. Falls dieser Abbau nicht initiiert wird, ist davon auszugehen, dass diese Variante ein funktionsunfähiges Protein entstehen lässt.

4.3 Sequenzbasierte genomweite Assoziationsstudie für das Merkmal direkte Totgeburten bei Holstein-Kühen

Nach der Anwendung von Qualitätsparametern auf den imputierten Datensatz lagen für 9013 Tiere der deutschen Holstein-Population (HOL und RED) Genotypen an 17.290.679 variablen Positionen für die Anwendung in der GWAS vor. Die genomweite Signifikanzgrenze lag nach der Bonferroni-Korrektur bei $P = 2,891 \times 10^{-9}$.

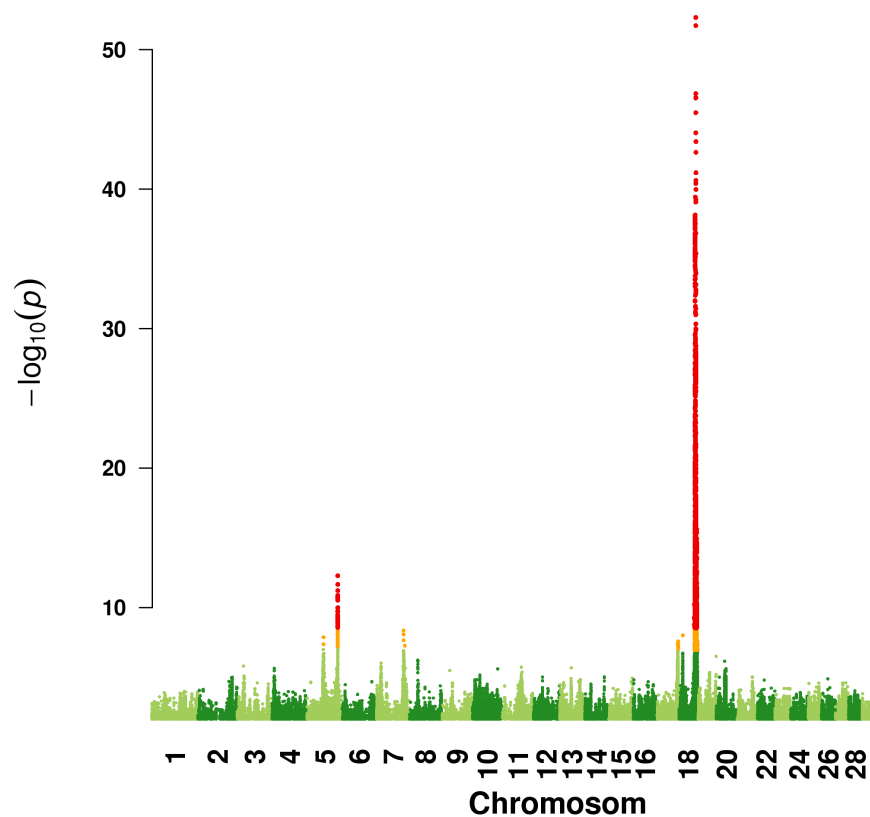


Abbildung 4.3.1: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen

Die Positionen der einzelnen Varianten, auf der x-Achse dargestellt, basieren auf dem UMD3.1-Assembly (ZIMM et al. 2009), mit Ausnahme der Geschlechtschromosomen. Auf der y-Achse ist der negativ-dekadische Logarithmus der P-Werte aus der GWAS aufgetragen. Chromosomweit signifikante Polymorphismen sind gelb markiert, genomweit signifikante Varianten sind in rot dargestellt ($P = 2,891 \times 10^{-9}$).

Die sequenzbasierte GWAS für den Relativzuchtwert der direkten Totgeburten bei Kühen (RZW dSBc) in der Holstein-Population brachte zwei genomweit signifikante QTL hervor (Chromosom

5 und 18) (Abbildung 4.3.1, Tabelle 4.7). Außerdem konnten vier weitere QTL identifiziert werden, die das chromosomale Signifikanzniveau erreichten (Chromosom 5, 7, 17, 18 und 29). Der Anteil der additiv genetischen Varianz an der phänotypischen Zuchtwert-Varianz des Merkmals direkte Totgeburten bei Kühen, der durch die in der GWAS angewandten Polymorphismen erklärt wird, beträgt 74,03%.

Tabelle 4.7: Auflistung der signifikantesten Polymorphismen für das Merkmal direkte Totgeburten bei Holsteinkühen

Chr	Position (bp)	MAF	P-Wert	α	Anteil QTL-Varianz	Kandidaten-Gen
5	57.377.229	0,0210	$1,3369 \times 10^{-8}$	8,2043	0,0769	
5	106.202.874	0,3514	$5,1841 \times 10^{-13}$	-1,1548	0,0168	<i>FGF23</i>
7	93.244.933	0,0952	$4,327 \times 10^{-9}$	1,3045	0,0081	<i>ARRDC3</i>
17	73.258.826	0,1307	$2,6154 \times 10^{-8}$	1,1959	0,0095	
18	15.038.275	0,3899	$9,7216 \times 10^{-9}$	-0,9224	0,0112	<i>VPS35</i>
18	59.586.457	0,1206	$5,0011 \times 10^{-53}$	4,0957	0,0984	<i>LOC787812</i>
29	50.126.590	0,0556	$9,8292 \times 10^{-8}$	16,873	0,8269	<i>H19</i>

Die Polymorphismen sind nach der Position im Genom basierend auf dem UMD3.1-Assembly (ZIMIN et al. 2009) geordnet. Die P-Werte und die Allelsubstitutionseffekte α resultieren aus der GWAS mit EMMAX (KANG et al. 2010). Der Anteil der QTL-Varianz ($2pq\alpha^2$) an der additiv-genetischen Varianz σ_a^2 wird mit $2pq\alpha^2/\sigma_a^2$ berechnet.

4.3.1.1 QTL auf Chromosom 5 (56,8 bis 57,9 Mb)

Einer der zwei identifizierten QTL auf Chromosom 5 ist im Bereich von 56 bis 58 Mb lokalisiert. Der Top-Marker mit einem P-Wert von $1,3369 \times 10^{-8}$ befindet sich an der Basenposition 57.377.299. In dieser QTL-Region zeigen zahlreiche Polymorphismen ein signifikantes Signal (Abbildung 4.3.2). Bei einer niedrigen Allelfrequenz weisen diese hoch signifikanten Varianten zudem einen hohen Allelsubstitutionseffekt auf. Das Kopplungsungleichgewicht dieser Mutationen zum Top-Marker (Chr5: 57.377.229) ist unterschiedlich, wobei nur wenige signifikante Varianten im hohen Kopplungsungleichgewicht (LD) mit dem Top-Marker stehen (Tabelle 4.8). In diesem Bereich sind viele Gene lokalisiert. Aufgrund des uneindeutigen Signals und der genreichen Region kann kein positionelles Kandidatengen lokalisiert werden.

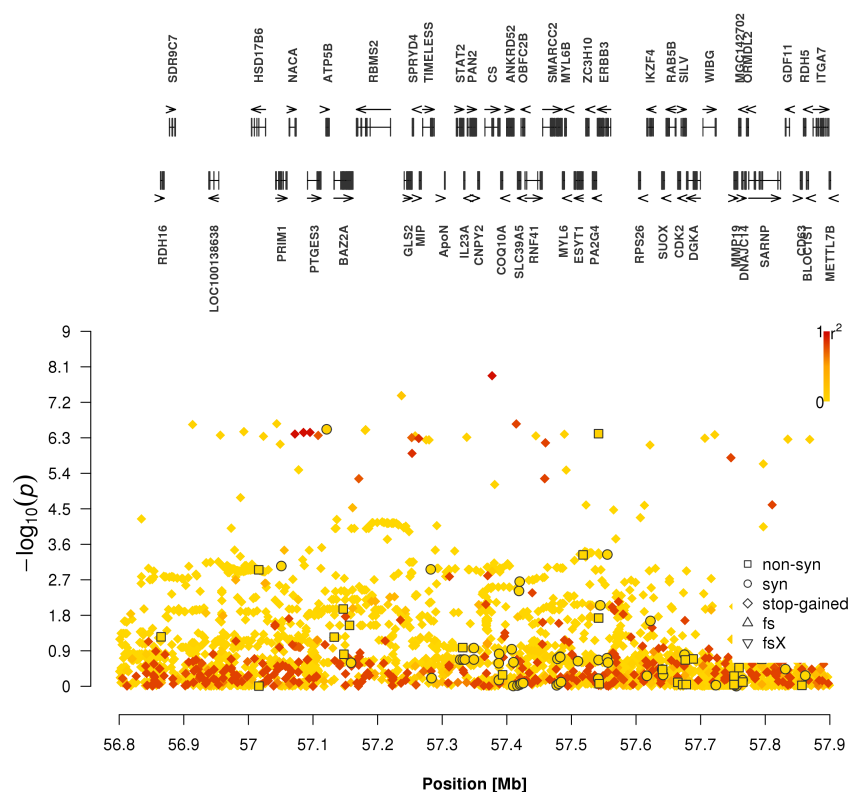


Abbildung 4.3.2: Region-Plot auf Chromosom 5 für direkte Totgeburten bei Holsteinkühen (56,8-57,9Mb)

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMIN et al. 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligen LD zum Top-Marker (Chr5: 57.377.229). Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Tabelle 4.8: Signifikante Polymorphismen auf Chromosom 5 (56,8 bis 57,9 Mb)

Position (bp)	REF	ALT	AF	P-Wert	α	LD als r^2
57.377.229*	C	A	0,021	$1,34 \times 10^{-8}$	8,20	0,95
57.236.936	A	T	0,040	$4,26 \times 10^{-8}$	4,84	0,16
57.043.783	G	A	0,037	$2,21 \times 10^{-7}$	4,40	0,05
57.414.671	T	C	0,005	$2,23 \times 10^{-7}$	9,91	0,8
56.913.641	C	G	0,040	$2,31 \times 10^{-7}$	3,22	0,05
57.121.113	C	T	0,020	$3,05 \times 10^{-7}$	2,79	0,06
57.181.577	G	A	0,021	$3,13 \times 10^{-7}$	2,75	0,05
57.180.690	A	T	0,020	$3,25 \times 10^{-7}$	2,74	0,05
56.992.921	T	G	0,015	$3,50 \times 10^{-7}$	4,23	0,05

Die Polymorphismen sind nach ihrem P-Wert aus der GWAS mit EMMAX geordnet. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; α = Allelsubstitutionseffekt; LD als r^2 gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr5: 57.377.229) an; * = Top-Marker.

4.3.1.2 QTL auf Chromosom 5 (105,7 bis 106,8 Mb)

Der signifikanteste Marker auf Chromosom 5 (Chr5: 106.202.874 bp) erreichte einen P-Wert von $5,1841 \times 10^{-13}$ (Abbildung 4.3.3). Es handelt sich dabei um eine intergenische Variante in der Nähe von *FGF23* und ist etwa 5 kb in 5'-Richtung vom Transkriptionsstart des Gens entfernt.

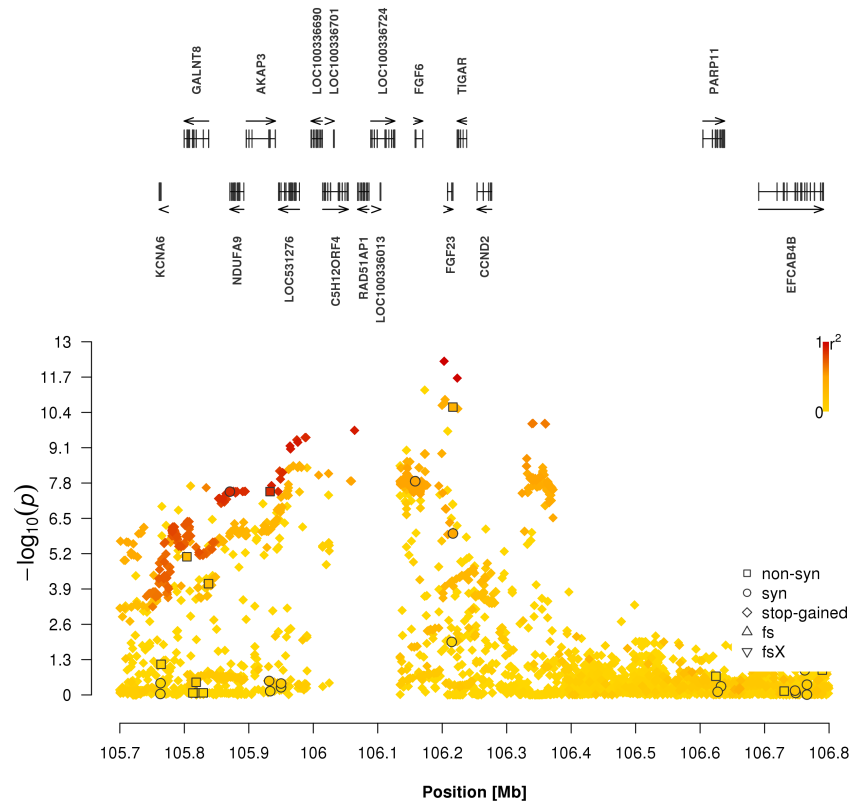


Abbildung 4.3.3: Region-Plot auf Chromosom 5 für direkte Totgeburten bei Holsteinkühen (105,7-106,8Mb)

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMLN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligen LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Charakterisierung von *FGF23*

Der Fibroblasten-Wachstumsfaktor 23 (fibroblast growth factor, FGF), der vom Gen *FGF23* kodiert wird, gehört zu einer Gruppe von Wachstumsfaktoren, die als FGF-Familie bezeichnet werden (ITO und ORNITZ 2004; JAKOB *et al.* 2006). FGF spielen eine wichtige Rolle beim Wachstum und bei der Differenzierung von Zellen (THOMAS 1987), insbesondere während der embryo-

nen Entwicklung (COHN *et al.* 1995). Im adulten Organismus sind diese Wachstumsfaktoren für die Wundheilung und Neubildung von Gefäßen verantwortlich (ORTEGA *et al.* 1998). An Zelloberflächen gehen die FGF-Proteine mit einem ihrer spezifischen Rezeptoren (FGFR) eine Bindung ein. Dabei dient das Heparansulfat-Proteoglykan (HSP) als Co-Faktor. Zusammen bilden FGF, FGFR und HSP den Dreifach-Komplex (*ternary complex*) und lösen im Innern der Zelle eine Signalkaskade aus (ORNITZ 2000).

FGF23 sorgt als wichtiger Regulator des Phosphat- und Vitamin-D-Haushalts für die Aufrechterhaltung eines konstanten Spiegels dieser beiden Komponenten im Blut (SHIMADA *et al.* 2004a). Der Wachstumsfaktor wird in Osteoblasten exprimiert (LIU *et al.* 2003, 2006) und anschließend zur Niere als Hauptzielort und zur Nebenschilddrüse transportiert. Dort bindet FGF23 an seinen spezifischen Rezeptorkomplex, der von Klotho und dem Rezeptor FGFR1 aktiviert wird (URAKAWA *et al.* 2006; KUROSU *et al.* 2006). Im proximalen Tubulus der Niere hemmt FGF23 die Expression von NPT2 (Natrium/Phosphat Transporter II) (BAUM *et al.* 2005), einem Kotransporter, der bei niedrigem Phosphatgehalt im Blut Phosphationen aus dem Urin in die Niere reabsorbiert und von dort aus ins Blut ausscheidet (BECK *et al.* 1998; MIYAMOTO *et al.* 2004). FGF23 dient als Gegenspieler von NPT2 und sorgt für eine reduzierte Wiederaufnahme von Phosphat in die Niere, sodass mehr Phosphat mit dem Urin ausgeschieden und der Phosphatspiegel im Blut gesenkt wird (BAUM *et al.* 2005). FGF23 spielt zusätzlich eine Rolle im *Plexus choroideus* im Gehirn und in der Hypophyse. Seine dortige Funktion ist jedoch noch unbekannt (LIU und QUARLES 2007).

Beim Menschen sind Mutationen im Gen *FGF23* mit zwei Erbdefekten assoziiert. Drei missense Mutationen sind als kausal für die autosomal dominante hypophosphatämische Rachitis (ADHR) (OMIM 193100) beschrieben (ADHR CONSORTIUM 2000). Durch die Überexpression von *FGF23* sinkt der Phosphatspiegel im Blut und die Produktion von 1,25-Dihydroxy-Vitamin D (1,25(OH)₂D) wird gehemmt. Betroffene Individuen leiden neben Rachitis auch an einer Erweichung der Knochen. Ein weiterer nicht-synonymer Aminosäureaustausch in *FGF23* ist verantwortlich für die hyperphosphatämische tumorale Calcinose (OMIM 211900) (BENET-PAGÈS *et al.* 2005), die sich in einem gegensätzlichen Phänotyp zur hypophosphatämischen Rachitis manifestiert. Der geringere FGF23-Gehalt im Blut lässt den Phosphatspiegel ansteigen und fördert die Produktion von 1,25(OH)₂D. Es kommt zur Verknöcherung von Weichteilen. Außerdem wird ein niedriger FGF23-Spiegel mit Hyperparathyreoidismus assoziiert, der durch eine vermehrte Sekretion des in der Nebenschilddrüse produzierten Parathormons charakterisiert ist (URAKAWA *et*

al. 2006). Es ist noch ungewiss, ob diese gesamten Veränderungen auf den Phosphat- oder Vitamin-D-Gehalt im Körper zurückzuführen sind oder auch direkt vom FGF23 ausgelöst werden können. Der Funktionsverlust des FGF23 aktivierenden Proteins Klotho führt bei Mäusen zu einem ähnlichen Phänotyp (KURO-O *et al.* 1997).

Das bovine *FGF23* ist auf dem positiven Strang des fünften Chromosoms des Rindergenoms lokalisiert und besteht aus drei Exons (Transkriptionsstart: 106.208.179; Transkriptionsende: 106.216.757). Das 245 Aminosäuren lange Protein setzt sich aus einem Signalpeptid (AS 1-19), einem N-terminalen Peptid (AS 20-175) und einem C-terminalen Peptid (AS 175-245) zusammen. Das N-terminale Fragment enthält die FGF-homologe Domäne, die für die Bindung mit den FGFR verantwortlich ist, während die C-terminale Domäne für die Adhäsion von Klotho an FGF23 zuständig ist (YAMAZAKI *et al.* 2008).

Tabelle 4.9: Polymorphismen im exonischen Bereich von FGF23

Position (bp)	Region	REF	ALT	AF	AA	SIFT	Polyphen	P-Wert	LD als r²
106.214.730	CDS	TCC	TCT	0,387	S100S			0,0112	0,003
106.216.557	CDS	AGC	GGC	0,598	S180G	tolerated	benign	2,5383x10 ⁻¹¹	0,372
106.216.613	CDS	GCC	GCT	0,206	A198A			1,1395x10 ⁻⁶	0,479

*Die Polymorphismen sind nach ihrer Lokalisation im Gen FGF23 basierend auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009) sortiert. Der Effekt der Mutationen, die einen Aminosäureaustausch herbeiführen, wurde mit SIFT (NG und HENKOFF 2003) und Polyphen (ADZHUBEI *et al.* 2010) vorhergesagt. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; AA = betroffene Aminosäuren; LD(r²) gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr5: 106.202.874) an.*

Der signifikanteste Polymorphismus für das Merkmal direkte Totgeburten bei Kühen ist 5.305 bp vom funktionellen Kandidatengen *FGF23* entfernt und liegt außerhalb des Promotorbereichs in einer intergenischen Region. Zahlreiche identifizierte Varianten (n=36) weisen zu diesem Top-Marker ein starkes LD ($r^2 > 0,8$) auf (Abbildung 4.3.3). Bei der Analyse des Kandidatengens wurden im flankierenden, regulatorischen und kodierenden Bereich insgesamt 26 Polymorphismen detektiert (vgl. Anhang Tabelle 9.8), von denen drei in der kodierenden Sequenz des Gens lokalisiert sind (Tabelle 4.9). Der signifikanteste kodierende SNP führt zu einem Aminosäureaustausch in einer hoch-konservierten Region der C-terminalen Domäne (p.S180G). *SIFT* und *Polyphen* haben diesen Austausch als „tolerated“ und „benign“ vorhergesagt. Mit dem Top-Marker liegt er im mittleren LD (0,372).

4.3.1.3 QTL auf Chromosom 7

Auf Chromosom 7 des bovinen Genoms wurde eine Region identifiziert, die mit dem Merkmal direkte Totgeburten bei Kühen assoziiert ist ($P\text{-Wert} = 4,327 \times 10^{-9}$). Der signifikanteste Marker ist im kodierenden Bereich des Gens *ARRDC3* lokalisiert (Abbildung 4.3.4).

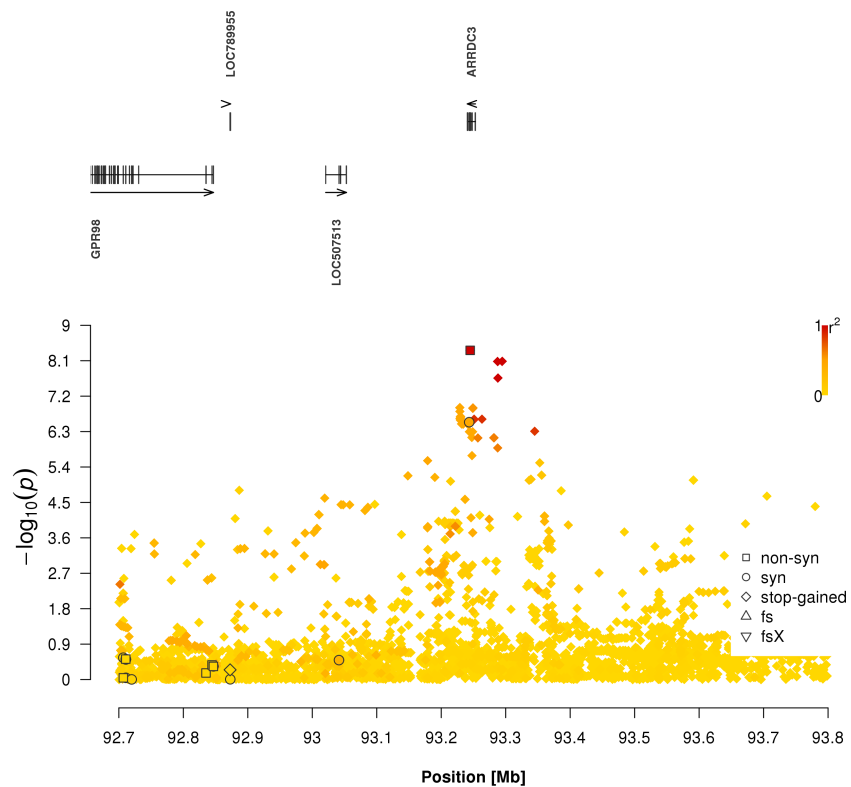


Abbildung 4.3.4: Region-Plot auf Chromosom 7 für direkte Totgeburten bei Holsteinkühen

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligen LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Charakterisierung *ARRDC3*

Das Gen *ARRDC3* kodiert für ein Arrestin domain-containing Protein 3 (*ARRDC3*), das die zwei N- und C-terminalen Arrestin-Domänen besitzt. Phylogenetisch gesehen und auch aufgrund der niedrigen Sequenzhomologie ist *ARRDC3* jedoch entfernt verwandt zu Arrestinen (AUBRY *et al.* 2009). Das Protein ist auch unter dem Namen Thioredoxin-bindendes Protein 2 ähnliches

induzierbares Membranprotein (Thioredoxin-binding protein-2-like inducible membrane protein, TLIMP) bekannt (OKA *et al.* 2006).

Beim Menschen wird *ARRDC3* ähnlich wie bei der Maus im subkutanen Gewebe, sowie in der Plazenta, Skelettmuskulatur, Milz, Lunge und Leber exprimiert (OKA *et al.* 2006; PATWARI *et al.* 2011). Das resultierende Protein *ARRDC3* ist in Wachstum und Vermehrung von Zellen involviert. *In vitro* und *in vivo* Studien zeigten, dass *ARRDC3* im Anfangsstadium der Brust-Karzinogenese runter reguliert wird und während der Tumorentwicklung sinkt. Dahingegen wirkt eine Überexpression des Gens hemmend auf das Zellwachstum, wobei die Expression von Integrin-beta-4 (*ITGB4*) beeinflusst wird (DRAHEIM *et al.* 2010). PATWARI *et al.* (2011) zeigten, dass das Gen *ARRDC3* mit Fettleibigkeit bei Männern assoziiert ist. Anschließende Studien mit Knockout-Mäusen ließen eine höhere perinatale Mortalität sowie ein geringeres Körpergewicht bei männlichen Tieren erkennen. In einer anderen Knockout-Studie wirkte sich das fehlende *ARRDC3*-Gen embryonal letal aus, sobald die Mutter kein lipidreiches Futter aufgenommen hat (SHEA *et al.* 2012). In genomweiten Studien beim Rind wurde *ARRDC3* mit dem Merkmal Geburtsgewicht assoziiert (SAATCHI 2013).

Im bovinen Genom konnte das Gen *ARRDC3* auf dem siebten Chromosom lokalisiert werden. Es liegt auf dem negativen Strang und ist aus acht Exons aufgebaut (Transkriptionsstart: 93.253.094; Transkriptionsende: 93.240.415). Das entstehende Protein setzt sich aus 414 Aminosäuren zusammen. Über die Spezies hinweg ist der Konservierungsgrad des *ARRDC3*-Proteins sehr hoch. An den Aminosäurepositionen 9 bis 165 des Proteins befindet sich die Arrestin-N-Domäne. Die Arrestin-C-Domäne ist an den Aminosäuren 187 bis 314 positioniert. Der C-terminale Schwanz hinter der C-Domäne enthält zwei *PPXY* Motive (*PPXY1*: AS 341-355; *PPXY2*: AS 384-399) (QI *et al.* 2014).

Basierend auf den Polymorphismen der GWAS wurden 68 Varianten im *ARRDC3*-Gen annotiert (vgl. Anhang Tabelle 9.9), von denen drei Polymorphismen im exonischen Bereich lokalisiert sind (Tabelle 4.10). Der Top-Marker führt zu einem nicht-synonymen Austausch eines Tyrosins zu Cystein (p.Y182C) zwischen der Arrestin-N- und Arrestin-C-Domäne. *SIFT* beschreibt die Auswirkung dieses Austausches als „tolerated“, während *Polyphen* die A>G-Mutation als „possibly damaging“ einstuft. Vier der insgesamt sieben Mutationen, die im hohen LD zum Top-Marker stehen ($r^2 > 0,8$), sind chromosomenweit signifikant.

Tabelle 4.10: Polymorphismen im exonischen Bereich von *ARRDC3*

Position (bp)	Region	REF	ALT	AF	Effekt	SIFT	Polyphen	P-Wert	LD als r^2
93.253.064	5'UTR	G	A	0,035				0,535	0,003
93.244.933*	CDS	TAT	TGT	0,905	Y182C	tolerated	possibly damaging	$4,327 \times 10^{-9}$	0,996
93.243.389	CDS	GGG	GGA	0,802	G377G			$2,8948 \times 10^{-7}$	0,458

Die Polymorphismen sind nach ihre Lokalisation im Gen *ARRDC3* basierend auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009) sortiert. Der Effekt der Mutationen, die einen Aminosäureaustausch herbeiführen, wurde mit SIFT (NG und HENIKOFF 2003) und Polyphen (ADZHUBEI *et al.* 2010) vorhergesagt. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; AA = betroffene Aminosäuren; LD(r^2) gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr7: 93.244.933) an; * Top-Marker.

4.3.1.4 QTL auf Chromosom 17

Die Assoziationsstudie für das Merkmal direkte Totgeburten bei Holsteinkühen brachte einen QTL auf Chromosom 17 hervor. Der Polymorphismus mit der höchsten Signifikanz (P-Wert = $2,6154 \times 10^{-8}$) ist im Gen *SLC2A11* lokalisiert, wobei weitere hoch signifikante Varianten mit einem hohen LD zum Top-Marker ($r^2 > 0,8$) in der genreichen Region lokalisiert sind (Abbildung 4.3.5).

In diesem assoziierten Bereich (73,1 bis 73,3 Mb) sind zwölf Gene lokalisiert, von denen vier bisher noch unbekannt bzw. nicht charakterisiert sind (*LOC100298791*, *LOC100296711*, *LOC789355*, *LOC506093*) (Abbildung 4.3.5). Das Zinkfingerprotein 70 (ZNF70) kodierende Gen *ZNF70* ist in Publikationen kaum beschrieben. Möglicherweise ist es wie die meisten ZNF-Proteine in der Lage, an spezifische DNA-Sequenzen zu binden und dort regulatorisch auf die Expression bestimmter Gene zu wirken (KLUG 1999).

Über die Funktion des *VPREB3*-Gen-Produkts ist bisher ebenfalls wenig bekannt. Die Expression von *VPREB3* findet im Knochenmark und in sekundären lymphatischen Organen statt (SHIRASAWA *et al.* 1993; RODIG *et al.* 2010). Vermutlich ist es an der Bildung und beim Transport von B-Lymphozyten im Immunsystem beteiligt (ROSNET *et al.* 2004). Insbesondere bei Tumoren der Lymphknoten wird *VPREB3* stark exprimiert (RODIG *et al.* 2010). Das Gen *CHCHD10* kodiert das Protein, das eine Coiled-Coil-Helix-Coiled-Coil-Helix-Domäne enthält. Es ist in Mitochondrien lokalisiert und wird in der quergestreiften Muskulatur, wie z.B. im Herzen, Zunge und Skelettmuskulatur, exprimiert (MARTHERUS *et al.* 2010; BANNWARTH *et al.* 2014). Das Protein wird für die Erzeugung von Energie in Form von Adenosintriphosphat (ATP) bei der oxidativen

Phosphorylierung in der Atmungskette benötigt (VAN KUILENBURG *et al.* 1992; MARTHERUS *et al.* 2010). Missense Mutationen in diesem Gen sind verantwortlich für eine autosomal dominante mitochondriale Myopathie (AJROUD-DRISS *et al.* 2012). Außerdem wird *CHCHD10* mit Erkrankungen des Nervensystems assoziiert (BANNWARTH *et al.* 2014).

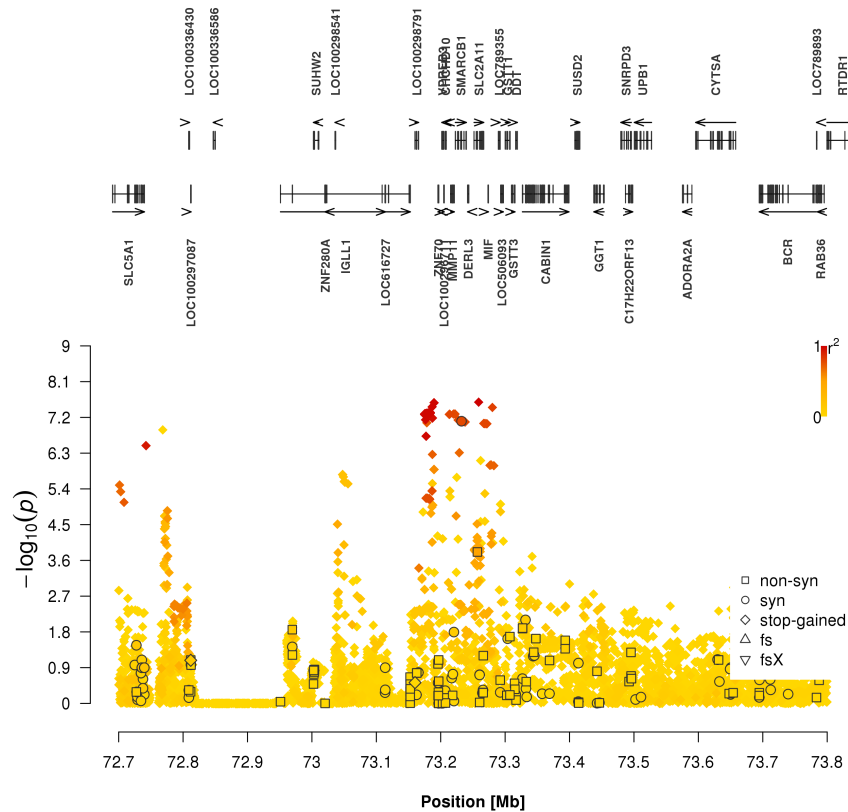


Abbildung 4.3.5: Region-Plot auf Chromosom 17 für direkte Totgeburten bei Holsteinkühen

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligem LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Die von *MMP11* kodierte Matrix Metalloproteinase 11 (MMP11, Stromelysin-3) wird in vielen Karzinomen beim Menschen entdeckt. In Tumor umgebenden, normalen Geweben wird sie selten exprimiert (KOSSAKOWSKA *et al.* 1996; TETU *et al.* 1998). Außerdem gibt es Hinweise, dass MMP11 in der Embryonalentwicklung involviert ist (DAMJANOVSKI *et al.* 2001; WEI und SHI 2005). SAVARIS *et al.* (2008) berichteten über eine Hochregulierung von MMP11 in der Gebärmutter während der Schwangerschaft. Ähnlich wie *MMP11* sind Mutationen im Gen *SMARCB1* mit verschiedenen Tumorarten assoziiert (EATON *et al.* 2011; KIM und ROBERTS 2014). Untersuchungen

zeigten, dass SMARCB1 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1) als Tumorsuppressor wirkt. Beim Menschen ist eine große Deletion auf Chromosom 22, einschließlich des Gens *SMARCB1*, mit einem erhöhten Risiko des schon im (Klein-) Kindesalter auftretenden rhabdoiden Tumors assoziiert (JACKSON *et al.* 2007). Das Transmembranprotein Derlin 3, vom Gen *DERL3* kodiert, ist am Mechanismus der mit dem Endoplasmatischen Retikulum assoziierten Degradation beteiligt (ODA *et al.* 2006; BELMONT *et al.* 2010). Dieser Abbauweg verhindert das Austreten von fehlgefalteten oder unvollendet zusammengebauten Polypeptiden aus dem Endoplasmatischen Retikulum und vermindert somit die Gefahr der Proteotoxizität und Beeinträchtigung der Zellfunktion. Durch Hypermethylierung bei verschiedenen Tumorarten wird *DERL3* inaktiviert (LOPEZ-SERRA *et al.* 2014). Das inaktivierte *DERL3* führt zu einer Überexpression von *SLC2A1*, welches durch die Glykolyse für die Energieversorgung zuständig ist. Während der Trächtigkeit und dem fetalem Wachstum ist die Energieversorgung durch *SLC2A1* von besonderer Bedeutung (YAMAGUCHI *et al.* 1996). Eine vergleichbare Rolle nimmt das Gen *SLC2A11* ein, welches in der assoziierten genreichen Region auf BTA5 lokalisiert ist. Der vom *SLC2A11*-Gen kodierte Glukosetransporter GLUT11 ist als Transmembranprotein für den Austausch von Glukose und Fruktose durch die Zellmembran zuständig (SASAKI *et al.* 2001; DOEGE *et al.* 2001; SCHEEPERS *et al.* 2005). GASTER *et al.* (2004) zeigten, dass GLUT11 in den langsam kontrahierenden, auf Dauerleistung ausgelegten Muskelzellen lokalisiert ist. Im Gegensatz zu anderen Glukosetransportern bleibt die Expression von *SLC2A11* vom physiologischen Befinden des Organismus unbeeinflusst. Während der Trächtigkeit ist der Glukosemetabolismus besonders wichtig für die Einnistung der Eizelle sowie für die Versorgung und Entwicklung des Embryos (VON WOLFF *et al.* 2003; KIM und MOLEY 2009). Als Bestandteil des angeborenen Immunsystems wirkt der vom Gen *MIF* verschlüsselte Makrophagen-Migrationshemmende Faktor (CALANDRA und ROGER 2003). Polymorphismen im *MIF*-Gen sind mit einem erhöhten Risiko für das Auftreten der am häufigsten entzündlichen Gelenkserkrankung rheumatoide Arthritis assoziiert (DONN *et al.* 2001). SINGH *et al.* (2012) vermuten, dass unterschiedliche *MIF*-Level in der Plazenta einen negativen Effekt auf die Geburt haben.

Die zehn signifikantesten Polymorphismen in der assoziierten QTL-Region weisen zueinander ein hohes Kopplungsungleichgewicht zueinander auf ($r^2 \geq 0,98$) (Tabelle 4.11). Die Allelsubstitutionseffekte dieser Varianten befinden sich auf einem etwa gleich hohen Level im Bereich von 1,04 bis 1,35.

Ergebnisse

Tabelle 4.11: Signifikante Polymorphismen auf Chromosom 17 (72,7 bis 73,8 Mb)

Position (bp)	REF	ALT	AF	P-Wert	α	LD als r^2
73.258.826*	A	G	0,13	$2,62 \times 10^{-8}$	1,19	1
73.189.624	A	C	0,15	$2,71 \times 10^{-8}$	1,13	0,97
73.186.754	T	C	0,12	$3,43 \times 10^{-8}$	1,35	0,97
73.186.764	T	C	0,12	$3,43 \times 10^{-8}$	1,35	0,97
73.280.318	G	A	0,15	$3,53 \times 10^{-8}$	1,04	0,8
73.183.467	A	G	0,13	$4,65 \times 10^{-8}$	1,11	0,98
73.181.967	T	C	0,13	$4,68 \times 10^{-8}$	1,11	0,98
73.183.818	T	C	0,13	$4,77 \times 10^{-8}$	1,13	0,98
73.183.949	A	C	0,13	$4,77 \times 10^{-8}$	1,13	0,98
73.182.642	A	G	0,13	$4,86 \times 10^{-8}$	1,11	0,98

*Die Polymorphismen sind nach ihrem P-Wert aus der GWAS mit EMMAX geordnet. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; α = Allelsubstitutionseffekt; LD als r^2 gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr17: 73.258.826) an; * = Top-Marker.*

4.3.1.5 QTL auf Chromosom 18 (14,5 bis 15,6 Mb)

Auf dem 18. bovinen Chromosom konnten zwei QTL für das Merkmal direkte Totgeburten bei Kühen lokalisiert werden. Der chromosomenweit signifikante QTL ist auf den Bereich von 14,5 bis 15,6 Mb beschränkt (15.038.275 bp; P-Wert = $9,7216 \times 10^{-9}$). Diese Variante ist im Gen *VPS35* annotiert (Abbildung 4.3.6).

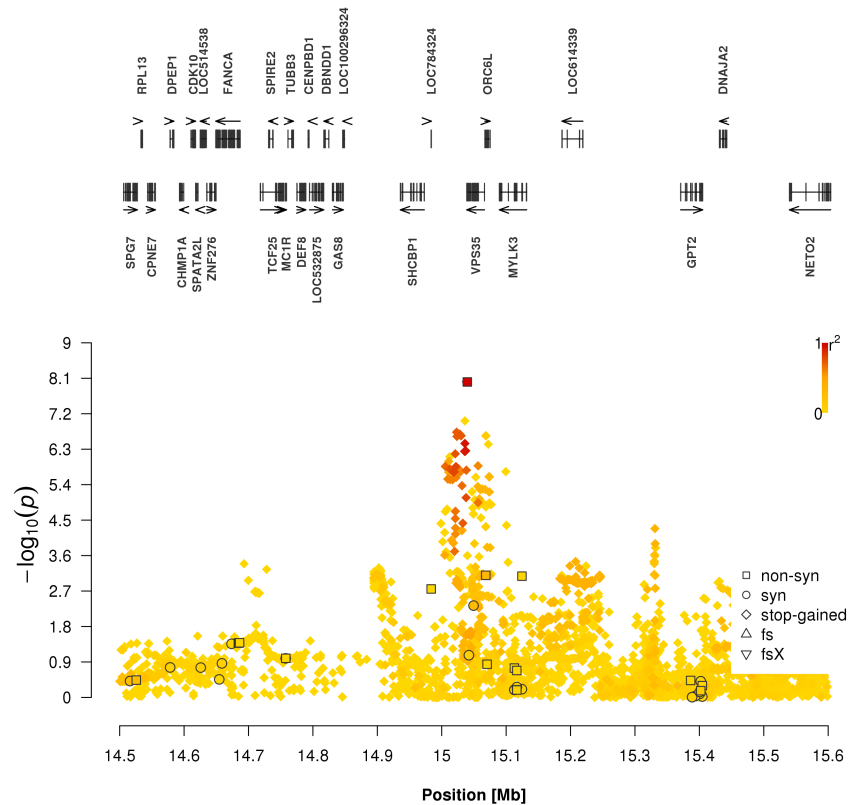


Abbildung 4.3.6: Region-Plot auf Chromosom 18 für direkte Totgeburten bei Holsteinkühen (14,5-15,6Mb)

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMLIN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligen LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Charakterisierung von *VPS35*

VPS35 kodiert für ein vakuoläres Protein sorting 35 (vacuolar protein sorting-associated protein 35, VPS35), ein Mitglied der *VPS*-Gene (EDGAR und POLAK 2000). Beim Menschen wird *VPS35* in zahlreichen Geweben exprimiert, insbesondere in Gehirn, Herz, Hoden, Ovarium, Dünndarm, Milz, Skelettmuskulatur und Plazenta (ZHANG *et al.* 2000). VPS35 ist Bestandteil eines

Proteinkomplexes, der die Rückgewinnung und den Transport von Membranproteinen aus den Endosomen zum Trans-Golgi-Netzwerk bewirkt (SHI *et al.* 2006; MCGOUGH und CULLEN 2011). Mit zwei weiteren VPS-Proteinen (VPS26 und VPS29) bildet es den Teil des Mechanismus, in dem die zu transportierenden Moleküle erkannt werden. Die N-terminale Domäne von VPS35 ist mit VPS26 assoziiert, während die C-terminale Region mit VPS29 verbunden wird. Die beiden VPS35-Domänen übernehmen die zentrale Rolle der Bindung des Retromers mit Endosomen (RESTREPO *et al.* 2007). Diese Bindung wird durch den Dimer Nexin als Teil des Komplexes zusätzlich verstärkt (HAFT *et al.* 2000; DENG *et al.* 2013). Dieser gesamte Retromerkomplex übernimmt essentielle Aufgaben in der Embryonalentwicklung und weiteren zellulären Aktivitäten, wie z.B. dem programmierten Zelltod (GRIFFIN *et al.* 2005; CHEN *et al.* 2010).

Mutationen in *VPS35* sind mit neurodegenerativen Erkrankungen wie die Parkinson- (OMIM 614203) und Alzheimer-Krankheit assoziiert (SMALL *et al.* 2005; VILARIÑO-GÜELL *et al.* 2011; ZIMPRICH *et al.* 2011; WEN *et al.* 2011).

Tabelle 4.12: Polymorphismen im exonischen Bereich von VPS35

Position (bp)	Region	REF	ALT	AF	Effekt	SIFT	Polyphen	P-Wert	LD als r^2
15.057.327	5'UTR	C	T	0,715				0,022	0,234
15.057.312	5'UTR	T	C	0,761				0,001	0,187
15.057.304	5'UTR	C	A	0,023				0,266	0,017
15.049.386	CDS	CTC	CTT	0,728	L383L			0,005	0,26
15.042.152	CDS	GCG	GCA	0,070	A610A			0,085	0,062
15.039.709*	CDS	AGT	AAT	0,390	S760N	tolerated	benign	9,8498x10 ⁻⁹	1
15.039.449	3'UTR	G	A	0,057				0,470	0,042

*Die Polymorphismen sind nach ihre Lokalisation im Gen VPS35 basierend auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009) sortiert. Der Effekt der Mutationen, die einen Aminosäureaustausch herbeiführen, wurde mit SIFT (NG und HENIKOFF 2003) und Polyphen (ADZHUBEI *et al.* 2010) vorhergesagt. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; AA = betroffene Aminosäuren; LD(r^2) gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr18: 15.039.709) an; * Top-Marker.*

VPS35 konnte im bovinen Genom auf dem 18. Chromosom lokalisiert werden. Es erstreckt sich über ~18,5 kb auf dem negativen Strang zwischen 15.038.816 bp bis 15.057.329 bp. Das aus 17 Exons bestehende Transkript mit einem vorgelagertem, untranslatierten Exon kodiert eine 796 Aminosäure lange Proteinsequenz. Das Protein ist beim Vergleich mit verschiedenen Spezies sehr hoch konserviert.

Insgesamt wurden 213 Mutationen in der Genregion von *VPS35* annotiert und in der GWAS analysiert. Im Promotor wurden 30, im 5'UTR drei, im 3'UTR eine und in der kodierenden Sequenz drei Varianten identifiziert. Unter den kodierenden Mutationen befindet sich der signifikanteste Marker dieses QTL für das Merkmal direkte Totgeburten bei Holsteinkühen auf Chromosom 18 (15.039.709 bp; P-Wert = $9,8498 \times 10^{-9}$). Dieser Allelaustausch wandelt in der C-terminalen Domäne die Aminosäure Serin zu Asparagin um (p.S760N). *SIFT* und *Polyphen* stufen diese Veränderung in der Proteinsequenz als harmlos ein (Tabelle 4.12). Keiner der genischen Varianten in *VPS35* liegt im hohen LD mit dem Top-Marker (MAX: 0,641; vgl. Anhang Tabelle 9.10).

4.3.1.6 QTL auf Chromosom 18 (59 bis 60,1 Mb)

Auf Chromosom 18 wurde der genomweit signifikanteste Polymorphismus für das Merkmal direkte Totgeburten bei Kühen lokalisiert (59.586.457 bp). Der P-Wert dieses Markers betrug $5,0011 \times 10^{-53}$. Dieser Bereich auf dem bovinen Chromosom 18 enthält zahlreiche unbekannte Gene („LOC“, Abbildung 4.3.7). Als positionelles Kandidatengen, dessen Transkriptionsstart 31 kb vom Top-Marker entfernt ist, konnte das Gen *LOC787812* identifiziert werden.

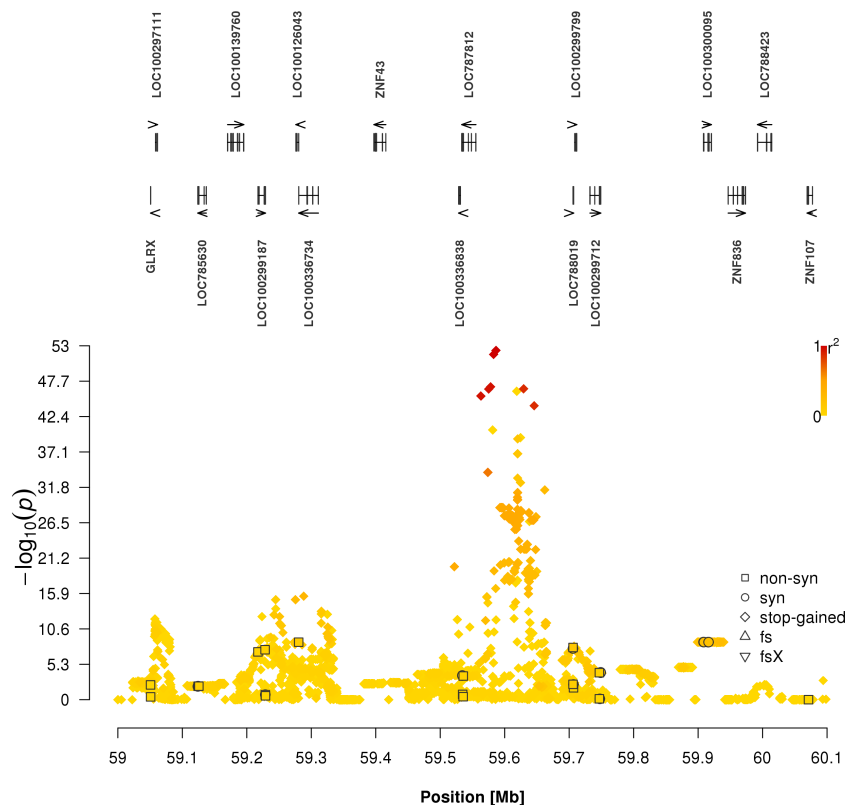


Abbildung 4.3.7: Region-Plot auf Chromosom 18 für direkte Totgeburten bei Holsteinkühen (59-60,1Mb)

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMIN et al. 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligem LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Charakterisierung von *LOC787812*

Das bisher nicht charakterisierte bovine Gen *LOC787812* liegt auf dem negativen Strang des 18. Chromosoms und erstreckt sich über eine Länge von ~22 kb, aufgeteilt in sechs Exons (Transkriptionsstart: 59.555.387; Transkriptionsende: 59.533.349). Mittels *BLAST*-Suche wurde

die 760 Aminosäure lange Sequenz des bovinen *LOC787812* (XP_002695345.3) mit allen Proteinsequenzen der RefSeq-Protein-Datenbank (PRUITT *et al.* 2014) abgeglichen, um basierend auf möglichen Sequenzhomologien ein geeignetes, bei einer andere Spezies bekanntes Protein zu identifizieren. Alle 760 vorgegebenen Aminosäuren stimmen lückenlos und vollständig mit dem bovinen Zinkfingerprotein 845 (ZNF845; XP_001255074.5) überein (vgl. Anhang Abbildung 9.2). Die Lage der beiden Gene ist nicht identisch. Das für ZNF835-Protein kodierende Gen befindet sich im Gegensatz zu *LOC787812* auf dem positiven Strang und ist ~280 kb von *LOC787812* entfernt (Chr18: 59.195.319 - 59.253.774).

Das von *LOC787812* kodierte Protein ist ähnlich wie zahlreiche Zinkfingerproteine (ZFP) durch eine N-terminale Krüppel-assoziierte Box (KRAB) charakterisiert (HUNTLEY *et al.* 2006), die eine typische Domäne für Transkriptionsfaktoren darstellt. Die KRAB-Domäne besteht aus einer A-Box und in einigen Fällen auch aus einer B-Box. Insbesondere die A-Box dient als Repressor der Transkription von Genen und verhindert somit die Entstehung des vom Gen kodierenden Proteins (WITZGALL *et al.* 1994). Dabei interagiert sie mit anderen Korepressoren (VISSING *et al.* 1995). Am C-terminalen Ende von Zinkfingerproteinen befindet sich die charakteristische Zink-Finger-Doppel-Domäne, die je nach Protein durch eine variierende Anzahl an Zink-Finger-Motiven (C₂H₂) definiert ist. Bei Proteinen mit einer KRAB-Domäne ist sie für die Bindung der DNA zuständig (LOOMAN *et al.* 2002).

Zinkfingerproteine mit einer KRAB-Domäne stellen eine der größten tetrapoden spezifischen Zinkfingerprotein-Familien dar und sind zusätzlich die größte Gruppe der Transkriptionsfaktoren in Säugetieren (LOOMAN *et al.* 2002). Sie sind meist in Gruppen lokalisiert, da sie wahrscheinlich durch Tandem-Duplikation entstanden sind (COLLINS *et al.* 2001). Für viele der Zinkfingerprotein-KRAB Gene ist die genaue Funktion noch ungeklärt. HUNTLEY *et al.* (2006) vermuten jedoch, dass viele in einer Region konzentrierte Zinkfingerprotein-KRAB-Paraloge an unterschiedlichen Prozessen beteiligt sind. Studien zeigten, dass Zinkfinger-KRAB-Proteine eine wichtige Rolle bei der Zelldifferenzierung und bei der Entwicklung von Organen, wie auch bei der Apoptose übernehmen (KATOH *et al.* 1998; JHEON *et al.* 2001; YANG 2003). Des Weiteren wurden sie mit Entwicklungsstörungen, neoplastischen Schäden und Schizophrenie assoziiert (TOMMERUP und VISSING 1995; TAKASE *et al.* 2001).

Die KRAB-A-Box wird beim *LOC787812*-kodierten Protein von den Aminosäuren 8 bis 46 festgelegt. Die Zinkfinger-Domäne erstreckt sich von den Aminosäuren 266 bis 688, wobei sich die C₂H₂-Motive im Bereich von 568 bis 593 befinden.

Ergebnisse

Tabelle 4.13: Polymorphismen im exonischen Bereich von LOC787812

Position (bp)	Region	REF	ALT	AF	Effekt	SIFT	Polyphen	P-Wert	LD als r^2
59.555.382	5'UTR	G	A	0,050				4,5202x10 ⁻⁵	0,007
59.555.340	5'UTR	TTTCTTT CTTCTA	TTTCTT TCTA	0,007				0,335	0
59.535.738	CDS	CCT	GCT	0,282	P194A	tolerated	benign	2,8x10 ⁻⁴	0,05
59.535.519	CDS	CCA	GCA	0,073	P267A	deleterious	possibly da- maging	0,357	0,01
59.534.816	CDS	TAT	TGT	0,113	Y501C		possibly da- maging	0,161	0,019
59.534.257	CDS	CCG	CCA	0,286	P634P			2,3x10 ⁻⁴	0,05

Die Polymorphismen sind nach ihre Lokalisation im Gen LOC787812 basierend auf dem UMD3.1-Assembly (ZIMIN et al. 2009) sortiert. Der Effekt der Mutationen, die einen Aminosäureaustausch herbeiführen, wurde mit SIFT (NG und HENIKOFF 2003) und Polyphen (ADZHUBEI et al. 2010) vorhergesagt. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; AA = betroffene Aminosäuren; LD als r^2 gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr18: 59.586.457) an.

Im Gen *LOC787812* wurden insgesamt 69 Polymorphismen bei Holstein identifiziert (vgl. Anhang Tabelle 9.11), von denen sechs im exonischen Bereich lokalisiert sind (Tabelle 4.13). Eine einzige Variante (59.555.382 bp) weist einen signifikanten P-Wert auf (P-Wert = 4,5202 x 10⁻⁵), jedoch steht sie, wie auch alle anderen genischen Varianten, in keinem oder einem nur sehr niedrigen Kopplungsungleichgewicht mit dem Top-Marker an Position 59.586.457. Somit können die genischen Polymorphismen in *LOC787812* als Kandidatenmutation ausgeschlossen werden. Vermutlich spielen regulatorische Varianten eine Rolle.

4.3.1.7 QTL auf Chromosom 29

Der QTL auf dem bovinen Chromosom 29 enthält zwei Polymorphismen, die die chromosomenweite Signifikanzgrenze von $P = 1,23827 \times 10^{-7}$ überschreiten. Der bei 50.126.590 bp lokalisierte Top-Marker erreicht einen P-Wert von $9,8292 \times 10^{-8}$. Der Transkriptionsstart des funktionellen Kandidatengens *H19* ist ~ 2 kb entfernt (Abbildung 4.3.8).

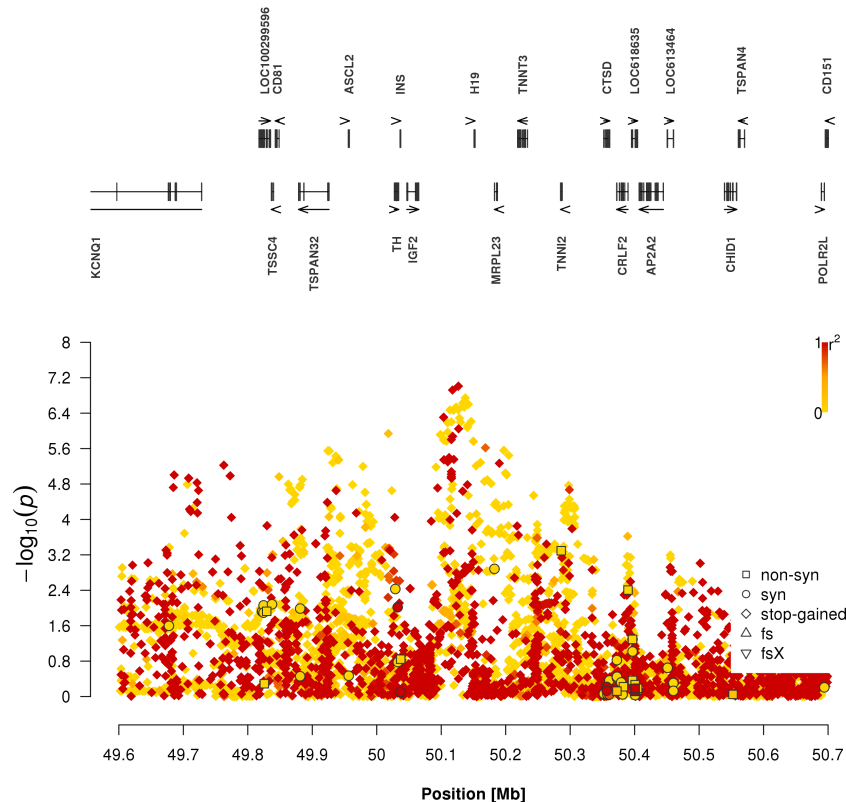


Abbildung 4.3.8: Region-Plot auf Chromosom 29 für direkte Totgeburten bei Holsteinkühen

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligen LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Charakterisierung von *H19*

Das Gen *H19* wird in ein RNA-Produkt transkribiert, welches nicht translatiert wird. Beim Menschen liegt *H19*-mRNA während der Embryonalentwicklung in der Plazenta und in verschiedenen Geweben des Embryos vor (GOSHEN *et al.* 1993). Eine Expression von *H19* ist mit einer Unterdrückung des Wachstums assoziiert. Während der frühen Trächtigkeit werden noch beide

parentale Allele (maternal und paternal) des *H19*-Gens exprimiert (ARIMA *et al.* 1997). Im späteren Verlauf der Trächtigkeit wird nur noch das maternale Allel exprimiert; das vom Vater ist durch Methylierung inaktiviert (BRANNAN *et al.* 1990; ZHANG *et al.* 2004). *H19* ist somit ein imprintetes Gen. Direkt nach der Geburt wird die Expression von *H19* mit Ausnahme der Skelettmuskulatur in allen Geweben herunter reguliert (BRUNKOW und TILGHMAN 1991). Postnatal wird *H19* während der Pubertät und der Schwangerschaft in der Brustdrüse transkribiert (ADRIAENSSENS *et al.* 1999).

LEIGHTON *et al.* (1995) zeigten, dass *H19* die Expression vom ~80 kb entfernten paternal exprimierten Insulin-ähnlichen Wachstumsfaktor 2 (*IGF2*) beeinflusst. Das Ausschalten des maternalen Allels von *H19* führt zu einer erhöhten Funktion von *IGF2*. Die Nachkommen, denen das maternale Allel von *H19* fehlte, wiesen ein höheres Körpergewicht auf. Bei Mäusen bewirkt das Ausschalten von *H19* ähnlich wie beim Menschen ein übermäßiges Wachstum (LEIGHTON *et al.* 1995; GABORY *et al.* 2009), während sich eine Überexpression von *H19* letal auswirkt (BRUNKOW and TILGHMAN 1991).

Tabelle 4.14: Polymorphismen im exonischen Bereich von H19

Position (bp)	Region	REF	ALT	AF	P-Wert	LD als r^2
50.150.909	EXON1	G	C	0,012	0,761	1
50.151.145	EXON1	G	A	0,100	0,907	1
50.151.244	EXON1	G	A	0,182	0,483	0,245
50.151.479	EXON1	G	A	0,871	0,962	1
50.151.678	EXON1	C	T	0,011	0,022	1
50.151.957	INTRON1	CCTCTCTCTCTCT	CCTCTCTCTCT	0,008	0,008	1
50.152.361	INTRON3	C	A	0,007	0,436	1
50.152.638	EXON5	G	A	0,183	0,485	0,24
50.152.969	EXON5	A	G	0,049	0,194	1

*Die Polymorphismen sind nach ihrer Lokalisation im Gen H19 basierend auf dem UMD3.1-Assembly (ZIMIN *et al.* 2009) sortiert. REF = Referenzallel in Leserichtung des Gens; ALT = Alternativallel in Leserichtung des Gens; AF = Allelfrequenz von ALT; LD(r^2) gibt das Kopplungsungleichgewicht der Varianten zum Top-Marker (Chr29: 50.148.761) an.*

Beim Menschen wird das Gen mit dem Beckwith-Wiedemann-Syndrom (OMIM 130650), dem Silver-Russel-Syndrom (OMIM 180860) und dem Wilms Tumor 2 (OMIM 194071) assoziiert. Das Beckwith-Wiedemann-Syndrom ist gekennzeichnet durch teilweise asymmetrischen Großwuchs verbunden mit bereits im Kindesalter vorhandenen Fehlbildungen und Tumoren (WEKSBERG *et al.* 2009). Dazu gehören embryonal auftretende bösartige Auswüchse, wie der in

der Niere erscheinende Wilms-Tumor. Das Silver-Russel-Syndrom ist eine Form des intrauterinen Kleinwuchses, einhergehend mit Fehlentwicklungen des Schädels und des Gesichts sowie einem niedrigeren Drang zur Nahrungsaufnahme (MARTÍNEZ NOGUEIRAS *et al.* 2001).

H19 befindet sich im bovinen Genom auf dem positiven Strang des 29. Chromosoms im Bereich von 50.150.775 bp bis 50.153.147 bp. Das aus fünf Exons bestehende Transkript hat eine Gesamtlänge von ~2,3 kb. Im *H19*-Gen sind 13 Varianten lokalisiert, die in der GWAS verwendet wurden (Tabelle 4.14), von denen sich neun Polymorphismen im vollständigen Kopplungsgleichgewicht mit dem Top-Marker befinden. Die gesamte Region um den Top-Marker weist ein sehr hohes LD auf (Abbildung 4.3.8). Mit Ausnahme einer Variante im Promotor von *H19* (50.149.817 bp; P-Wert = $1,72 \times 10^{-6}$) sind die genischen Polymorphismen dieses Gens gering signifikant.

5 Diskussion

Die Einführung von NGS-Technologien ermöglicht die Re-Sequenzierung kompletter Genome. Aufgrund stetig sinkender Kosten dieser Methode kann die komplette genomische Variation einzelner Individuen preiswert analysiert werden und maßgeblich zur Aufdeckung kausaler Mutationen beitragen (EICHLER *et al.* 2010).

In dieser Arbeit wurden genomweite Re-Sequenzierungsdaten beim Rind basierend auf 214 Genomen analysiert, um potentiell kausale DNA-Varianten zu identifizieren. Dazu wurden vorzeitige Stopp-Codons systematisch identifiziert und analysiert sowie eine sequenzbasierte GWAS für das Merkmal Totgeburten in der Rasse Holstein durchgeführt.

5.1 Sichtung der genomischen Variation

Die genomweite Re-Sequenzierung von 214 Individuen aus sechs unterschiedlichen Rinderrassen brachte mehr als 22 Mio. variable Positionen hervor. Aufgrund der niedrigen und immer weiter sinkenden N_e in Rinderpopulationen (NOMURA *et al.* 2001; HAYES *et al.* 2003; THE BOVINE HAPMAP CONSORTIUM 2009) ist die hohe Variation zunächst erstaunlich, denn Populationen mit geringer N_e weisen ein hohes LD und somit eine geringere genetische Variation auf. Das Ausmaß der identifizierten Polymorphismen ist durch die unterschiedlichen LD-Strukturen in den verschiedenen Populationen zu erklären (THE BOVINE HAPMAP CONSORTIUM 2009). Trotz der sinkenden N_e in den einzelnen Populationen ist die genetische Diversität beim Rind und somit auch innerhalb der sechs analysierten Rassen hoch. Des Weiteren ist die Einbeziehung von Schlüssel-tieren, die Bestandteil der N_e sind und einen großen genetischen Anteil an der aktuellen Population haben, von entscheidender Bedeutung für den Umfang detektierter Polymorphismen (GODDARD und HAYES 2009a).

In vorangegangenen Studien beim Rind, in denen einzelne Tiere genomweit analysiert wurden, lag die Anzahl detektierter Polymorphismen bei etwa drei Mio. Positionen (ECK *et al.* 2009; STOTHARD *et al.* 2011). Im vorliegenden Datensatz konnten unabhängig von der genomischen Abdeckung und Rassezugehörigkeit des einzelnen Tieres etwa sechs Mio. variable biallelische Polymorphismen in jedem Individuum detektiert werden, vergleichbar mit CHOI *et al.* (2014). Die Anwendung der multi-sample calling Methode (LI 2011) und die anschließende

Beagle-Imputation (BROWNING und BROWNING 2009) erklären die höhere Anzahl an Polymorphismen je Tier. Auch Tiere ohne Abdeckung erhalten einen Genotyp für eine im Datensatz identifizierte variable Position. Sobald Individuen mit einem hoch abgedeckten Genom in die Analyse einbezogen sind, können unabhängig von der Abdeckung für alle Proben zuverlässige Genotypen ermittelt werden (DEPRISTO *et al.* 2011). Neben der in dieser Arbeit angewandten multi-sample calling Methode und der Imputation mit *Beagle* erschwert die Verwendung unterschiedlicher Programme und Filter den Vergleich zwischen den einzelnen Studien.

Die 214 re-sequenzierten Individuen weisen durchschnittlich eine ~ 7 -fache Abdeckung des Genoms auf. Studien belegen, dass diese Genomabdeckung ausreichend ist, um eine umfangreiche und zuverlässige Identifizierung der Varianten und ihrer Genotypen zu gewährleisten (KAWAHARA-MIKI *et al.* 2011; STOTHARD *et al.* 2011; JANSEN *et al.* 2013). Basierend auf diesen Ergebnissen kann davon ausgegangen werden, dass die genomische Variation in Form von SNP und kurzen InDels in jedem vorliegend untersuchten Tier mit einer hohen Sensitivität und Spezifität abgefragt wurde, was durch die hohe Übereinstimmung der Genotypen, die hohe NRS- und die geringe NRD-Rate bestätigt wird. Eine Vielzahl dieser beiden Mutationsformen liegt dank verschiedener Studien in öffentlichen Datenbanken vor (ECK *et al.* 2009; THE BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM *et al.* 2009; THE BOVINE HAPMAP CONSORTIUM 2009; STOTHARD *et al.* 2011; DAETWYLER *et al.* 2014). Jedoch ist der Katalog der genetischen Variation bis dato noch nicht vollständig. Strukturelle Varianten, die vor allem aufgrund ihrer Komplexität bisher wenig Berücksichtigung für die genetische Aufklärung eines Phänotyps fanden, sind in öffentlichen Datenbanken wenig präsent. Dies bedeutet jedoch nicht, dass sie nicht im Rindergenom auftreten und bei der Aufdeckung von Erbdefekten keine Bedeutung haben. Beim Rind sind bereits große Deletionen bekannt, die kausal für genetisch bedingte Fehler sind, wie z.B. Brachyspina (CHARLIER *et al.* 2012) sowie Aborte und Totgeburten (FLSIKOWSKI *et al.* 2010). Daher sollte zur Vervollständigung der gesamten genomischen Variation die Entdeckung komplexer Mutationen in zukünftigen Studien angestrebt werden. Dazu existieren bereits verschiedenste Methoden, deren Anwendung in NGS-Daten die Detektion von strukturellen Varianten vereinfachen können (NORD *et al.* 2011; ABYZOV *et al.* 2011).

5.2 Identifizierung kausaler Mutationen

Die Ergebnisse der funktionellen Annotation der identifizierten Varianten zeigen, dass sich die Verteilung der Varianten zwischen genischen und intergenischen Bereichen bzw. zwischen kodierender und nicht-kodierender Region enorm unterscheidet, was vor allem durch den geringen Anteil der kodierenden Region am Gesamtgenom zu erklären ist. Besonders niedrig ist der Anteil an nicht-synonymen Polymorphismen. Da nicht-synonyme Austausche im Vergleich zu synonymen und nicht-kodierenden Polymorphismen seltener auftreten (YAMAGUCHI-KABATA *et al.* 2008), ist von einem schwerwiegendem Effekt auszugehen (LI *et al.* 2010a). Es ist wahrscheinlich, dass Varianten mit funktioneller Auswirkung den größten Effekt auf ein Merkmal haben. Nicht-synonyme Austausche stehen aufgrund ihrer möglichen schädlichen verändernden Wirkung auf das Protein oft unter dem negativen Selektionsdruck (CARGILL *et al.* 1999; BOYKO *et al.* 2008). In der vorliegenden Arbeit wurde nur ein geringer Anteil der nicht-synonymen Austausche als schädlich eingestuft. SUNYAEV *et al.* (2001) konnten beispielsweise zeigen, dass nicht alle nicht-synonymen Aminosäureaustausche einen negativen Effekt auf das resultierende Peptid haben. Da die Effekte der Aminosäureaustausche auf Vorhersage-Programme beruhen, sollten die Ergebnisse in weiteren Studien genauer analysiert werden.

Neben den nicht-synonymen Austauschen kommen insbesondere vorzeitige Stopp-Codons selten im Genom vor. Auch wenn die Auswirkungen von Stoppmutationen stark von ihrer Position im Protein sowie vom betroffenen Gen abhängen (THEIN *et al.* 1990; INOUE *et al.* 2004; SAVAS *et al.* 2006), lässt deren geringe Frequenz im Vergleich zu Aminosäureaustauschen einen schädlicheren Effekt vermuten (YAMAGUCHI-KABATA *et al.* 2008; BALASUBRAMANIAN *et al.* 2011; MACARTHUR *et al.* 2012).

Die Gesamtheit der genomischen Variation ist ein wichtiger Bestandteil bei der Identifizierung und dem Management von Mutationen, die für quantitative Leistungsmerkmale, Erbfehler oder sonstige spezielle Phänotypen kausal sind. Liegen Genomsequenzen vor, können bereits bekannte Varianten, z.B. aus der OMIA-Datenbank (ONLINE MENDELIAN INHERITANCE IN ANIMALS 2014), direkt abgefragt werden, auch wenn kein Tier mit dem dazugehörigem Phänotyp im Datensatz enthalten ist. Die Überprüfung erfolgt nur über Genotypinformationen, wobei die Präsenz von Anlageträgern einer Mutation ausreicht, um die Segregation dieser Variante in der Population zu beurteilen. Mit Hilfe der Re-Sequenzierung von Schlüsseltieren kann der Ursprung des Defektallels ermittelt werden. Bullenlinien, die von einem Trägartier abstammen, sollten genauer beobachtet und überprüft werden. Die Kenntnis über den Trägerstatus eines Defekts kann durch

die Vermeidung von Risikoanpaarungen direkt im Zuchtprogrammen genutzt werden und auf diese Weise einen wichtigen Beitrag zur Verbesserung der Leistung und des Gesundheitszustandes von Rinderpopulationen leisten.

Die postulierte Kausalität von Mutationen kann durch die Verfügbarkeit von Genomsequenzen und durch deren direkte Abfrage ebenso in Frage gestellt oder sogar widerlegt werden. Der für die Thrombopathie in Simmental ursächliche Aminosäureaustausch (BOUDREAU *et al.* 2007) ist womöglich nicht die einzige Mutation, die kausal für den Phänotyp ist. Basierend auf der Information von Genomsequenzen vermutet JUNG (2013) eine weitere kausale Variante für die Blutgerinnungsstörung beim Rind. Des Weiteren konnte das als embryonal letal bezeichnete vorzeitige Stopp-Codon im Gen *SHBG* (FRITZ *et al.* 2013) in der vorliegenden Datenanalyse widerlegt werden. Die mittels NGS identifizierte Segregation der Mutation in BFV veranlasste eine populationsweite Genotypisierung an dieser Position. Durch das Auffinden homozygoter Tiere konnte nachgewiesen werden, dass ein embryonal letaler Effekt des SNP nicht zutrifft.

5.2.1 Vorzeitiger Abbruch der Translation

Mutationen, die den Leserahmen des Gens unterbrechen, wurden häufig als potentielle Kandidaten für die Kausalität von Erbkrankheiten thematisiert. Basierend auf den 214 Rindergenomdaten konnten 100 vorzeitige Stopp-Codons in 94 verschiedenen Genen zuverlässig identifiziert und bestätigt werden. Ein Großteil der vorhergesagten Stoppmutationen ist jedoch aufgrund falscher Genannotation, Mapping- oder Sequenzierfehler verworfen worden (MACARTHUR *et al.* 2012). Ähnlich wie in anderen Studien weisen einzelne Gene mehrere Stoppmutationen auf (MACARTHUR *et al.* 2012; ALSALEM *et al.* 2013). Vermutlich handelt es sich hierbei um Pseudogene ohne essentielle Funktion (MACARTHUR *et al.* 2012).

Die detektierten Varianten wurden mit Hilfe der *Sanger-Re-Sequenzierung* überprüft. Von den 39 Stoppmutationen mit fehlenden Homozygoten konnten so für zwölf Varianten ALT-homozygote Tiere detektiert werden. Diese Überprüfung sowie die Anwendung weiterer Filter könnten die geringere Anzahl an vorzeitigen Stopp-Codons beim Rind im Vergleich zu anderen Untersuchungen beim Menschen (MACARTHUR *et al.* 2012) sowohl innerhalb des Panels als auch je Individuum erklären. Außerdem zeigen die Ergebnisse die enorme Bedeutung der anschließenden Validierung.

Für 30 Gene, in denen eine hoch qualitative Stoppmutation identifiziert wurde, liegen Phänotypen bei Mensch und Maus vor. Für acht dieser Varianten, deren Gene im Immunsystem relevant sind, wurde nur der heterozygote Genotyp beobachtet. Das Fehlen von Homozygoten an diesen Positionen mit Phänotypinformation lässt auf ein potentiell schädliches rezessives Allel schließen, insbesondere bei einer mittleren Allelfrequenz (VANRADEN *et al.* 2011b). Jedoch können homozygote Genotypen auch durchaus aufgrund der geringen Allelfrequenz dieser Polymorphismen nicht beobachtet worden sein (SAVAS *et al.* 2006).

Jedes der 214 untersuchten Tiere trägt durchschnittlich zwölf Stoppmutationen. Kein Tier ist im Besitz aller vollständig funktionellen Gene (BALASUBRAMANIAN *et al.* 2011). Eine Proteinverkürzung oder ein Proteinverlust scheint demnach kompatibel und in keinem hohen Maße beeinträchtigend für ausgewachsene Rinder zu sein. Auch im humanen Genom sind sowohl heterozygote als auch homozygote vorzeitige Stopp-Codons nicht außergewöhnlich, selbst in gesunden Individuen (NG *et al.* 2008; YNGVADOTTIR *et al.* 2009; MACARTHUR und TYLER-SMITH 2010; MACARTHUR *et al.* 2012). Dabei sollten Stoppmutationen in gesunden Organismen jedoch funktionell kategorisiert werden, denn auch hier können kausale Allele schwerwiegender Krankheiten vorliegen. Als heterozygoter Anlageträger einer solchen Mutation, die verantwortlich für rezessive Erkrankungen ist, sind die Individuen jedoch nicht gefährdet.

Basierend auf den Regelfällen in Säugetieren (NAGY und MAQUAT 1998) wird durch 30 verifizierte Stopp-Codons möglicherweise ein NMD initiiert, u.a. in *VCAM1* und *FREMI1*. Der NMD dient als Schutzmechanismus, um die Entstehung verkürzter, veränderter Proteine zu verhindern (FRISCHMEYER und DIETZ 1999; SCHELL *et al.* 2002). Wird der Abbau der mRNA durch ein vorzeitiges Stopp-Codon veranlasst und sind gleichzeitig alle Isoformen von der Stoppmutation betroffen, kann es zu einem vollständigen Verlust der Proteinfunktion kommen (MATSUDA *et al.* 2008). Ausnahmen von dieser NMD-Regel (FRISCHMEYER und DIETZ 1999; ISKEN und MAQUAT 2007) zeigen, dass vorzeitige Stopp-Codons nicht zwangsläufig zu einem vollständigen Verlust der Proteinfunktion führen muss. Andere Stoppmutationen verkürzen das Protein und können einen noch schwerwiegenderen Effekt ausüben, insbesondere wenn eine strukturelle oder funktionelle Domäne betroffen ist (SCHELL *et al.* 2002; SAVAS *et al.* 2006; BALASUBRAMANIAN *et al.* 2011), ähnlich wie bei der Variante in *TAS2R38*. Die Stoppmutation in *TAS2R38* wird vermutlich den Verlust der letzten beiden Transmembrandomänen hervorrufen. Das Fehlen dieser Domänen, die für die Signaltransduktion der G-Proteine im intrazellulären Raum zuständig sind (VAIDEHI *et al.* 2002; FLORIANO *et al.* 2006), könnte die Funktion des Proteins extrem beeinträchtigen. Auch die Pankre-

aslipase, deren Gen *PNLIP* von einer Stoppmutation betroffen ist, wird möglicherweise ihre hochfunktionelle PLAT-Domäne verlieren. Demzufolge könnte die Lipase nicht mehr an den Kofaktor Kolipase binden (BATEMAN und SANDFORD 1999; CHAHINIAN *et al.* 2000; LOWE 2002) und ihre Funktion bei der Hydrolyse von Triglyceriden ausüben (LOWE *et al.* 1989; DAVIS *et al.* 1991; LOWE 2002). Ein gestörter Fettstoffwechsel wirkt sich gegebenenfalls auf die Qualität der Milch und des Fleisches aus (KEMPSTER 1981; NICKERSON 1995; WOOD *et al.* 2004). TANOMURA *et al.* (2011) konnten bereits Mutationen im Gen *PNLIP* mit der Fleischmarmorierung assoziieren.

Das verifizierte vorzeitige Stopp-Codon in *FREMI*, das die Proteinfunktion aufgrund des Verlusts der C-Typ-Lektin-Domäne durch einen NMD oder durch die Proteinverkürzung stören kann, kann analog zu Studien bei Maus und Mensch (AL-GAZALI *et al.* 2002; ALAZAMI *et al.* 2009; SLAVOTINEK *et al.* 2011) einen vergleichbaren Defekt beim Rind auslösen. Diese Schädigung der Proteinfunktion äußert sich in Syndaktylie, renalen und anorektalen Missbildungen sowie Deformationen des Auges (VARNUM und FOX 1981; SMYTH *et al.* 2004).

Durchschnittlich gehen ~42% der Aminosäuresequenz durch die Stoppmutation verloren. Es existieren zahlreiche Varianten, die am Anfang als auch am Ende der Sequenz lokalisiert sind. Gegebenenfalls sind die Proteinfunktionen der kodierenden Gene, bei denen die Stoppmutation am 3'-Ende des Transkripts lokalisiert ist und die Wirksamkeit des kodierten Proteins dementsprechend in geringem Ausmaß beeinträchtigt wird, toleranter gegenüber der Entstehung eines vorzeitigen Stopp-Codons und noch voll funktionsfähig (vgl. *QRFPR*) (SAVAS *et al.* 2006; MACARTHUR *et al.* 2012). Andererseits kann auch bei der Lokalisation des Stopp-Codons nahe des 5'-Endes die Funktion der Proteine erhalten bleiben, falls sich hinter dem vorzeitigen Stopp-Codon ein alternatives Start-Codon befindet (OZISIK *et al.* 2003; HOWARD *et al.* 2004; MACARTHUR *et al.* 2012).

Die Allelfrequenzverteilung der Stoppmutationen ist vergleichbar mit der Allelfrequenzverteilung der als schädlich vorhergesagten nicht-synonymen Austausche, sowohl über alle Rassen hinweg als auch bei der Unterscheidung zwischen den einzelnen Rassen. Varianten, die ein vorzeitiges Stopp-Codon entstehen lassen, sind im niedrigen Allelfrequenzbereich vorzufinden. Auch beim Menschen ist eine Anreicherung von gering frequenten Stoppmutationen erkennbar (BALASUBRAMANIAN *et al.* 2011; MACARTHUR *et al.* 2012; KIEZUN *et al.* 2013). Ein Großteil identifizierter vorzeitiger Stopp-Codons im Humangenom wird als kausal für autosomal-rezessive Erbkrankheiten beschrieben (FRISCHMEYER und DIETZ 1999; FAY und WU 2003; MORT *et al.* 2008), was auch beim Rind denkbar wäre (z.B. Stoppmutation in *FREMI*). Die Anhäufung der LoF-Varian-

ten für seltene Allele deutet auf deren schädlichen Wirkung und demzufolge auf eine negative Selektion gegen diese Allele hin (FAY und WU 2003; BALASUBRAMANIAN *et al.* 2011). Die niedrige Allelfrequenz der Stoppmutation in *VCAM1* in Verbindung mit dem letalen Phänotyp, der durch den Funktionsverlust des vaskulären Adhäsionsmoleküls herbeigeführt wird (GURTNER *et al.* 1995; KWEE *et al.* 1995a), lässt auf einen vergleichbaren, letalen Effekt beim Rind schließen. Möglich ist auch, dass die Variante eine Bedeutung bei arteriosklerotischen und pro-inflammatorischen Erkrankungen hat (CYBULSKY *et al.* 2001; DANSKY *et al.* 2001).

Gene, die in der Geruchs- und Geschmackswahrnehmung involviert sind, werden oft von vorzeitigen Stopp-Codons affektiert (ALSALEM *et al.* 2013). Je nach Umgebung und Ernährungspräferenzen stehen Mutationen in diesen Genen unter einem Selektionsdruck. Die Stoppmutation im Bittergeschmacksrezeptorgen *TAS2R38* (HOON *et al.* 1999; ADLER *et al.* 2000; KIM *et al.* 2003; WOODING *et al.* 2004) könnte beim Rind das Geschmacksvermögen für bittere Stoffe in ähnlicher Weise wie beim Menschen beeinflussen. Die geringe Allelfrequenz und die fehlenden ALT-homozygoten Genotypen deuten auf eine negative Selektion hin. Man kann vermuten, dass Tiere mit einem funktionierenden Rezeptor einen ausgeprägten bitteren Geschmackssinn haben und somit die Aufnahme von bitteren, oft auch giftigen Pflanzen verhindert werden könnte (MEYERHOF 2005).

In den 214 Tieren wurden neben den gering frequenten Stoppmutationen auch solche mit hoher Allelfrequenz und der Anwesenheit von homozygoten Tieren für beide Genotypen detektiert. Die Existenz homozygoter Genotypen bedeutet jedoch nicht, dass die beeinträchtigte Proteineigenschaft des homozygot inaktivierten Gens keinen Einfluss auf einen Phänotyp hat. Das Protein kann durch die Einführung des vorzeitigen Stopp-Codons womöglich stark geschädigt sein und einen bestimmten Phänotyp verändern. Der Phänotyp unterliegt jedoch keinem natürlichen Selektionsdruck, sodass das Überleben des Individuums davon unbeeinflusst bleibt (SAVAS *et al.* 2006). Diese Mutationen zeigen demzufolge keine Evidenz für einen Effekt auf die Überlebensfähigkeit und Fitness der Tiere (SAVAS *et al.* 2006; YNGVADOTTIR *et al.* 2009). Tiere mit homozygotem Genotyp werden nicht natürlich ausselektiert.

Die Stoppmutation im Gen *PNLIP* deutet auf einen stark schädigenden Effekt auf das Protein hin. Die vermutlich resultierende Störung des Fettmetabolismus und der Absorption von Fetten scheint die betroffenen Individuen jedoch nicht so stark zu benachteiligen, dass sie nicht überlebensfähig sind. Individuen mit einer verkürzten Pankreaslipase könnten jedoch unerwünschte

Qualitätsmerkmale der Milch und des Fleisches aufweisen. Die geringere Anzahl an ALT-homozygoten Tieren könnte durch künstliche Selektion begründet sein.

Das vorzeitige Stopp-Codon im Gen *STRC* wird die Funktionalität des kodierten Proteins im Rind sehr wahrscheinlich beeinträchtigen und könnte ähnlich wie bei Mensch und Maus eine Störung des Hörvermögens verursachen (VERPY *et al.* 2001, 2008; AVIDAN *et al.* 2003; ZHANG *et al.* 2007). Rinder, die akustische Reize nicht wahrnehmen können, sind keiner negativen Selektion ausgesetzt, was durch die Anwesenheit von ALT-homozygoten Tieren bestätigt werden kann. Für den Tierhalter kann sich ein beeinträchtigtes Hörvermögen seiner Tiere positiv auswirken, da sie möglicherweise stressfreier sind und der Umgang mit ihnen leichter ist (GRANDIN 1998). Daher kann die Allelfrequenz der Variante ansteigen.

Hoch frequente Varianten können auf einen positiven Einfluss hindeuten und entsprechen somit der „less is more“-Hypothese (OLSON 1999; MACARTHUR und TYLER-SMITH 2010). Bei einer positiven Selektion besteht die Möglichkeit einer Fixierung der Allelfrequenz in der Population (XUE *et al.* 2006; MACARTHUR *et al.* 2007). Ein vorzeitiges Stopp-Codon in einem Gen kann die Funktion eines Proteins beeinträchtigen ohne Nachteile für ein spezifisches Merkmal zu haben. Das Ausschalten der Proteinfunktion kann sich darüber hinaus positiv auf die Gesundheit des Organismus auswirken. Homozygote Träger einer solchen Mutation haben Vorteile gegenüber anderen Individuen. Beim Menschen sind einige dieser Mutationen bekannt, wie z.B. das Gen *CASP12*, das bei einem Knockout eine Resistenz gegenüber Sepsis hervorruft (XUE *et al.* 2006). Mutationen mit einem ähnlichen Effekt kommen auch im Rindengenom in Betracht. In einem augenscheinlich gesundem Tier unter den 214 re-sequenzierten Individuen sind sieben Stoppmutationen im homozygoten Zustand vorzufinden. Es ist nicht auszuschließen, dass mindestens eine Variante vorteilhaft ist und dem positiven Selektionsdruck ausgesetzt ist. Die Stoppmutation im Gen *VTCNI* stellt trotz ihrer niedrigen Allelfrequenz eine möglicherweise positiv auswirkende Kandidatenmutation dar. Bei Mäusen wirkt sich ein ausgeschaltetes *VTCNI* in Form einer Resistenz gegenüber Infektionen vorteilhaft aus (ZHU *et al.* 2009). Beim Rind könnte das vorzeitige Stopp-Codon ähnliche positive Effekte bewirken.

Einige Stoppmutationen haben nur einen geringen Effekt auf einen Phänotyp, wie möglicherweise das vorzeitige Stopp-Codon im Gen *QRFPR*. Die Allelfrequenz des mutierten Allels liegt im mittleren bis hohen Bereich. Individuen, die Träger solcher Mutationen sind, weisen keine Veränderung in einem Merkmal gegenüber Nicht-Trägertieren auf, trotz einer potentiell starken Schädigung der Funktion des betroffenen Proteins (YNGVADOTTIR *et al.* 2009). Dieses Phänomen

kann unterschiedlichste Gründe haben. Die Funktionen von Proteinen werden, falls sie einer Proteinfamilie angehören, gegebenenfalls von anderen Mitgliedern dieser Gruppe oder auch von anderen Genen mit ähnlicher Funktion übernommen (GU *et al.* 2003; MACARTHUR *et al.* 2012). TAS2R38 könnte entsprechend wie beim Menschen nicht den einzigen Rezeptor für bitteren Geschmack darstellen (CHANDRASHEKAR *et al.* 2000, 2006; BROCKHOFF *et al.* 2011). Zahlreiche weitere Rezeptoren befähigen einem Lebewesen zu einem bitteren Geschmacksempfinden. Das Ausschalten dieses Rezeptors muss somit nicht notwendigerweise den bitteren Geschmackssinn beim Rind beeinträchtigen. Möglich ist auch, dass es sich bei den betroffenen Genen bereits um Pseudogene handelt, die keine Funktion ausüben (YAMAGUCHI-KABATA *et al.* 2008). Außerdem können alternative Isoformen des betroffenen Gens, die nicht von der Stoppmutation betroffen sind, weiterhin die volle Funktion des Proteins ausüben (u.a. *VCAMI*) (TAKEDA *et al.* 2006; BALASUBRAMANIAN *et al.* 2011; MACARTHUR *et al.* 2012). Daraus resultierend kann die Allelfrequenz einzelner vorzeitiger Stopp-Codons in nicht essentiellen Genen bei Anwesenheit von homozygoten Tieren ansteigen und sich möglicherweise fixieren (YNGVADOTTIR *et al.* 2009; MACARTHUR und TYLER-SMITH 2010; MACARTHUR *et al.* 2012). Sind nur einzelne Transkripte eines Gens betroffen, kann die Stoppmutation jedoch auch gewebsspezifische Auswirkungen haben (FRASER und XIE 2009) und entsprechend ursächlich für Mendelsche Erbkrankheiten oder für komplexe Merkmale sein (UZUMCU *et al.* 2006).

Die identifizierten Varianten segregieren nicht immer in allen Rassen, wie z.B. die Stoppmutation in *VCAMI*, die ausschließlich bei BSW detektiert wurde, oder das in HOL und RED identifizierte Stopp-Codon in *FREMI*. Bei einzelnen Mutationen kommt es zwischen den Rassen zu Abweichungen der Allelfrequenz. Diese unterschiedlichen Allelfrequenzen könnten aufgrund rassespezifischer Zuchtziele entstanden sein (LIEN *et al.* 1999; QANBARI *et al.* 2014). Des Weiteren sind Tiere aus unterschiedlichen Rassen oder Populationen je nach Umweltbedingung und Zuchtziel verschiedenen Selektionsdrücken ausgesetzt, sodass Unterschiede in der Häufigkeit der Allele zu beobachten sind (MACARTHUR *et al.* 2007; BALASUBRAMANIAN *et al.* 2011). Die Variante im Gen für den bitteren Geschmacksrezeptor *TAS2R38* segregiert ausschließlich in Rassen, die überwiegend auf der Weide gehalten werden. Eben diese Populationen sind darauf angewiesen, bitter schmeckende und dementsprechend auch giftige Pflanzen meiden zu können (MORTON und CAMPBELL 1997).

5.2.1.1 Validierung der Auswirkungen der Stoppmutationen

Eine hohe Allelfrequenz allein ist kein guter Indikator für den vorteilhaften Effekt einer Stoppmutation. Die Frequenz eines Polymorphismus kann in ähnlicher Weise ansteigen, wenn dieser z.B. aufgrund der Pseudogenisierung kaum einen Effekt hat. Zur Unterscheidung zwischen Pseudogen und positivem Effekt müssen weitere Parameter herangezogen werden. Auch niedrige Allelfrequenzen können nicht als einziger Hinweis auf einen schädlichen Effekt dienen. Zur funktionellen Aufklärung der Stoppmutationen ist die Untersuchung betroffener Individuen, insbesondere derer mit ALT-homozygotem Genotyp, unerlässlich. Hilfreich ist vor allem ein phänotypischer Vergleich von Trägertieren mit Nicht-Trägertieren. Ergebnisse aus Knockout-Studien bei Mäusen oder bereits beschriebene Erkrankungen und Phänotypen beim Menschen (Tabelle 4.5) können Hinweise auf das mögliche Erscheinungsbild beim Rind geben.

Um die Kausalität der Stoppmutation in *STRC* für die Taubheit beim Rind zu analysieren, sind indirekte Indikatoren, wie z.B. leichter Umgang, geringe Reizbarkeit und Stressfreiheit, für die Überprüfung eines gestörten Hörvermögens nützlich. Das vorzeitige Stopp-Codon in *PNLIP* könnte sich auf Unterschiede in der Fleisch- und Milchqualität bezüglich der Marmorierung und des MilCHFettgehalts auswirken. Zur Validierung der Vermutung, dass der SNP in *TAS2R38* den bitteren Geschmackssinn beeinträchtigt, könnten Fütterungsversuche an Individuen mit unterschiedlichem Genotyp durchgeführt werden.

Da die untersuchten Varianten teilweise rassespezifisch segregieren, sollten die Genome zusätzlicher Tiere und insbesondere auch Phänotypdaten der Rassen analysiert werden, um den Katalog der Stoppmutationen beim Rind zu vervollständigen und bei der Aufklärung rassetypischer Unterschiede im Phänotyp Fortschritte zu erzielen. Informationen über Selektionssignaturen können außerdem bei der Detektion von rassecharakteristischen Varianten hilfreich sein.

Polymorphismen, die aufgrund einer Kombination an Indizien auf einen letalen Effekt hindeuten, wie z.B. geringe Allelfrequenz, fehlende Homozygote und eine letale Wirkung bei anderen Spezies, sollten populationsweit typisiert werden (VANRADEN *et al.* 2011b). Die Anzahl der zu analysierenden Tiere in einer Population hängt von der populationsspezifischen Allelfrequenz des mutierten Allels ab (Abbildung 5.2.1). Die Genotypisierung ist ein einfaches Mittel zur Überprüfung, ob die fehlenden Homozygoten aufgrund der geringen Allelfrequenz oder aufgrund eines letalen Effekts der Mutation nicht detektiert wurden.

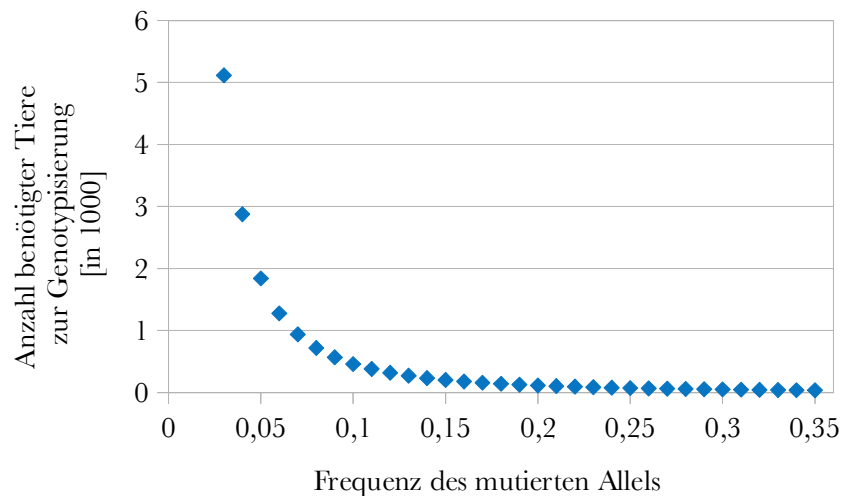


Abbildung 5.2.1: Benötigte Tieranzahl zur Überprüfung der Stoppmutationen

Zur Überprüfung eines letalen Effekts einer Stoppmutation müssen in Abhängigkeit von der populationsspezifischen Allelfrequenz der Variante eine bestimmte Anzahl an Individuen in der Population genotypisiert werden ($P=0,01$).

Zur Überprüfung, welche Transkripte in welcher Weise von den Mutationen betroffen sind, sollte in zukünftigen Studien eine gewebspezifische quantitative PCR für jede einzelne Mutation durchgeführt werden. Die Inklusion von RNA-Sequenzierungsdaten aus verschiedenen Geweben mehrerer Tiere ermöglicht eine Abschätzung des Effekts auf die Genexpression, der von den Stoppmutationen ausgeht (MACARTHUR *et al.* 2012). Insbesondere der von den Stoppmutationen potentiell induzierte Abbau der mRNA kann durch das Fehlen der mRNA einen zusätzlichen Hinweis für den NMD liefern (PISKOL *et al.* 2013).

Der vorliegende Katalog der tatsächlichen LoF-Varianten im Rindengenom wird nicht vollständig sein, da in dieser Analyse z.B. Frameshift-Mutationen und Polymorphismen an Spleißstellen nicht berücksichtigt wurden (MACARTHUR *et al.* 2012). Auch diese Varianten können zu einem vorzeitigen Abbruch der Translation führen und einen Verlust oder eine Veränderung der Proteinfunktion verursachen. Dennoch stellt die Liste mit hoch qualitativen vorzeitigen Stopp-Codons eine umfassende Einschätzung zu den Stoppmutationen im Rindengenom dar, auch wenn zahlreiche potentiell schädliche Varianten durch angewandte Filter ausgeschlossen wurden.

Zusätzlich zu den sieben ausführlich diskutierten Stoppmutationen existieren noch weitere Gene, deren Inaktivierung phänotypische Konsequenzen beim Rind haben könnte (Tabelle 4.5).

Darüber hinaus liegen Daten von vorzeitigen Stopp-Codons vor, bei denen bisher noch keine Aussage über einen möglichen Effekt getroffen werden konnte. In zukünftigen Studien sollten diese Varianten weiter untersucht werden, um deren funktionelle Auswirkungen und eine mögliche Kausalität für Erbkrankheiten zu klären.

5.2.2 Kausalität der identifizierten QTL für das Merkmal direkte Totgeburten bei Kühen

Für das Merkmal direkte Totgeburten bei Kühen wurde eine genomweite Assoziationsstudie auf Basis von 9013 HOL- und RED-Tieren an ~17 Mio. genomweiten polymorphen Positionen durchgeführt. Dabei wurden sieben QTL auf fünf verschiedenen bovinen Autosomen identifiziert. Für vier der identifizierten QTL sind Mutationen in den analysierten Kandidatengenomen als kausale Varianten für das Merkmal direkte Totgeburten bei Kühen sehr wahrscheinlich.

Das für den Fibroblasten-Wachstumsfaktor kodierende Gen *FGF23*, das im ersten QTL auf Chromosom 5 lokalisiert ist, reguliert den Phosphat- und Vitamin-D-Gehalt im Blut (SHIMADA *et al.* 2004b). Ein funktionell geschädigter Wachstumsfaktor könnte den Phosphat- und Vitamin-D-Metabolismus erheblich beeinträchtigen und sich auf die Totgeburtenrate auswirken. Vitamin-D nimmt eine Schlüsselrolle im Kalziummetabolismus und somit auch bei der Entwicklung und beim Aufbau der Knochensubstanz ein (ADHR CONSORTIUM 2000; BENET-PAGÈS *et al.* 2005). Daher kann sich eine Vitamin-D-Defizienz in einer kongenitalen Rachitis, in einer Knochenerweichung des Schädels sowie in einer Osteopenie des Neugeborenen manifestieren (ORBAK *et al.* 2007). Ein Vitamin-D-Mangel während der Schwangerschaft steht beim Menschen im Verdacht die Präeklampsie zu verursachen. Eine erhöhte maternale und neonatale Sterblichkeit sind die Folge dieser hypertensiven Erkrankung (MACKAY *et al.* 2001; SEELY 2007). Außerdem beeinflusst die Vitamin-D-Versorgung der Mutter die Geburtsgröße ihres Nachkommens (SABOUR *et al.* 2006), die einen erheblichen Einfluss auf die Totgeburtenrate hat. Es gibt Hinweise, dass eine Vitamin-D-Defizienz für Geburtsfehler verantwortlich ist, die aufgrund einer gestörten Zelldifferenzierung und -reifung im Herzen und Gehirn entstehen (BROWN *et al.* 2003; NIBBELINK *et al.* 2007). KENDRICK *et al.* (2011) konnten den gleichen, direkten Einfluss von FGF23-Konzentrationen im Blut auf kardiovaskuläre Erkrankungen und eine erhöhte Sterblichkeit feststellen. Der in dieser GWAS signifikanteste Polymorphismus, der vor dem Gen *FGF23* lokalisiert ist, könnte dessen Regulation steuern. Möglicherweise beeinträchtigt auch der nicht-synonyme Austausch in dem hoch konservierten Bereich dieses Gens (p.S180G) die Funktionalität des Wachstumsfaktors.

Der auf dem siebten Chromosom positionierte QTL ist durch einen nicht-synonymen Austausch im Kandidatengen *ARRDC3* charakterisiert (p.Y182C). PATWARI *et al.* (2011) zeigten, dass das kodierte Protein als Mitglied der Arrestinfamilie die Fettleibigkeit bei Mensch und Maus beeinflusst. Ein nicht-funktionelles *ARRDC3* beeinflusst nicht nur das Körpergewicht, sondern erhöht auch die perinatale Mortalität (PATWARI *et al.* 2011; SHEA *et al.* 2012). Der Aminosäureaustausch in *ARRDC3* ist ein Polymorphismus, der kausal für die Totgeburtenrate sein könnte, sowohl direkt als auch indirekt aufgrund des potentiell höheren Geburtsgewichtes.

Ein signifikanter Polymorphismus auf Chromosom 18 bewirkt einen nicht-synonymen Austausch im Gen *VPS35*. GRIFFIN *et al.* (2005) konnten zeigen, dass ein gestörter Retromerkomplex, dessen Herzstück von *VPS35* gebildet wird, den FrühTod von Embryonen hervorruft. Die Expression dieses Gens findet hauptsächlich in Osteoklasten und Osteoblasten statt (XIA *et al.* 2013). Das kodierte vakuoläre Protein *VPS35* ist vermutlich für die Bildung und Formung von Knochen, insbesondere bei der neonatalen Entwicklung, zuständig. Trotz des vorhergesagten harmlosen Effekts des Aminosäureaustausches (p.S760N) könnte die Proteinfunktion und vor allem die Bindung der C-terminalen Region von *VPS35* an den Kofaktor *VPS26* beeinträchtigt sein und die Knochenbildung beeinflussen.

Auf Chromosom 29 konnte im ~2kb entfernten Bereich vom Gen *H19* ein QTL für das Merkmal direkte Totgeburten bei Kühen identifiziert werden. Das resultierende Transkript, das nicht in ein Protein translatiert wird, hat insbesondere durch die Regulation der Expression von *IGF2* während der Embryonalentwicklung einen bedeutenden Einfluss auf das Wachstum des Organismus (LEIGHTON *et al.* 1995). Beim Menschen werden verschiedene Mutationen in *H19* mit Wachstumsstörungen assoziiert (MARTÍNEZ NOGUEIRAS *et al.* 2001; WEKSBERG *et al.* 2009). Aufgrund der hohen Korrelation der Geburtsgröße mit einem schweren Geburtsverlauf und Totgeburten (JOHANSON *et al.* 2011) repräsentiert die signifikante Variante eine potentiell kausale Mutation für das Merkmal Totgeburten. Möglicherweise wirkt sie aufgrund ihrer Position zu *H19* regulatorisch auf die Transkription. BRUNKOW und TILGHMAN (1991) konnten zeigen, dass eine Veränderung der Regulation von *H19* in Form einer Hochregulierung sich in einer pränatalen Letalität manifestierte.

Für die verbleibenden QTL auf den Chromosomen 5, 17 und 18 konnte kein eindeutiges Kandidatengen lokalisiert werden. Diese Regionen sollten in weiteren Studien genauer analysiert werden, insbesondere der QTL auf Chromosom 18 (59,0 bis 60,1 Mb). Dieser bringt ein sehr klares

Signal hervor, der aus dem assoziierten Polymorphismus mit dem niedrigsten P-Wert der durchgeführten GWAS resultiert.

5.2.3 Vergleich der identifizierten QTL mit anderen Studien

Zahlreiche Studien der Vergangenheit haben sich mit der genetischen Aufklärung und den Beziehungen der unterschiedlichen Kalbmerkmale untereinander, wie dem Kalbeverlauf, der Totgeburten und dem Geburtsgewicht, beschäftigt.

Die in dieser Arbeit identifizierten QTL auf Chromosom 5 sowie auf Chromosom 17 sind bisher in keiner Studie im direkten Zusammenhang mit Totgeburten identifiziert worden. Auf Chromosom 5 konnten andere Studien jedoch Regionen identifizieren, die mit dem direkten als auch maternalen Kalbeverlauf assoziiert sind. SCHROOTEN *et al.* (2000) beschreiben in der niederländischen Holsteinpopulation im Bereich von 110 Mb einen QTL für Körpergrößen-abhängige Merkmale, die den Geburtsverlauf und demnach auch die Totgeburten beeinflussen könnten. Im Bereich von 106 Mb konnten SAATCHI *et al.* (2014) eine kausale Region für das Geburtsgewicht bei Hereford identifizieren. Sie vermuteten einen nicht-synonymen Aminosäureaustausch (p.V94A) im von *FGF23* benachbarten Gen *TIGAR* als kausale Variante. Das Gen wurde bereits mit dem Glukosemetabolismus und mit der Apoptose in Verbindung gebracht (BENSAAD *et al.* 2006; KIMATA *et al.* 2010). Dieser postulierte Polymorphismus konnte im vorliegenden Datensatz nicht bestätigt werden. Es ist denkbar, dass SAATCHI *et al.* (2014) aufgrund der geringen Markerdichte die Variante in *FGF23* nicht identifizieren konnten und ein Signal an einer anderen Position erhielten. Bei den Rassen Angus und Brangus brachte die Analyse von SAATCHI *et al.* (2014) außerdem eine signifikante Region bei ~50 Mb für die Merkmale Geburtsgewicht und Kalbeverlauf hervor, ähnlich wie in dieser Studie für Totgeburten. McCLURE *et al.* (2010) berichteten über einen vergleichbaren QTL für den Kalbeverlauf. Neben dem QTL auf Chromosom 5 wurden auf Chromosom 17 Regionen für den Kalbeverlauf beim Gelbvieh (~55 Mb) und bei Limousin (~66 Mb) ermittelt (McCLURE *et al.* 2010; SAATCHI *et al.* 2014). Jedoch wurden für das Merkmal Totgeburten bisher noch keine QTL auf diesem bovinen Autosom berichtet.

In Übereinstimmung mit dieser Studie erhielten KÜHN *et al.* (2003) und THOMASEN *et al.* (2008) bei ihrer Analyse für das Merkmal direkte Totgeburten bei HOL-Tieren Hinweise für einen QTL in der gleichen Region auf Chromosom 7. Andere Studien assoziierten diesen Locus außerdem mit Geburtsverlauf und Geburtsgewicht (McCLURE *et al.* 2010; SAHANA *et al.* 2011; SAATCHI *et al.* 2014). SAATCHI *et al.* (2014) postulierten *ARRDC3* bereits als Kandidatengen, was die Ergebnisse dieser

Analyse bekräftigt. Auf Basis der vorliegenden Daten konnte ein Aminosäureaustausch im Gen *ARRDC3* vermutlich als kausale Variante identifiziert werden, der bisher noch nicht beschrieben wurde. Dieses Ergebnis verdeutlicht die Effizienz der genomweiten Re-Sequenzierung.

Auf dem 18. Chromosom konnten zwei Regionen identifiziert werden, die mit den direkten Totgeburten bei Kühen assoziiert sind. Darunter befand sich der aus dieser GWAS hervorgehende signifikanteste QTL am distalen Ende dieses Autosoms. Auch zahlreiche Studien konnten die Umgebung dieser signifikanten Region mit dem Merkmal Totgeburten sowie mit weiteren Kalbmerkmalen assoziieren (u.a. KÜHN *et al.* 2003; SEIDENSPINNER *et al.* 2009; BRAND *et al.* 2010; SAHANA *et al.* 2011). Möglicherweise weichen die Bereiche aufgrund der nicht durchgeführten Feinkartierung und der geringeren Markerdichte in vorangegangenen Studien von dem QTL ab. Die Ergebnisse gelten entsprechend auch für den weniger signifikanten Bereich zu Beginn des 18. Chromosoms (SCHULMAN *et al.* 2008; SAHANA *et al.* 2011), sowie für den QTL auf Chromosom 29, der neben dem Merkmal Totgeburten auch mit dem Kalbeverlauf und der Größe des Kalbes assoziiert wurde (MAGEE *et al.* 2010; SAHANA *et al.* 2011; SAATCHI *et al.* 2014).

Die meisten QTL mit großem Effekt, die bereits in anderen Studien beschrieben wurden, konnten in dieser Arbeit identifiziert werden. Mit Hilfe der Genomsequenzinformationen gelang die genauere Eingrenzung sowie die verbesserte Feinkartierung der kausalen Bereiche, teilweise bis hin zur potentiell kausalen Mutation. Neben den in dieser GWAS detektierten Loci für Totgeburten existieren aus den Ergebnissen anderer Studien weitere QTL für das untersuchte Merkmal (u.a. SCHULMAN *et al.* 2008; PAUSCH *et al.* 2011; SAHANA *et al.* 2011). Der von PAUSCH *et al.* (2011) auf Chromosom 14 lokalisierte QTL (~24 Mb) erreichte in der vorliegenden Analyse weder das genomweite, noch das chromosomenweite Signifikanzniveau. Es sind jedoch leicht niedrigere P-Werte in diesem Bereich zu erkennen (Abbildung 5.2.2), was auf einen potentiellen QTL hindeuten könnten. Eine nachfolgende Untersuchung mit einer größeren Tieranzahl könnte zur Aufklärung dieser Region beitragen. Denkbar ist auch, dass der von PAUSCH *et al.* (2011) erkannte QTL eine für die BFV populationspezifische Variante darstellt und nicht in HOL segregiert. Ähnliches gilt für die von SCHULMAN *et al.* (2008) erfassten Regionen bei Finnish Ayrshire Tieren auf Chromosom 4 und 15.

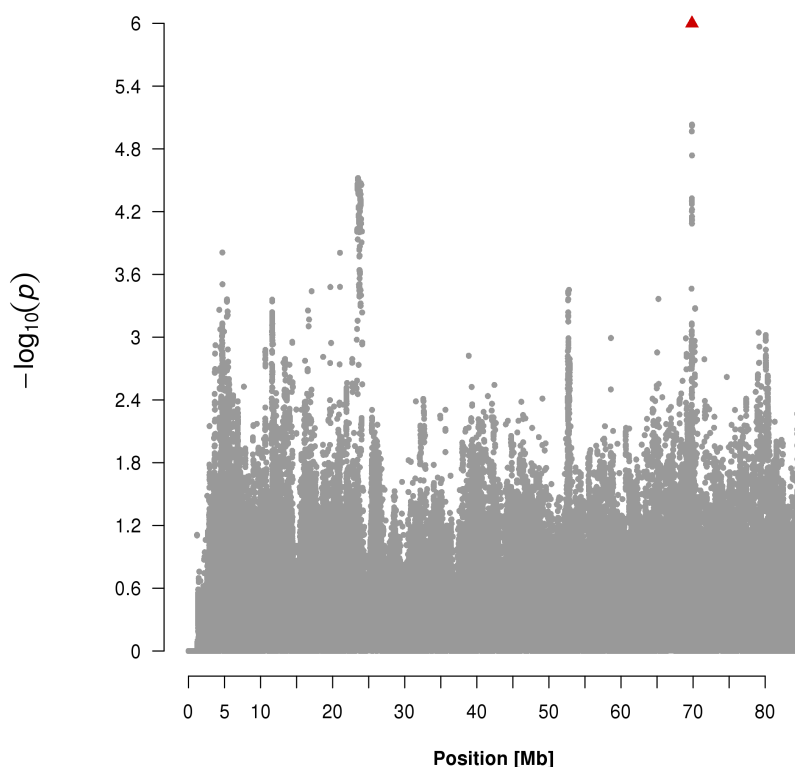


Abbildung 5.2.2: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen auf Chromosom 14

Die Positionen der einzelnen Varianten, auf der x-Achse dargestellt, basieren auf dem UMD3.1-Assembly (ZIMM *et al.* 2009). Auf der y-Achse ist der negativ dekadische Logarithmus der P-Werte aus der GWAS aufgetragen. Jeder SNP ist mit einem grauen Punkt dargestellt. Mit einem roten Dreieck ist die Variante mit dem niedrigsten P-Wert des Chromosoms visualisiert.

Viele der Bereiche, die in dieser Arbeit mit Totgeburten assoziiert sind, weisen in anderen Studien Assoziationen zu weiteren, mit der Kalbung in Verbindung stehenden Merkmalen auf. Begründet wird dies durch die hohen genetischen Korrelationen zwischen Kalbeverlaufsmerkmalen und der Totgeburt (STEINBOCK *et al.* 2003). Insbesondere die Körpergröße des Kalbes als direktes Merkmal scheint eine Rolle zu spielen. Drei der assoziierten Kandidatengene (*ARRDC3*, *FGF23* und *H19*) scheinen die Größe des Kalbes zu beeinflussen. Einerseits kann die Größe des Kalbes positiv beeinflusst werden. Jedoch können zu große Kälber eine Schweregeburt nach sich ziehen und somit die Wahrscheinlichkeit für eine Totgeburt erhöhen. Andererseits können Mutationen in den drei Genen das Gegenteil bewirken und kleinwüchsige Kälber entstehen lassen. Eine Unterentwicklung dieser Kälber ist denkbar. Möglicherweise verenden diese Tiere bereits im Mutterleib oder sind nach der Geburt nicht überlebensfähig. Symptome einer schweren

Kalbung müssen dabei nicht notwendigerweise vorliegen. Unabhängig von der Körpergröße des Kalbes und vom Kalbeverlauf können jedoch auch andere Ursachen vorliegen (BERGLUND *et al.* 2003). Pleiotrope Effekte von Mutationen, die beide Merkmale unabhängig voneinander beeinflussen, können dafür zuständig sein (SAATCHI *et al.* 2014).

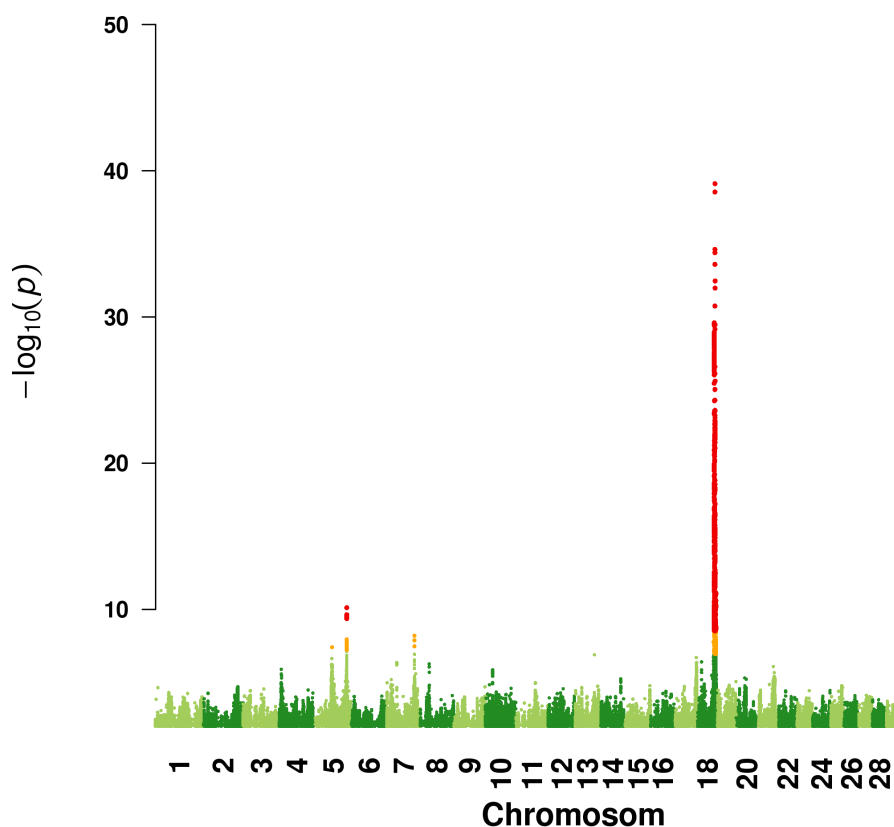


Abbildung 5.2.3: Manhattan-Plot der Ergebnisse der sequenzbasierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Färsen

Die Positionen der einzelnen Varianten, auf der x-Achse dargestellt, basieren auf dem UMD3.1-Assembly (ZIMLN *et al.* 2009), mit Ausnahme der Geschlechtschromosomen. Auf der y-Achse ist der negativ dekadische Logarithmus der P-Werte der GWAS aufgetragen. Chromosomenweit signifikante Polymorphismen sind gelb markiert, genomweit signifikante Varianten sind in rot dargestellt.

Die Anzahl an signifikanten QTL für das Merkmal direkte Totgeburten zwischen Erstlaktierenden (direct Stillbirth heifer, dSBh) und Kühen in späterer Laktation (dSBc) ist in der vorliegenden Arbeit ähnlich (Abbildung 4.3.1 und Abbildung 5.2.3). Die hohen genetischen Korrelationen der mit der Kalbung assoziierten Merkmale zwischen Färsen und Kühen in Höhe von 0,4 bis 0,7

(LUO *et al.* 2002; STEINBOCK *et al.* 2003) unterstreichen dieses Resultat. In der Literatur wird zwischen Färsen und Kühen bezüglich der Totgeburten ein enormer Unterschied beschrieben. Andere Studien beschreiben mehr QTL für Färsen, da die mit dem Kalbeverlauf assoziierten Merkmale, wie auch Totgeburten, bei erstlaktierenden Kühen eine höhere Heritabilität aufweisen als bei Kühen in höheren Laktationen (STEINBOCK *et al.* 2003; ERIKSSON *et al.* 2004). Die einzelnen kausalen Regionen bei Färsen könnten im Vergleich zu den wenigen QTL bei Kühen jedoch einen kleineren Effekt auf das Merkmal ausüben, was das geringere Signifikanzniveau in der vorliegenden Assoziationsstudie bei Färsen rechtfertigen würde. In ihrer Summe erklären die vielen QTL bei Färsen jedoch einen größeren Anteil der phänotypischen Varianz als bei Kühen, sodass bei Erstlaktierenden eine höhere Heritabilität erreicht wird. Des Weiteren sind vermutlich aufgrund des niedrigeren Effekts je QTL einzelne, in anderen Studien identifizierte Regionen in dieser Analyse nicht entdeckt worden. Diese Ergebnisse widersprechen den Resultaten von SAHANA *et al.* (2011). Ihre Analyse brachte bei Kühen weniger QTL hervor als bei Färsen, was sie mit der geringeren Heritabilität in späteren Laktationen begründen.

Um den genauen Effekt und die Anzahl der kausalen Regionen der Laktation auf die Totgeburten bzw. auf die Kalbmerkmale zu analysieren, sollten genomweite Assoziationsstudien mit Phänotypen zu jeder einzelnen Laktation auf Basis der gleichen Tieranzahl durchgeführt werden.

5.3 Herausforderungen bei der Identifizierung kausaler Varianten

5.3.1 Dichte der Markerkarte

Die Anzahl an Polymorphismen und deren Verteilung über das Genom bestimmen die von den vorhandenen Markern erklärte Heritabilität. Beim Milchvieh können je nach Merkmal 5 bis 88% der genetischen Varianz durch den 54K SNP-Chip erklärt werden (JENSEN *et al.* 2012; HAILE-MARIAM *et al.* 2013). Die benötigte Dichte der Markerkarten hängt erheblich vom Ausmaß des LD ab, bedingt durch die N_e einer Population. Populationen mit kleiner N_e , wie z.B. Holstein (KHATKAR *et al.* 2008; QANBARI *et al.* 2010), benötigen im Gegensatz zu anderen Populationen, wie Bayrisch Fleckvieh, weniger Polymorphismen, um die unabhängigen Chromosomensegmente abgreifen zu können (DAETWYLER *et al.* 2010). Die auf den SNP-Chips vorhandenen Polymorphismen wurden aufgrund ihrer mittleren Allelfrequenz ausgewählt, sodass seltene kausale Varianten, mit möglicherweise geringem Kopplungsungleichgewicht zu den Chip-basierten Markern nicht

in Erwägung gezogen werden. Die kausalen Regionen können zwar bereits mit geringeren Markerdichten lokalisiert werden (Abbildung 9.3), jedoch sind bis zur Identifizierung der kausalen Mutation noch viele aufwendige Schritte notwendig. Die Verfügbarkeit der gesamten genomischen Variation und die Kenntnis über deren Wirkung auf das Merkmal ist die Voraussetzung, um die genetische Varianz vollständig zu erfassen. Die in dieser Studie durchgeführte Analyse mit der genomischen Variation aus genomweiter Re-Sequenzierung benötigt nicht mehr das Kopplungsungleichgewicht zwischen Marker und kausaler Variante, da die ursächliche Mutation direkt im Datensatz enthalten ist. Man kann so direkt zur kausalen Variante gelangen. Die Polymorphismen-basierte Heritabilität wird die nahezu identischen Werte für die Erbllichkeit des Merkmals erreichen können. Die sequenzbasierten Varianten aus der genomweiten Assoziationsstudie für direkte Totgeburten bei Kühen erklären laut Analyseprogrammen etwa 75% der Zuchtwert-Varianz. Es scheinen aber noch weitere kausale Mutationen im Genom des Rindes vorhanden zu sein, die in dieser Analyse nicht eingeschlossen sind und die verbleibende genetische Varianz auffüllen würden. Vermutlich sind bereits während der Imputation sowie beim Filtern sehr seltener Varianten ($MAF < 0,005$) kausale Polymorphismen entfernt worden.

Das Erzielen einer höheren Markerdichte erhöht die Chance, eine echte Assoziation zu erhalten. Trotz dieser kosteneffizienten Methode zur Vorhersage der gesamten genomischen Variation werden Polymorphismen mit einer geringen Allelfrequenz von dieser Methodik benachteiligt. In dieser Studie lagen einige Stoppmutation bei der Überprüfung im imputierten Datensatz monomorph für das Referenzallel vor. Sie sind nicht mehr informativ für weitere Analysen. Zur erfolgreichen Vorhersage dieser seltenen Mutationen sollten im Optimalfall alle Tiere genomweit re-sequenziert oder alternativ die Tieranzahl im Referenzdatensatz der Imputation erhöht werden, um somit die Allelfrequenz ansteigen zu lassen (LI *et al.* 2011; DRUET *et al.* 2014). Außerdem können Polymorphismen mit geringer Allelfrequenz, die eigentlich kausal für ein Merkmal sind, „synthetische“ Assoziationen entstehen lassen (DICKSON *et al.* 2010). Die zahlreichen ursächlichen, seltenen Varianten in einer Region können das Top-Signal auf eine häufigere Variante übertragen und eine falsch-positive Assoziation in einem anderen, nahegelegenen Bereich entstehen lassen. Möglicherweise entstehen so Top-Signale im regulatorischen Bereich, obwohl die kausale Mutationen im kodierenden Bereich lokalisiert ist. Dennoch sollten niedrig frequente, regulatorische Varianten als potentielle kausale Mutationen für komplexe Merkmale nicht ausgeschlossen und der Fokus nicht nur auf Mutationen im kodierenden Bereich gelegt werden (CIRULLI und GOLDSTEIN 2010; KOUFARIOTIS *et al.* 2014). Sowohl Polymorphismen im UTR als auch im Intron

scheinen Phänotypen in Säugern zu beeinflussen (SPILIANAKIS *et al.* 2005; WU *et al.* 2012). Diese Varianten beeinflussen die Expression eines Gens, die als kausale Variante jedoch schwieriger zu identifizieren sind (MEYER *et al.* 2008), insbesondere da die Problematik in der Festlegung des regulatorischen Bereichs liegt (LI und MONTGOMERY 2013). Für die Überprüfung der Kausalität regulatorischer Varianten, wie z.B. der Mutation kurz vor dem Gen *FGF23*, sind Expressionsstudien erforderlich, die im Rahmen dieser Studie nicht durchgeführt werden konnten.

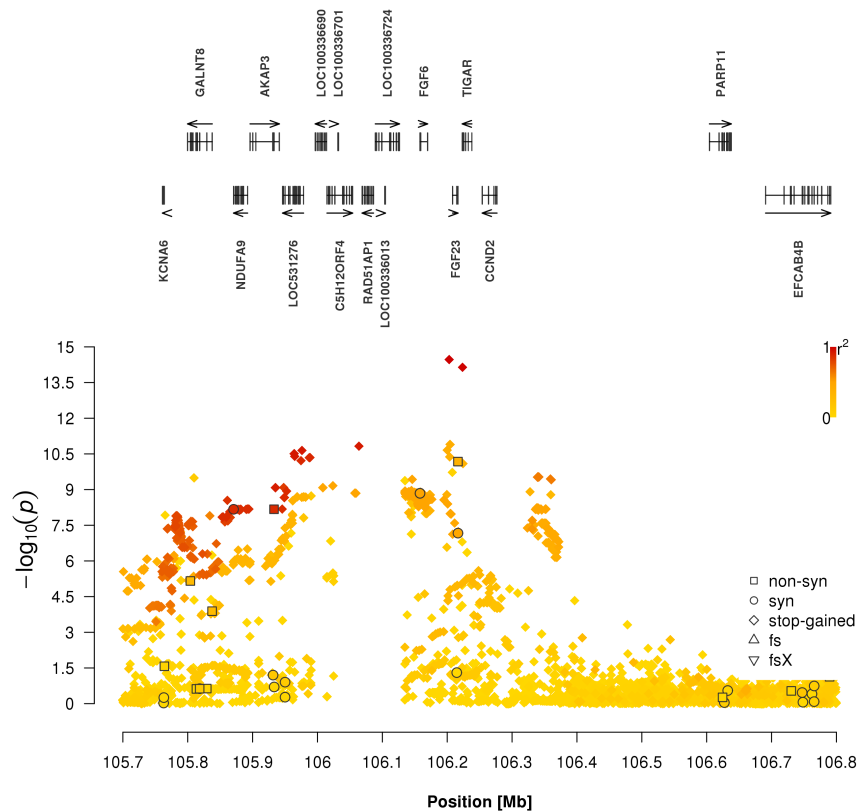


Abbildung 5.3.1: Region-Plot auf Chromosom 5 für den direkten Kalbeverlauf bei Kühen

Die physikalische Position der Varianten basiert auf dem UMD3.1-Assembly (ZIMLIN *et al.* 2009). Die Varianten sind mit einer Raute dargestellt. Die Farbe der Raute entspricht dem jeweiligem LD zum Top-Marker. Positionen, die im kodierenden Bereich eines Gens annotiert wurden, sind mit verschiedenen Symbolen markiert.

Der in dieser Arbeit für Totgeburten lokalisierte QTN auf Chromosom 5 liegt in einer Region, in der SAATCHI *et al.* (2014) einen SNP als kausal für den Kalbeverlauf postuliert haben. Während der für Totgeburten potentiell kausale Polymorphismus im Gen *FGF23* liegt, beschreiben SAATCHI *et al.* (2014) das benachbarte Gen *TIGAR* als Kandidatengen. Als Vermutung liegt nahe, dass die Feinkartierung des Merkmals Kalbeverlauf aufgrund der geringen Markerdichte nicht umsetzbar

war. Die sequenzbasierte GWAS für Totgeburten und Kalbeverlauf liefern den identischen Top-Marker (Abbildung 5.3.1). Die hohen genetischen Korrelationen zwischen Kalbeverlauf und Totgeburten (STEINBOCK *et al.* 2003) lassen auf die Kausalität des gleichen Polymorphismus mit pleiotroper Wirkung schließen, was durch die sequenzbasierte GWAS für den Kalbeverlauf erhärtet wird.

Genomsequenz-basierte Studien haben einen Vorteil gegenüber SNP-Chip basierten Analysen, da die kausale Mutation in den Daten vorliegen sollte. Für einige Regionen, die ein extrem hohes Kopplungsungleichgewicht aufweisen, bringt die Verfügbarkeit der gesamten genomischen Variation jedoch keinen zusätzlichen Nutzen für die Detektion der ursächlichen Variante. Alle Varianten dieses Haplotypen kommen als kausale Mutation in Frage, wie beim QTL auf Chr17 ersichtlich (UDLER *et al.* 2010; FAYE *et al.* 2013; DRUET *et al.* 2014). Um zwischen kausalen und gekoppelten Polymorphismen zu unterscheiden sind Informationen über den Konservierungsgrad oder die Transkriptionsfaktorbindestellen hilfreich. Eine populationsweite Genotypisierung der Kandidaten-SNP ist dennoch unerlässlich.

5.3.2 Qualität der Phänotypinformationen

Um kausale Varianten erfolgreich kartieren zu können, insbesondere für Merkmale mit geringer Heritabilität wie die Totgeburtenrate (STEINBOCK *et al.* 2003; SEIDENSPINNER *et al.* 2009), ist es wichtig, möglichst zuverlässige und genaue Phänotypdaten in der genomweiten Assoziationsstudie anzuwenden (PAUSCH *et al.* 2011). Zuchtwerte aus Nachkommenprüfung eignen sich aufgrund der wiederholten Leistungen gut für die Aufklärung komplexer, niedrig heritabler Merkmale (GEORGES *et al.* 1995; PAUSCH *et al.* 2011), jedoch enthalten die Zuchtwerte nur Informationen über additiv-genetische Effekte (FALCONER und MACKAY 1996). Wie diese Studie andeuteten, können vermutlich auch Imprinting-Effekte wie auf Chromosom 29 auf Basis von Zuchtwerten erfasst werden, jedoch gestaltet sich das Abschätzen der Effektgröße schwierig. Die durch Epistasie und Dominanz erklärte genetische Varianz bleibt außerdem unbeachtet, sodass Zuchtwertinformationen nicht zur vollständigen Aufklärung der genetischen Architektur komplexer Merkmale beitragen können.

Des Weiteren ist für die detaillierte Analyse der Ursache für Totgeburten eine genauere Definition der Phänotypen notwendig. Nur so kann unterschieden werden, ob der letale Effekt durch einen schweren Geburtsverlauf bedingt ist, oder ob das Kalb bereits pränatal in der Gebärmutter verstorben ist (BERGLUND *et al.* 2003). Auch um den Einfluss der Laktationsanzahl auf die

Totgeburtenrate bzw. die unterschiedlichen kausalen Varianten in den einzelnen Laktationen bestimmen zu können, sollten die Phänotypdaten nicht nur für die Färsen und Kühen aus allen Laktationen unterteilt, sondern nach der jeweiligen Laktationsanzahl differenziert werden.

Um schließlich die Kausalität der potentiellen ursächlichen Polymorphismen zu überprüfen und zu bestätigen, ist die Implementierung direkter Phänotypen, wie z.B. die Expression des Gens bzw. die Menge eines bestimmten Proteins, notwendig. Nur so kann neben der Verwendung von Vorhersagetoools die Auswirkung einer Mutation bestimmt werden.

5.3.3 Effekt der Polymorphismen

Drei der sieben identifizierten QTL können als maßgeblich für das Merkmal Totgeburtenrate bestimmt werden, da sie einen Großteil der additiv-genetischen Varianz erklären ($> 7\%$). Diese genetische Architektur ist charakteristisch für komplexe Merkmale. Laut GODDARD und HAYES (2009b) existieren im Genom neben zahlreichen Loci mit kleinem Effekt auf ein komplexes Merkmal auch vereinzelte Genorte mit großem Effekt. Dennoch liegt wenig Evidenz für Allele mit einem so immensen Einfluss vor, wie es bei dem QTL auf Chromosom 5 und 29 der Fall ist (MAHER 2008). Der Top-Marker des identifizierten QTL auf Chromosom 29 erklärt mehr als 80% der additiv-genetischen Varianz des Merkmals direkte Totgeburten bei Holsteinkühen (Tabelle 4.7). Möglicherweise ist die Schätzung dieses hohen Anteils der erklärten Varianz aufgrund eines Stichprobenfehlers entstanden (HAYES *et al.* 2010). In der Analyse von MACKAY *et al.* (2012) erklärten von mehr als einer Million vorgegebener SNP je nach Merkmal nur wenige mehr als 50%. Eine Kreuzvalidierung unter Verwendung aller SNP stellte schließlich nur 6% fest (OBER *et al.* 2012).

Bei komplexen Merkmalen liegen wenige QTL vor, die einen so großen Anteil an der Varianz haben (MAHER 2008), wobei *DGAT1* ein Beleg für solche Loci darstellt (WINTER *et al.* 2002; THALLER *et al.* 2003). Der Effekt einiger Allele kann sehr hoch sein, jedoch ist die Frequenz dieser Varianten meist sehr niedrig, sodass sie nur einen kleinen Beitrag zur Varianz leisten und vermutlich nicht in GWAS entdeckt werden können (HAYES *et al.* 2010). Nach EYRE-WALKER (2010) könnte ein großer Effekt eines Allels vorliegen, falls die Allele aus einer natürlichen Selektion hervorgegangen sind. HAYES *et al.* (2010) erklären sich den großen Anteil eines QTL an der Varianz außerdem durch das Auffangen des Effekts eines anderen, nahegelegenen QTL. Ausgeprägte LD-Blöcke beim Rind aufgrund einer geringen N_e bedeuten mehr Varianz je wirkendem Locus (HAYES *et al.* 2010), was auch auf den QTL auf Chromosom 29 zutreffen könnte. Eine Erhöhung

der Stichprobengröße kann die Genauigkeit der Schätzung des Effekts erhöhen, jedoch kann dieser selbst bei gering heritablen Merkmalen wahrscheinlich keine Verbesserung herbeiführen (W_{RAY} *et al.* 2013).

Schließlich können zusätzliche Mechanismen, wie Genotyp-Umwelt-Interaktionen (B_{HATTACHARJEE} *et al.* 2010) und Epistasie, von Bedeutung sein und zur Überschätzung des QTL auf Chromosom 29 führen (M_{ACKAY} 2001; C_{ORVA} und M_{EDRANO} 2001). Insbesondere auch die Imprinting-Effekte (M_{OTT} *et al.* 2014), wie bei *IGF2* und *H19*, dürfen bei dieser Region nicht vernachlässigt werden. Unter Einbeziehung von Methylierungsdaten könnten die Effekte der Polymorphismen in dieser Region abgeschätzt werden. Um den tatsächlichen Effekt der Variante zu schätzen, sollten diese in der Population und in zusätzlichen Tieren validiert werden.

5.3.4 Qualität der Imputation

Da die populationsweite Re-Sequenzierung noch mit hohen Kosten verbunden ist, wird die genomische Variation von Schlüsseltieren auf deren Nachkommen mit Hilfe niedrig- und mitteldichter Markerarten hochgerechnet (J_{ANSEN} *et al.* 2013; D_{AE}TWYLER *et al.* 2014). Die Imputation von mitteldichten Genotypen auf hochdichte, sequenzbasierte Genotypen (L_I *et al.* 2009b; P_{AUSCH} *et al.* 2013; V_{AN BINSBERGEN} *et al.* 2014) ist ein effizientes und kostengünstiges Verfahren, um die populationsweite Sequenzinformation zu erhalten. Jedoch bringt die imputierte Sequenzinformation keinen zusätzlichen Nutzen, falls die Qualität der Imputation ungenügend ist. Die Imputationsgenauigkeit kann von zahlreichen Faktoren erhöht werden, z.B. durch den Anstieg der Markerdichte bei geringer werdendem Abstand zwischen den Markern, durch die Zunahme des Kopplungsungleichgewichts, durch eine Erhöhung der MAF oder durch eine erhöhte Verwandtschaftsstruktur zwischen den Tieren im Referenz- und im imputierten Datensatz (u.a. I_{WATA} und J_{ANNINK} 2010; D_{RUET} *et al.* 2010; V_{ANRADEN} *et al.* 2011a; H_{AYES} *et al.* 2012; H_{ICKEY} *et al.* 2012). Insbesondere die Imputationsgenauigkeit von SNP mit geringer Allelfrequenz werden von diesen Parametern beeinflusst (L_I *et al.* 2011).

In dieser Studie wurden innerhalb der HOL-Population zunächst die niedrigdichten auf die hochdichten Genotypen imputiert und anschließend auf Basis von 122 HOL- und RED-Genomsequenzen des 1000-Bullen-Genom-Projekts die Sequenzinformation vorhergesagt. Für die Sequenzimputation ist die stufenweise Imputation vom 50K über den 777K auf die Genomsequenz die zuverlässigste Methode, vor allem wenn zwischen den Referenz- und den imputierten Tieren eine hoher Verwandtschaftsgrad besteht (V_{AN BINSBERGEN} *et al.* 2014). Die 122 genomweit

sequenzierten Tiere werden als Schlüsseltiere der HOL-Population beschrieben und weisen eine geringe Verwandtschaft zueinander auf, um auf diese Weise die verschiedenen Chromosomen-segmente abzugreifen. Die Verwendung dieser Tiere als Referenzpopulation würde eine hohe Imputationsgenauigkeit bei der Ableitung der Genotypen ihrer Nachkommen oder ihrer nahen Verwandten vermuten lassen. Jedoch erklären die 122 HOL- und RED-Tiere je nach landesspezifischer HOL-Populationen einen unterschiedlichen Anteil der genetischen Varianz in den einzelnen Populationen (Australien: ~72,3%; Frankreich: ~43,3%; Dänemark: ~55,3%) (DAETWYLER *et al.* 2014).

Trotz der hohen Ähnlichkeit innerhalb der internationalen HOL-Populationen (Abbildung 5.3.2), die auf einen engen Verwandtschaftsgrad bei niedriger N_e zurückzuführen ist, werden die einzelnen Populationen von verschiedenen Gründertieren beeinflusst. Die 122 HOL- und RED-Genomsequenzen sind für die in dieser Analyse durchgeführten Sequenzimputation nicht optimal ausgewählt, da sie vermutlich in der deutschen HOL- und RED-Population einen vergleichbar niedrigen Anteil der genetischen Diversität wie in Frankreich und Dänemark repräsentieren.

Des Weiteren gestaltet sich die Imputation seltener Varianten schwierig (ASIMIT und ZEGGINI 2012). LI *et al.* (2011) zeigten, dass für eine erfolgreiche Imputation von Polymorphismen mit einer MAF kleiner als 0,005 eine Referenzpopulation von mehr als 1200 Individuen benötigt wird, was in dieser Analyse durch die Verwendung der Tiere aus dem 1000-Bullen-Genom-Projekt nicht erfüllt werden konnte. Da die 122 HOL- und RED-Tiere Schlüsseltiere darstellen, werden seltene Varianten höchstwahrscheinlich nicht erfasst. Auch wenn auf Basis der 54K- und 777K-Daten Polymorphismen mit einer geringen Allelfrequenz erscheinen, können diese nicht imputiert werden, wenn sie nicht innerhalb der 122 Genomsequenzen segregieren.

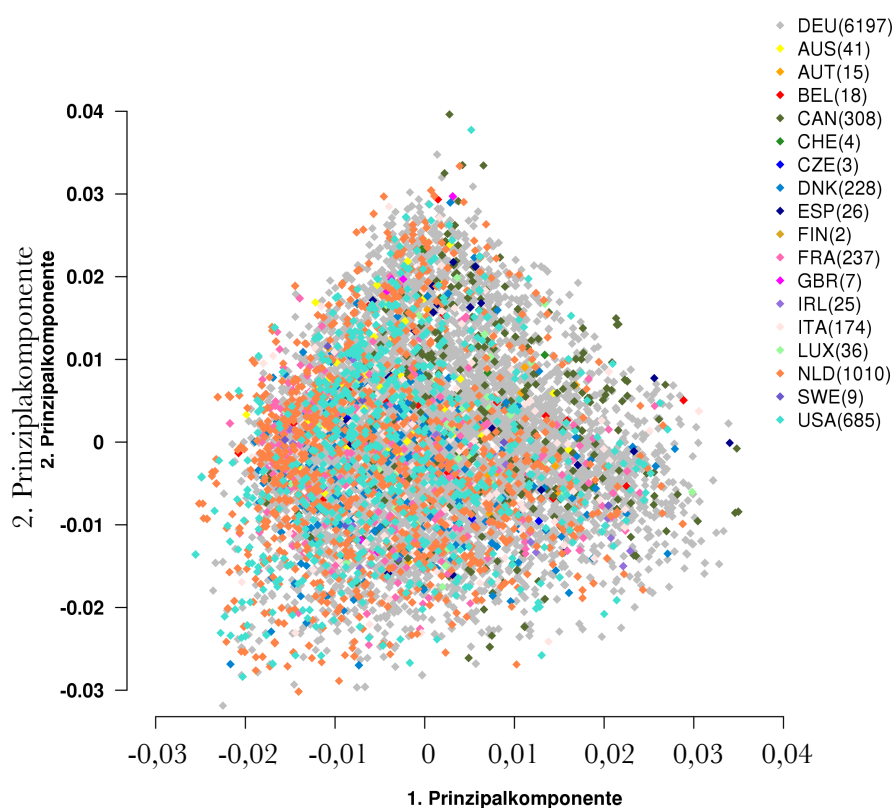


Abbildung 5.3.2: Ergebnisse der Prinzipalkomponentenanalyse innerhalb der Holsteintiere

Für die Prinzipalkomponentenanalyse wurde die genomische Verwandtschaftsmatrix mit PLINK v1.9 (PURCELL *et al.* 2007) aufgestellt. Anschließend wurden die Prinzipalkomponenten mit GCTA (version 1.24) (YANG *et al.* 2011) geschätzt. Die Ziffer in den Klammern hinter den Länderkürzeln gibt die Anzahl der analysierten Tiere je Land an. DEU = Deutschland; AUS = Australien; AUT = Österreich; BEL = Belgien; CAN = Kanada; CHE = Schweiz; CZE = Tschechien; DNK = Dänemark; ESP = Spanien; FIN = Finnland; FRA = Frankreich; GBR = Großbritannien; IRL = Irland; ITA = Italien; LUX = Luxemburg; NLD = Niederlande; SWE = Schweden; USA = Vereinigte Staaten von Amerika

Die Zuverlässigkeit der imputierten Varianten, insbesondere der seltenen Mutationen, ist scheinbar beeinträchtigt. Die Verwendung anderer Imputationsprogramme, wie z.B. MaCH (LI *et al.* 2010b) oder IMPUTE (MARCHINI *et al.* 2007), könnten zu einer höheren Genauigkeit führen, vor allem bei der Verfügbarkeit einer kleinen Referenzpopulation (BROWNING 2008; NOTHNAGEL *et al.* 2009). Allerdings ist der Rechenaufwand bei diesen beiden Programmen im Vergleich zu Beagle höher, was sich nachteilig auf die Analyse größerer Datensätze auswirkt (NOTHNAGEL *et al.* 2009; PAUSCH *et al.* 2013). Eine Erhöhung der Referenzpopulation für die Imputation der mitteldichten sowie für die Sequenzimputation könnte sich außerdem vorteilhaft auswirken (PAUSCH *et al.* 2013), sodass auch Polymorphismen mit niedriger MAF möglichst genau abgeleitet werden können.

Insbesondere diese Mutationen machen in ihrer Summe einen Großteil der Varianz bei komplexen Merkmalen aus (GODDARD und HAYES 2009b).

5.3.5 Falsch-positive Varianten

Durch multiples Testen (HOGGART *et al.* 2008) können zahlreiche falsch-positive Ergebnisse in GWAS entstehen. Um diese Fehler zu minimieren und zuverlässigere Resultate zu erhalten, wurde bei der durchgeführten GWAS die Signifikanzgrenze nach Bonferroni korrigiert.

Auch aufgrund der engen Verwandtschaftsstrukturen in Nutztierpopulationen (NOTTER 1999), wie auch bei HOL, können in GWAS falsch-positive Assoziationen auftreten (MARCHINI *et al.* 2004; ROSENBERG und NORDBORG 2006). Zur Berücksichtigung dieser Verwandtschaftseffekte wurde die Studie mittels Varianzkomponenten-basiertem Ansatz auf Populationsstratifikation korrigiert. Die Verwandtschaftsbeziehungen wurden mit Hilfe einer Verwandtschaftsmatrix nach VANRADEN *et al.* (2011a) in die Analyse eingeschlossen.

Trotz der stetigen Weiterentwicklung der NGS-Technologien sind diese nicht frei von falsch-positiven Sequenzinformationen (DURBIN *et al.* 2010; DEPRISTO *et al.* 2011; KIEZUN *et al.* 2012). Zur Abschätzung der Sensitivität und Genauigkeit der Genotypen wurden in dieser Analyse die NRD- und NRS-Rate sowie die Übereinstimmung der Genotypen nach DEPRISTO *et al.* (2011) verwendet. Jedoch liegen noch keine geeigneten Methoden vor, falsch-positive Varianten zu filtern, ohne dabei echte Polymorphismen zu eliminieren. Daher wurden in dieser Studie alle relevanten Polymorphismen mittels *Sanger*-Re-Sequenzierung überprüft, um die bei NGS-Technologien auftretenden Sequenzierartefakte (RUFFALO *et al.* 2011) auszuschließen. Insbesondere bei der Analyse der Stoppmutationen war diese Anwendung sinnvoll. Andernfalls wäre die Anzahl der vorzeitigen Stopp-Codons extrem überschätzt und möglicherweise falsch interpretiert worden.

5.3.6 Qualität des UMD3.1-Assemblys

Bei der Verifizierung der vorzeitigen Stopp-Codons sind 65 Varianten aufgrund fehlerhafter Annotation, sowie 15 weitere Mutationen wegen Mapping-Fehlern aus nachfolgenden Analysen ausgeschlossen worden. Diese Ergebnisse spiegeln die unzureichende Qualität der Referenzsequenz und ihrer Genannotation wieder.

Trotz der Verbesserung des Rindergenoms auf UMD3.1 (ZIMIN *et al.* 2009) weist das Assembly noch Lücken und Fehler auf (BOHMANOVA *et al.* 2010; PARTIPILO *et al.* 2011). Einzelne

Polymorphismen werden an falsche Positionen kartiert. Vermutlich entsprechen die Positionen der Marker aus einer problematischen Region nicht ihrer physikalischen Position, was zu falsch-positiven Ergebnissen führen wird. Trotz der Verwendung von paired-end Reads, vorwiegend mit einer Länge von 100 bp je Read, die im Vergleich zu single-end Reads vorteilhafter beim Alignment sind (WANG *et al.* 2010), können Sequenzen falschen Bereichen zugeordnet werden. Hinweise dafür ergeben sich aus einer unterdurchschnittlichen Abdeckung der Region sowie niedrige Genotypqualitäten der identifizierten Varianten in diesem Bereich. Die verschiedenen Alignmentprogramme bieten unterschiedliche Parameter zur Erhöhung der Genauigkeit und geben mehrere Werte für die Qualität des Alignments an (LI und DURBIN 2009; FRITH *et al.* 2010). Das Filtern von gering qualitativen Sequenzfragmenten könnte jedoch neben einem Anstieg der Genauigkeit auch zu einem Verlust von richtig kartierten Reads führen und den Informationsgehalt mindern. Daher wurden in dieser Analyse nur PCR-Duplikate markiert, um mit hoher Sicherheit eine Reduktion inkorrekt Information zu bewirken.

Ein weiterer Nachteil des Rinder-Assemblys ist die Genannotation. Bisher konzentrierte sich die Genannotation vor allem auf Protein-kodierende Gene, wobei insbesondere die Annotation von Pseudogenen ignoriert wurde (BALASUBRAMANIAN *et al.* 2011). Ein Problem liegt vor, falls laut Genannotation ein Transkript als Pseudogen beschrieben wird. Identifizierte Stoppmutationen in diesen Genen werden vermutlich für weitere Analysen nicht weiter berücksichtigt. Als Beispiel ist das Gen *ACTN3* zu nennen, das in der Ensembl Datenbank als polymorphes Pseudogen bezeichnet wird (ENSG00000248746.1). Tatsächlich segregiert im humanen Genom eine Stoppmutation in diesem Gen, die durch einen vollständigen Proteinverlust zu einer verbesserten Muskelfunktion führt (MACARTHUR *et al.* 2007). Außerdem wäre eine Beachtung der unterschiedlichen Isoformen im Assembly sinnvoll. Alternatives Spleißen von Genen kann dazu führen, dass ein detektierter Polymorphismus nicht den kodierenden Bereich aller Transkripte betrifft (Kapitel 5.2.1). Die Funktion des Proteins kann durch andere, nicht betroffene Isoformen weiterhin ausgeführt werden. Diese Mutationen scheiden als kausale Kandidaten aus. In dieser Untersuchung wurde die Genstruktur jedes Kandidatengens manuell reannotiert, um falsch-positive Mutationen aufzudecken und aus nachfolgenden Analysen auszuschließen. Aufgrund der Genomprojekte beim Rind (u.a. ECK *et al.* 2009; STOTHARD *et al.* 2011; JANSEN *et al.* 2013; DAETWYLER *et al.* 2014) ist in naher Zukunft von einer Erneuerung und Verbesserung des Assemblys auszugehen. Die Implementierung der Ergebnisse aus Projekten der Transkriptom-Sequenzierung im Assembly wird zudem höchstwahrscheinlich eine Verbesserung der Annotation hervorrufen.

5.4 Anwendung in genomischer Selektion

Die genomisch geschätzten Zuchtwerte zur frühestmöglichen Selektion der besten Tiere basieren zur Zeit auf 54K-SNP Chip-Daten. Die Schätzung der Effekte je SNP erfolgt mit Hilfe der Informationen des LD zwischen kausaler Mutation und Marker (HABIER *et al.* 2007). Bei Populationen mit geringer N_e nimmt die Genauigkeit der Zuchtwertschätzung aufgrund von Rekombinationsereignissen, die die LD-Blöcke aufbrechen, über Generationen hinweg ab. Mit einer genomischen Zuchtwertschätzung auf Basis von Genomsequenzen kann dieses Problem bewältigt werden, da die kausale Variante eingeschlossen ist und das Kopplungsungleichgewicht nicht mehr als Informationsquelle benötigt wird (MEUWISSEN und GODDARD 2010). Bei Populationen, in denen sich die ursächliche Variante mit anderen Markern im vollständigem Kopplungsungleichgewicht liegt, kann sich jedoch der große Effekt der kausalen Mutation in viele kleine Effekte auf die gekoppelten Varianten verteilen (ERBE *et al.* 2012).

Falls die kausale Variante für ein Merkmal bekannt ist, kann diese Information direkt in die Zuchtprogramme implementiert werden. Die Größe der Effekte der zur ursächlichen Mutation gekoppelten Marker wird nicht mehr benötigt. Ein weiterer Vorteil ist die Anwendung des Wissens über eine kausale Mutation in anderen Rassen, falls diese auch dort segregiert. Dementsprechend wäre eine Zuchtwertschätzung über Rassen hinweg möglich, was bisher nicht umzusetzen war (DAETWYLER *et al.* 2012; ERBE *et al.* 2012; OLSON *et al.* 2012; HAYES *et al.* 2013).

Die Genauigkeit der genomischen Zuchtwertschätzung hängt davon ab, wie groß der Anteil der genetischen Varianz ist, der durch die SNP erklärt werden kann (JENSEN *et al.* 2012; HAILE-MARIAM *et al.* 2013). Ein Großteil der genetischen Varianz komplexer Merkmale wird durch zahlreiche seltene Varianten erklärt, die nicht auf SNP-Chips implementiert sind. Mit Hilfe der Mutationen mit geringer Allelfrequenz, die aus der Genomsequenz resultieren, kann eine höhere Genauigkeit genomisch geschätzten Zuchtwerte erreicht werden (HAYES *et al.* 2013; DRUET *et al.* 2014).

Darüber hinaus kann in den Rinderpopulationen genauer auf bestimmte Merkmale selektiert und ein höherer Zuchterfolg erreicht werden, wenn die kausale Variante bekannt ist. Bisher wurden womöglich zwei QTN unterschiedlicher Merkmale aufgrund des hohen Kopplungsungleichgewichts zueinander durch einen SNP auf dem SNP-Chip abgefragt. THOMASEN *et al.* (2008) identifizierten auf BTA7 einen QTL, der vermutlich nur aufgrund von Kopplung das Merkmal Totgeburten und auch die Größe des Kalbes beeinflusst. So würde eine Selektion auf eine reduzierte Totgeburtenrate direkt auch ein Verringerung der Größe des Kalbes bedeuten. Der Effekt des

QTL auf die jeweiligen Merkmale war womöglich fehlgeschätzt, falls nicht der gleiche QTN mit pleiotroper Wirkung zu Grunde liegt. Bei negativ korrelierten Merkmalen dieser beiden QTN würde bei der Selektion auf das eine Merkmal das andere automatisch in Mitleidenschaft gezogen. Sind die exakten Positionen und Effekte der QTN bekannt, ist es möglich, nur rekombinante Haplotypen mit dem jeweiligen gewünschten positiven Effekt zu selektieren. Somit kann die Totgeburtenrate gesenkt werden ohne andere Merkmale negativ zu beeinflussen. Jedoch können sich die QTN auch pleiotrop auswirken, sodass mehrere Phänotypen durch eine Variante beeinflusst werden (SAATCHI *et al.* 2014).

6 Zusammenfassung

Die Aufklärung der genetischen Architektur komplexer Merkmale und spezieller Phänotypen kann einen enormen Beitrag zur Steigerung des Zuchtfortschritts und zur Verhinderung der Ausbreitung von Erbfehlern in Rinderpopulationen leisten. Die mit NGS-Technologien erfassten Informationen über die vollständige genomische Variation ermöglichen eine Beschleunigung des Prozesses der Identifizierung kausaler Varianten.

In dieser Studie wurden auf Basis von Genomsequenzen mit mittlerer Abdeckung (~7-fach) von 214 Individuen sechs verschiedener Rinderrassen etwa 23 Mio. variable Positionen in Form von SNP und InDels identifiziert. Beim Vergleich der sequenzbasierten Genotypen mit den Genotypen, die aus der Typisierung mit dem IlluminaHD BeachChip resultierten, konnte über alle Tiere hinweg eine hohe Sensitivität von mindestens 99,63% und eine große Übereinstimmung der Genotypen von mindestens 97,21% unabhängig von der jeweiligen Genomabdeckung des Tieres verzeichnet werden. Die funktionelle Annotation der Varianten deckte 66.532 nicht-synonyme Basenaustausche sowie 1038 Mutationen, die zu einer vorzeitigen Beendigung der Gentranslation führen, auf. Bei der Überprüfung nach bereits bekannten kausalen Mutationen beim Rind konnten elf OMIA-Varianten und fünf für quantitative Merkmale kausale Nukleotide (QTN) identifiziert werden, für die mindestens eines der 214 Tiere das mutierte Allel trägt.

Die Analyse der vorzeitigen Stopp-Codons verifizierte 100 hoch qualitative Varianten, von denen jedes Tier etwa zwölf Polymorphismen trägt. Ein Großteil dieser Varianten erscheint bei einer geringen Allelfrequenz, was auf ihren schädlichen Effekt hindeutet, insbesondere wenn keine homozygoten Tiere entdeckt werden konnten (z.B. bei *FREMI* und *VCAMI*). Diese Positionen sollten populationsweit genotypisiert werden, um zu klären, ob der Polymorphismus einen letalen Effekt hat oder ob die homozygoten Tiere aufgrund der geringen Allelfrequenz bisher unentdeckt blieben. Neben diesen für rezessiv vererbte Krankheiten potentiell kausalen Mutationen scheinen im Genom des Rindes Stoppmutationen zu existieren, bei denen das betroffene Protein tolerant gegenüber eine durch die Stoppmutation verursachte Veränderung oder einen Verlust der Proteinfunktion steht (vgl. *QRFPR*). Möglicherweise wirkt sich die Stoppmutation sogar vorteilhaft aus, wie z.B. beim Gen *VTCN1*. Mit Hilfe von Informationen aus Knockout-Studien bei Mäusen oder Genanalysen beim Menschen kann ein potentieller Phänotyp der vorzeitigen Stopp-Codons beim Rind vorhergesagt werden. Die Identifizierung von Trägertieren solcher

Stoppmutationen und deren systematische Phänotypisierung, insbesondere von homozygoten Individuen, kann einen wertvollen Einblick in die Funktionalität vieler Gene bringen und zur Aufklärung genetischer Defekte beitragen.

Bei Holstein konnten mit einer sequenzbasierten GWAS für das Merkmal Totgeburten insgesamt sieben QTL detektiert werden. Als positionelle und funktionelle Kandidatengene wurden *FGF23*, *ARRDC3*, *VPS35*, *LOC787812* und *H19* identifiziert. Für die QTL auf Chromosom 5 und 17 konnte aufgrund der genreichen Region kein eindeutiges Gen bestimmt werden. Die Funktionalität der Gene wurde in der Literatur bereits direkt mit dem Merkmal Totgeburten oder auch mit den hoch korrelierten Kalbeverlaufsmerkmalen in Verbindung gebracht. Auf Chromosom 5 konnte ein potentieller QTN im regulatorischen Bereich des für einen Wachstumsfaktor kodierenden Gens *FGF23* festgestellt werden (P-Wert = $5,184 \times 10^{-13}$). Der QTL auf Chromosom 7 liefert einen nicht-synonymen Austausch (p.Y182C) im *ARRDC3*-Gen als signifikantesten Polymorphismus (P-Wert = $4,327 \times 10^{-9}$). Ein weiterer Aminosäureaustausch im Gen *VPS35* (p.S760N) ist möglicherweise ursächlich für den QTL auf Chromosom 18 (P-Wert = $9,849 \times 10^{-9}$).

Die GWAS basierend auf Genomsequenzen ermöglicht die direkte Assoziation der kausalen Varianten. Es besteht keine Notwendigkeit mehr ein bestehendes Kopplungsungleichgewicht zwischen Marker und kausaler Variante zu haben. Die in dieser Arbeit identifizierten, potentiell ursächlichen Polymorphismen für das Merkmal Totgeburten können umgehend in die genomische Selektion der Rinderzuchtprogramme implementiert werden. Im Übrigen profitieren auch andere, insbesondere kleine Populationen von dieser Information, bei denen die genomische Selektion aufgrund der kleinen Referenzpopulation nicht durchgeführt werden kann. Mit der Berücksichtigung dieser QTN in der genomischen Selektion kann eine Reduzierung der Anzahl totgeborener Kälber bei Rindern erzielt werden.

7 Summary

The elucidation of the genetic architecture of complex traits and of special phenotypes can considerably contribute to an increase of breeding success and may prevent the expansion of genetic defects. The availability of the total genomic variation based on NGS-technology enables the rapid identification of causal variants.

In this study nearly 23 million variable positions were identified, in terms of SNPs and InDels based on whole genome sequence data (7-fold average coverage) of 214 individuals originating from six different breeds. The comparison of sequence-based genotypes with data resulting from genotyping with IlluminaHD BeadChip among all analysed animals revealed a high sensitivity of at least 99.63% and a great genotype concordance not less than 97.21%, independent of the average genome coverage of an individual. The functional annotation of the variants yielded 66,532 non-synonymous exchanges and 1038 mutations predicted to cause a premature stop codon. The examination for known causal mutations in the bovine genome detected eleven OMIA-variants and five nucleotides underlying a quantitative trait (QTN). At least one of the 214 re-sequenced animals carries the mutated allele for these known causal mutations.

The analysis of premature stop codons resulted in 100 high quality variants. Every animal carries on average twelve of these variants. Stop mutations are particularly overrepresented among rare alleles indicating their deleterious effect, especially if homozygous animals are absent, e.g. *FREMI* and *VCAMI*. Population-wide genotyping of the positions with missing homozygotes is required for elucidating whether the mutations are deleterious or homozygous animals are unidentified based on rare alleles. Along with these mutations representing potentially causal variants for recessive inherited traits, the bovine genome comprises additional stop mutations affecting proteins which are putatively tolerant to the modification or loss of protein function such as *QRFPR*. Stop mutations may even have an advantageous effect (e.g. *VTCNI*). Considering information from knockout studies in mice or gene analyses in humans facilitates the prediction of a feasible phenotype caused by the premature stop codon in cattle. The identification and detailed inspection of animals carrying such mutations, particularly homozygous individuals, may give an insight to the functionality of affected genes and will contribute to the clarification of genetic defects.

The GWAS in Holstein revealed seven QTL associated with stillbirth in total. The postulated positional and functional candidate genes were *FGF23*, *ARRDC3*, *VPS35*, *LOC787812* and *H19*. Due to the gene density, no gene could be explicitly determined for the QTL on chromosome 5 and 17. The functionality of the postulated candidate genes has already been directly associated with stillbirth or with the highly correlated calving traits. On chromosome 5 a putative QTN was identified in the regulatory region of *FGF23* gene, encoding a growth factor (P value = 5.184×10^{-13}). The most significant polymorphism of the QTL in chromosome 7 was a non-synonymous mutation (p.Y182C) in gene *ARRDC3* (P value = 4.327×10^{-9}). An additional amino acid exchange in the gene *VPS35* (p.S760N) represents a putative causal mutation for the QTL on chromosome 18 (P value = 9.849×10^{-9}).

The GWAS based on whole genome sequence data allows for the direct association of causal trait variants. The LD between marker and causal mutation is not required anymore. The putative causal polymorphism for stillbirth identified in this study can be implemented in genomic selection of breeding programmes in cattle. Particularly small cattle populations which are not able to perform reliable genomic selection due to their small reference population size can benefit from this information. The use and consideration of the identified QTN in genomic selection programmes could reduce the amount of stillborn calves in cattle.

8 Literaturverzeichnis

- ABYZOV A., URBAN A. E., SNYDER M., GERSTEIN M., 2011 CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**: 974–984.
- ADAMS H. A., 2012 Identification of a Nonsense Mutation in APAF1 that is Causal for a Decrease in Reproductive Efficiency in Dairy Cattle. In: *Plant and Animal Genome*.
- ADHR CONSORTIUM, 2000 Autosomal dominant hypophosphataemic rickets is associated with mutations in FGF23. *Nat. Genet.* **26**: 345–348.
- ADLER E., HOON M. A., MUELLER K. L., CHANDRASHEKAR J., RYBA N. J., ZUKER C. S., 2000 A novel family of mammalian taste receptors. *Cell* **100**: 693–702.
- ADRIAENSSENS E., LOTTIN S., DUGIMONT T., FAUQUETTE W., COLL J., DUPOUY J. P., BOILLY B., CURGY J. J., 1999 Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene* **18**: 4460–4473.
- ADZHUBEI I. A., SCHMIDT S., PESHKIN L., RAMENSKY V. E., GERASIMOVA A., BORK P., KONDRASHOV A. S., SUNYAEV S. R., 2010 A method and server for predicting damaging missense mutations. *Nat. Methods* **7**: 248–249.
- AJROUD-DRISS S., FECTO F., AJROUD K., SIDDIQUE T., 2012 Mutations in the Nuclear Encoded Novel Mitochondrial Protein CHCHD10 Cause an Autosomal Dominant Mitochondrial Myopathy (IN7-2.002). *Neurology* **78**: IN7-2.002–IN7-2.002.
- ALAZAMI A. M., SHAHEEN R., ALZAHIRANI F., SNAPE K., SAGGAR A., BRINKMANN B., BAVI P., AL-GAZALI L. I., ALKURAYA F. S., 2009 FREM1 mutations cause bifid nose, renal agenesis, and anorectal malformations syndrome. *Am. J. Hum. Genet.* **85**: 414–418.
- ALBERS H. M., REINARDS T. H. C. M., BRINKMAN D. M. C., KAMPHUIS S. S. M., ROSSUM M. A. J. van, HOPPENREIJS E. P. H., GIRSCHICK H. J., WOUTERS C., SAURENMANN R. K., BAKKER E., VERDUIJN W., SLAGBOOM P., HUIZINGA T. W. J., TOES R. E. M., HOUWING-DUISTERMAAT J. J., CATE R. ten, SCHILHAM M. W., 2014 Genetic variation in VTCN1 (B7-H4) is associated with course of disease in juvenile idiopathic arthritis. *Ann. Rheum. Dis.* **73**: 1198–1201.
- ALSALEM A. B., HALEES A. S., ANAZI S., ALSHAMEKH S., ALKURAYA F. S., 2013 Autozygome sequencing expands the horizon of human knockout research and provides novel insights into human phenotypic variation. *PLoS Genet.* **9**: e1004030.
- ALTSCHUL S. F., GISH W., MILLER W., MYERS E. W., LIPMAN D. J., 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- ARIMA T., MATSUDA T., TAKAGI N., WAKE N., 1997 Association of IGF2 and H19 imprinting with choriocarcinoma development. *Cancer Genet. Cytogenet.* **93**: 39–47.
- ASIMIT J. L., ZEGGINI E., 2012 Imputation of rare variants in next-generation association studies. *Hum. Hered.* **74**: 196–204.
- AUBRY L., GUETTA D., KLEIN G., 2009 The Arrestin Fold: Variations on a Theme. *Curr. Genomics* **10**: 133–142.
- AVIDAN N., TAMARY H., DGANY O., CATTAN D., PARIENTE A., THULLIEZ M., BOROT N., MOATI L.,

- BARTHELME A., SHALMON L., KRASNOV T., BEN-ASHER E., OLENDER T., KHEN M., YANIV I., ZAIZOV R., SHALEV H., DELAUNAY J., FELLOUS M., LANCET D., BECKMANN J. S., 2003 CATSPER2, a human autosomal nonsyndromic male infertility gene. *Eur. J. Hum. Genet. EJHG* **11**: 497–502.
- AWAD M. M., DALAL D., CHO E., AMAT-ALARCON N., JAMES C., TICHNELL C., TUCKER A., RUSSELL S. D., BLUEMKE D. A., DIETZ H. C., CALKINS H., JUDGE D. P., 2006 DSG2 mutations contribute to arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Am. J. Hum. Genet.* **79**: 136–142.
- AZUMA Y.-T., MATSUO Y., KUWAMURA M., YANCOPOULOS G. D., VALENZUELA D. M., MURPHY A. J., NAKAJIMA H., KAROW M., TAKEUCHI T., 2010 Interleukin-19 protects mice from innate-mediated colonic inflammation. *Inflamm. Bowel Dis.* **16**: 1017–1028.
- BACHMANOV A. A., BEAUCHAMP G. K., 2007 Taste Receptor Genes. *Annu. Rev. Nutr.* **27**: 389–414.
- BALASUBRAMANIAN S., HABEGGER L., FRANKISH A., MACARTHUR D. G., HARTE R., TYLER-SMITH C., HARROW J., GERSTEIN M., 2011 Gene inactivation and its implications for annotation in the era of personal genomics. *Genes Dev.* **25**: 1–10.
- BANNWARTH S., AIT-EL-MKADEM S., CHAUSSENOT A., GENIN E. C., LACAS-GERVAIS S., FRAGAKI K., BERG-ALONSO L., KAGEYAMA Y., SERRE V., MOORE D. G., VERSCHUEREN A., ROUZIER C., BER I. L., AUGÉ G., COCHAUD C., LESPINASSE F., N’GUYEN K., SEPTENVILLE A. de, BRICE A., YU-WAI-MAN P., SESAKI H., POUGET J., PAQUIS-FLUCKLINGER V., 2014 A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain*: awu138.
- BARIBAULT H., DANA O., GUPTA J., YANG L., SUN B., RICHARDS W., TIAN H., 2006 The G-protein-coupled receptor GPR103 regulates bone formation. *Mol. Cell. Biol.* **26**: 709–717.
- BATEMAN A., SANDFORD R., 1999 The PLAT domain: a new piece in the PKD1 puzzle. *Curr. Biol.* **9**: R588–S2.
- BAUM M., SCHIAVI S., DWARAKANATH V., QUIGLEY R., 2005 Effect of fibroblast growth factor-23 on phosphate transport in proximal tubules. *Kidney Int.* **68**: 1148–1153.
- BECK L., KARAPLIS A. C., AMIZUKA N., HEWSON A. S., OZAWA H., TENENHOUSE H. S., 1998 Targeted inactivation of Npt2 in mice leads to severe renal phosphate wasting, hypercalciuria, and skeletal abnormalities. *Proc. Natl. Acad. Sci.* **95**: 5372–5377.
- BELMONT P. J., CHEN W. J., PEDRO M. N. S., THURAU D. J., LOWE N. G., GUDE N., HILTON B., WOLKOWICZ R., SUSSMAN M. A., GLEMBOTSKI C. C., 2010 Roles for Endoplasmic Reticulum–Associated Degradation and the Novel Endoplasmic Reticulum Stress Response Gene Derlin-3 in the Ischemic Heart. *Circ. Res.* **106**: 307–316.
- BENET-PAGÈS A., ORLIK P., STROM T. M., LORENZ-DEPIEREUX B., 2005 An FGF23 missense mutation causes familial tumoral calcinosis with hyperphosphatemia. *Hum. Mol. Genet.* **14**: 385–390.
- BENNEKUM A. M. VAN, FISHER E. A., BLANER W. S., HARRISON E. H., 2000 Hydrolysis of retinyl esters by pancreatic triglyceride lipase. *Biochemistry (Mosc.)* **39**: 4900–4906.
- BENSAAD K., TSURUTA A., SELAK M. A., VIDAL M. N. C., NAKANO K., BARTRONS R., GOTTLIEB E., VOUSDEN K. H., 2006 TIGAR, a p53-inducible regulator of glycolysis and apoptosis. *Cell* **126**: 107–120.
- BERGER P. J., CUBAS A. C., KOEHLER K. J., HEALEY M. H., 1992 Factors affecting dystocia and

- early calf mortality in Angus cows and heifers. *J. Anim. Sci.* **70**: 1775–1786.
- BERGLUND B., STEINBOCK L., ELVANDER M., 2003 Causes of Stillbirth and Time of Death in Swedish Holstein Calves Examined Post Mortem. *Acta Vet. Scand.* **44**: 111–120.
- BERRY S. D., DAVIS S. R., BEATTIE E. M., THOMAS N. L., BURRETT A. K., WARD H. E., STANFIELD A. M., BISWAS M., ANKERSMIT-UDY A. E., OXLEY P. E., BARNETT J. L., PEARSON J. F., DOES Y. VAN DER, MACGIBBON A. H. K., SPELMAN R. J., LEHNERT K., SNELL R. G., 2009 Mutation in bovine beta-carotene oxygenase 2 affects milk color. *Genetics* **182**: 923–926.
- BHATTACHARJEE S., WANG Z., CIAMPA J., KRAFT P., CHANOCK S., YU K., CHATTERJEE N., 2010 Using Principal Components of Genetic Variation for Robust and Powerful Detection of Gene-Gene Interactions in Case-Control and Case-Only Studies. *Am. J. Hum. Genet.* **86**: 331–342.
- BINSBERGEN R. VAN, BINK M. C., CALUS M. P., EEUWIJK F. A. VAN, HAYES B. J., HULSEGGI I., VEERKAMP R. F., 2014 Accuracy of imputation to whole-genome sequence data in Holstein Friesian cattle. *Genet. Sel. Evol. GSE* **46**: 41.
- BLAKE J. A., BULT C. J., EPPIG J. T., KADIN J. A., RICHARDSON J. E., MOUSE GENOME DATABASE GROUP, 2014 The Mouse Genome Database: integration of and access to knowledge about the laboratory mouse. *Nucleic Acids Res.* **42**: D810–817.
- BLOTT S., KIM J.-J., MOISIO S., SCHMIDT-KÜNTZEL A., CORNET A., BERZI P., CAMBISANO N., FORD C., GRISART B., JOHNSON D., KARIM L., SIMON P., SNELL R., SPELMAN R., WONG J., VILKKI J., GEORGES M., FARNIR F., COPPIETERS W., 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.
- BODMER W., BONILLA C., 2008 Common and rare variants in multifactorial susceptibility to common diseases. *Nat. Genet.* **40**: 695–701.
- BOHMANOVA J., SARGOLZAEI M., SCHENKEL F. S., 2010 Characteristics of linkage disequilibrium in North American Holsteins. *BMC Genomics* **11**: 421.
- BOICHARD D., CHUNG H., DASSONNEVILLE R., DAVID X., EGGEN A., FRITZ S., GIETZEN K. J., HAYES B. J., LAWLEY C. T., SONSTEGARD T. S., TASSELL C. P. VAN, VANRADEN P. M., VIAUD-MARTINEZ K. A., WIGGANS G. R., FOR THE BOVINE LD CONSORTIUM, 2012 Design of a Bovine Low-Density SNP Array Optimized for Imputation. *PLoS ONE* **7**: e34130.
- BOUDREAUX M. K., SCHMUTZ S. M., FRENCH P. S., 2007 Calcium diacylglycerol guanine nucleotide exchange factor I (CalDAG-GEFI) gene mutations in a thrombopathic Simmental calf. *Vet. Pathol.* **44**: 932–935.
- BOYKO A. R., WILLIAMSON S. H., INDAP A. R., DEGENHARDT J. D., HERNANDEZ R. D., LOHMUELLER K. E., ADAMS M. D., SCHMIDT S., SNINSKY J. J., SUNYAEV S. R., WHITE T. J., NIELSEN R., CLARK A. G., BUSTAMANTE C. D., 2008 Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* **4**: e1000083.
- BRAND B., BAES C., MAYER M., REINSCH N., SEIDENSPINNER T., THALLER G., KÜHN C., 2010 Quantitative trait loci mapping of calving and conformation traits on Bos taurus autosome 18 in the German Holstein population. *J. Dairy Sci.* **93**: 1205–1215.
- BRANNAN C. I., DEES E. C., INGRAM R. S., TILGHMAN S. M., 1990 The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* **10**: 28–36.

- BROCKHOFF A., BEHRENS M., ROUDNITZKY N., APPENDINO G., AVONTO C., MEYERHOF W., 2011 Receptor Agonism and Antagonism of Dietary Bitter Compounds. *J. Neurosci.* **31**: 14775–14782.
- BROWN J., BIANCO J. I., McGRATH J. J., EYLES D. W., 2003 1,25-dihydroxyvitamin D3 induces nerve growth factor, promotes neurite outgrowth and inhibits mitosis in embryonic rat hippocampal neurons. *Neurosci. Lett.* **343**: 139–143.
- BROWN J. P., BULLWINKEL J., BARON-LÜHR B., BILLUR M., SCHNEIDER P., WINKING H., SINGH P. B., 2010 HP1gamma function is required for male germ cell survival and spermatogenesis. *Epigenetics Chromatin* **3**: 9.
- BROWNING S. R., 2008 Missing data imputation and haplotype phase inference for genome-wide association studies. *Hum. Genet.* **124**: 439–450.
- 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.
- BRUNKOW M. E., TILGHMAN S. M., 1991 Ectopic expression of the H19 gene in mice causes prenatal lethality. *Genes Dev.* **5**: 1092–1101.
- BUITKAMP J., SEMMER J., GÖTZ K.-U., 2011 Arachnomelia syndrome in Simmental cattle is caused by a homozygous 2-bp deletion in the molybdenum cofactor synthesis step 1 gene (MOCS1). *BMC Genet.* **12**: 11.
- CALANDRA T., ROGER T., 2003 Macrophage migration inhibitory factor: a regulator of innate immunity. *Nat. Rev. Immunol.* **3**: 791–800.
- CAMBI A., FIGDOR C., 2009 Necrosis: C-type lectins sense cell death. *Curr. Biol. CB* **19**: R375–378.
- CARGILL M., ALTSHULER D., IRELAND J., SKLAR P., ARDLIE K., PATIL N., SHAW N., LANE C. R., LIM E. P., KALYANARAMAN N., NEMESH J., ZIAUGRA L., FRIEDLAND L., ROLFE A., WARRINGTON J., LIPSHUTZ R., DALEY G. Q., LANDER E. S., 1999 Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* **22**: 231–238.
- CARLOS T. M., SCHWARTZ B. R., KOVACH N. L., YEE E., ROSA M., OSBORN L., CHI-ROSSO G., NEWMAN B., LOBB R., ROSA M. R. M. to, 1990 Vascular cell adhesion molecule-1 mediates lymphocyte adherence to cytokine-activated cultured human endothelial cells [published erratum appears in *Blood* 1990 Dec 1;76(11):2420]. *Blood* **76**: 965–970.
- CARRENO B. M., COLLINS M., 2003 BTLA: a new inhibitory receptor with a B7-like ligand. *Trends Immunol.* **24**: 524–527.
- CHAHINIAN H., SIAS B., CARRIÈRE F., 2000 The C-terminal domain of pancreatic lipase: functional and structural analogies with c2 domains. *Curr. Protein Pept. Sci.* **1**: 91–103.
- CHAN J. R., BLUMENSCHEN W., MURPHY E., DIVEU C., WIEKOWSKI M., ABBONDANZO S., LUCIAN L., GEISLER R., BRODIE S., KIMBALL A. B., GORMAN D. M., SMITH K., MALEFYT R. de W., KASTELEIN R. A., McCLANAHAN T. K., BOWMAN E. P., 2006 IL-23 stimulates epidermal hyperplasia via TNF and IL-20R2-dependent mechanisms with implications for psoriasis pathogenesis. *J. Exp. Med.* **203**: 2577–2587.
- CHANDRASHEKAR J., HOON M. A., RYBA N. J. P., ZUKER C. S., 2006 The receptors and cells for mammalian taste. *Nature* **444**: 288–294.

- CHANDRASHEKAR J., MUELLER K. L., HOON M. A., ADLER E., FENG L., GUO W., ZUKER C. S., RYBA N. J., 2000 T2Rs function as bitter taste receptors. *Cell* **100**: 703–711.
- CHARLIER C., AGERHOLM J. S., COPPIETERS W., KARLSKOV-MORTENSEN P., LI W., JONG G. DE, FASQUELLE C., KARIM L., CIRERA S., CAMBISANO N., AHARIZ N., MULLAART E., GEORGES M., FREDHOLM M., 2012 A deletion in the bovine FANCI gene compromises fertility by causing fetal death and brachyspina. *PloS One* **7**: e43085.
- CHARLIER C., COPPIETERS W., ROLLIN F., DESMECHT D., AGERHOLM J. S., CAMBISANO N., CARTA E., DARDANO S., DIVE M., FASQUELLE C., FRENNET J.-C., HANSET R., HUBIN X., JORGENSEN C., KARIM L., KENT M., HARVEY K., PEARCE B. R., SIMON P., TAMA N., NIE H., VANDEPUTTE S., LIEN S., LONGERI M., FREDHOLM M., HARVEY R. J., GEORGES M., 2008 Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat. Genet.* **40**: 449–454.
- CHEN L., 2004 Co-inhibitory molecules of the B7–CD28 family in the control of T-cell immunity. *Nat. Rev. Immunol.* **4**: 336–347.
- CHEN D., XIAO H., ZHANG K., WANG B., GAO Z., JIAN Y., QI X., SUN J., MIAO L., YANG C., 2010 Retromer is required for apoptotic cell clearance by phagocytic receptor recycling. *Science* **327**: 1261–1264.
- CHOI J.-W., LIAO X., STOTHARD P., CHUNG W.-H., JEON H.-J., MILLER S. P., CHOI S.-Y., LEE J.-K., YANG B., LEE K.-T., HAN K.-J., KIM H.-C., JEONG D., OH J.-D., KIM N., KIM T.-H., LEE H.-K., LEE S.-J., 2014 Whole-Genome Analyses of Korean Native and Holstein Cattle Breeds by Massively Parallel Sequencing. *PLoS ONE* **9**: e101127.
- CIRULLI E. T., GOLDSTEIN D. B., 2010 Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* **11**: 415–425.
- COCK P. J. A., FIELDS C. J., GOTO N., HEUER M. L., RICE P. M., 2010 The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **38**: 1767–1771.
- COHEN-ZINDER M., SEROUSSI E., LARKIN D. M., LOOR J. J., EVERTS-VAN DER WIND A., LEE J.-H., DRACKLEY J. K., BAND M. R., HERNANDEZ A. G., SHANI M., LEWIN H. A., WELLER J. I., RON M., 2005 Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* **15**: 936–944.
- COHN M. J., IZPISÚA-BELMONTE J. C., ABUD H., HEATH J. K., TICKLE C., 1995 Fibroblast growth factors induce additional limb development from the flank of chick embryos. *Cell* **80**: 739–746.
- COLE J. B., WAURICH B., WENSCH-DORENDORF M., BICKHART D. M., SWALVE H. H., 2014 A genome-wide association study of calf birth weight in Holstein cattle using single nucleotide polymorphisms and phenotypes predicted from auxiliary traits. *J. Dairy Sci.* **97**: 3156–3172.
- COLE J. B., WIGGANS G. R., VANRADEN P. M., 2007 Genetic Evaluation of Stillbirth in United States Holsteins Using a Sire-Maternal Grandsire Threshold Model. *J. Dairy Sci.* **90**: 2480–2488.
- COLLINS T., STONE J. R., WILLIAMS A. J., 2001 All in the Family: the BTB/POZ, KRAB, and SCAN Domains. *Mol. Cell. Biol.* **21**: 3609–3615.

- CONRAD D. F., PINTO D., REDON R., FEUK L., GOKCUMEN O., ZHANG Y., AERTS J., ANDREWS T. D., BARNES C., CAMPBELL P., FITZGERALD T., HU M., IHM C. H., KRISTIANSSON K., MACARTHUR D. G., MACDONALD J. R., ONYIAH I., PANG A. W. C., ROBSON S., STIRRUPS K., VALSesia A., WALTER K., WEI J., TYLER-SMITH C., CARTER N. P., LEE C., SCHERER S. W., HURLES M. E., 2010 Origins and functional impact of copy number variation in the human genome. *Nature* **464**: 704–712.
- CONSTANS J., CONRI C., 2006 Circulating markers of endothelial function in cardiovascular disease. *Clin. Chim. Acta Int. J. Clin. Chem.* **368**: 33–47.
- CONTE C., EBELING M., MARCUZ A., NEF P., ANDRES-BARQUIN P. J., 2002 Identification and characterization of human taste receptor genes belonging to the TAS2R family. *Cytogenet. Genome Res.* **98**: 45–53.
- CORVA P. M., MEDRANO J. F., 2001 Quantitative trait loci (QTLs) mapping for growth traits in the mouse: a review. *Genet. Sel. Evol. GSE* **33**: 105–132.
- CYBULSKY M. I., IYAMA K., LI H., ZHU S., CHEN M., IYAMA M., DAVIS V., GUTIERREZ-RAMOS J. C., CONNELLY P. W., MILSTONE D. S., 2001 A major role for VCAM-1, but not ICAM-1, in early atherosclerosis. *J. Clin. Invest.* **107**: 1255–1262.
- DAETWYLER H. D., CAPITAN A., PAUSCH H., STOTHARD P., BINSBERGEN R. VAN, BRØNDUM R. F., LIAO X., DJARI A., RODRIGUEZ S. C., GROHS C., ESQUERRÉ D., BOUCHEZ O., ROSSIGNOL M.-N., KLOPP C., ROCHA D., FRITZ S., EGGEN A., BOWMAN P. J., COOTE D., CHAMBERLAIN A. J., ANDERSON C., VAN TASSELL C. P., HULSEGGE I., GODDARD M. E., GULDBRANDTSEN B., LUND M. S., VEERKAMP R. F., BOICHARD D. A., FRIES R., HAYES B. J., 2014 Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* **advance online publication**.
- DAETWYLER H. D., KEMPER K. E., WERF J. H. J. VAN DER, HAYES B. J., 2012 Components of the Accuracy of Genomic Prediction in a Multi-Breed Sheep Population. *J. Anim. Sci.*
- DAETWYLER H. D., PONG-WONG R., VILLANUEVA B., WOOLLIAMS J. A., 2010 The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. *Genetics* **185**: 1021–1031.
- DAHA N. A., LIE B. A., TROUW L. A., STOOKEN G., SCHONKEREN J. J. M., DING B., KVIEN T. K., SCHILHAM M. W., PADYUKOV L., HUIZINGA T. W. J., TOES R., 2012 Novel genetic association of the VTCN1 region with rheumatoid arthritis. *Ann. Rheum. Dis.* **71**: 567–571.
- DAMJANOVSKI S., AMANO T., LI Q., PEI D., SHI Y. B., 2001 Overexpression of matrix metalloproteinases leads to lethality in transgenic *Xenopus laevis*: implications for tissue-dependent functions of matrix metalloproteinases during late embryonic development. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **221**: 37–47.
- DANSKY H. M., BARLOW C. B., LOMINSKA C., SIKES J. L., KAO C., WEINSAFT J., CYBULSKY M. I., SMITH J. D., 2001 Adhesion of monocytes to arterial endothelium and initiation of atherosclerosis are critically dependent on vascular cell adhesion molecule-1 gene dosage. *Arterioscler. Thromb. Vasc. Biol.* **21**: 1662–1667.
- DAVEY J. W., HOHENLOHE P. A., ETTER P. D., BOONE J. Q., CATCHEN J. M., BLAXTER M. L., 2011 Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **12**: 499–510.
- DAVIS R. C., DIEP A., HUNZIKER W., KLISAK I., MOHANDAS T., SCHOTZ M. C., SPARKES R. S., LUSIS A.

- J., 1991 Assignment of human pancreatic lipase gene (PNLIP) to chromosome 10q24-q26. *Genomics* **11**: 1164–1166.
- DENG H., GAO K., JANKOVIC J., 2013 The VPS35 gene and Parkinson's disease. *Mov. Disord.* **28**: 569–575.
- DEPRISTO M. A., BANKS E., POPLIN R., GARIMELLA K. V., MAGUIRE J. R., HARTL C., PHILIPPAKIS A. A., ANGEL G. DEL, RIVAS M. A., HANNA M., MCKENNA A., FENNEL T. J., KERNYTSKY A. M., SIVACHENKO A. Y., CIBULSKIS K., GABRIEL S. B., ALTSHULER D., DALY M. J., 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**: 491–498.
- DICKSON S. P., WANG K., KRANTZ I., HAKONARSON H., GOLDSTEIN D. B., 2010 Rare Variants Create Synthetic Genome-Wide Associations. *PLoS Biol* **8**: e1000294.
- DIPOLO R., BEAUGÉ L., 2006 Sodium/calcium exchanger: influence of metabolic regulation on ion carrier interactions. *Physiol. Rev.* **86**: 155–203.
- DOEGE H., BOCIANSKI A., SCHEEPERS A., AXER H., ECKEL J., JOOST H. G., SCHÜRMAN A., 2001 Characterization of human glucose transporter (GLUT) 11 (encoded by SLC2A11), a novel sugar-transport facilitator specifically expressed in heart and skeletal muscle. *Biochem. J.* **359**: 443–449.
- DONN R. P., SHELLEY E., OLLIER W. E., THOMSON W., BRITISH PAEDIATRIC RHEUMATOLOGY STUDY GROUP, 2001 A novel 5'-flanking region polymorphism of macrophage migration inhibitory factor is associated with systemic-onset juvenile idiopathic arthritis. *Arthritis Rheum.* **44**: 1782–1785.
- DRAHEIM K. M., CHEN H.-B., TAO Q., MOORE N., ROCHE M., LYLE S., 2010 ARRDC3 suppresses breast cancer progression by negatively regulating integrin beta4. *Oncogene* **29**: 5032–5047.
- DRICKAMER K., 1999 C-type lectin-like domains. *Curr. Opin. Struct. Biol.* **9**: 585–590.
- DRÖGEMÜLLER C., TETENS J., SIGURDSSON S., GENTILE A., TESTONI S., LINDBLAD-TOH K., LEEB T., 2010 Identification of the bovine Arachnomelia mutation by massively parallel sequencing implicates sulfite oxidase (SUOX) in bone development. *PLoS Genet.* **6**.
- DRUET T., MACLEOD I. M., HAYES B. J., 2014 Toward genomic prediction from whole-genome sequence data: impact of sequencing design on genotype imputation and accuracy of predictions. *Heredity* **112**: 39–47.
- DRUET T., SCHROOTEN C., ROOS A. P. W. DE, 2010 Imputation of genotypes from different single nucleotide polymorphism panels in dairy cattle. *J. Dairy Sci.* **93**: 5443–5454.
- DURBIN R., ABECASIS GR, ALTSHULER DL, AUTON A, BROOKS LD, DURBIN RM, 2010 A map of human genome variation from population-scale sequencing. *Nature* **467**: 1061–1073.
- EAGLEN S. A. E., COFFEY M. P., WOOLLIAMS J. A., WALL E., 2012 Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. *Genet. Sel. Evol.* **44**: 23.
- EATON K. W., TOOKE L. S., WAINWRIGHT L. M., JUDKINS A. R., BIEGEL J. A., 2011 Spectrum of SMARCB1/INI1 mutations in familial and sporadic rhabdoid tumors. *Pediatr. Blood Cancer* **56**: 7–15.
- EBERLEIN A., TAKASUGA A., SETOGUCHI K., PFUHL R., FLISIKOWSKI K., FRIES R., KLOPP N., FÜRBASS R.,

- WEIKARD R., KÜHN C., 2009 Dissection of genetic factors modulating fetal growth in cattle indicates a substantial role of the non-SMC condensin I complex, subunit G (NCAPG) gene. *Genetics* **183**: 951–964.
- ECK S. H., BENET-PAGÈS A., FLISIKOWSKI K., MEITINGER T., FRIES R., STROM T. M., 2009 Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery. *Genome Biol.* **10**: R82.
- EDGAR A. J., POLAK J. M., 2000 Human homologues of yeast vacuolar protein sorting 29 and 35. *Biochem. Biophys. Res. Commun.* **277**: 622–630.
- EICHLER E. E., FLINT J., GIBSON G., KONG A., LEAL S. M., MOORE J. H., NADEAU J. H., 2010 Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* **11**: 446–450.
- EKINE C. C., ROWE S. J., BISHOP S. C., KONING D.-J. de, 2014 Why Breeding Values Estimated Using Familial Data Should Not Be Used for Genome-Wide Association Studies. *G3 GenesGenomesGenetics* **4**: 341–347.
- ELICES M. J., OSBORN L., TAKADA Y., CROUSE C., LUHOWSKYJ S., HEMLER M. E., LOBB R. R., 1990 VCAM-1 on activated endothelium interacts with the leukocyte integrin VLA-4 at a site distinct from the VLA-4/Fibronectin binding site. *Cell* **60**: 577–584.
- ERBE M., HAYES B. J., MATUKUMALLI L. K., GOSWAMI S., BOWMAN P. J., REICH C. M., MASON B. A., GODDARD M. E., 2012 Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci.* **95**: 4114–4129.
- ERIKSSON S., NÄSHOLM A., JOHANSSON K., PHILIPSSON J., 2004 Genetic parameters for calving difficulty, stillbirth, and birth weight for Hereford and Charolais at first and later parities. *J. Anim. Sci.* **82**: 375–383.
- ESHKIND L., TIAN Q., SCHMIDT A., FRANKE W. W., WINDOFFER R., LEUBE R. E., 2002 Loss of desmoglein 2 suggests essential functions for early embryonic development and proliferation of embryonal stem cells. *Eur. J. Cell Biol.* **81**: 592–598.
- EWING B., GREEN P., 1998 Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Res.* **8**: 186–194.
- EWING B., HILLIER L., WENDL M. C., GREEN P., 1998 Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. *Genome Res.* **8**: 175–185.
- EYRE-WALKER A., 2010 Evolution in health and medicine Sackler colloquium: Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proc. Natl. Acad. Sci. U. S. A.* **107 Suppl 1**: 1752–1756.
- FALCONER D. S., MACKAY T. F. C., 1996 *Introduction fo Quantitative Genetics*. 4th ed. Benjamin Cummings.
- FAY J. C., WU C.-I., 2003 Sequence divergence, functional constraint, and selection in protein evolution. *Annu. Rev. Genomics Hum. Genet.* **4**: 213–235.
- FAYE L. L., MACHIELA M. J., KRAFT P., BULL S. B., SUN L., 2013 Re-Ranking Sequencing Variants in the Post-GWAS Era for Accurate Causal Variant Identification. *PLoS Genet.* **9**.
- FIGARELLA C., CARO A. DE, LEUPOLD D., POLEY J. R., 1980 Congenital pancreatic lipase deficiency. *J. Pediatr.* **96**: 412–416.

- FIRAT E., SAVEANU L., AICHELE P., STAEHELI P., HUAI J., GAEDICKE S., NIL A., BESIN G., KANZLER B., ENDERT P. VAN, NIEDERMANN G., 2007 The role of endoplasmic reticulum-associated aminopeptidase 1 in immunity to infection and in cross-presentation. *J. Immunol. Baltim. Md 1950* **178**: 2241–2248.
- FLIES D. B., CHEN L., 2007 The new B7s: playing a pivotal role in tumor immunity. *J. Immunother. Hagerstown Md 1997* **30**: 251–260.
- FLISIKOWSKI K., VENHORANTA H., NOWACKA-WOSZUK J., MCKAY S. D., FLYCKT A., TAPONEN J., SCHNABEL R., SCHWARZENBACHER H., SZCZERBAL I., LOHI H., FRIES R., TAYLOR J. F., SWITONSKI M., ANDERSSON M., 2010 A novel mutation in the maternally imprinted PEG3 domain results in a loss of MIMT1 expression and causes abortions and stillbirths in cattle (*Bos taurus*). *PloS One* **5**: e15116.
- FLORIANO W. B., HALL S., VAIDEHI N., KIM U., DRAYNA D., III W. A. G., 2006 Modeling the human PTC bitter-taste receptor interactions with bitter tastants. *J. Mol. Model.* **12**: 931–941.
- FONTANESI L., SCOTTI E., RUSSO V., 2012 Haplotype variability in the bovine MITF gene and association with piebaldism in Holstein and Simmental cattle breeds. *Anim. Genet.* **43**: 250–256.
- FRASER H. B., XIE X., 2009 Common polymorphic transcript variation in human disease. *Genome Res.* **19**: 567–575.
- FREEDMAN A. S., MUNRO J. M., RICE G. E., BEVILACQUA M. P., MORIMOTO C., MCINTYRE B. W., RHYNHART K., POBER J. S., NADLER L. M., 1990 Adhesion of human B cells to germinal centers in vitro involves VLA-4 and INCAM-110. *Science* **249**: 1030–1033.
- FRISCHMEYER P. A., DIETZ H. C., 1999 Nonsense-mediated mRNA decay in health and disease. *Hum. Mol. Genet.* **8**: 1893–1900.
- FRITH M. C., HAMADA M., HORTON P., 2010 Parameters for accurate genome alignment. *BMC Bioinformatics* **11**: 80.
- FRTZ S., CAPITAN A., DJARI A., RODRIGUEZ S. C., BARBAT A., BAUR A., GROHS C., WEISS B., BOUSSAHA M., ESQUERRÉ D., KLOPP C., ROCHA D., BOICHARD D., 2013 Detection of Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. *PLoS ONE* **8**: e65550.
- FUKUSUMI S., YOSHIDA H., FUJII R., MARUYAMA M., KOMATSU H., HABATA Y., SHINTANI Y., HINUMA S., FUJINO M., 2003 A New Peptidic Ligand and Its Receptor Regulating Adrenal Function in Rats. *J. Biol. Chem.* **278**: 46387–46395.
- GABORY A., RIPOCHE M.-A., DIGARCHER A. L., WATRIN F., ZIYYAT A., FORNÉ T., JAMMES H., AINSCOUGH J. F. X., SURANI M. A., JOURNOT L., DANDOLO L., 2009 H19 acts as a trans regulator of the imprinted gene network controlling growth in mice. *Development* **136**: 3413–3421.
- GAO Y., CHAUDHURI J., ZHU C., DAVIDSON L., WEAVER D. T., ALT F. W., 1998 A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. *Immunity* **9**: 367–376.
- GASTER M., HANDBERG A., SCHÜRMAN A., JOOST H.-G., BECK-NIELSEN H., SCHRÖDER H. D., 2004 GLUT11, but not GLUT8 or GLUT12, is expressed in human skeletal muscle in a fibre type-specific pattern. *Pflüg. Arch. Eur. J. Physiol.* **448**: 105–113.
- AL-GAZALI L. I., BAKIR M., HAMUD O. A., GERAMI S., 2002 An autosomal recessive syndrome of nasal anomalies associated with renal and anorectal malformations. *Clin. Dysmorphol.*

11: 33–38.

- GELDERMANN H., 1975 Investigations on inheritance of quantitative characters in animals by gene markers I. Methods. Theor. Appl. Genet. **46**: 319–330.
- GEORGES M., NIELSEN D., MACKINNON M., MISHRA A., OKIMOTO R., PASQUINO A. T., SARGEANT L. S., SORENSEN A., STEELE M. R., ZHAO X., WOMACK J. E., HOESCHELE I., 1995 Mapping Quantitative Trait Loci Controlling Milk Production in Dairy Cattle by Exploiting Progeny Testing. Genetics **139**: 907–920.
- GHAVI HOSSEIN-ZADEH N., NEJATI-JAVAREMI A., MIRAEI-ASHTIANI S. R., KOHRAM H., 2008 An Observational Analysis of Twin Births, Calf Stillbirth, Calf Sex Ratio, and Abortion in Iranian Holsteins. J. Dairy Sci. **91**: 4198–4205.
- GILHAM D., LABONTÉ E. D., ROJAS J. C., JANDACEK R. J., HOWLES P. N., HUI D. Y., 2007 Carboxyl ester lipase deficiency exacerbates dietary lipid absorption abnormalities and resistance to diet-induced obesity in pancreatic triglyceride lipase knockout mice. J. Biol. Chem. **282**: 24642–24649.
- GODDARD M., HAYES B., 2009a Genomic selection based on dense genotypes inferred from sparse genotypes. Proc Assoc Advmt Anim Breed Genet **18**: 26–29.
- GODDARD M. E., HAYES B. J., 2009b Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet **10**: 381–391.
- GORDON D., ABAJIAN C., GREEN P., 1998 Consed: A Graphical Tool for Sequence Finishing. Genome Res. **8**: 195–202.
- GOSHEN R., RACHMILEWITZ J., SCHNEIDER T., DE-GROOT N., ARIEL I., PALT Z., HOCHBERG A. A., 1993 The expression of the H-19 and IGF-2 genes during human embryogenesis and placental development. Mol. Reprod. Dev. **34**: 374–379.
- GRANDIN T., 1998 Review: Reducing Handling Stress Improves Both Productivity and Welfare. Prof. Anim. Sci. **14**: 1–10.
- GREMME G., BRENDL V., SPARKS M. E., KURTZ S., 2005 Engineering a software tool for gene structure prediction in higher organisms. Inf. Softw. Technol. **47**: 965–978.
- GRIFFIN C. T., TREJO J., MAGNUSON T., 2005 Genetic evidence for a mammalian retromer complex containing sorting nexins 1 and 2. Proc. Natl. Acad. Sci. U. S. A. **102**: 15173–15177.
- GRISART B., COPPIETERS W., FARNIR F., KARIM L., FORD C., BERZI P., CAMBISANO N., MNI M., REID S., SIMON P., SPELMAN R., GEORGES M., SNELL R., 2002 Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Res. **12**: 222–231.
- GU Z., STEINMETZ L. M., GU X., SCHARFE C., DAVIS R. W., LI W.-H., 2003 Role of duplicate genes in genetic robustness against null mutations. Nature **421**: 63–66.
- GURTNER G. C., DAVIS V., LI H., MCCOY M. J., SHARPE A., CYBULSKY M. I., 1995 Targeted disruption of the murine VCAM1 gene: essential role of VCAM-1 in chorioallantoic fusion and placentation. Genes Dev. **9**: 1–14.
- HABIER D., FERNANDO R. L., DEKKERS J. C. M., 2007 The impact of genetic relationship information on genome-assisted breeding values. Genetics **177**: 2389–2397.
- HAFT C. R., SIERRA M. de la L., BAFFORD R., LESNIAK M. A., BARR V. A., TAYLOR S. I., 2000 Human Orthologs of Yeast Vacuolar Protein Sorting Proteins Vps26, 29, and 35:

- Assembly into Multimeric Complexes. *Mol. Biol. Cell* **11**: 4105–4116.
- HAILE-MARIAM M., NIEUWHOF G. J., BEARD K. T., KONSTATINOV K. V., HAYES B. J., 2013 Comparison of heritabilities of dairy traits in Australian Holstein-Friesian cattle from genomic and pedigree data and implications for genomic evaluations. *J. Anim. Breed. Genet. Z. Für Tierz. Zücht.* **130**: 20–31.
- HANSEN M., MISZTAL I., LUND M. S., PEDERSEN J., CHRISTENSEN L. G., 2004 Undesired phenotypic and genetic trend for stillbirth in Danish Holsteins. *J. Dairy Sci.* **87**: 1477–1486.
- HARVEY R. J., DEPNER U. B., WÄSSLE H., AHMADI S., HEINDL C., REINOLD H., SMART T. G., HARVEY K., SCHÜTZ B., ABO-SALEM O. M., ZIMMER A., POISBEAU P., WELZL H., WOLFER D. P., BETZ H., ZEILHOFER H. U., MÜLLER U., 2004 GlyR alpha3: an essential target for spinal PGE2-mediated inflammatory pain sensitization. *Science* **304**: 884–887.
- HAYES B. J., BOWMAN P. J., DAETWYLER H. D., KIJAS J. W., WERF J. H. J. VAN DER, 2012 Accuracy of genotype imputation in sheep breeds. *Anim. Genet.* **43**: 72–80.
- HAYES B. J., LEWIN H. A., GODDARD M. E., 2013 The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. *Trends Genet.* **29**: 206–214.
- HAYES B. J., PRYCE J., CHAMBERLAIN A. J., BOWMAN P. J., GODDARD M. E., 2010 Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genet.* **6**: e1001139.
- HAYES B. J., VISSCHER P. M., MCPARTLAN H. C., GODDARD M. E., 2003 Novel Multilocus Measure of Linkage Disequilibrium to Estimate Past Effective Population Size. *Genome Res.* **13**: 635–643.
- HEINS B. J., HANSEN L. B., SEYKORA A. J., 2006 Calving Difficulty and Stillbirths of Pure Holsteins versus Crossbreds of Holstein with Normande, Montbeliarde, and Scandinavian Red. *J. Dairy Sci.* **89**: 2805–2810.
- HICKEY J. M., CROSSA J., BABU R., CAMPOS G. DE LOS, 2012 Factors Affecting the Accuracy of Genotype Imputation in Populations from Several Maize Breeding Programs. *Crop Sci.* **52**: 654.
- HOGGART C. J., CLARK T. G., IORIO M. DE, WHITTAKER J. C., BALDING D. J., 2008 Genome-wide significance for dense SNP and resequencing data. *Genet. Epidemiol.* **32**: 179–185.
- HOLM H., GUDBJARTSSON D. F., SULEM P., MASSON G., HELGADOTTIR H. T., ZANON C., MAGNUSSON O. T., HELGASON A., SAEMUNDSDOTTIR J., GYLFASSON A., STEFANSDOTTIR H., GRETARSDOTTIR S., MATTHIASSEN S. E., THORGEIRSSON G., JONASDOTTIR A., SIGURDSSON A., STEFANSSON H., WERGE T., RAFNAR T., KIEMENEY L. A., PARVEZ B., MUHAMMAD R., RODEN D. M., DARBAR D., THORLEIFSSON G., WALTERS G. B., KONG A., THORSTEINSDOTTIR U., ARNAR D. O., STEFANSSON K., 2011 A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet* **43**: 316–320.
- HOON M. A., ADLER E., LINDEMEIER J., BATTEY J. F., RYBA N. J. P., ZUKER C. S., 1999 Putative Mammalian Taste Receptors: A Class of Taste-Specific GPCRs with Distinct Topographic Selectivity. *Cell* **96**: 541–551.
- HOWARD M. T., MALIK N., ANDERSON C. B., VOSKUIL J. L. A., ATKINS J. F., GIBBONS R. J., 2004 Attenuation of an amino-terminal premature stop codon mutation in the ATRX gene by

- an alternative mode of translational initiation. *J. Med. Genet.* **41**: 951–956.
- HOWIE B., FUCHSBERGER C., STEPHENS M., MARCHINI J., ABECASIS G. R., 2012 Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**: 955–959.
- <http://picard.sourceforge.net>,
- HUANG L., WANG C., ROSENBERG N. A., 2009 The Relationship between Imputation Error and Statistical Power in Genetic Association Studies in Diverse Populations. *Am. J. Hum. Genet.* **85**: 692–698.
- HUGGINS K. W., CAMAROTA L. M., HOWLES P. N., HUI D. Y., 2003 Pancreatic triglyceride lipase deficiency minimally affects dietary fat absorption but dramatically decreases dietary cholesterol absorption in mice. *J. Biol. Chem.* **278**: 42899–42905.
- HUNTLEY S., BAGGOTT D. M., HAMILTON A. T., TRAN-GYAMFI M., YANG S., KIM J., GORDON L., BRANSCOMB E., STUBBS L., 2006 A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res.* **16**: 669–677.
- INOUE K., KHAJAVI M., OHYAMA T., HIRABAYASHI S., WILSON J., REGGIN J. D., MANCIAS P., BUTLER I. J., WILKINSON M. F., WEGNER M., LUPSKI J. R., 2004 Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nat. Genet.* **36**: 361–369.
- ISHIHARA A., KANATANI A., MASHIKO S., TANAKA T., HIDAKA M., GOMORI A., IWAASA H., MURAI N., EGASHIRA S., MURAI T., MITOBE Y., MATSUSHITA H., OKAMOTO O., SATO N., JITSUOKA M., FUKURODA T., OHE T., GUAN X., MACNEIL D. J., PLOEG L. H. T. V. der, NISHIKIBE M., ISHII Y., IHARA M., FUKAMI T., 2006 A neuropeptide Y Y5 antagonist selectively ameliorates body weight gain and associated parameters in diet-induced obese mice. *Proc. Natl. Acad. Sci.* **103**: 7154–7158.
- ISKEN O., MAQUAT L. E., 2007 Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev.* **21**: 1833–3856.
- ITOH N., ORNITZ D. M., 2004 Evolution of the Fgf and Fgfr gene families. *Trends Genet.* **20**: 563–569.
- IWATA H., JANNINK J.-L., 2010 Marker Genotype Imputation in a Low-Marker-Density Panel with a High-Marker-Density Reference Panel: Accuracy Evaluation in Barley Breeding Lines. *Crop Sci.* **50**: 1269.
- JACKSON E. M., SHAIKH T. H., GURURANGAN S., JONES M. C., MALKIN D., NIKKEL S. M., ZUPPAN C. W., WAINWRIGHT L. M., ZHANG F., BIEGEL J. A., 2007 High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor. *Hum. Genet.* **122**: 117–127.
- JACOBSON D. A., KUZNETSOV A., LOPEZ J. P., KASH S., ÄMMÄLÄ C. E., PHILIPSON L. H., 2007 Kv2.1 Ablation Alters Glucose-Induced Islet Electrical Activity, Enhancing Insulin Secretion. *Cell Metab.* **6**: 229–235.
- JAKOB F., EBERT R., SEUFERT J., 2006 Fgf23. *AfCS-Nat. Mol. Pages.*
- JANSEN S., AIGNER B., PAUSCH H., WYSOCKI M., ECK S., BENET-PAGÈS A., GRAF E., WIELAND T., STROM T. M., MEITINGER T., FRIES R., 2013 Assessment of the genomic variation in a cattle population by re-sequencing of key animals at low to medium coverage. *BMC*

- Genomics **14**: 446.
- JENSEN J., SU G., MADSEN P., 2012 Partitioning additive genetic variance into genomic and remaining polygenic components for complex traits in dairy cattle. *BMC Genet.* **13**: 44.
- JEROME-MAJEWSKA L. A., ACHKAR T., LUO L., LUPU F., LACY E., 2010 The trafficking protein Tmed2/p24beta(1) is required for morphogenesis of the mouse embryo and placenta. *Dev. Biol.* **341**: 154–166.
- JHEON A. H., GANSS B., CHEIFETZ S., SODEK J., 2001 Characterization of a novel KRAB/C2H2 zinc finger transcription factor involved in bone development. *J. Biol. Chem.* **276**: 18282–18289.
- JIANG Y., LUO L., GUSTAFSON E. L., YADAV D., LAVERTY M., MURGOLO N., VASSILEVA G., ZENG M., LAZ T. M., BEHAN J., QIU P., WANG L., WANG S., BAYNE M., GREENE J., MONSMA F., ZHANG F. L., 2003 Identification and Characterization of a Novel RF-amide Peptide Ligand for Orphan G-protein-coupled Receptor SP9155. *J. Biol. Chem.* **278**: 27652–27657.
- JOHANSON J. M., BERGER P. J., TSURUTA S., MISZTAL I., 2011 A Bayesian threshold-linear model evaluation of perinatal mortality, dystocia, birth weight, and gestation length in a Holstein herd. *J. Dairy Sci.* **94**: 450–460.
- JUNG S., 2013 Molekulargenetische Aufklärung von zwei Erbkrankheiten der Fleckvieh-Population.
- KANATANI A., MASHIKO S., MURAI N., SUGIMOTO N., ITO J., FUKURODA T., FUKAMI T., MORIN N., MACNEIL D. J., PLOEG L. H. VAN DER, SAGA Y., NISHIMURA S., IHARA M., 2000 Role of the Y1 receptor in the regulation of neuropeptide Y-mediated feeding: comparison of wild-type, Y1 receptor-deficient, and Y5 receptor-deficient mice. *Endocrinology* **141**: 1011–1016.
- KANG H. M., SUL J. H., SERVICE S. K., ZAITLEN N. A., KONG S., FREIMER N. B., SABATTI C., ESKIN E., 2010 Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**: 348–354.
- KATO H., TAKEUCHI O., SATO S., YONEYAMA M., YAMAMOTO M., MATSUI K., UEMATSU S., JUNG A., KAWAI T., ISHII K. J., YAMAGUCHI O., OTSU K., TSUJIMURA T., KOH C.-S., REIS E SOUSA C., MATSUURA Y., FUJITA T., AKIRA S., 2006 Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**: 101–105.
- KATOH O., OGURI T., TAKAHASHI T., TAKAI S., FUJIWARA Y., WATANABE H., 1998 ZK1, a novel Krüppel-type zinc finger gene, is induced following exposure to ionizing radiation and enhances apoptotic cell death on hematopoietic cells. *Biochem. Biophys. Res. Commun.* **249**: 595–600.
- KAWAHARA-MIKI R., TSUDA K., SHIWA Y., ARAI-KICHISE Y., MATSUMOTO T., KANESAKI Y., ODA S., EBIHARA S., YAJIMA S., YOSHIKAWA H., KONO T., 2011 Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi. *BMC Genomics* **12**: 103.
- KEMPSTER A. J., 1981 Fat partition and distribution in the carcasses of cattle, sheep and pigs: A review. *Meat Sci.* **5**: 83–98.
- KENDRICK J., CHEUNG A. K., KAUFMAN J. S., GREENE T., ROBERTS W. L., SMITS G., CHONCHOL M., HOST INVESTIGATORS, 2011 FGF-23 associates with death, cardiovascular events, and initiation of chronic dialysis. *J. Am. Soc. Nephrol. JASN* **22**: 1913–1922.

- KHATKAR M. S., NICHOLAS F. W., COLLINS A. R., ZENGER K. R., CAVANAGH J. A., BARRIS W., SCHNABEL R. D., TAYLOR J. F., RAADSMA H. W., 2008 Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC Genomics* **9**: 187.
- KIEZUN A., GARIMELLA K., DO R., STITZIEL N. O., NEALE B. M., MCLAREN P. J., GUPTA N., SKLAR P., SULLIVAN P. F., MORAN J. L., HULTMAN C. M., LICHTENSTEIN P., MAGNUSSON P., LEHNER T., SHUGART Y. Y., PRICE A. L., BAKKER P. I. W. de, PURCELL S. M., SUNYAEV S. R., 2012 Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44**: 623–630.
- KIEZUN A., PULIT S. L., FRANCIOLI L. C., DIJK F. VAN, SWERTZ M., BOOMSMA D. I., DUIJN C. M. VAN, SLAGBOOM P. E., OMMEN G. J. B. VAN, WIJMENGA C., GENOME OF THE NETHERLANDS CONSORTIUM, BAKKER P. I. W. DE, SUNYAEV S. R., 2013 Deleterious alleles in the human genome are on average younger than neutral alleles of the same frequency. *PLoS Genet.* **9**: e1003301.
- KIM U., JORGENSON E., COON H., LEPPERT M., RISCH N., DRAYNA D., 2003 Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* **299**: 1221–1225.
- KIM S. T., MOLEY K. H., 2009 Regulation of Facilitative Glucose Transporters and AKT/MAPK/PRKAA Signaling via Estradiol and Progesterone in the Mouse Uterine Epithelium. *Biol. Reprod.* **81**: 188–198.
- KIM K. H., ROBERTS C. W. M., 2014 Mechanisms by which SMARCB1 loss drives rhabdoid tumor growth. *Cancer Genet.*
- KIMATA M., MATOBA S., IWAI-KANAI E., NAKAMURA H., HOSHINO A., NAKAOKA M., KATAMURA M., OKAWA Y., MITA Y., OKIGAKI M., IKEDA K., TATSUMI T., MATSUBARA H., 2010 p53 and TIGAR regulate cardiac myocyte energy homeostasis under hypoxic stress. *Am. J. Physiol. Heart Circ. Physiol.* **299**: H1908–1916.
- KIYOZUMI D., TAKEICHI M., NAKANO I., SATO Y., FUKUDA T., SEKIGUCHI K., 2012 Basement membrane assembly of the integrin $\alpha 8\beta 1$ ligand nephronectin requires Fraser syndrome-associated proteins. *J. Cell Biol.* **197**: 677–689.
- KLUG A., 1999 Zinc finger peptides for the regulation of gene expression. *J. Mol. Biol.* **293**: 215–218.
- KLUNGLAND H., VÅGE 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 Off. J. Int. Mamm. Genome Soc.* **6**: 636–639.
- KOLTES J. E., MISHRA B. P., KUMAR D., KATARIA R. S., TOTIR L. R., FERNANDO R. L., COBBOLD R., STEFFEN D., COPPIETERS W., GEORGES M., REECY J. M., 2009 A nonsense mutation in cGMP-dependent type II protein kinase (PRKG2) causes dwarfism in American Angus cattle. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 19250–19255.
- KOSSAKOWSKA A. E., HUCHCROFT S. A., URBANSKI S. J., EDWARDS D. R., 1996 Comparative analysis of the expression patterns of metalloproteinases and their inhibitors in breast neoplasia, sporadic colorectal neoplasia, pulmonary carcinomas and malignant non-Hodgkin's lymphomas in humans. *Br. J. Cancer* **73**: 1401–1408.
- KOUFARIOTIS L., CHEN Y.-P. P., BOLORMAA S., HAYES B. J., 2014 Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. *BMC Genomics*

15: 436.

- KREBS S., MEDUGORAC I., RÖTHER S., STRÄSSER K., FÖRSTER M., 2007 A missense mutation in the 3-ketodihydrosphingosine reductase FVT1 as candidate causal mutation for bovine spinal muscular atrophy. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 6746–6751.
- KROEZE W. K., SHEFFLER D. J., ROTH B. L., 2003 G-protein-coupled receptors at a glance. *J. Cell Sci.* **116**: 4867–4869.
- KRYCZEK I., WEI S., ZOU L., ZHU G., MOTTRAM P., XU H., CHEN L., ZOU W., 2006 Cutting Edge: Induction of B7-H4 on APCs through IL-10: Novel Suppressive Mode for Regulatory T Cells. *J. Immunol.* **177**: 40–44.
- KÜHN C., BENNEWITZ J., REINSCH N., XU N., THOMSEN H., LOOFT C., BROCKMANN G. A., SCHWERIN M., WEIMANN C., HIENDLEDER S., ERHARDT G., MEDJUGORAC I., FÖRSTER M., BREINIG B., REINHARDT F., REENTS R., RUSS I., AVERDUNK G., BLÜMEL J., KALM E., 2003 Quantitative Trait Loci Mapping of Functional Traits in the German Holstein Cattle Population. *J. Dairy Sci.* **86**: 360–368.
- KUILENBURG A. B. P. VAN, BEEUMEN J. J. VAN, DER MEER N. M. VAN, MUIJSERS A. O., 1992 Subunits VIIa,b,c of human cytochrome c oxidase. *Eur. J. Biochem.* **203**: 193–199.
- KURIMASA A., OUYANG H., DONG L. J., WANG S., LI X., CORDON-CARDO C., CHEN D. J., LI G. C., 1999 Catalytic subunit of DNA-dependent protein kinase: impact on lymphocyte development and tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 1403–1408.
- KURO-O M., MATSUMURA Y., AIZAWA H., KAWAGUCHI H., SUGA T., UTSUGI T., OHYAMA Y., KURABAYASHI M., KANAME T., KUME E., IWASAKI H., IIDA A., SHIRAKI-IIDA T., NISHIKAWA S., NAGAI R., NABESHIMA Y. I., 1997 Mutation of the mouse klotho gene leads to a syndrome resembling ageing. *Nature* **390**: 45–51.
- KUROSU H., OGAWA Y., MIYOSHI M., YAMAMOTO M., NANDI A., ROSENBLATT K. P., BAUM M. G., SCHIAVI S., HU M.-C., MOE O. W., KURO-O M., 2006 Regulation of Fibroblast Growth Factor-23 Signaling by Klotho. *J. Biol. Chem.* **281**: 6120–6123.
- KWEE L., BALDWIN H. S., SHEN H. M., STEWART C. L., BUCK C., BUCK C. A., LABOW M. A., 1995a Defective development of the embryonic and extraembryonic circulatory systems in vascular cell adhesion molecule (VCAM-1) deficient mice. *Dev. Camb. Engl.* **121**: 489–503.
- KWEE L., BURNS D. K., RUMBERGER J. M., NORTON C., WOLITZKY B., TERRY R., LOMBARD-GILLOOLY K. M., SHUSTER D. J., KONTGEN F., STEWART C., 1995b Creation and characterization of E-selectin- and VCAM-1-deficient mice. *Ciba Found. Symp.* **189**: 17–28; discussion 28–34, 77–78.
- LANDER E. S., BOTSTEIN D., 1987 Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* **236**: 1567–1570.
- LARKIN M. A., BLACKSHIELDS G., BROWN N. P., CHENNA R., MCGETTIGAN P. A., MCGWILLIAM H., VALENTIN F., WALLACE I. M., WILM A., LOPEZ R., THOMPSON J. D., GIBSON T. J., HIGGINS D. G., 2007 Clustal W and Clustal X version 2.0. *Bioinforma. Oxf. Engl.* **23**: 2947–2948.
- LEE D. K., NGUYEN T., LYNCH K. R., CHENG R., VANTI W. B., ARKHITKO O., LEWIS T., EVANS J. F., GEORGE S. R., O'DOWD B. F., 2001 Discovery and mapping of ten novel G protein-coupled receptor genes. *Gene* **275**: 83–91.
- LEIGHTON P. A., INGRAM R. S., EGGENSCHWILER J., EFSTRATIADIS A., TILGHMAN S. M., 1995 Disruption

- of imprinting caused by deletion of the H19 gene region in mice. *Nature* **375**: 34–39.
- LEIPPRANDT J. R., CHEN H., HORVATH J. E., QIAO X. T., JONES M. Z., FRIDERICI K. H., 1999 Identification of a bovine beta-mannosidosis mutation and detection of two beta-mannosidase pseudogenes. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **10**: 1137–1141.
- LEWIS S. E., SEARLE S. M. J., HARRIS N., GIBSON M., LYER V., RICHTER J., WIEL C., BAYRAKTAROGLIR L., BIRNEY E., CROSBY M. A., KAMINKER J. S., MATTHEWS B. B., PROCHNIK S. E., SMITHY C. D., TUPY J. L., RUBIN G. M., MISRA S., MUNGALL C. J., CLAMP M. E., 2002 Apollo: a sequence annotation editor. *Genome Biol.* **3**: 12.
- LI H., 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993.
- LI H., DURBIN R., 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.
- LI H., HANDSAKER B., WYSOKER A., FENNEL T., RUAN J., HOMER N., MARTH G., ABECASIS G., DURBIN R., 1000 GENOME PROJECT DATA PROCESSING SUBGROUP, 2009a The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- LI L., LI Y., BROWNING S. R., BROWNING B. L., SLATER A. J., KONG X., APONTE J. L., MOOSER V. E., CHISSOE S. L., WHITTAKER J. C., NELSON M. R., EHM M. G., 2011 Performance of Genotype Imputation for Rare Variants Identified in Exons and Flanking Regions of Genes. *PLoS ONE* **6**: e24945.
- LI X., MONTGOMERY S. B., 2013 Detection and Impact of Rare Regulatory Variants in Human Disease. *Front. Genet.* **4**.
- LI Y., VINCKENBOSCH N., TIAN G., HUERTA-SANCHEZ E., JIANG T., JIANG H., ALBRECHTSEN A., ANDERSEN G., CAO H., KORNELIUSSEN T., GRARUP N., GUO Y., HELLMAN I., JIN X., LI Q., LIU J., LIU X., SPARSØ T., TANG M., WU H., WU R., YU C., ZHENG H., ASTRUP A., BOLUND L., HOLMKVIST J., JØRGENSEN T., KRISTIANSEN K., SCHMITZ O., SCHWARTZ T. W., ZHANG X., LI R., YANG H., WANG J., HANSEN T., PEDERSEN O., NIELSEN R., WANG J., 2010a Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nat. Genet.* **42**: 969–972.
- LI Y., WILLER C. J., DING J., SCHEET P., ABECASIS G. R., 2010b MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**: 816–834.
- LI Y., WILLER C., SANNA S., ABECASIS G., 2009b Genotype Imputation. *Annu. Rev. Genomics Hum. Genet.* **10**: 387–406.
- LIEN S., KANTANEN J., OLSAKER I., HOLM L.-E., EYTHORSDDOTTIR E., SANDBERG K., DALSGARD B., ADALSTEINSSON S., 1999 Comparison of milk protein allele frequencies in Nordic cattle breeds. *Anim. Genet.* **30**: 85–91.
- LIU S., GUO R., SIMPSON L. G., XIAO Z.-S., BURNHAM C. E., QUARLES L. D., 2003 Regulation of Fibroblastic Growth Factor 23 Expression but Not Degradation by PHEX. *J. Biol. Chem.* **278**: 37419–37426.
- LIU S., QUARLES L. D., 2007 How Fibroblast Growth Factor 23 Works. *J. Am. Soc. Nephrol.* **18**: 1637–1647.

- LIU S., ZHOU J., TANG W., JIANG X., ROWE D. W., QUARLES L. D., 2006 Pathogenic role of Fgf23 in Hyp mice. *Am. J. Physiol. - Endocrinol. Metab.* **291**: E38–E49.
- LOOMAN C., ÅBRINK M., MARK C., HELLMAN L., 2002 KRAB Zinc Finger Proteins: An Analysis of the Molecular Mechanisms Governing Their Increase in Numbers and Complexity During Evolution. *Mol. Biol. Evol.* **19**: 2118–2130.
- LÓPEZ DE MATURANA E., UGARTE E., KOMEN J., ARENDONK J. A. M. VAN, 2007 Consequences of Selection for Yield Traits on Calving Ease Performance. *J. Dairy Sci.* **90**: 2497–2505.
- LOPEZ-SERRA P., MARCILLA M., VILLANUEVA A., RAMOS-FERNANDEZ A., PALAU A., LEAL L., WAHI J. E., SETIEN-BARANDA F., SZCZESNA K., MOUTINHO C., MARTINEZ-CARDUS A., HEYN H., SANDOVAL J., PUERTAS S., VIDAL A., SANJUAN X., MARTINEZ-BALIBREA E., VIÑALS F., PERALES J. C., BRAMSEM J. B., ØRNTTOFT T. F., ANDERSEN C. L., TABERNEIRO J., McDERMOTT U., BOXER M. B., HEIDEN M. G. V., ALBAR J. P., ESTELLER M., 2014 A DERL3-associated defect in the degradation of SLC2A1 mediates the Warburg effect. *Nat. Commun.* **5**.
- LOWE M. E., 2002 The triglyceride lipases of the pancreas. *J. Lipid Res.* **43**: 2007–2016.
- LOWE M. E., ROSENBLUM J. L., STRAUSS A. W., 1989 Cloning and characterization of human pancreatic lipase cDNA. *J. Biol. Chem.* **264**: 20042–20048.
- LUO M. F., BOETTCHER P. J., SCHAEFFER L. R., DEKKERS J. C. M., 2002 Estimation of genetic parameters of calving ease in first and second parities of Canadian Holsteins using Bayesian methods. *Livest. Prod. Sci.* **74**: 175–184.
- MACARTHUR D. G., BALASUBRAMANIAN S., FRANKISH A., HUANG N., MORRIS J., WALTER K., JOSTINS L., HABEGGER L., PICKRELL J. K., MONTGOMERY S. B., ALBERS C. A., ZHANG Z. D., CONRAD D. F., LUNTER G., ZHENG H., AYUB Q., DEPRISTO M. A., BANKS E., HU M., HANDSAKER R. E., ROSENFELD J. A., FROMER M., JIN M., MU X. J., KHURANA E., YE K., KAY M., SAUNDERS G. I., SUNER M.-M., HUNT T., BARNES I. H. A., AMID C., CARVALHO-SILVA D. R., BIGNELL A. H., SNOW C., YNGVADOTTIR B., BUMPSTEAD S., COOPER D. N., XUE Y., ROMERO I. G., WANG J., LI Y., GIBBS R. A., MCCARROLL S. A., DERMITZAKIS E. T., PRITCHARD J. K., BARRETT J. C., HARROW J., HURLES M. E., GERSTEIN M. B., TYLER-SMITH C., 2012 A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes. *Science* **335**: 823–828.
- MACARTHUR D. G., SETO J. T., RAFTERY J. M., QUINLAN K. G., HUTTLEY G. A., HOOK J. W., LEMCKERT F. A., KEE A. J., EDWARDS M. R., BERMAN Y., HARDEMAN E. C., GUNNING P. W., EASTEAL S., YANG N., NORTH K. N., 2007 Loss of ACTN3 gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nat. Genet.* **39**: 1261–1265.
- MACARTHUR D. G., TYLER-SMITH C., 2010 Loss-of-function variants in the genomes of healthy humans. *Hum. Mol. Genet.* **19**: R125–R130.
- MACKAY T. F. C., 2001 Quantitative trait loci in *Drosophila*. *Nat. Rev. Genet.* **2**: 11–20.
- MACKAY A. P., BERG C. J., ATRASH H. K., 2001 Pregnancy-related mortality from preeclampsia and eclampsia. *Obstet. Gynecol.* **97**: 533–538.
- MACKAY T. F. C., RICHARDS S., STONE E. A., BARBADILLA A., AYROLES J. F., ZHU D., CASILLAS S., HAN Y., MAGWIRE M. M., CRIDLAND J. M., RICHARDSON M. F., ANHOLT R. R. H., BARRÓN M., BESS C., BLANKENBURG K. P., CARBONE M. A., CASTELLANO D., CHABOUB L., DUNCAN L., HARRIS Z., JAVAID M., JAYASEELAN J. C., JHANGIANI S. N., JORDAN K. W., LARA F., LAWRENCE F., LEE S. L., LIBRADO P., LINHEIRO R. S., LYMAN R. F., MACKEY A. J., MUNIDASA M., MUZNY D. M.,

- NAZARETH L., NEWSHAM I., PERALES L., PU L.-L., QU C., RÀMIA M., REID J. G., ROLLMANN S. M., ROZAS J., SAADA N., TURLAPATI L., WORLEY K. C., WU Y.-Q., YAMAMOTO A., ZHU Y., BERGMAN C. M., THORNTON K. R., MITTELMAN D., GIBBS R. A., 2012 The *Drosophila melanogaster* Genetic Reference Panel. *Nature* **482**: 173–178.
- MAGEE D. A., SIKORA K. M., BERKOWICZ E. W., BERRY D. P., HOWARD D. J., MULLEN M. P., EVANS R. D., SPILLANE C., MACHUGH D. E., 2010 DNA sequence polymorphisms in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle. *BMC Genet.* **11**: 93.
- MAHER B., 2008 Personal genomes: The case of the missing heritability. *Nature* **456**: 18–21.
- MARCHINI J., CARDON L. R., PHILLIPS M. S., DONNELLY P., 2004 The effects of human population structure on large genetic association studies. *Nat. Genet.* **36**: 512–517.
- MARCHINI J., HOWIE B., MYERS S., McVEAN G., DONNELLY P., 2007 A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**: 906–913.
- MARTHERUS R. S. R. M., SLUITER W., TIMMER E. D. J., VANHERLE S. J. V., SMEETS H. J. M., AYOUBI T. A. Y., 2010 Functional annotation of heart enriched mitochondrial genes GBAS and CHCHD10 through guilt by association. *Biochem. Biophys. Res. Commun.* **402**: 203–208.
- MARTÍNEZ NOGUEIRAS A., TEIXEIRA COSTEIRA M., SARAIVA MOREIRA H., ARAUJO ANTUNES H., 2001 [Russel-Silver syndrome]. *An. Esp. Pediatría* **54**: 591–594.
- MATSUDA D., SATO H., MAQUAT L. E., 2008 Chapter 9. Studying nonsense-mediated mRNA decay in mammalian cells. *Methods Enzymol.* **449**: 177–201.
- MATUKUMALLI L. K., LAWLEY C. T., SCHNABEL R. D., TAYLOR J. F., ALLAN M. F., HEATON M. P., O'CONNELL J., MOORE S. S., SMITH T. P. L., SONSTEGARD T. S., TASSELL C. P. VAN, 2009 Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PLoS ONE* **4**: e5350.
- McCLURE M. C., BICKHART D., NULL D., VANRADEN P., XU L., WIGGANS G., LIU G., SCHROEDER S., GLASSCOCK J., ARMSTRONG J., COLE J. B., TASSELL C. P. VAN, SONSTEGARD T. S., 2014 Bovine Exome Sequence Analysis and Targeted SNP Genotyping of Recessive Fertility Defects BH1, HH2, and HH3 Reveal a Putative Causative Mutation in SMC2 for HH3. *PLoS ONE* **9**: e92769.
- McCLURE M. C., MORSCI N. S., SCHNABEL R. D., KIM J. W., YAO P., ROLF M. M., MCKAY S. D., GREGG S. J., CHAPPLE R. H., NORTHUTT S. L., TAYLOR J. F., 2010 A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim. Genet.* **41**: 597–607.
- MCGOUGH I. J., CULLEN P. J., 2011 Recent Advances in Retromer Biology. *Traffic* **12**: 963–971.
- McKUSICK-NATHANS INSTITUTE OF GENETIC MEDICINE, JOHNS HOPKINS UNIVERSITY (BALTIMORE, MD), 2014 Online Mendelian Inheritance in Man, OMIM. <http://omim.org>.
- MEDUGORAC I., KEMTER J., RUSS I., PIETROWSKI D., NÜSKE S., REICHENBACH H.-D., SCHMAHL W., FÖRSTER M., 2003 Mapping of the bovine spinal muscular atrophy locus to Chromosome 24. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **14**: 383–391.
- MEIJERING A., 1984 Dystocia and stillbirth in cattle — A review of causes, relations and implications. *Livest. Prod. Sci.* **11**: 143–177.

- MEUWISSEN T., GODDARD M., 2010 Accurate Prediction of Genetic Values for Complex Traits by Whole-Genome Resequencing. *Genetics* **185**: 623–631.
- MEYER C. L., BERGER P. J., KOEHLER K. J., 2000 Interactions among factors affecting stillbirths in Holstein cattle in the United States. *J. Dairy Sci.* **83**: 2657–2663.
- MEYER C. L., BERGER P. J., KOEHLER K. J., THOMPSON J. R., SATTLER C. G., 2001 Phenotypic trends in incidence of stillbirth for Holsteins in the United States. *J. Dairy Sci.* **84**: 515–523.
- MEYER K. B., MAIA A.-T., O'REILLY M., TESCHENDORFF A. E., CHIN S.-F., CALDAS C., PONDER B. A. J., 2008 Allele-Specific Up-Regulation of FGFR2 Increases Susceptibility to Breast Cancer. *PLoS Biol* **6**: e108.
- MEYERHOF W., 2005 Elucidation of mammalian bitter taste. *Rev. Physiol. Biochem. Pharmacol.* **154**: 37–72.
- MIDDLETON M. K., ZUKAS A. M., RUBINSTEIN T., JACOB M., ZHU P., ZHAO L., BLAIR I., PURÉ E., 2006 Identification of 12/15-lipoxygenase as a suppressor of myeloproliferative disease. *J. Exp. Med.* **203**: 2529–2540.
- MILLS R. E., WALTER K., STEWART C., HANDSAKER R. E., CHEN K., ALKAN C., ABYZOV A., YOON S. C., YE K., CHEETHAM R. K., CHINWALLA A., CONRAD D. F., FU Y., GRUBERT F., HAJIRASOULIHA I., HORMOZDIARI F., IAKOUCHEVA L. M., IQBAL Z., KANG S., KIDD J. M., KONKEL M. K., KORN J., KHURANA E., KURAL D., LAM H. Y. K., LENG J., LI R., LI Y., LIN C.-Y., LUO R., MU X. J., NEMESH J., PECKHAM H. E., RAUSCH T., SCALLY A., SHI X., STROMBERG M. P., STUTZ A. M., URBAN A. E., WALKER J. A., WU J., ZHANG Y., ZHANG Z. D., BATZER M. A., DING L., MARTH G. T., McVEAN G., SEBAT J., SNYDER M., WANG J., YE K., EICHLER E. E., GERSTEIN M. B., HURLES M. E., LEE C., McCARROLL S. A., KORBEL J. O., 2011 Mapping copy number variation by population-scale genome sequencing. *Nature* **470**: 59–65.
- MIYAKE K., MEDINA K., ISHIHARA K., KIMOTO M., AUERBACH R., KINCADE P. W., 1991 A VCAM-like adhesion molecule on murine bone marrow stromal cells mediates binding of lymphocyte precursors in culture. *J. Cell Biol.* **114**: 557–565.
- MIYAMOTO K., SEGAWA H., ITO M., KUWAHATA M., 2004 Physiological Regulation of Renal Sodium-Dependent Phosphate Cotransporters. *Jpn. J. Physiol.* **54**: 93–102.
- MIYAWAKI K., YAMADA Y., BAN N., IHARA Y., TSUKIYAMA K., ZHOU H., FUJIMOTO S., OKU A., TSUDA K., TOYOKUNI S., HIAI H., MIZUNOYA W., FUSHIKI T., HOLST J. J., MAKINO M., TASHITA A., KOBARA Y., TSUBAMOTO Y., JINNOUCHI T., JOMORI T., SEINO Y., 2002 Inhibition of gastric inhibitory polypeptide signaling prevents obesity. *Nat. Med.* **8**: 738–742.
- MORT M., IVANOV D., COOPER D. N., CHUZHANOVA N. A., 2008 A meta-analysis of nonsense mutations causing human genetic disease. *Hum. Mutat.* **29**: 1037–1047.
- MORTON J. M., CAMPBELL P. H., 1997 Disease signs reported in south-eastern Australian dairy cattle while grazing Brassica species. *Aust. Vet. J.* **75**: 109–113.
- MOTT R., YUAN W., KAISAKI P., GAN X., CLEAK J., EDWARDS A., BAUD A., FLINT J., 2014 The Architecture of Parent-of-Origin Effects in Mice. *Cell* **156**: 332–342.
- NADEAU J. H., 2001 Modifier genes in mice and humans. *Nat. Rev. Genet.* **2**: 165–174.
- NAGY E., MAQUAT L. E., 1998 A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**: 198–199.
- NAPIREI M., KARSUNKY H., ZEVNIK B., STEPHAN H., MANNHERZ H. G., MÖRÖY T., 2000 Features of

- systemic lupus erythematosus in Dnase1-deficient mice. *Nat. Genet.* **25**: 177–181.
- NEGRI L., LATTANZI R., GIANNINI E., COLUCCI M., MARGHERITI F., MELCHIORRI P., VELLANI V., TIAN H., FELICE M. DE, PORRECA F., 2006 Impaired nociception and inflammatory pain sensation in mice lacking the prokineticin receptor PKR1: focus on interaction between PKR1 and the capsaicin receptor TRPV1 in pain behavior. *J. Neurosci. Off. J. Soc. Neurosci.* **26**: 6716–6727.
- NG P. C., HENIKOFF S., 2003 SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**: 3812–3814.
- NG P. C., LEVY S., HUANG J., STOCKWELL T. B., WALENZ B. P., LI K., AXELROD N., BUSAM D. A., STRAUSBERG R. L., VENTER J. C., 2008 Genetic Variation in an Individual Human Exome. *PLoS Genet* **4**: e1000160.
- NIBBELINK K. A., TISHKOFF D. X., HERSHEY S. D., RAHMAN A., SIMPSON R. U., 2007 1,25(OH)₂-vitamin D₃ actions on cell proliferation, size, gene expression, and receptor localization, in the HL-1 cardiac myocyte. *J. Steroid Biochem. Mol. Biol.* **103**: 533–537.
- NICHOLAS F. W., HOBBS M., 2014 Mutation discovery for Mendelian traits in non-laboratory animals: a review of achievements up to 2012. *Anim. Genet.* **45**: 157–170.
- NICKERSON S. C., 1995 Milk production: Factors affecting milk composition. In: Harding F (Ed.), *Milk Quality*, Springer US, pp. 3–24.
- NICKERSON D. A., TOBE V. O., TAYLOR S. L., 1997 PolyPhred: Automating the Detection and Genotyping of Single Nucleotide Substitutions Using Fluorescence-Based Resequencing. *Nucleic Acids Res.* **25**: 2745–2751.
- NOMURA T., HONDA T., MUKAI F., 2001 Inbreeding and effective population size of Japanese Black cattle. *J. Anim. Sci.* **79**: 366–370.
- NORD A., LEE M., KING M.-C., WALSH T., 2011 Accurate and exact CNV identification from targeted high-throughput sequence data. *BMC Genomics* **12**: 184.
- NOTHNAGEL M., ELLINGHAUS D., SCHREIBER S., KRAWCZAK M., FRANKE A., 2009 A comprehensive evaluation of SNP genotype imputation. *Hum. Genet.* **125**: 163–171.
- NOTTER D. R., 1999 The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci.* **77**: 61–69.
- OBER U., AYROLES J. F., STONE E. A., RICHARDS S., ZHU D., GIBBS R. A., STRICKER C., GIANOLA D., SCHLATHER M., MACKAY T. F. C., SIMIANER H., 2012 Using Whole-Genome Sequence Data to Predict Quantitative Trait Phenotypes in *Drosophila melanogaster*. *PLoS Genet* **8**: e1002685.
- ODA Y., OKADA T., YOSHIDA H., KAUFMAN R. J., NAGATA K., MORI K., 2006 Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation. *J. Cell Biol.* **172**: 383–393.
- OHKUMO T., KONDO Y., YOKOI M., TSUKAMOTO T., YAMADA A., SUGIMOTO T., KANAO R., HIGASHI Y., KONDOH H., TATEMATSU M., MASUTANI C., HANAOKA F., 2006 UV-B radiation induces epithelial tumors in mice lacking DNA polymerase eta and mesenchymal tumors in mice deficient for DNA polymerase iota. *Mol. Cell. Biol.* **26**: 7696–7706.
- OKA S., MASUTANI H., LIU W., HORITA H., WANG D., KIZAKA-KONDOH S., YODOI J., 2006 Thioredoxin-binding protein-2-like inducible membrane protein is a novel vitamin D₃

- and peroxisome proliferator-activated receptor (PPAR)gamma ligand target protein that regulates PPARgamma signaling. *Endocrinology* **147**: 733–743.
- OLSON M. V., 1999 When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* **64**: 18–23.
- OLSON K. M., VANRADEN P. M., TOOKER M. E., 2012 Multibreed genomic evaluations using purebred Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* **95**: 5378–5383.
- OMASITS U., AHRENS C. H., MÜLLER S., WOLLSCHIED B., 2013 Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinforma. Oxf. Engl.*
- ONLINE MENDELIAN INHERITANCE IN ANIMALS, OMIA, FACULTY OF VETERINARY SCIENCE OF SYDNEY, University of Sydney, 4. Juni 2014 . World Wide Web URL: <http://omia.angis.org.au/>.
- ORBAK Z., KARACAN M., DONERAY H., KARAKELLEOGLU C., 2007 Congenital rickets presenting with hypocalcaemic seizures. *West Indian Med. J.* **56**: 364–367.
- ORNITZ D. M., 2000 FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *BioEssays* **22**: 108–112.
- ORTEGA S., ITTMANN M., TSANG S. H., EHRLICH M., BASILICO C., 1998 Neuronal defects and delayed wound healing in mice lacking fibroblast growth factor 2. *Proc. Natl. Acad. Sci.* **95**: 5672–5677.
- OSBORN L., HESSION C., TIZARD R., VASSALLO C., LUHOWSKYJ S., CHI-ROSSO G., LOBB R., 1989 Direct expression cloning of vascular cell adhesion molecule 1, a cytokine-induced endothelial protein that binds to lymphocytes. *Cell* **59**: 1203–1211.
- OSBORN L., VASSALLO C., BROWNING B. G., TIZARD R., HASKARD D. O., BENJAMIN C. D., DOUGAS I., KIRCHHAUSEN T., 1994 Arrangement of domains, and amino acid residues required for binding of vascular cell adhesion molecule-1 to its counter-receptor VLA-4 (alpha 4 beta 1). *J. Cell Biol.* **124**: 601–608.
- OZISIK G., MANTOVANI G., ACHERMANN J. C., PERSANI L., SPADA A., WEISS J., BECK-PECCOZ P., JAMESON J. L., 2003 An alternate translation initiation site circumvents an amino-terminal DAX1 nonsense mutation leading to a mild form of X-linked adrenal hypoplasia congenita. *J. Clin. Endocrinol. Metab.* **88**: 417–423.
- PARTIPILO G., D'ADDABBO P., LACALANDRA G. M., LIU G. E., ROCCHI M., 2011 Refinement of *Bos taurus* sequence assembly based on BAC-FISH experiments. *BMC Genomics* **12**: 639.
- PATWARI P., EMILSSON V., SCHADT E. E., CHUTKOW W. A., LEE S., MARSILI A., ZHANG Y., DOBRIN R., COHEN D. E., LARSEN P. R., ZAVACKI A. M., FONG L. G., YOUNG S. G., LEE R. T., 2011 The Arrestin Domain Containing 3 (ARRDC3) Protein Regulates Body Mass and Energy Expenditure. *Cell Metab.* **14**: 671–683.
- PAUSCH H., AIGNER B., EMMERLING R., EDEL C., GÖTZ K.-U., FRIES R., 2013 Imputation of high-density genotypes in the Fleckvieh cattle population. *Genet. Sel. Evol.* **45**: 3.
- PAUSCH H., FLISIKOWSKI K., JUNG S., EMMERLING R., EDEL C., GÖTZ K.-U., FRIES R., 2011 Genome-Wide Association Study Identifies Two Major Loci Affecting Calving Ease and Growth-Related Traits in Cattle. *Genetics* **187**: 289–297.
- PAUSCH H., KÖLLE S., WURMSER C., SCHWARZENBACHER H., EMMERLING R., JANSEN S., TROTTMANN M., FUERST C., GÖTZ K.-U., FRIES R., 2014 A Nonsense Mutation in TMEM95 Encoding a Nondescript Transmembrane Protein Causes Idiopathic Male Subfertility in Cattle. *PLoS*

- Genet **10**: e1004044.
- PAUSCH H., WANG X., JUNG S., KROGMEIER D., EDEL C., EMMERLING R., GÖTZ K.-U., FRIES R., 2012 Identification of QTL for UV-Protective Eye Area Pigmentation in Cattle by Progeny Phenotyping and Genome-Wide Association Analysis. *PLoS ONE* **7**: e36346.
- PETERS S. O., KIZILKAYA K., GARRICK D. J., FERNANDO R. L., REECY J. M., WEABER R. L., SILVER G. A., THOMAS M. G., 2012 Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. *J. Anim. Sci.* **90**: 3398–3409.
- PHILIPSSON J., 1996 Strategies to reduce problems in calving performance and stillbirths by selection and differential use of bulls. *Interbull Bull.* **0**.
- PIERCE K. L., PREMONT R. T., LEFKOWITZ R. J., 2002 Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* **3**: 639–650.
- PILICHOU K., NAVA A., BASSO C., BEFFAGNA G., BAUCE B., LORENZON A., FRIGO G., VETTORI A., VALENTE M., TOWBIN J., THIENE G., DANIELI G. A., RAMPAZZO A., 2006 Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy. *Circulation* **113**: 1171–1179.
- PISKOL R., RAMASWAMI G., LI J. B., 2013 Reliable identification of genomic variants from RNA-seq data. *Am. J. Hum. Genet.* **93**: 641–651.
- POLTE T., NEWMAN W., RAGHUNATHAN G., GOPAL T. V., 1991 Structural and functional studies of full-length vascular cell adhesion molecule-1: internal duplication and homology to several adhesion proteins. *DNA Cell Biol.* **10**: 349–357.
- POSCH M. G., POSCH M. J., GEIER C., ERDMANN B., MUELLER W., RICHTER A., RUPPERT V., PANKUWEIT S., MAISCH B., PERROT A., BUTTGEREIT J., DIETZ R., HAVERKAMP W., OZCELIK C., 2008 A missense variant in desmoglein-2 predisposes to dilated cardiomyopathy. *Mol. Genet. Metab.* **95**: 74–80.
- PRUITT K. D., BROWN G. R., HIATT S. M., THIBAUD-NISSEN F., ASTASHYN A., ERMOLAEVA O., FARRELL C. M., HART J., LANDRUM M. J., MCGARVEY K. M., MURPHY M. R., O'LEARY N. A., PUJAR S., RAJPUT B., RANGWALA S. H., RIDDICK L. D., SHKEDA A., SUN H., TAMEZ P., TULLY R. E., WALLIN C., WEBB D., WEBER J., WU W., DICUCCIO M., KITTS P., MAGLOTT D. R., MURPHY T. D., OSTEL J. M., 2014 RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* **42**: D756–763.
- PURCELL S., NEALE B., TODD-BROWN K., THOMAS L., FERREIRA M. A. R., BENDER D., MALLER J., SKLAR P., BAKKER P. I. W. DE, 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., PAUSCH H., JANSEN S., SOMEL M., STROM T. M., FRIES R., NIELSEN R., SIMIANER H., 2014 Classic Selective Sweeps Revealed by Massive Sequencing in Cattle. *PLoS Genet* **10**: e1004148.
- QANBARI S., PIMENTEL E. C. G., TETENS J., THALLER G., LICHTNER P., SHARIFI A. R., SIMIANER H., 2010 The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* **41**: 346–356.
- QI S., O'HAYRE M., GUTKIND J. S., HURLEY J. H., 2014 Structural and Biochemical Basis for Ubiquitin Ligase Recruitment by Arrestin-related Domain-containing Protein-3

- (ARRDC3). *J. Biol. Chem.* **289**: 4743–4752.
- QIAN Y., HONG B., SHEN L., WU Z., YAO H., ZHANG L., 2013 B7-H4 enhances oncogenicity and inhibits apoptosis in pancreatic cancer cells. *Cell Tissue Res.* **353**: 139–151.
- QIAN Y., SHEN L., CHENG L., WU Z., YAO H., 2011 B7-H4 expression in various tumors determined using a novel developed monoclonal antibody. *Clin. Exp. Med.* **11**: 163–170.
- QIU Q., ZHANG G., MA T., QIAN W., WANG J., YE Z., CAO C., HU Q., KIM J., LARKIN D. M., AUVEL L., CAPITANU B., MA J., LEWIN H. A., QIAN X., LANG Y., ZHOU R., WANG L., WANG K., XIA J., LIAO S., PAN S., LU X., HOU H., WANG Y., ZANG X., YIN Y., MA H., ZHANG J., WANG Z., ZHANG Y., ZHANG D., YONEZAWA T., HASEGAWA M., ZHONG Y., LIU W., ZHANG Y., HUANG Z., ZHANG S., LONG R., YANG H., WANG J., LENSTRA J. A., COOPER D. N., WU Y., WANG J., SHI P., WANG J., LIU J., 2012 The yak genome and adaptation to life at high altitude. *Nat. Genet.*
- RESTREPO R., ZHAO X., PETER H., ZHANG B., ARVAN P., NOTHWEHR S. F., 2007 Structural Features of Vps35p Involved in Interaction with Other Subunits of the Retromer Complex. *Traffic* **8**: 1841–1853.
- REY J., FREZAL J., ROYER P., LAMY M., 1966 [Congenital absence of pancreatic lipase]. *Arch. Fr. Pédiatrie* **23**: 5–14.
- RISCH N. J., 2000 Searching for genetic determinants in the new millennium. *Nature* **405**: 847–856.
- RODIG S. J., KUTOK J. L., PATERSON J. C., NITTA H., ZHANG W., CHAPUY B., TUMWINE L. K., MONTES-MORENO S., AGOSTINELLI C., JOHNSON N. A., BEN-NERIAH S., FARINHA P., SHIPP M. A., PIRIS M. A., GROGAN T. M., PILERI S. A., GASCOYNE R. D., MARAFIOTI T., 2010 The pre-B-cell receptor associated protein VpreB3 is a useful diagnostic marker for identifying c-MYC translocated lymphomas. *Haematologica* **95**: 2056–2062.
- ROSENBERG N. A., NORDBORG M., 2006 A general population-genetic model for the production by population structure of spurious genotype-phenotype associations in discrete, admixed or spatially distributed populations. *Genetics* **173**: 1665–1678.
- ROSNET O., BLANCO-BETANCOURT C., GRIVEL K., RICHTER K., SCHIFF C., 2004 Binding of Free Immunoglobulin Light Chains to VpreB3 Inhibits Their Maturation and Secretion in Chicken B Cells. *J. Biol. Chem.* **279**: 10228–10236.
- ROZEN S., SKALETSKY H., 1999 Primer3 on the WWW for General Users and for Biologist Programmers. In: Misener S, Krawetz SA (Eds.), *Bioinformatics Methods and Protocols, Methods in Molecular Biology*TM. Humana Press, pp. 365–386.
- RUFFALO M., LAFRAMBOISE T., KOYUTÜRK M., 2011 Comparative analysis of algorithms for next-generation sequencing read alignment. *Bioinformatics* **27**: 2790–2796.
- RYBICKI B. A., WALEWSKI J. L., MALIARIK M. J., KIAN H., IANNUZZI M. C., 2005 The BTNL2 Gene and Sarcoidosis Susceptibility in African Americans and Whites. *Am. J. Hum. Genet.* **77**: 491–499.
- SAATCHI M., 2013 QTL Mapping for Body Weights and Calving Ease in Hereford Cattle Using 1Mb SNP Windows and Bayesian Inference. In: *Plant and Animal Genome*.
- SAATCHI M., SCHNABEL R. D., TAYLOR J. F., GARRICK D. J., 2014 Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. *BMC Genomics* **15**: 442.

- SABOUR H., HOSSEIN-NEZHAD A., MAGHBOOLI Z., MADANI F., MIR E., LARIJANI B., 2006 Relationship between pregnancy outcomes and maternal vitamin D and calcium intake: A cross-sectional study. *Gynecol. Endocrinol. Off. J. Int. Soc. Gynecol. Endocrinol.* **22**: 585–589.
- SAHANA G., GULDBRANDTSEN B., LUND M. S., 2011 Genome-wide association study for calving traits in Danish and Swedish Holstein cattle. *J. Dairy Sci.* **94**: 479–486.
- SALCEDA S., TANG T., KMET M., MUNTEANU A., GHOSH M., MACINA R., LIU W., PILKINGTON G., PAPKOFF J., 2005 The immunomodulatory protein B7-H4 is overexpressed in breast and ovarian cancers and promotes epithelial cell transformation. *Exp. Cell Res.* **306**: 128–141.
- SANGER F., NICKLEN S., COULSON A. R., 1977 DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 5463–5467.
- SASAKI T., MINOSHIMA S., SHIOHAMA A., SHINTANI A., SHIMIZU A., ASAKAWA S., KAWASAKI K., SHIMIZU N., 2001 Molecular Cloning of a Member of the Facilitative Glucose Transporter Gene Family GLUT11 (SLC2A11) and Identification of Transcription Variants. *Biochem. Biophys. Res. Commun.* **289**: 1218–1224.
- SATHYANARAYANA B. K., HAHN Y., PATANKAR M. S., PASTAN I., LEE B., 2009 Mesothelin, Stereocilin, and Otoancorin are predicted to have superhelical structures with ARM-type repeats. *BMC Struct. Biol.* **9**: 1.
- SATO H., TAKETOMI Y., ISOGAI Y., MIKI Y., YAMAMOTO K., MASUDA S., HOSONO T., ARATA S., ISHIKAWA Y., ISHII T., KOBAYASHI T., NAKANISHI H., IKEDA K., TAGUCHI R., HARA S., KUDO I., MURAKAMI M., 2010 Group III secreted phospholipase A2 regulates epididymal sperm maturation and fertility in mice. *J. Clin. Invest.* **120**: 1400–1414.
- SAVARIS R. F., HAMILTON A. E., LESSEY B. A., GIUDICE L. C., 2008 Endometrial gene expression in early pregnancy: lessons from human ectopic pregnancy. *Reprod. Sci. Thousand Oaks Calif* **15**: 797–816.
- SAVAS S., TUZMEN S., OZCELIK H., 2006 Human SNPs resulting in premature stop codons and protein truncation. *Hum. Genomics* **2**: 274.
- SCHAEFFER L. R., 2006 Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet. Z. Für Tierz. Zücht.* **123**: 218–223.
- SCHEEPERS A., SCHMIDT S., MANOLESCU A., CHEESEMAN C. I., BELL A., ZAHN C., JOOST H.-G., SCHÜRMAN A., 2005 Characterization of the human SLC2A11 (GLUT11) gene: alternative promoter usage, function, expression, and subcellular distribution of three isoforms, and lack of mouse orthologue. *Mol. Membr. Biol.* **22**: 339–351.
- SHELL T., KULOZIK A. E., HENTZE M. W., 2002 Integration of splicing, transport and translation to achieve mRNA quality control by the nonsense-mediated decay pathway. *Genome Biol.* **3**: REVIEWS1006.
- SCHORK N. J., MURRAY S. S., FRAZER K. A., TOPOL E. J., 2009 Common vs. rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.* **19**: 212–219.
- SCHROOTEN C., BOVENHUIS H., COPPIETERS W., ARENDONK J. A. VAN, 2000 Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *J. Dairy Sci.* **83**: 795–806.
- SCHULMAN N. F., SAHANA G., LUND M. S., VIITALA S. M., VILKKI J. H., 2008 Quantitative trait loci for fertility traits in Finnish Ayrshire cattle. *Genet. Sel. Evol.* **40**: 195.

- SCHWARZ E. M., BENZER S., 1997 Calx, a Na-Ca exchanger gene of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 10249–10254.
- SCHWEINGRUBER C., RUFENER S. C., ZÜND D., YAMASHITA A., MÜHLEMANN O., 2013 Nonsense-mediated mRNA decay — Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1829**: 612–623.
- SEELY E. W., 2007 Calcitropic hormones in preeclampsia: a renewal of interest. *J. Clin. Endocrinol. Metab.* **92**: 3402–3403.
- SEIDENSPINNER T., BENNEWITZ J., REINHARDT F., THALLER G., 2009 Need for sharp phenotypes in QTL detection for calving traits in dairy cattle. *J. Anim. Breed. Genet.* **126**: 455–462.
- SEVILLA L. M., NACHAT R., GROOT K. R., KLEMENT J. F., UITTO J., DJIAN P., MAATTA A., WATT F. M., 2007 Mice deficient in involucrin, envoplakin, and periplakin have a defective epidermal barrier. *J. Cell Biol.* **179**: 1599–1612.
- SHARIAT-MADAR Z., MAHDI F., WARNOCK M., HOMEISTER J. W., SRIKANTH S., KRIVANOVSKI Y., MURPHEY L. J., JAFFA A. A., SCHMAIER A. H., 2006 Bradykinin B2 receptor knockout mice are protected from thrombosis by increased nitric oxide and prostacyclin. *Blood* **108**: 192–199.
- SHEA F. F., ROWELL J. L., LI Y., CHANG T.-H., ALVAREZ C. E., 2012 Mammalian Alpha Arrestins Link Activated Seven Transmembrane Receptors to Nedd4 Family E3 Ubiquitin Ligases and Interact with Beta Arrestins. *PLoS ONE* **7**: e50557.
- SHELDON W., 1964 Congenital Pancreatic Lipase Deficiency. *Arch. Dis. Child.* **39**: 268–271.
- SHI H., ROJAS R., BONIFACINO J. S., HURLEY J. H., 2006 The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain. *Nat. Struct. Mol. Biol.* **13**: 540–548.
- SHIMADA T., HASEGAWA H., YAMAZAKI Y., MUTO T., HINO R., TAKEUCHI Y., FUJITA T., NAKAHARA K., FUKUMOTO S., YAMASHITA T., 2004a FGF-23 Is a Potent Regulator of Vitamin D Metabolism and Phosphate Homeostasis. *J. Bone Miner. Res.* **19**: 429–435.
- SHIMADA T., KAKITANI M., YAMAZAKI Y., HASEGAWA H., TAKEUCHI Y., FUJITA T., FUKUMOTO S., TOMIZUKA K., YAMASHITA T., 2004b Targeted ablation of Fgf23 demonstrates an essential physiological role of FGF23 in phosphate and vitamin D metabolism. *J. Clin. Invest.* **113**: 561–568.
- SHIMOMURA Y., WAJID M., PETUKHOVA L., KURBAN M., CHRISTIANO A. M., 2010 Autosomal-dominant woolly hair resulting from disruption of keratin 74 (KRT74), a potential determinant of human hair texture. *Am. J. Hum. Genet.* **86**: 632–638.
- SHIN H. D., PARK B. L., KIM L. H., LEE H.-S., KIM T.-Y., BAE S.-C., 2004 Common DNase I polymorphism associated with autoantibody production among systemic lupus erythematosus patients. *Hum. Mol. Genet.* **13**: 2343–2350.
- SHIRASAWA T., OHNISHI K., HAGIWARA S., SHIGEMOTO K., TAKEBE Y., RAJEWSKY K., TAKEMORI T., 1993 A novel gene product associated with mu chains in immature B cells. *EMBO J.* **12**: 1827–1834.
- SICA G. L., CHOI I.-H., ZHU G., TAMADA K., WANG S.-D., TAMURA H., CHAPOVAL A. I., FLIES D. B., BAJORATH J., CHEN L., 2003 B7-H4, a Molecule of the B7 Family, Negatively Regulates T Cell Immunity. *Immunity* **18**: 849–861.

- SINGH P. P., LUCCHI N. W., BLACKSTOCK A., UDHAYAKUMAR V., SINGH N., 2012 Intervillous macrophage migration inhibitory factor is associated with adverse birth outcomes in a study population in Central India. *PLoS One* **7**: e51678.
- SLAVOTINEK A. M., BARANZINI S. E., SCHANZE D., LABELLE-DUMAIS C., SHORT K. M., CHAO R., YAHYAVI M., BIJLSMA E. K., CHU C., MUSONE S., WHEATLEY A., KWOK P.-Y., MARLES S., FRYNS J.-P., MAGA A. M., HASSAN M. G., GOULD D. B., MADIREDDY L., LI C., COX T. C., SMYTH I., CHUDLEY A. E., ZENKER M., 2011 Manitoba-oculo-tricho-anal (MOTA) syndrome is caused by mutations in *FREM1*. *J. Med. Genet.* **48**: 375–382.
- SMALL S. A., KENT K., PIERCE A., LEUNG C., KANG M. S., OKADA H., HONIG L., VONSATTEL J.-P., KIM T.-W., 2005 Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Ann. Neurol.* **58**: 909–919.
- SMYTH I., DU X., TAYLOR M. S., JUSTICE M. J., BEUTLER B., JACKSON I. J., 2004 The extracellular matrix gene *Frem1* is essential for the normal adhesion of the embryonic epidermis. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 13560–13565.
- SNELLING W. M., ALLAN M. F., KEELE J. W., KUEHN L. A., MCDANELD T., SMITH T. P. L., SONSTEGARD T. S., THALLMAN R. M., BENNETT G. L., 2010 Genome-wide association study of growth in crossbred beef cattle. *J. Anim. Sci.* **88**: 837–848.
- SONSTEGARD T. S., COLE J. B., VANRADEN P. M., TASSELL C. P. VAN, NULL D. J., SCHROEDER S. G., BICKHART D., MCCLURE M. C., 2013 Identification of a Nonsense Mutation in *CWC15* Associated with Decreased Reproductive Efficiency in Jersey Cattle. *PLoS ONE* **8**: e54872.
- SPILIANAKIS C. G., LALIOU M. D., TOWN T., LEE G. R., FLAVELL R. A., 2005 Interchromosomal associations between alternatively expressed loci. *Nature* **435**: 637–645.
- STAUB E., HINZMANN B., ROSENTHAL A., 2002 A novel repeat in the melanoma-associated chondroitin sulfate proteoglycan defines a new protein family. *FEBS Lett.* **527**: 114–118.
- STEINBOCK L., NÄSHOLM A., BERGLUND B., JOHANSSON K., PHILIPSSON J., 2003 Genetic effects on stillbirth and calving difficulty in Swedish Holsteins at first and second calving. *J. Dairy Sci.* **86**: 2228–2235.
- STOTHARD P., CHOI J.-W., BASU U., SUMNER-THOMSON J. M., MENG Y., LIAO X., MOORE S. S., 2011 Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. *BMC Genomics* **12**: 559.
- SUNYAEV S., RAMENSKY V., KOCH I., LATHE W. 3rd, KONDRASHOV A. S., BORK P., 2001 Prediction of deleterious human alleles. *Hum. Mol. Genet.* **10**: 591–597.
- TAKADA Y., NARUSE C., COSTA Y., SHIRAKAWA T., TACHIBANA M., SHARIF J., KEZUKA-SHIOTANI F., KAKIUCHI D., MASUMOTO H., SHINKAI Y., OHBO K., PETERS A. H. F. M., TURNER J. M. A., ASANO M., KOSEKI H., 2011 HP1 γ links histone methylation marks to meiotic synapsis in mice. *Development* **138**: 4207–4217.
- TAKASE K., OHTSUKI T., MIGITA O., TORU M., INADA T., YAMAKAWA-KOBAYASHI K., ARINAMI T., 2001 Association of *ZNF74* gene genotypes with age-at-onset of schizophrenia. *Schizophr. Res.* **52**: 161–165.
- TAKAYANAGI Y., MATSUMOTO H., NAKATA M., MERA T., FUKUSUMI S., HINUMA S., UETA Y., YADA T., LENG G., ONAKA T., 2008 Endogenous prolactin-releasing peptide regulates food intake in rodents. *J. Clin. Invest.* **118**: 4014–4024.

- TAKAYASU S., SAKURAI T., IWASAKI S., TERANISHI H., YAMANAKA A., WILLIAMS S. C., IGUCHI H., KAWASAWA Y. I., IKEDA Y., SAKAKIBARA I., OHNO K., IOKA R. X., MURAKAMI S., DOHMAE N., XIE J., SUDA T., MOTOIKE T., OHUCHI T., YANAGISAWA M., SAKAI J., 2006 A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice. *Proc. Natl. Acad. Sci.* **103**: 7438–7443.
- TAKEDA J., SUZUKI Y., NAKAO M., BARRERO R. A., KOYANAGI K. O., JIN L., MOTONO C., HATA H., ISOGAI T., NAGAI K., OTSUKI T., KURYSHEV V., SHIONYU M., YURA K., GO M., THIERRY-MIEG J., THIERRY-MIEG D., WIEMANN S., NOMURA N., SUGANO S., GOJOBORI T., IMANISHI T., 2006 Large-scale identification and characterization of alternative splicing variants of human gene transcripts using 56,419 completely sequenced and manually annotated full-length cDNAs. *Nucleic Acids Res.* **34**: 3917–3928.
- TAKETOMI Y., UENO N., KOJIMA T., SATO H., MURASE R., YAMAMOTO K., TANAKA S., SAKANAKA M., NAKAMURA M., NISHITO Y., KAWANA M., KAMBE N., IKEDA K., TAGUCHI R., NAKAMIZO S., KABASHIMA K., GELB M. H., ARITA M., YOKOMIZO T., NAKAMURA M., WATANABE K., HIRAI H., NAKAMURA M., OKAYAMA Y., RA C., ARITAKE K., URADE Y., MORIMOTO K., SUGIMOTO Y., SHIMIZU T., NARUMIYA S., HARA S., MURAKAMI M., 2013 Mast cell maturation is driven via a group III phospholipase A2-prostaglandin D2-DP1 receptor paracrine axis. *Nat. Immunol.* **14**: 554–563.
- TANOMURA H., MURAMATSU Y., YAMAMOTO T., OHTA T., KOSE H., YAMADA T., 2011 Difference in PNLIP Allele Frequency Distribution Between High-Marbled and Low-Marbled Cattle. *J. Anim. Vet. Adv.* **10**: 1714–1718.
- TAYLOR P. R., HEYDECK D., JONES G. W., KRONKE G., FUNK C. D., KNAPPER S., ADAMS D., KUHN H., O'DONNELL V. B., 2012 Development of myeloproliferative disease in 12/15-lipoxygenase deficiency. *Blood* **119**: 6173–6174.
- TETU B., BRISSON J., LAPOINTE H., BERNARD P., 1998 Prognostic significance of stromelysin 3, gelatinase A, and urokinase expression in breast cancer. *Hum. Pathol.* **29**: 979–985.
- THALLER G., KÜHN C., WINTER A., EWALD G., BELLMANN O., WEGNER J., ZÜHLKE H., FRIES R., 2003 DGAT1, a new positional and functional candidate gene for intramuscular fat deposition in cattle. *Anim. Genet.* **34**: 354–357.
- THE BOVINE GENOME SEQUENCING AND ANALYSIS CONSORTIUM, ELSIK C. G., TELLAM R. L., WORLEY K. C., 2009 The Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution. *Science* **324**: 522–528.
- THE BOVINE HAPMAP CONSORTIUM, 2009 Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds. *Science* **324**: 528–532.
- THEIN S. L., HESKETH C., TAYLOR P., TEMPERLEY I. J., HUTCHINSON R. M., OLD J. M., WOOD W. G., CLEGG J. B., WEATHERALL D. J., 1990 Molecular basis for dominantly inherited inclusion body beta-thalassemia. *Proc. Natl. Acad. Sci. U. S. A.* **87**: 3924–3928.
- THOMAS K. A., 1987 Fibroblast growth factors. *FASEB J.* **1**: 434–440.
- THOMASEN J. R., GULDBRANDTSEN B., SØRENSEN P., THOMSEN B., LUND M. S., 2008 Quantitative Trait Loci Affecting Calving Traits in Danish Holstein Cattle. *J. Dairy Sci.* **91**: 2098–2105.
- THOMSEN B., HORN P., PANITZ F., BENDIXEN E., PETERSEN A. H., HOLM L.-E., NIELSEN V. H., AGERHOLM J. S., ARNBJERG J., BENDIXEN C., 2006 A missense mutation in the bovine SLC35A3 gene,

- encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. *Genome Res.* **16**: 97–105.
- THORVALDSDÓTTIR H., ROBINSON J. T., MESIROV J. P., 2013 Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**: 178–192.
- TILBEURGH H. VAN, EGLOFF M. P., MARTINEZ C., RUGANI N., VERGER R., CABBILLAU C., 1993 Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature* **362**: 814–820.
- TOMMERUP N., VISSING H., 1995 Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identify putative candidate genes for developmental and malignant disorders. *Genomics* **27**: 259–264.
- UDLER M. S., TYRER J., EASTON D. F., 2010 Evaluating the power to discriminate between highly correlated SNPs in genetic association studies. *Genet. Epidemiol.* **34**: 463–468.
- URAKAWA I., YAMAZAKI Y., SHIMADA T., IJIMA K., HASEGAWA H., OKAWA K., FUJITA T., FUKUMOTO S., YAMASHITA T., 2006 Klotho converts canonical FGF receptor into a specific receptor for FGF23. *Nature* **444**: 770–774.
- URAYAMA K., GUILINI C., TURKERI G., TAKIR S., KUROSE H., MESSADDEQ N., DIERICH A., NEBIGIL C. G., 2008 Prokineticin receptor-1 induces neovascularization and epicardial-derived progenitor cell differentiation. *Arterioscler. Thromb. Vasc. Biol.* **28**: 841–849.
- UZUMCU A., NORGETT E. E., DINDAR A., UYGUNER O., NISLI K., KAYSERILI H., SAHIN S. E., DUPONT E., SEVERNS N. J., LEIGH I. M., YUKSEL-APAK M., KELSELL D. P., WOLLNIK B., 2006 Loss of desmoplakin isoform I causes early onset cardiomyopathy and heart failure in a Naxos-like syndrome. *J. Med. Genet.* **43**: e05–e05.
- VAIDEHI N., FLORIANO W. B., TRABANINO R., HALL S. E., FREDDOLINO P., CHOI E. J., ZAMANAKOS G., GODDARD W. A., 2002 Prediction of structure and function of G protein-coupled receptors. *Proc. Natl. Acad. Sci.* **99**: 12622–12627.
- VALENTONYTE R., HAMPE J., HUSE K., ROSENSTIEL P., ALBRECHT M., STENZEL A., NAGY M., GAEDE K. I., FRANKE A., HAESLER R., KOCH A., LENGAUER T., SEEGERT D., REILING N., EHLERS S., SCHWINGER E., PLATZER M., KRAWCZAK M., MÜLLER-QUERNHEIM J., SCHÜRSMANN M., SCHREIBER S., 2005 Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat. Genet.* **37**: 357–364.
- VANRADEN P. M., 2008 Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* **91**: 4414–4423.
- VANRADEN P. M., MILLER R. H., 2006 Effects of Nonadditive Genetic Interactions, Inbreeding, and Recessive Defects on Embryo and Fetal Loss by Seventy Days. *J. Dairy Sci.* **89**: 2716–2721.
- VANRADEN P. M., O'CONNELL J. R., WIGGANS G. R., WEIGEL K. A., 2011a Genomic evaluations with many more genotypes. *Genet. Sel. Evol. GSE* **43**: 10.
- VANRADEN P. M., OLSON K. M., NULL D. J., HUTCHISON J. L., 2011b Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J. Dairy Sci.* **94**: 6153–6161.
- VARNUM D. S., FOX S. C., 1981 Head blebs: a new mutation on chromosome 4 of the mouse. *J. Hered.* **72**: 293–293.

- VERPY E., LEIBOVICI M., MICHALSKI N., GOODYEAR R. J., HOUDON C., WEIL D., RICHARDSON G. P., PETIT C., 2011 Stereocilin connects outer hair cell stereocilia to one another and to the tectorial membrane. *J. Comp. Neurol.* **519**: 194–210.
- VERPY E., MASMOUDI S., ZWAENEPOEL I., LEIBOVICI M., HUTCHIN T. P., CASTILLO I. DEL, NOUAILLE S., BLANCHARD S., LAINÉ S., POPOT J. L., MORENO F., MUELLER R. F., PETIT C., 2001 Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. *Nat. Genet.* **29**: 345–349.
- VERPY E., WEIL D., LEIBOVICI M., GOODYEAR R. J., HAMARD G., HOUDON C., LEFÈVRE G. M., HARDELIN J.-P., RICHARDSON G. P., AVAN P., PETIT C., 2008 Stereocilin-deficient mice reveal the origin of cochlear waveform distortions. *Nature* **456**: 255–258.
- VIITALA S., SZYDA J., BLOTT S., SCHULMAN N., LIDAUER M., MÄKI-TANIILA A., GEORGES M., VILKKI J., 2006 The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle. *Genetics* **173**: 2151–2164.
- VILARIÑO-GÜELL C., WIDER C., ROSS O. A., DACHSEL J. C., KACHERGUS J. M., LINCOLN S. J., SOTO-ORTOLAZA A. I., COBB S. A., WILHOITE G. J., BACON J. A., BEHROUZ B., MELROSE H. L., HENTATI E., PUSCHMANN A., EVANS D. M., CONIBEAR E., WASSERMAN W. W., AASLY J. O., BURKHARD P. R., DJALDETTI R., GHICA J., HENTATI F., KRYGOWSKA-WAJS A., LYNCH T., MELAMED E., RAJPUT A., RAJPUT A. H., SOLIDA A., WU R.-M., UTTI R. J., WSZOLEK Z. K., VINGERHOETS F., FARRER M. J., 2011 VPS35 Mutations in Parkinson Disease. *Am. J. Hum. Genet.* **89**: 162–167.
- VISSCHER P. M., GODDARD M. E., 2004 Prediction of the confidence interval of quantitative trait Loci location. *Behav. Genet.* **34**: 477–482.
- VISSERS L. E. L. M., COX T. C., MAGA A. M., SHORT K. M., WIRADJAJA F., JANSSEN I. M., JEHEE F., BERTOLA D., LIU J., YAGNIK G., SEKIGUCHI K., KIYOZUMI D., BOKHOVEN H. VAN, MARCELIS C., CUNNINGHAM M. L., ANDERSON P. J., BOYADJIEV S. A., PASSOS-BUENO M. R., VELTMAN J. A., SMYTH I., BUCKLEY M. F., ROSCIOLI T., 2011 Heterozygous mutations of FREM1 are associated with an increased risk of isolated metopic craniosynostosis in humans and mice. *PLoS Genet.* **7**: e1002278.
- VISSING H., MEYER W. K.-H., AAGAARD L., TOMMERUP N., THIESEN H.-J., 1995 Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Lett.* **369**: 153–157.
- VONDERHEIDE R. H., TEDDER T. F., SPRINGER T. A., STAUNTON D. E., 1994 Residues within a conserved amino acid motif of domains 1 and 4 of VCAM-1 are required for binding to VLA-4. *J. Cell Biol.* **125**: 215–222.
- WANG Z., FANG B., CHEN J., ZHANG X., LUO Z., HUANG L., CHEN X., LI Y., 2010 De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). *BMC Genomics* **11**: 726.
- WANG X., WURMSER C., PAUSCH H., JUNG S., REINHARDT F., TETENS J., THALLER G., FRIES R., 2012 Identification and Dissection of Four Major QTL Affecting Milk Fat Content in the German Holstein-Friesian Population. *PLoS ONE* **7**: e40711.
- WANG Y., ZHANG H.-X., SUN Y.-P., LIU Z.-X., LIU X.-S., WANG L., LU S.-Y., KONG H., LIU Q.-L., LI X.-H., LU Z.-Y., CHEN S.-J., CHEN Z., BAO S.-S., DAI W., WANG Z.-G., 2007 Rig-I^{-/-} mice

- develop colitis associated with downregulation of G alpha i2. *Cell Res.* **17**: 858–868.
- WASIF N., NAQVI S. K. ul-H., BASIT S., ALI N., ANSAR M., AHMAD W., 2011 Novel mutations in the keratin-74 (KRT74) gene underlie autosomal dominant woolly hair/hypotrichosis in Pakistani families. *Hum. Genet.* **129**: 419–424.
- WEI L., SHI Y.-B., 2005 Matrix metalloproteinase stromelysin-3 in development and pathogenesis. *Histol. Histopathol.* **20**: 177–185.
- WEKSBERG R., SHUMAN C., BECKWITH J. B., 2009 Beckwith–Wiedemann syndrome. *Eur. J. Hum. Genet.* **18**: 8–14.
- WELLER J. I., KASHI Y., SOLLER M., 1990 Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* **73**: 2525–2537.
- WEN L., TANG F.-L., HONG Y., LUO S.-W., WANG C.-L., HE W., SHEN C., JUNG J.-U., XIONG F., LEE D., ZHANG Q.-G., BRANN D., KIM T.-W., YAN R., MEI L., XIONG W.-C., 2011 VPS35 haploinsufficiency increases Alzheimer's disease neuropathology. *J. Cell Biol.* **195**: 765–779.
- WETTERSTRAND K., 2014 DNA Sequencing costs: data from the NHGRI Large-Scale Genome Sequencing Program. <http://www.genome.gov/sequencingcosts/>.
- WINKLER F. K., D'ARCY A., HUNZIKER W., 1990 Structure of human pancreatic lipase. *Nature* **343**: 771–774.
- WINTER A., KRÄMER W., WERNER F. A. O., KOLLERS S., KATA S., DURSTEWITZ G., BUITKAMP J., WOMACK J. E., THALLER G., FRIES R., 2002 Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. *Proc. Natl. Acad. Sci.* **99**: 9300–9305.
- WITZGALL R., O'LEARY E., LEAF A., ONALDI D., BONVENTRE J. V., 1994 The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl. Acad. Sci. U. S. A.* **91**: 4514–4518.
- WOLFF M. VON, URSEL S., HAHN U., STELDINGER R., STROWITZKI T., 2003 Glucose transporter proteins (GLUT) in human endometrium: expression, regulation, and function throughout the menstrual cycle and in early pregnancy. *J. Clin. Endocrinol. Metab.* **88**: 3885–3892.
- WOOD J. D., RICHARDSON R. I., NUTE G. R., FISHER A. V., CAMPO M. M., KASAPIDOU E., SHEARD P. R., ENSER M., 2004 Effects of fatty acids on meat quality: a review. *Meat Sci.* **66**: 21–32.
- WOODING S., KIM U.-K., BAMSHAD M. J., LARSEN J., JORDE L. B., DRAYNA D., 2004 Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. *Am. J. Hum. Genet.* **74**: 637–646.
- WRAY N. R., YANG J., HAYES B. J., PRICE A. L., GODDARD M. E., VISSCHER P. M., 2013 Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* **14**: 507–515.
- WU Y., PI J. S., PAN A. L., PU Y. J., DU J. P., SHEN J., LIANG Z. H., ZHANG J. R., 2012 An SNP in the MyoD1 gene intron 2 associated with growth and carcass traits in three duck populations. *Biochem. Genet.* **50**: 898–907.
- XIA W.-F., TANG F.-L., XIONG L., XIONG S., JUNG J.-U., LEE D.-H., LI X.-S., FENG X., MEI L., XIONG W.-C., 2013 Vps35 loss promotes hyperresorptive osteoclastogenesis and osteoporosis

- via sustained RANKL signaling. *J. Cell Biol.* **200**: 821–837.
- XUE Y., DALY A., YNGVADOTTIR B., LIU M., COOP G., KIM Y., SABETI P., CHEN Y., STALKER J., HUCKLE E., BURTON J., LEONARD S., ROGERS J., TYLER-SMITH C., 2006 Spread of an Inactive Form of Caspase-12 in Humans Is Due to Recent Positive Selection. *Am. J. Hum. Genet.* **78**.
- YAMAGUCHI M., SAKATA D. M., OGURA K., MIYAKE A., 1996 Gestational changes of glucose transporter gene expression in the mouse placenta and decidua. *J. Endocrinol. Invest.* **19**: 567–569.
- YAMAGUCHI-KABATA Y., SHIMADA M. K., HAYAKAWA Y., MINOSHIMA S., CHAKRABORTY R., GOJOBORI T., IMANISHI T., 2008 Distribution and Effects of Nonsense Polymorphisms in Human Genes. *PLoS ONE* **3**: e3393.
- YAMAZAKI Y., TAMADA T., KASAI N., URAKAWA I., AONO Y., HASEGAWA H., FUJITA T., KUROKI R., YAMASHITA T., FUKUMOTO S., SHIMADA T., 2008 Anti-FGF23 neutralizing antibodies show the physiological role and structural features of FGF23. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **23**: 1509–1518.
- YANG J.-J., 2003 A novel zinc finger protein, ZZaPK, interacts with ZAK and stimulates the ZAK-expressing cells re-entering the cell cycle. *Biochem. Biophys. Res. Commun.* **301**: 71–77.
- YANG J., LEE S. H., GODDARD M. E., VISSCHER P. M., 2011 GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* **88**: 76–82.
- YASUTOMO K., HORIUCHI T., KAGAMI S., TSUKAMOTO H., HASHIMURA C., URUSHIHARA M., KURODA Y., 2001 Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat. Genet.* **28**: 313–314.
- YNGVADOTTIR B., XUE Y., SEARLE S., HUNT S., DELGADO M., MORRISON J., WHITTAKER P., DELOUKAS P., TYLER-SMITH C., 2009 A Genome-wide Survey of the Prevalence and Evolutionary Forces Acting on Human Nonsense SNPs. *Am. J. Hum. Genet.* **84**: 224–234.
- YOCHUM C. L., BHATTACHARYA P., PATTI L., MIROCHNITCHENKO O., WAGNER G. C., 2010 Animal model of autism using GSTM1 knockout mice and early post-natal sodium valproate treatment. *Behav. Brain Res.* **210**: 202–210.
- YORK I. A., BREHM M. A., ZENDZIAN S., TOWNE C. F., ROCK K. L., 2006 Endoplasmic reticulum aminopeptidase 1 (ERAP1) trims MHC class I-presented peptides in vivo and plays an important role in immunodominance. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 9202–9207.
- ZANG X., LOKE P., KIM J., MURPHY K., WAITZ R., ALLISON J. P., 2003 B7x: A widely expressed B7 family member that inhibits T cell activation. *Proc. Natl. Acad. Sci.* **100**: 10388–10392.
- ZEGGINI E., 2011 Next-generation association studies for complex traits. *Nat Genet* **43**: 287–288.
- ZHAN B., FADISTA J., THOMSEN B., HEDEGAARD J., PANITZ F., BENDIXEN C., 2011 Global assessment of genomic variation in cattle by genome resequencing and high-throughput genotyping. *BMC Genomics* **12**: 557.
- ZHANG S., KUBOTA C., YANG L., ZHANG Y., PAGE R., O'NEILL M., YANG X., TIAN X. C., 2004 Genomic Imprinting of H19 in Naturally Reproduced and Cloned Cattle. *Biol. Reprod.* **71**: 1540–1544.
- ZHANG Y., MALEKPOUR M., AL-MADANI N., KAHRIZI K., ZANGANEH M., LOHR N. J., MOHSENI M., MOJAHEDI F., DANESHI A., NAJMABADI H., SMITH R. J. H., 2007 Sensorineural deafness and

- male infertility: a contiguous gene deletion syndrome. *J. Med. Genet.* **44**: 233–240.
- ZHANG Z., XIN D., WANG P., ZHOU L., HU L., KONG X., HURST L. D., 2009 Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay. *BMC Biol.* **7**: 23.
- ZHANG P., YU L., GAO J., FU Q., DAI F., ZHAO Y., ZHENG L., ZHAO S., 2000 Cloning and characterization of human VPS35 and mouse Vps35 and mapping of VPS35 to human chromosome 16q13-q21. *Genomics* **70**: 253–257.
- ZHU G., AUGUSTINE M. M., AZUMA T., LUO L., YAO S., ANAND S., RIETZ A. C., HUANG J., XU H., FLIES A. S., FLIES S. J., TAMADA K., COLONNA M., DEURSEN J. M. A. VAN, CHEN L., 2009 B7-H4-deficient mice display augmented neutrophil-mediated innate immunity. *Blood* **113**: 1759–1767.
- ZIMIN A. V., DELCHER A. L., FLOREA L., KELLEY D. R., SCHATZ M. C., PUIU D., HANRAHAN F., PERTEA G., TASSELL C. P. VAN, SONSTEGARD T. S., MARÇAIS G., ROBERTS M., SUBRAMANIAN P., YORKE J. A., SALZBERG S. L., 2009 A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* **10**: R42.
- ZIMPRICH A., BENET-PAGÈS A., STRUHAL W., GRAF E., ECK S. H., OFFMAN M. N., HAUBENBERGER D., SPIELBERGER S., SCHULTE E. C., LICHTNER P., ROSSLE S. C., KLOPP N., WOLF E., SEPPI K., PIRKER W., PRESSLAUER S., MOLLENHAUER B., KATZENSCHLAGER R., FOKI T., HOTZY C., REINTHALER E., HARUTYUNYAN A., KRALOVICS R., PETERS A., ZIMPRICH F., BRÜCKE T., POEWE W., AUFF E., TRENKWALDER C., ROST B., RANSMAYR G., WINKELMANN J., MEITINGER T., STROM T. M., 2011 A Mutation in VPS35, Encoding a Subunit of the Retromer Complex, Causes Late-Onset Parkinson Disease. *Am. J. Hum. Genet.* **89**: 168–175.

9 Anhang

Tabelle 9.1: Ausgewählte Tiere zur genomweiten Re-Sequenzierung

Bam-ID	Lab-ID	Hit-Nr.	Interbull-ID	Tiername	Rasse	Geb.jahr	Interbull-ID Vater
55069	7515	756711620016730	SIMCHEM711620016730	REDAD	BFV	1973	REDUSAM000001599157
55070	8145	276000929276244	SIMDEUM000929276244	STREIF	BFV	1978	SIMDEUM000929081239
55071	8147	276000929189864	SIMDEUM000929189864	ROMULUS	BFV	1975	SIMDEUM000929015959
55072	8186	276000912851233	SIMDEUM000912851233	HORWEIN	BFV	1986	SIMDEUM000809706945
55073	8146	276000925265555	SIMDEUM000925265555	BALBO	BFV	1974	SIMDEUM000926800806
55074	8148	276000910915308	SIMDEUM000910915308	EGOL	BFV	1991	SIMDEUM000915512806
55075	7556	276000928504510	SIMDEUM000928504510	RENNER	BFV	1981	SIMCHEM711620016730
55076	8149	276000913892370	SIMDEUM000913892370	SAMURAI	BFV	1992	SIMDEUM000914240004
55077	8232	276000926428551	SIMDEUM000926428551	POSTILLION	BFV	1981	SIMDEUM000925714495
55078	8198	276000910794565	SIMDEUM000910794565	GEBAL	BFV	1992	SIMDEUM000916835898
55687	8243	276000910865689	SIMDEUM000910865689	ROMANEK	BFV	1989	SIMDEUM000927583034
55689	8246	276007600026785	SIMDEUM007600026785	HOLZER	BFV	1985	SIMDEUM007600020383
55690	8247	276000918174246	SIMDEUM000918174246	REGIO	BFV	1996	SIMDEUM000915732780
55691	8248	276000911474933	SIMDEUM000911474933	BOSS	BFV	1993	SIMDEUM000914411965
55692	8249	276000913008210	SIMDEUM000913008210	REXON	BFV	1989	SIMDEUM000925973831
55693	8250	276000912971290	SIMDEUM000912971290	ROMEN	BFV	1988	SIMDEUM000925978871
55694	8251	276000926189400	SIMDEUM000926189400	DIRTECK	BFV	1983	SIMDEUM000924326746
55695	8252	276000911825633	SIMDEUM000911825633	RALBO	BFV	1988	SIMDEUM000925952548
55696	8254	276000919587582	SIMDEUM000919587582	HALLING	BFV	1988	SIMDEUM000925964575
55697	8255	276000913325437	SIMDEUM000913325437	POLDI	BFV	1996	SIMDEUM000917355651
55698	8256	276000910950070	SIMDEUM000910950070	SPORT	BFV	1992	SIMDEUM000919554268
56503	8421	276000915581022	SIMDEUM000915581022	HORB	BFV	1986	SIMDEUM000809706945
56504	8422	276000918912889	SIMDEUM000918912889	HUMID	BFV	1996	SIMDEUM000915262687
56505	8423	276000915040032	SIMDEUM000915040032	HUMLANG	BFV	1994	SIMDEUM000915262687
56506	8424	276000920747734	SIMDEUM000920747734	PROPELLER	BFV	1981	SIMDEUM000925714495
56507	8425	276000919598352	SIMDEUM000919598352	REPORT	BFV	1989	SIMDEUM000928504510
56508	8426	276000919253926	SIMDEUM000919253926	UTNACH	BFV	1991	SIMFRAM001083000111
56509	8427	276000915732780	SIMDEUM000915732780	RENGER	BFV	1989	SIMDEUM000928504510
58941	8510	276000809706945	SIMDEUM000809706945	HORROR	BFV	1979	SIMDEUM000809094147
58943	8511	276000804179455	SIMDEUM000804179455	HODSCHA	BFV	1983	SIMDEUM000805465523
58944	8512	276000912851741	SIMDEUM000912851741	MORWEL	BFV	1989	SIMAUTM000842871443
58945	8674	276000979317838	SIMDEUM000979317838	HAXL	BFV	1966	SIMDEUM000970388123
58946	8785	276000932739095	SIMDEUM000932739095	WATERBERG	BFV	1999	SIMDEUM000918035013
58947	8786	276000935684041	SIMDEUM000935684041	MANDELA	BFV	2001	SIMDEUM000915079575
58948	8787	276000931098195	SIMDEUM000931098195	MARTL-DE	BFV	1999	SIMAUTM000040568233
58949	8788	276000933038755	SIMDEUM000933038755	ZAHNER	BFV	2000	SIMDEUM000915393011
58950	8789	276000938263111	SIMDEUM000938263111	DAGO	BFV	2004	SIMAUTM000447242233
58951	9065	276000933663105	SIMDEUM000933663105	WEINOLD	BFV	1999	SIMDEUM000919322930
58952	8791	276000940823906	SIMDEUM000940823906	MORGENGRAU	BFV	2006	SIMDEUM000934225983
58953	G4	276000943707915	SIMDEUM000943707915	G4	BFV	2009	SIMDEUM000941809001
65649	9241	276000911331078	SIMDEUM000911331078	HODACH	BFV	1989	SIMDEUM000804179455
65650	9242	276000911485840	SIMDEUM000911485840	ZASTER	BFV	1994	SIMDEUM000917845051
65651	9243	276000574600534	SIMAUTM000574600534	MAXI	BFV	1999	SIMAUTM000040568233
65652	9244	276000934127155	SIMDEUM000934127155	HILAB	BFV	2000	SIMDEUM000919350741
65653	9245	40000927146534	SIMAUTM000927146534	LEO	BFV	1998	SIMAUTM000541950446
65654	9246	276000918555090	SIMDEUM000918555090	RANDY	BFV	1994	SIMDEUM000911825633
65655	9247	276000918892486	SIMDEUM000918892486	MALIST	BFV	1996	SIMAUTM000040568233
65656	9248	276000941115399	SIMDEUM000941115399	HEROCK	BFV	2007	SIMAUTM000671844642
65657	TCP_13	276000946485312	SIMDEUM000946485312	TCP_13	BFV	2011	SIMDEUM000936487534

Anhang

Bam-ID	Lab-ID	Hit-Nr.	Interbull-ID	Tiername	Rasse	Geb.jahr	Interbull-ID Vater
65658	9306	276000340174036	REDDEUM000340174036	FABER	RED	1996	HOLUSAM000002129913
65659	9299	276001015194625	HOLDEUM001015194625	SIoux	HOL	1991	HOLUSAM000001964484
65660	NA	276000577745977	REDDEUM000577745977	GOLDSTAR	RED	1997	REDNLDM000706570603
65661	9300	276001015216836	HOLDEUM001015216836	MAGNUM	HOL	1992	HOLUSAM000002020049
65663	NA	124000006961162	HOLCANM000006961162	CHAMPION	HOL	1997	HOLCANM000005470579
66498	9425	276000578891748	HOLDEUM000578891748	MASCOL	HOL	2000	HOLITAM006001001962
66499	9289	276002261530135	HOLDEUM002261530135	MANAT	HOL	1995	HOLUSAM000002119526
66500	9290	528000255206543	HOLNLDM000255206543	ALVES	HOL	1999	HOLUSAM000002231596
66501	9291	276001020186127	HOLDEUM001020186127	TOBER	HOL	1984	HOLUSAM000001721333
66502	9292	276001021098657	HOLDEUM001021098657	BONATUS	HOL	1991	HOLUSAM000001929410
66503	9293	276001020477549	HOLDEUM001020477549	STATUS	HOL	1986	HOLCANM000000352790
66504	9294	276000345785578	REDDEUM000345785578	LICHTBLICK	RED	2000	HOLDEUM001012337003
66505	9295	276000348565591	HOLDEUM000348565591	MERGIM	HOL	2004	HOLCHEM310037833550
66506	9296	276000345461818	HOLDEUM000345461818	JOSE	HOL	2001	HOLFRAM005694028588
66507	9297	276000347213836	HOLDEUM000347213836	WIZZARD	HOL	2000	HOLNLDM000159659261
66508	9298	276001015148130	HOLDEUM001015148130	PIT	HOL	1989	HOLUSAM000001878472
66511	9301	276000578448776	HOLDEUM000578448776	LAUDAN	HOL	1998	HOLDEUM001012104370
66512	9302	840000017088070	HOLUSAM000017088070	TICKET	HOL	1996	HOLUSAM000002122404
66513	9303	276000341621945	HOLDEUM000341621945	BLACKMAN	HOL	1997	HOLDEUM001020708373
66514	9304	276000577604324	HOLDEUM000577604324	ZECHER	HOL	1997	HOLUSAM000002137511
66515	9305	250007297006288	HOLFRAM007297006288	GIBOR	HOL	1997	HOLFRAM002991000305
66517	9307	276000578720303	REDDEUM000578720303	LUDOX	RED	1999	HOLUSAM000002228267
66518	9308	276000576536679	REDDEUM000576536679	ORIGIN	RED	1995	REDDEUM000505639209
66519	9309	276000578889436	REDDEUM000578889436	CARMANO	RED	2001	HOLUSAM000002252648
66520	9310	276001012261613	HOLDEUM001012261613	UBBO	HOL	1994	HOLUSAM000001920807
66521	9311	276001012056808	HOLDEUM001012056808	LASSO	HOL	1990	HOLUSAM000001926224
66522	NA	756710488074272	SIMCHEM710488074272	JOYEUX	SIM	1977	MONFRAM000000160839
66524	NA	756712006074573	SIMCHEM712006074573	FLEURON	SIM	1994	SIMCHEM712820066389
68262	9312	276001500058043	HOLDEUM001500058043	ZUNDER	HOL	1996	HOLUSAM000002156938
68263	9313	276000102650777	HOLDEUM000102650777	LADIN	HOL	1994	HOLUSAM000002071864
68264	9314	276000101544835	HOLDEUM000101544835	PATRICK	HOL	1990	HOLUSAM000001929410
68265	9316	276000340930144	HOLDEUM000340930144	DOLCH	HOL	1996	HOLNLDM000316418898
68266	9317	276001013278183	HOLDEUM001013278183	TRAILOR	HOL	1989	HOLUSAM000001882797
68267	9318	276001013340602	HOLDEUM001013340602	AMARETTO	HOL	1991	HOLCANM000000383622
68268	9320	756110480021128	BSWCHEM110480021128	VERO	BSW	1970	BSWCHEM110191090826
68269	9321	756111512030200	BSWCHEM111512030200	HORST	BSW	1972	BSWCHEM110023045482
68270	9322	756110706016501	BSWCHEM110706016501	ALBO	BSW	1973	BSWUSAM000000138750
68271	9323	756110471084194	BSWCHEM110471084194	OREON	BSW	1974	BSWUSAM000000150198
68272	9324	756110269048407	BSWCHEM110269048407	AXIUS	BSW	1975	BSWUSAM000000139618
68273	9325	756110518030504	BSWCHEM110518030504	NORDAL	BSW	1977	BSWUSAM000000152118
68274	9326	756110097189518	BSWCHEM110097189518	NOVIO	BSW	1977	BSWUSAM000000156006
68275	9327	756110556008244	BSWCHEM110556008244	ZOLL	BSW	1977	BSWUSAM000000147207
68276	9328	756111030024286	BSWCHEM111030024286	BRYAN	BSW	1979	BSWUSAM000000159771
68277	9329	756110008038300	BSWCHEM110008038300	WINNETOU	BSW	1979	BSWCHEM110487006746
68278	9330	756110098063770	BSWCHEM110098063770	NATO	BSW	1980	BSWUSAM000000153272
68279	9331	756110482036847	BSWCHEM110482036847	ASCOT	BSW	1982	BSWUSAM000000172466
68280	9332	756110203063244	BSWCHEM110203063244	JOSS	BSW	1982	BSWUSAM000000161170
68281	9333	756110501069856	BSWCHEM110501069856	LEADER	BSW	1982	BSWCHEM110044064448
68282	9334	756110098071188	BSWCHEM110098071188	JORAT	BSW	1985	BSWUSAM000000170157
68283	9335	756110374046077	BSWCHEM110374046077	BELL-ET	BSW	1987	BSWUSAM000000182143
68284	9336	840000001841381	BSWUSAM000000184138	ENSIGN-ET	BSW	1987	BSWUSAM000000176173
68285	9337	756110919005798	BSWCHEM110919005798	IMPEROR	BSW	1987	BSWUSAM000000179556
68286	9338	840000001840872	BSWUSAM000000184087	STARBUCK	BSW	1987	BSWUSAM000000175751
68287	9339	756110502100268	BSWCHEM110502100268	ENORM	BSW	1988	BSWUSAM000000182043
68288	9340	840000001862768	BSWUSAM000000186276	PROPHET-ET	BSW	1989	BSWUSAM000000179801
68289	9341	840000001873610	BSWUSAM000000187361	EMERALD-DE	BSW	1990	BSWUSAM000000181329
68290	9342	756110106053793	BSWCHEM110106053793	MEDOR	BSW	1990	BSWCHEM110060127370

Anhang

Bam-ID	Lab-ID	Hit-Nr.	Interbull-ID	Tiername	Rasse	Geb.jahr	Interbull-ID Vater
68291	9343	756111453045080	BSWCHEM111453045080	NIPPEL	BSW	1990	BSWCHEM110716027719
68292	9344	840000001891829	BSWUSAM000000189182	COLLECTION	BSW	1992	BSWUSAM000000181217
68293	9345	840000001897975	BSWUSAM000000189797	CHIME-ET	BSW	1993	BSWUSAM000000185301
68294	9346	756110023148411	BSWCHEM110023148411	CHRISTO	BSW	1993	BSWUSAM000000181972
68295	9347	756110040047292	BSWCHEM110040047292	VOERIS	BSW	1994	BSWCHEM110719023268
68296	9348	756110043090653	BSWCHEM110043090653	GRAL	BSW	1995	BSWCHEM110054337594
68297	9349	756110057253457	BSWCHEM110057253457	KASTOR	BSW	1998	BSWCHEM110120063693
68298	9350	756110121201483	BSWCHEM110121201483	ZEUS	BSW	1999	BSWTTAMVI00000121852
68299	9351	756120000414675	BSWCHEM120000414675	JOLDEN-ET	BSW	2000	BSWUSAM000000189181
68300	9352	840000001956184	BSWUSAM000000195618	VIGOR-ET	BSW	2001	BSWUSAM000000191215
68301	9353	756713388027539	SIMCHEM713388027539	AELPLER	SIM	1966	SIMCHEM714800844437
68303	9354	756713065011578	SIMCHEM713065011578	VERGIL	SIM	1969	SIMCHEM715585015531
68305	9355	840000016293915	REDUSAM000001629391	TRIPLE	RED	1972	HOLCANM000000288790
68306	9356	756713775055619	SIMCHEM713775055619	AELPLER	SIM	1973	SIMCHEM716685030974
68307	9357	840000016509436	REDUSAM000001650943	FIRESTAR	RED	1973	HOLUSAM000001430145
68309	9358	756716395019139	SIMCHEM716395019139	TIM	SIM	1983	REDCANM000000311569
68310	9359	756716918033765	SIMCHEM716918033765	RIVAROL	SIM	1974	REDCANM000000311569
68311	9360	756713620026917	SIMCHEM713620026917	MARS	SIM	1974	SIMCHEM713300019239
68312	9361	756714610076769	REDCHEM714610076769	JEROME	RED	1976	REDUSAM000001599157
68313	9362	840000000222358	REDUSAM000000022235	TONTO	RED	1979	HOLUSAM000001620273
68314	9363	840000018423716	REDUSAM000001842371	CAVEMAN	RED	1980	HOLUSAM000001620273
68315	9364	756712154013875	SIMCHEM712154013875	KANIS	SIM	1980	SIMCHEM712856045518
68316	9365	756716955044748	SIMCHEM716955044748	CONDOR	SIM	1982	SIMCHEM713190071003
68317	9366	756712237013952	REDCHEM712237013952	BEMOL	RED	1982	REDUSAM000001826808
68318	9367	276000002463460	HOLDEUM001004565147	ORKAN	HOL	1982	HOLUSAM000001672151
68319	9368	840000018807783	REDCANM000000392953	FIRE	RED	1982	REDCANM000000352882
68320	9369	756714495058218	REDCHEM714495058218	JONN-ET	RED	1983	HOLUSAM000001695963
68321	9370	124000003764557	HOLCANM000000376455	JUBILANT-E	HOL	1983	REDUSAM000001629391
68322	9371	840000019737119	REDUSAM000001973711	SABER-ET	RED	1984	HOLUSAM000001672325
68323	9372	840000003871781	REDCANM000000387178	ROCHESTER	RED	1985	HOLUSAM000001710057
68324	9373	840000020901554	REDUSAM000002090155	SWATCH-ET	RED	1988	HOLUSAM000001857259
68325	9374	840000021760617	REDUSAM000002176061	HANS-ET	RED	1991	REDCANM000000381413
68326	9375	756310048190307	HOLCHEM310048190307	RIOSTAR	HOL	1991	HOLUSAM000001929410
68327	9376	756716955055126	SIMCHEM716955055126	FILOU	SIM	1991	SIMCHEM713190091124
68328	9377	756716575029323	SIMCHEM716575029323	TIPIC	SIM	1993	REDCHEM712856067770
68329	9378	756713191046185	SIMCHEM713191046185	BELLO	SIM	1993	SIMCHEM713191039323
68330	9379	756714800037525	REDCHEM714800037525	BACCALA-ET	RED	1994	REDNLDM000316124490
68331	9380	756712455090575	SIMCHEM712455090575	AELPLER_3	SIM	1995	SIMCHEM712139067572
68332	9381	756120019078516	REDCHEM120019078516	INCAS	RED	2001	REDUSAM000128291546
9315	9315	528000295614731	HOLNLDM000295614731	STYLIST	HOL	2002	HOLUSAM000002282997
9610	9610	276000930988845	SIMDEUM000930988845	SAMUT	BFV	1998	SIMDEUM000913892370
9611	9611	276000931055961	SIMDEUM000931055961	SAFIR	BFV	1998	SIMDEUM000913892370
9612	9612	276000932408546	SIMDEUM000932408546	SAMPRAS	BFV	1998	SIMDEUM000913892370
9613	9613	276000932059928	SIMDEUM000932059928	REPTEIT	BFV	1998	SIMDEUM000919598352
9614	9614	276000811017046	SIMDEUM000811017046	ROMSEL	BFV	1998	SIMDEUM000912971290
9615	9615	276000932204616	SIMDEUM000932204616	ROBERTO	BFV	1998	SIMDEUM000919598352
9616	9616	276000932627221	SIMDEUM000932627221	RAINER	BFV	1999	SIMDEUM000911455235
9617	9617	276000918923365	SIMDEUM000918923365	MADERA	BFV	1997	SIMAUTM000040568233
9618	9618	276000933223030	SIMDEUM000933223030	MORBO	BFV	2000	SIMDEUM000808038600
9619	9619	276000915508121	SIMDEUM000915508121	HATZ	BFV	1996	SIMDEUM000911331078
9620	9620	276000932878499	SIMDEUM000932878499	HULOCK	BFV	1999	SIMDEUM000910789350
9621	9621	276000932785883	SIMDEUM000932785883	GEBALOT	BFV	1999	SIMDEUM000910794565
9623	9623	276000918162560	SIMDEUM000918162560	RUBEL	BFV	1996	SIMDEUM000915732780
9624	9624	276000937295834	SIMDEUM000937295834	POLK	BFV	2003	SIMDEUM000913325437
9625	9625	276000933855292	SIMDEUM000933855292	RALSTER	BFV	2000	SIMDEUM000912121855
9626	9626	276000918197923	SIMDEUM000918197923	WAGRAIN PS	BFV	2006	SIMDEUM000934492505
9705	9705	276000932934199	SIMDEUM000932934199	BOSBO	BFV	2000	SIMDEUM000911474933

Anhang

Bam-ID	Lab-ID	Hit-Nr.	Interbull-ID	Tiername	Rasse	Geb.jahr	Interbull-ID Vater
9706	9706	276000918966927	SIMDEUM000918966927	ERGO	BFV	1997	SIMDEUM000914750324
9707	9707	276000916200528	SIMDEUM000916200528	ENRICO	BFV	1997	SIMDEUM000910915308
9708	9708	276000932657399	SIMDEUM000932657399	HIRMER	BFV	1999	SIMDEUM000919350741
9709	9709	276000936580580	SIMDEUM000936580580	HIMALAYA	BFV	2002	SIMDEUM000912392042
9715	9715	276000913163490	SIMDEUM000913163490	RAPID	BFV	1993	SIMDEUM000925952548
9717	9717	276000917327047	SIMDEUM000917327047	STREUF	BFV	1988	SIMDEUM000927550527
9718	9718	40000681337645	SIMAUTM000681337645	STIER	BFV	2001	SIMAUTM000474437441
9722	9722	276000934694674	SIMDEUM000934694674	VAGABUND	BFV	2000	SIMDEUM000918555090
9724	9724	40000277755233	SIMAUTM000277755233	WAXIN	BFV	1993	SIMDEUM000912851233
9725	9725	276000919879431	SIMDEUM000919879431	REMONT	BFV	1997	SIMDEUM000919598352
9726	9726	756712018039935	SIMCHEM712018039935	ALDOR	SIM	1990	SIMCHEM712929011174
9727	9727	276000932878729	SIMDEUM000932878729	WEBAL	BFV	1999	SIMDEUM000919322930
9728	9728	276000936487481	SIMDEUM000936487481	MANITOBA	BFV	2002	SIMDEUM000915079575
9729	9729	276000937694254	SIMDEUM000937694254	RESOLUT	BFV	2003	SIMDEUM000919879431
9730	9730	276000936189219	SIMDEUM000936189219	RUREX	BFV	2002	SIMDEUM000918105400
9731	9731	276000940100513	SIMDEUM000940100513	WALDBRAND	BFV	2006	SIMDEUM000934492505
9732	9732	276000940930385	SIMDEUM000940930385	WITZBOLD	BFV	2006	SIMDEUM000934492505
9733	9733	276000936284807	SIMDEUM000936284807	ILION	BFV	2002	SIMDEUM000918174246
9734	9734	276000937187135	SIMDEUM000937187135	RUAKANA	BFV	2002	SIMDEUM000918105400
9735	9735	276000918105400	SIMDEUM000918105400	RUAP	BFV	1995	SIMDEUM000912971290
9736	9736	276000912392042	SIMDEUM000912392042	HIPPO	BFV	1996	SIMDEUM000911331078
9737	9737	276000937793170	SIMDEUM000937793170	HUPSOL	BFV	2003	SIMDEUM000918912889
9738	9738	276000935904510	SIMDEUM000935904510	IMPOSIUM	BFV	2002	SIMDEUM000918174246
9739	9739	276000911043667	SIMDEUM000911043667	ROMEL	BFV	1995	SIMDEUM000912971290
9740	9740	CZ61790058	CZ61790058	RAD-GURU	BFV	2010	
9741	9741	756120099980013	BSWCHEM120099980013	----	BSW	2013	
9742	9742	276000936822851	SIMDEUM000936822851	INDER	BFV	2002	SIMDEUM000918174246
9743	9743	276000938595455	SIMDEUM000938595455	MOMO	BFV	2004	SIMAUTM000410617633
9754	9754	276000936319703	SIMDEUM000936319703	RALMESBACH	BFV	2001	SIMDEUM000915949901
9755	9755	276000937608774	SIMDEUM000937608774	RAMIG	BFV	2004	SIMDEUM000918151241
9756	9756	276000937911240	SIMDEUM000937911240	RAMKAN	BFV	2004	SIMDEUM000918151241
9766	9766	40000963638719	SIMAUTM000963638719	----	BFV	2013	SIMDEUM000813516428
9767	9767	276000803608138	SIMDEUM000803608138	POLZER	BFV	1959	SIMAUTM00000002320
9768	9768	276000937746688	SIMDEUM000937746688	ETTAL	BFV	2004	SIMDEUM000916200528
9769	9769	276000937169790	SIMDEUM000937169790	HOLZMICHL	BFV	2003	SIMDEUM000912392042
9770	9770	276000935463615	SIMDEUM000935463615	MANGFALL	BFV	2001	SIMDEUM000915079575
9771	9771	276000938063849	SIMDEUM000938063849	MANGOPE	BFV	2004	SIMAUTM000410617633
9772	9772	276000936487534	SIMDEUM000936487534	ROUND UP	BFV	2002	SIMDEUM000912291736
9773	9773	276000938322620	SIMDEUM000938322620	SAMLAND	BFV	2005	SIMDEUM000932408546
9774	9774	276000939560629	SIMDEUM000939560629	VALUTA	BFV	2006	SIMDEUM000934586859
9775	9775	276000940220038	SIMDEUM000940220038	WALDHOER	BFV	2006	SIMDEUM000934492505
9776	9776	276000933179301	SIMDEUM000933179301	ZASPORT	BFV	2000	SIMDEUM000911485840
9777	9777	276000940777732	SIMDEUM000940777732	ZAUBER	BFV	2006	SIMDEUM000933038755
9778	9778	276000935704715	SIMDEUM000935704715	DON JUAN	BFV	2002	SIMDEUM000915122401
9779	9779	276000942238113	SIMDEUM000942238113	FIDELIS	BFV	2007	SIMDEUM000935703443
esmeralda	8144	276000938065319	SIMDEUM000938065319	ESMERALDA	BFV	2004	SIMDEUM000918151241
valero	8143	276000941364903	SIMDEUM000941364903	VALERO	BFV	2007	SIMDEUM000934586859
vanstein	9046	276000934586859	SIMDEUM000934586859	VANSTEIN	BFV	2000	SIMDEUM000918555090
100002	100002	PSK_128F	PSK_128F	female1	NFC	NA	
100003	100003	PSK_326F	PSK_326F	female2	NFC	NA	
100004	100004	PSK_370F	PSK_370F	female3	NFC	NA	

Die für die Re-Sequenzierung selektierten Tiere sind nach ihrer Bam-ID sortiert. Bam-IDs entsprechen der internen Namen der Sequenzen je Tier; Lab-IDs entsprechen der internen Labornummer; Hit-Nr. ist die Nummer der Lebendohrmarke; Interbull-ID ist die internationale Nummer des Tieres; die zugehörigen Rassen der re-sequenzier-ten Tiere sind mit ihren Kürzeln dargestellt; Geb.jahr entspricht dem Geburtsjahr des Tieres.

Anhang

Tabelle 9.2: Verteilung der Readlängen je Tier [in %]

Bam-ID	36bp	76bp	96bp	100bp	101bp
55069	-	-	-	-	100,00
55070	-	-	-	100,00	-
55071	-	-	-	100,00	-
55072	-	-	-	100,00	-
55073	-	-	-	-	100,00
55074	-	-	-	-	100,00
55075	-	-	-	100,00	-
55076	-	-	22,37	22,37	55,26
55077	-	-	16,50	16,50	66,99
55078	-	-	17,97	17,98	64,05
55687	-	-	-	-	100,00
55689	-	-	15,69	84,31	-
55690	-	-	49,99	50,01	-
55691	-	-	-	100,00	-
55692	-	-	-	100,00	-
55693	-	-	-	100,00	-
55694	-	-	-	100,00	-
55695	-	-	-	36,54	63,46
55696	-	-	-	37,29	62,71
55697	-	-	-	100,00	-
55698	-	-	-	100,00	-
56503	-	-	-	100,00	-
56504	-	-	-	100,00	-
56505	-	-	-	100,00	-
56506	-	-	-	100,00	-
56507	-	-	-	100,00	-
56508	-	-	-	100,00	-
56509	-	-	-	100,00	-
58941	-	-	-	-	100,00
58943	-	-	-	100,00	-
58944	-	-	-	100,00	-
58945	-	-	-	100,00	-
58946	-	-	-	100,00	-
58947	-	-	-	100,00	-
58948	-	-	-	100,00	-
58949	-	-	-	-	100,00
58950	-	-	-	100,00	-
58951	-	-	-	-	100,00
58952	-	-	-	100,00	-
58953	-	-	-	100,00	-
65649	-	-	-	-	100,00
65650	-	-	-	-	100,00
65651	-	-	-	-	100,00
65652	-	-	-	-	100,00
65653	-	-	-	-	100,00
65654	-	-	-	-	100,00
65655	-	-	-	-	100,00
65656	-	-	-	-	100,00
65657	-	-	-	-	100,00
65658	-	-	-	-	100,00
65659	-	-	-	-	100,00
65660	-	-	-	-	100,00
65661	-	-	-	-	100,00
65663	-	-	-	-	100,00
66498	-	-	-	-	100,00
66499	-	-	-	-	100,00

Anhang

Bam-ID	36bp	76bp	96bp	100bp	101bp
66500	-	-	-	-	100,00
66501	-	-	-	-	100,00
66502	-	-	-	-	100,00
66503	-	-	-	-	100,00
66504	-	-	-	-	100,00
66505	-	-	-	-	100,00
66506	-	-	-	-	100,00
66507	-	-	-	-	100,00
66508	-	-	-	-	100,00
66511	-	-	-	-	100,00
66512	-	-	-	-	100,00
66513	-	-	-	-	100,00
66514	-	-	-	-	100,00
66515	-	-	-	-	100,00
66517	-	-	-	-	100,00
66518	-	-	-	-	100,00
66519	-	-	-	-	100,00
66520	-	-	-	-	100,00
66521	-	-	-	-	100,00
66522	-	-	-	-	100,00
66524	-	-	-	-	100,00
68262	-	-	-	-	100,00
68263	-	-	-	-	100,00
68264	-	-	-	-	100,00
68265	-	-	-	-	100,00
68266	-	-	-	-	100,00
68267	-	-	-	-	100,00
68268	-	-	-	-	100,00
68269	-	-	-	-	100,00
68270	-	-	-	-	100,00
68271	-	-	-	-	100,00
68272	-	-	-	-	100,00
68273	-	-	-	-	100,00
68274	-	-	-	-	100,00
68275	-	-	-	-	100,00
68276	-	-	-	-	100,00
68277	-	-	-	-	100,00
68278	-	-	-	-	100,00
68279	-	-	-	-	100,00
68280	-	-	-	-	100,00
68281	-	-	-	-	100,00
68282	-	-	-	-	100,00
68283	-	-	-	-	100,00
68284	-	-	-	-	100,00
68285	-	-	-	-	100,00
68286	-	-	-	-	100,00
68287	-	-	-	-	100,00
68288	-	-	-	-	100,00
68289	-	-	-	-	100,00
68290	-	-	-	-	100,00
68291	-	-	-	-	100,00
68292	-	-	-	-	100,00
68293	-	-	-	-	100,00
68294	-	-	-	-	100,00
68295	-	-	-	-	100,00
68296	-	-	-	-	100,00
68297	-	-	-	-	100,00
68298	-	-	-	-	100,00

Anhang

Bam-ID	36bp	76bp	96bp	100bp	101bp
68299	-	-	-	-	100,00
68300	-	-	-	-	100,00
68301	-	-	-	-	100,00
68303	-	-	-	-	100,00
68305	-	-	-	-	100,00
68306	-	-	-	-	100,00
68307	-	-	-	-	100,00
68309	-	-	-	-	100,00
68310	-	-	-	-	100,00
68311	-	-	-	-	100,00
68312	-	-	-	-	100,00
68313	-	-	-	-	100,00
68314	-	-	-	-	100,00
68315	-	-	-	-	100,00
68316	-	-	-	-	100,00
68317	-	-	-	-	100,00
68318	-	-	-	-	100,00
68319	-	-	-	-	100,00
68320	-	-	-	-	100,00
68321	-	-	-	-	100,00
68322	-	-	-	-	100,00
68323	-	-	-	-	100,00
68324	-	-	-	-	100,00
68325	-	-	-	-	100,00
68326	-	-	-	-	100,00
68327	-	-	-	-	100,00
68328	-	-	-	-	100,00
68329	-	-	-	-	100,00
68330	-	-	-	-	100,00
68331	-	-	-	-	100,00
68332	-	-	-	-	100,00
9315	-	-	-	-	100,00
9610	-	-	-	-	100,00
9611	-	-	-	-	100,00
9612	-	-	-	-	100,00
9613	-	-	-	-	100,00
9614	-	-	-	-	100,00
9615	-	-	-	-	100,00
9616	-	-	-	-	100,00
9617	-	-	-	-	100,00
9618	-	-	-	-	100,00
9619	-	-	-	-	100,00
9620	-	-	-	-	100,00
9621	-	-	-	-	100,00
9622	-	-	-	-	100,00
9623	-	-	-	-	100,00
9624	-	-	-	-	100,00
9625	-	-	-	-	100,00
9626	-	-	-	-	100,00
9704	-	-	-	-	100,00
9705	-	-	-	-	100,00
9706	-	-	-	-	100,00
9707	-	-	-	-	100,00
9708	-	-	-	-	100,00
9709	-	-	-	-	100,00
9710	-	-	-	-	100,00
9711	-	-	-	-	100,00
9712	-	-	-	-	100,00

Anhang

Bam-ID	36bp	76bp	96bp	100bp	101bp
9713	-	-	-	-	100,00
9715	-	-	-	-	100,00
9717	-	-	-	-	100,00
9718	-	-	-	-	100,00
9719	-	-	-	-	100,00
9720	-	-	-	-	100,00
9721	-	-	-	-	100,00
9722	-	-	-	-	100,00
9724	-	-	-	-	100,00
9725	-	-	-	-	100,00
9726	-	-	-	-	100,00
9727	-	-	-	-	100,00
9728	-	-	-	-	100,00
9729	-	-	-	-	100,00
9730	-	-	-	-	100,00
9731	-	-	-	-	100,00
9732	-	-	-	-	100,00
9733	-	-	-	-	100,00
9734	-	-	-	-	100,00
9735	-	-	-	-	100,00
9736	-	-	-	-	100,00
9737	-	-	-	-	100,00
9738	-	-	-	-	100,00
9739	-	-	-	-	100,00
9740	-	-	-	-	100,00
9741	-	-	-	-	100,00
9742	-	-	-	-	100,00
9743	-	-	-	-	100,00
9744	-	-	-	-	100,00
9745	-	-	-	-	100,00
9746	-	-	-	-	100,00
9747	-	-	-	-	100,00
9748	-	-	-	-	100,00
9749	-	-	-	-	100,00
9750	-	-	-	-	100,00
9751	-	-	-	-	100,00
9752	-	-	-	-	100,00
9753	-	-	-	-	100,00
9754	-	-	-	-	100,00
9755	-	-	-	-	100,00
9756	-	-	-	-	100,00
9757	-	-	-	-	100,00
9758	-	-	-	-	100,00
9759	-	-	-	-	100,00
9760	-	-	-	-	100,00
9761	-	-	-	-	100,00
9762	-	-	-	-	100,00
9763	-	-	-	-	100,00
9764	-	-	-	-	100,00
9766	-	-	-	-	100,00
9767	-	-	-	-	100,00
9768	-	-	-	-	100,00
9769	-	-	-	-	100,00
9770	-	-	-	-	100,00
9771	-	-	-	-	100,00
9772	-	-	-	-	100,00
9773	-	-	-	-	100,00
9774	-	-	-	-	100,00

Anhang

Bam-ID	36bp	76bp	96bp	100bp	101bp
9775	-	-	-	-	100,00
9776	-	-	-	-	100,00
9777	-	-	-	-	100,00
9778	-	-	-	-	100,00
9779	-	-	-	-	100,00
9780	-	-	-	-	100,00
9781	-	-	-	-	100,00
9782	-	-	-	-	100,00
9783	-	-	-	-	100,00
9784	-	-	-	-	100,00
9785	-	-	-	-	100,00
9786	-	-	-	-	100,00
9787	-	-	-	-	100,00
9788	-	-	-	-	100,00
9789	-	-	-	-	100,00
9790	-	-	-	-	100,00
9791	-	-	-	-	100,00
9792	-	-	-	-	100,00
9793	-	-	-	-	100,00
9800	-	-	-	-	100,00
9801	-	-	-	-	100,00
9802	-	-	-	-	100,00
9803	-	-	-	-	100,00
9804	-	-	-	-	100,00
9807	-	-	-	-	100,00
9811	-	-	-	-	100,00
esmeralda	-	69,73	-	30,27	-
valero	-	-	-	100,00	-
vanstein	100,00	-	-	-	-
100002	-	-	-	-	100,00
100003	-	-	-	-	100,00
100004	-	-	-	-	100,00
100023	-	-	-	-	100,00

Die für die Re-Sequenzierung selektierten Tiere sind nach ihrer Bam-ID sortiert. Bam-IDs entsprechen der internen Namen der Sequenzen je Tier; bp entspricht der Basenpaarlänge der sequenzierten Fragmente.

Tabelle 9.3: Mapping Statistik

Bam-ID	Name	Gesamte Reads	Duplikate [%]	Kartierte Reads [%]	Kartierte Reads Netto[%]	Kartierte Basen Netto	Durchschnittliche Abdeckung Netto	
55069	REDAD	163.482.913	5,17	96,43	149.489.938	91,44	15,10GB	5,68
55070	STREIF	248.209.072	6,50	97,90	227.199.055	91,54	22,72GB	8,54
55071	ROMULUS	275.347.611	6,38	98,04	252.713.383	91,78	25,27GB	9,50
55072	HORWEIN	218.095.293	5,92	97,72	200.490.494	91,93	20,05GB	7,54
55073	BALBO	173.087.784	5,26	96,37	158.034.806	91,30	15,96GB	6,00
55074	EGOL	156.705.901	5,86	96,28	142.031.377	90,64	14,35GB	5,39
55075	RENNER	257.059.233	7,06	97,79	233.629.804	90,89	23,36GB	8,78
55076	SAMURAI	295.394.459	4,65	97,04	273.327.445	92,53	27,24GB	10,24
55077	POSTILLION	251.461.184	5,40	96,89	230.495.110	91,66	23,05GB	8,67
55078	GEBAL	284.445.379	5,65	96,90	260.051.305	91,42	25,98GB	9,77
55687	ROMANEK	162.759.413	7,00	96,14	145.527.098	89,41	14,70GB	5,53
55689	HOLZER	287.537.748	7,35	97,72	260.332.777	90,54	25,87GB	9,73
55690	REGIO	160.290.585	4,83	98,05	149.570.539	93,31	14,66GB	5,51
55691	BOSS	123.602.395	6,43	95,93	110.944.671	89,76	11,09GB	4,17
55692	REXON	138.843.685	6,86	95,81	123.887.746	89,23	12,39GB	4,66
55693	ROMEN	131.466.898	7,41	96,80	117.828.683	89,63	11,78GB	4,43
55694	DIRTECK	150.863.976	5,59	97,29	138.578.560	91,86	13,86GB	5,21
55695	RALBO	261.211.185	8,41	96,46	230.777.234	88,35	23,22GB	8,73
55696	HALLING	174.223.713	8,02	97,64	156.464.263	89,81	15,74GB	5,92
55697	POLDI	245.108.260	8,48	97,18	217.996.621	88,94	21,80GB	8,20
55698	SPORT	270.847.013	8,83	95,63	236.135.825	87,18	23,61GB	8,88
56503	HORB	181.507.121	4,61	97,99	169.656.055	93,47	16,97GB	6,38
56504	HUMID	160.842.497	4,16	97,84	150.815.249	93,77	15,08GB	5,67
56505	HUMLANG	227.151.631	4,29	98,35	213.806.924	94,13	21,38GB	8,04
56506	PROPELLER	190.028.977	4,09	98,18	178.943.367	94,17	17,89GB	6,73
56507	REPORT	208.459.066	4,91	98,23	194.715.327	93,41	19,47GB	7,32
56508	UTNACH	207.068.014	4,59	98,23	194.066.215	93,72	19,41GB	7,30
56509	RENGER	176.164.262	4,71	97,99	164.482.114	93,37	16,45GB	6,18
58941	HORROR	161.865.174	5,16	96,24	147.729.603	91,27	14,92GB	5,61
58943	HODSCHA	226.335.891	4,35	95,54	206.832.039	91,38	20,68GB	7,78
58944	MORWEL	216.683.526	3,59	96,85	202.333.113	93,38	20,23GB	7,61
58945	HAXL	207.869.816	4,05	96,97	193.404.741	93,04	19,34GB	7,27
58946	WATERBERG	208.384.577	4,45	97,03	193.209.439	92,72	19,32GB	7,26
58947	MANDELA	182.548.630	3,72	97,06	170.592.510	93,45	17,06GB	6,41
58948	MARTL-DE	160.719.456	4,32	97,13	149.353.352	92,93	14,94GB	5,61
58949	ZAHNER	160.661.919	6,47	96,48	144.980.005	90,24	14,64GB	5,50
58950	DAGO	212.036.734	5,66	96,65	193.340.167	91,18	19,33GB	7,27
58951	WEINOLD	160.626.127	6,89	96,41	144.198.385	89,77	14,56GB	5,48
58952	MORGENGRAU	178.349.820	4,57	96,93	164.966.017	92,50	16,50GB	6,20
58953	G4	225.268.008	5,15	97,06	207.389.836	92,06	20,74GB	7,80
65649	HODACH	476.417.926	7,07	96,22	425.999.503	89,42	43,03GB	16,18
65650	ZASTER	436.077.734	8,28	96,25	384.983.689	88,28	38,88GB	14,62
65651	MAXI	370.004.056	8,33	96,94	328.814.495	88,87	33,21GB	12,49
65652	HILAB	332.296.442	6,85	97,84	302.841.675	91,14	30,59GB	11,50
65653	LEO	279.960.716	6,48	97,94	256.416.561	91,59	25,90GB	9,74
65654	RANDY	307.150.566	7,51	97,62	277.332.888	90,29	28,01GB	10,53
65655	MALIST	356.635.310	7,94	97,73	320.867.204	89,97	32,41GB	12,18
65656	HEROCK	302.724.706	6,56	98,00	277.214.206	91,57	28,00GB	10,53
65657	TCP_13	330.956.026	8,70	97,34	294.116.309	88,87	29,71GB	11,17
65658	FABER	385.934.050	19,66	96,57	299.437.165	77,59	30,24GB	11,37
65659	SIoux	349.139.760	8,43	97,11	310.456.819	88,92	31,36GB	11,79
65660	GOLDSTAR	349.958.124	7,65	97,14	313.957.671	89,71	31,71GB	11,92
65661	MAGNUM	370.499.222	17,83	97,03	295.420.809	79,74	29,84GB	11,22

Anhang

Bam-ID	Name	Gesamte Reads	Duplikate [%]	Kartierte Reads [%]	Kartierte Reads Netto[%]	Kartierte Basen Netto	Durchschnittliche Abdeckung Netto	
65663	CHAMPION	291.319.858	15,53	97,55	240.053.858	82,40	24,25GB	9,11
66498	MASCOL	364.105.874	8,21	97,90	327.193.859	89,86	33,05GB	12,42
66499	MANAT	358.163.168	6,40	98,23	329.288.559	91,94	33,26GB	12,50
66500	ALVES	323.079.028	8,13	98,10	291.188.280	90,13	29,41GB	11,06
66501	TOBER	341.640.244	8,23	98,15	307.725.382	90,07	31,08GB	11,68
66502	BONATUS	287.849.746	7,60	97,94	260.489.728	90,50	26,31GB	9,89
66503	STATUS	368.275.714	7,39	98,06	334.452.448	90,82	33,78GB	12,70
66504	LICHTBLICK	374.217.170	7,01	98,03	341.132.723	91,16	34,45GB	12,95
66505	MERGIM	369.705.482	6,69	98,25	338.920.144	91,67	34,23GB	12,87
66506	JOSE	347.335.464	11,81	98,26	300.983.321	86,65	30,40GB	11,43
66507	WIZZARD	361.751.712	6,36	97,95	331.803.110	91,72	33,51GB	12,60
66508	PIT	343.083.406	5,81	98,17	317.246.102	92,47	32,04GB	12,05
66511	LAUDAN	323.788.768	7,01	98,19	295.644.938	91,31	29,86GB	11,23
66512	TICKET	298.182.896	9,00	98,11	266.220.729	89,28	26,89GB	10,11
66513	BLACKMAN	319.095.456	9,36	98,12	283.802.886	88,94	28,66GB	10,78
66514	ZECHER	341.534.150	6,93	98,22	312.186.020	91,41	31,53GB	11,85
66515	GIBOR	342.026.096	6,97	97,97	311.743.946	91,15	31,49GB	11,84
66517	LUDOX	453.639.284	15,85	97,26	371.275.926	81,84	37,50GB	14,10
66518	ORIGIN	323.423.486	6,63	97,99	295.894.460	91,49	29,89GB	11,24
66519	CARMANO	315.035.442	9,26	98,01	280.177.341	88,94	28,30GB	10,64
66520	UBBO	327.882.672	7,72	98,06	296.702.616	90,49	29,97GB	11,27
66521	LASSO	324.795.148	8,59	97,96	290.852.070	89,55	29,38GB	11,04
66522	JOYEUX	295.580.486	4,20	98,22	278.118.276	94,09	28,09GB	10,56
66524	FLEURON	295.561.924	5,58	98,13	273.865.827	92,66	27,66GB	10,40
68262	ZUNDER	313.608.110	7,73	97,97	283.493.398	90,40	28,63GB	10,76
68263	LADIN	354.131.970	6,59	96,02	317.608.934	89,69	32,08GB	12,06
68264	PATRICK	289.241.146	7,09	96,33	258.875.104	89,50	26,15GB	9,83
68265	DOLCH	307.739.478	8,82	98,09	275.236.881	89,44	27,80GB	10,45
68266	TRAILOR	374.652.726	7,06	95,60	332.876.802	88,85	33,62GB	12,64
68267	AMARETTO	323.518.854	9,43	97,94	286.988.180	88,71	28,99GB	10,90
68268	VERO	322.860.286	7,22	98,13	293.955.487	91,05	29,69GB	11,16
68269	HORST	350.949.790	6,87	95,28	311.429.750	88,74	31,45GB	11,82
68270	ALBO	367.136.492	8,40	98,08	329.814.180	89,83	33,31GB	12,52
68271	OREON	295.848.034	6,99	97,35	267.902.024	90,55	27,06GB	10,17
68272	AXIUS	321.547.492	7,99	98,04	290.054.338	90,21	29,30GB	11,01
68273	NORDAL	358.267.014	9,10	98,09	319.448.312	89,16	32,26GB	12,13
68274	NOVIO	323.459.806	9,27	97,89	287.277.711	88,81	29,02GB	10,91
68275	ZOLL	380.505.218	7,50	95,29	335.390.514	88,14	33,87GB	12,73
68276	BRYAN	374.123.018	6,96	95,83	333.575.472	89,16	33,69GB	12,67
68277	WINNETOU	346.571.890	8,26	95,93	304.989.984	88,00	30,80GB	11,58
68278	NATO	357.882.212	6,65	95,80	320.059.697	89,43	32,33GB	12,15
68279	ASCOT	429.031.870	8,94	97,43	380.659.539	88,73	38,45GB	14,45
68280	JOSS	334.530.342	7,79	97,95	302.159.303	90,32	30,52GB	11,47
68281	LEADER	394.265.364	8,57	96,50	347.832.393	88,22	35,13GB	13,21
68282	JORAT	402.835.694	7,08	96,05	359.508.450	89,24	36,31GB	13,65
68283	BELL-ET	353.899.990	7,69	96,82	316.311.909	89,38	31,95GB	12,01
68284	ENSIGN-ET	371.803.348	8,12	96,84	330.825.976	88,98	33,41GB	12,56
68285	IMPEROR	378.140.382	9,53	96,77	331.073.444	87,55	33,44GB	12,57
68286	STARBUCK	380.555.278	8,50	96,26	335.185.247	88,08	33,85GB	12,73
68287	ENORM	377.837.180	7,85	96,53	336.099.498	88,95	33,95GB	12,76
68288	PROPHET-ET	440.501.220	7,93	96,74	392.374.437	89,07	39,63GB	14,90
68289	EMERALD-DE	397.596.086	8,11	86,60	316.372.728	79,57	31,95GB	12,01
68290	MEDOR	332.608.130	8,19	96,69	295.251.076	88,77	29,82GB	11,21
68291	NIPPEL	337.537.066	6,80	96,54	303.713.002	89,98	30,68GB	11,53
68292	COLLECTION	326.241.602	7,78	97,18	292.361.946	89,62	29,53GB	11,10
68293	CHIME-ET	338.979.966	6,89	97,04	306.282.442	90,35	30,93GB	11,63

Anhang

Bam-ID	Name	Gesamte Reads	Duplikate [%]	Kartierte Reads [%]	Kartierte Reads Netto[%]	Kartierte Basen Netto	Durchschnittliche Abdeckung Netto	
68294	CHRISTO	395.443.786	8,36	93,36	338.330.085	85,56	34,17GB	12,85
68295	VOERIS	369.308.954	6,27	97,05	335.939.566	90,96	33,93GB	12,76
68296	GRAL	412.164.190	5,31	96,83	377.936.397	91,70	38,17GB	14,35
68297	KASTOR	371.220.770	6,14	95,81	333.831.756	89,93	33,72GB	12,68
68298	ZEUS	390.876.014	6,52	96,55	352.785.348	90,26	35,63GB	13,40
68299	JOLDEN-ET	432.540.294	5,92	96,00	390.642.664	90,31	39,45GB	14,83
68300	VIGOR-ET	338.164.548	5,57	93,02	297.040.466	87,84	30,00GB	11,28
68301	AELPLER	440.147.704	5,93	96,10	397.930.739	90,41	40,19GB	15,11
68303	VERGIL	411.227.552	6,13	96,36	371.962.612	90,45	37,57GB	14,12
68305	TRIPLE	388.308.270	6,73	98,22	355.725.044	91,61	35,93GB	13,51
68306	AELPLER	367.952.396	5,81	96,57	334.686.209	90,96	33,80GB	12,71
68307	FIRESTAR	401.623.036	6,61	96,32	361.277.016	89,95	36,49GB	13,72
68309	TIM	337.396.372	6,04	98,07	310.904.973	92,15	31,40GB	11,81
68310	RIVAROL	320.012.770	6,78	97,07	289.564.250	90,49	29,25GB	10,99
68311	MARS	332.501.826	6,92	97,20	300.836.833	90,48	30,38GB	11,42
68312	JEROME	346.256.450	6,64	96,99	313.542.012	90,55	31,67GB	11,91
68313	TONTO	341.124.056	6,66	96,91	308.591.588	90,46	31,17GB	11,72
68314	CAVEMAN	349.217.388	7,02	96,82	314.374.007	90,02	31,75GB	11,94
68315	KANIS	348.846.776	6,89	97,07	315.280.057	90,38	31,84GB	11,97
68316	CONDOR	350.885.702	6,55	96,97	317.965.506	90,62	32,11GB	12,07
68317	BEMOL	330.470.486	6,72	96,83	298.506.487	90,33	30,15GB	11,33
68318	ORKAN	362.116.512	6,60	96,81	327.427.707	90,42	33,07GB	12,43
68319	FIRE	385.380.604	7,97	96,81	343.326.285	89,09	34,68GB	13,04
68320	JONN-ET	348.083.344	6,93	96,95	314.073.818	90,23	31,72GB	11,93
68321	JUBILANT-E	329.744.000	8,76	98,05	294.980.345	89,46	29,79GB	11,20
68322	SABER-ET	356.372.446	6,73	97,00	322.417.465	90,47	32,56GB	12,24
68323	ROCHESTER	386.075.902	6,78	96,63	347.758.147	90,08	35,12GB	13,20
68324	SWATCH-ET	343.160.138	6,40	96,98	311.484.477	90,77	31,46GB	11,83
68325	HANS-ET	326.230.154	7,13	98,07	297.147.877	91,09	30,01GB	11,28
68326	RIOSTAR	346.927.160	7,49	98,14	314.973.111	90,79	31,81GB	11,96
68327	FILOU	342.210.888	7,02	97,93	311.596.937	91,05	31,47GB	11,83
68328	TIPIC	294.467.464	6,27	97,79	269.923.600	91,66	27,26GB	10,25
68329	BELLO	349.138.122	10,59	98,18	306.474.447	87,78	30,95GB	11,64
68330	BACCALA-ET	297.626.546	8,02	98,10	268.554.892	90,23	27,12GB	10,20
68331	AELPLER_3	305.946.644	6,33	97,92	280.594.641	91,71	28,34GB	10,65
68332	INCAS	350.562.616	6,66	98,13	321.105.576	91,60	32,43GB	12,19
9315	STYLIST	329.997.996	10,09	98,17	291.278.102	88,27	29,42GB	11,06
9610	SAMUT	546.235.398	14,01	97,68	458.822.148	84,00	46,34GB	17,42
9611	SAFIR	504.803.222	13,60	97,49	425.194.240	84,23	42,94GB	16,14
9612	SAMPRAS	542.196.834	10,15	97,01	472.596.895	87,16	47,73GB	17,94
9613	REPTEIT	522.065.280	12,74	97,54	444.355.355	85,11	44,88GB	16,87
9614	ROMSEL	516.875.066	13,51	97,63	436.454.957	84,44	44,08GB	16,57
9615	ROBERTO	524.243.828	17,22	97,22	421.904.226	80,48	42,61GB	16,02
9616	RAINER	410.488.326	11,55	98,01	355.868.259	86,69	35,94GB	13,51
9617	MADERA	461.214.490	16,11	98,04	379.358.308	82,25	38,32GB	14,40
9618	MORBO	413.271.730	13,45	97,95	350.351.770	84,78	35,39GB	13,30
9619	HATZ	426.239.958	14,68	97,97	356.281.129	83,59	35,98GB	13,53
9620	HULOCK	361.168.818	16,28	97,91	296.059.091	81,97	29,90GB	11,24
9621	GEBALOT	384.427.542	9,40	98,27	342.274.956	89,03	34,57GB	13,00
9623	RUBEL	314.078.890	13,10	98,05	267.594.193	85,20	27,03GB	10,16
9624	POLK	338.775.672	16,93	97,80	275.215.587	81,24	27,80GB	10,45
9625	RALSTER	327.901.852	9,43	98,09	291.304.622	88,84	29,42GB	11,06
9626	WAGRAIN PS	329.962.022	13,82	97,98	278.603.570	84,44	28,14GB	10,58
9705	BOSBO	245.211.960	15,55	96,74	200.327.534	81,70	20,23GB	7,61
9706	ERGO	272.449.886	10,18	98,42	240.851.505	88,40	24,33GB	9,15
9707	ENRICO	214.729.212	8,38	98,44	193.664.034	90,19	19,56GB	7,35

Anhang

Bam-ID	Name	Gesamte Reads	Duplikate [%]	Kartierte Reads [%]	Kartierte Reads Netto[%]	Kartierte Basen Netto	Durchschnittliche Abdeckung Netto	
9708	HIRMER	187.159.666	27,96	96,96	130.729.675	69,85	13,20GB	4,96
9709	HIMALAYA	856.19.218	28,05	96,68	59.560.216	69,56	6,02GB	2,26
9715	RAPID	108.054.632	5,01	95,28	97.800.696	90,51	9,88GB	3,71
9717	STREUF	516.538.742	10,53	97,72	451.635.014	87,43	45,62GB	17,15
9718	STIER	518.984.250	8,74	97,66	462.520.591	89,12	46,71GB	17,56
9722	VAGABUND	325.808.698	10,17	98,26	287.559.688	88,26	29,04GB	10,92
9724	WAXIN	502.005.218	8,00	95,67	441.832.224	88,01	44,63GB	16,78
9725	REMONT	300.806.310	9,92	98,09	265.772.185	88,35	26,84GB	10,09
9726	ALDOR	518.006.588	9,21	97,61	459.049.974	88,62	46,36GB	17,43
9727	WEBAL	520.156.810	8,80	97,69	463.396.443	89,09	46,80GB	17,60
9728	MANITOBA	127.492.362	32,34	95,98	82.796.044	64,94	8,36GB	3,14
9729	RESOLUT	120.397.884	33,84	96,32	76.732.359	63,73	7,75GB	2,91
9730	RUREX	129.521.018	26,57	96,56	91.840.209	70,91	9,28GB	3,49
9731	WALDBRAND	132.546.696	33,45	96,59	85.199.367	64,28	8,61GB	3,24
9732	WITZBOLD	133.902.476	28,03	95,88	92.399.562	69,01	9,33GB	3,51
9733	ILION	138.217.538	27,54	96,68	96.831.196	70,06	9,78GB	3,68
9734	RUAKANA	164.428.208	25,18	96,73	119.010.771	72,38	12,02GB	4,52
9735	RUAP	145.966.502	24,40	96,74	106.746.879	73,13	10,78GB	4,05
9736	HIPPO	126.766.774	28,63	96,87	87.638.372	69,13	8,85GB	3,33
9737	HUPSOL	170.015.956	25,55	96,85	122.597.052	72,11	12,38GB	4,66
9738	IMPOSIUM	131.025.706	28,43	96,73	90.715.618	69,23	9,16GB	3,44
9739	ROMEL	141.518.182	26,28	96,71	100.893.778	71,29	10,19GB	3,83
9740	RAD-GURU	520.750.152	8,54	97,04	462.160.978	88,75	46,68GB	17,55
9741	----	522.521.028	9,67	97,25	459.005.278	87,84	46,36GB	17,43
9742	INDER	336.067.716	8,87	98,39	301.321.111	89,66	30,43GB	11,44
9743	MOMO	352.511.988	9,38	98,34	314.151.330	89,12	31,73GB	11,93
9754	RALMESBACH	351.033.034	9,20	98,31	313.369.660	89,27	31,65GB	11,90
9755	RAMIG	369.428.530	9,44	98,30	328.887.486	89,03	33,22GB	12,49
9756	RAMKAN	361.267.998	9,26	98,38	322.493.340	89,27	32,57GB	12,25
9766	----	375.866.338	6,27	97,65	344.017.525	91,53	34,75GB	13,06
9767	POLZER	373.135.632	4,05	98,19	351.544.678	94,21	35,51GB	13,35
9768	ETTAL	322.551.924	10,56	98,27	283.515.160	87,90	28,64GB	10,77
9769	HOLZMICHL	302.308.986	9,17	98,13	269.459.616	89,13	27,22GB	10,23
9770	MANGFALL	297.007.658	8,88	98,12	265.540.379	89,41	26,82GB	10,08
9771	MANGOPE	326.091.110	9,50	98,02	289.265.018	88,71	29,22GB	10,98
9772	ROUND UP	329.254.580	9,62	98,07	291.848.294	88,64	29,48GB	11,08
9773	SAMLAND	324.049.462	10,04	98,24	286.373.977	88,37	28,92GB	10,87
9774	VALUTA	369.824.844	10,48	98,33	325.567.876	88,03	32,88GB	12,36
9775	WALDHOER	367.310.078	11,88	98,31	318.207.905	86,63	32,14GB	12,08
9776	ZASPORT	326.163.568	10,45	98,33	287.219.363	88,06	29,01GB	10,91
9777	ZAUBER	307.016.308	9,54	98,12	272.487.654	88,75	27,52GB	10,35
9778	DON JUAN	344.161.176	9,66	98,20	305.311.400	88,71	30,84GB	11,59
9779	FIDELIS	292.612.752	9,42	98,18	260.217.020	88,93	26,28GB	9,88
esmeralda	ESMERALDA	999.157.620	14,56	98,60	841.664.522	84,24	70,08GB	26,35
valero	VALERO	342.410.932	14,44	93,83	274.888.063	80,28	27,49GB	10,33
vanstein	VANSTEIN	716.137.590	13,11	94,46	587.820.476	82,08	21,16GB	7,96
100002	female1	330.922.722	9,16	98,48	296.043.524	89,46	29,90GB	11,24
100003	female2	361.831.682	8,36	98,38	326.213.329	90,16	32,95GB	12,39
100004	female3	319.620.860	8,23	98,37	288.526.391	90,27	29,14GB	10,96

Die für die Re-Sequenzierung selektierten Tiere sind nach ihrer Bam-ID sortiert. Bam-IDs entsprechen der internen Namen der Sequenzen je Tier; Duplikate entsprechen den PCR-Duplikaten, die während der Amplifikation bei der Libraryaufbereitung entstehen und mit MarkDuplicates von PICARD (<http://picard.sourceforge.net>) markiert wurden; Kartierte Reads entsprechen dem Anteil der Reads, die an das UMD3.1-Assembly (ZIMIN et al. 2009) mit BWA (LI und DURBIN et al. 2009) kartiert werden konnten; Kartierte Reads Netto sind die kartierten Reads ohne Duplikate; Kartierte Basen Netto ergeben sich aus der Multiplikation der kartierten Reads Netto mit der Readlänge (Tabelle 9.2); Durchschnittliche Abdeckung Netto ergibt sich aus der Division der kartierten Basen Netto mit der Genomlänge (~2,67 GB).

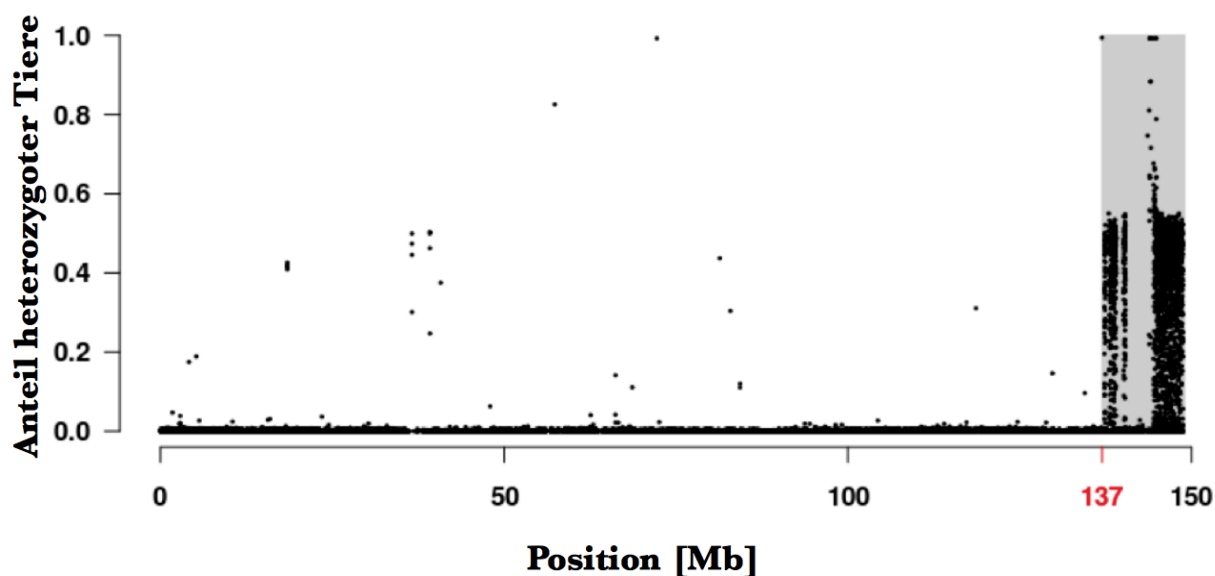


Abbildung 9.1: Bestimmung der Grenze zwischen pseudoautosomaler und nicht-pseudoautosomaler Region des bovinen X-Chromosoms

Für 896 männliche Tiere der BFV Population standen 777K-Genotypen zur Verfügung. Die chromosomale Position der SNPs wurde basierend auf dem Referenzgenom UMD3.1 (ZIMM et al. 2009) festgelegt. Der Anteil an heterozygoten Tieren auf dem X-Chromosom basiert auf 38.868 SNP. Der schattierte Kasten hebt die pseudoautosomale Region hervor, in der eine erhöhte Heterozygotie zu verzeichnen war. Die Grenze wurde auf 137.000.000 Basen festgelegt und erstreckt sich bis zum Ende des Chromosoms. Die pseudoautosomale Region des X-Chromosoms wurde wie ein autosomales Chromosom behandelt.

Tabelle 9.4: Anzahl Varianten je Chromosom

Chromosom	Varianten	Biallelische SNP	Triallelische SNP	InDels
Chr1	1.412.027	1.279.865	2111	130.051
Chr2	1.154.970	1.046.758	1814	106.398
Chr3	1.011.206	918.809	1432	90.965
Chr4	1.076.327	975.913	1370	99.044
Chr5	1.068.358	970.460	1554	96.344
Chr6	1.036.432	938.399	1491	96.542
Chr7	944.720	855.725	1356	87.639
Chr8	961.673	871.546	1330	88.797
Chr9	893.083	807.281	1295	84.507
Chr10	909.268	825.212	1369	82.687
Chr11	901.604	819.187	1362	81.055
Chr12	965.969	878.750	1237	85.982
Chr13	695.656	631.560	1510	62.586
Chr14	732.046	664.014	1161	66.871
Chr15	822.658	748.776	1001	72.881
Chr16	719.803	653.778	1267	64.758
Chr17	671.159	608.407	810	61.942
Chr18	566.701	515.015	909	50.777
Chr19	546.290	495.930	933	49.427
Chr20	653.761	593.676	896	59.189
Chr21	603.126	547.846	871	54.409
Chr22	520.837	473.433	649	46.755
Chr23	639.723	584.753	810	54.160
Chr24	555.325	503.917	853	50.555
Chr25	402.076	367.595	502	33.979
Chr26	461.556	419.463	709	41.384
Chr27	424.399	384.430	608	39.361
Chr28	432.808	394.125	603	38.080
Chr29	515.558	470.820	714	44.024
ChrX	524.501	456.705	1309	66.487
ChrX_PAR	114.302	104.132	197	9973

Tabellarische Übersicht der identifizierten Polymorphismen je Chromosom, bei denen mindestens ein Tier von der Referenzsequenz UMD3.1 (ZIMLN et al. 2009) abweicht.

Tabelle 9.5: Übersicht der Allelaustausche

		Alternativ			
		A	C	T	G
Referenz	A	0	805.390	646.435	3.037.087
	C	952.783	0	4.134.552	818.718
	T	643.610	3.031.584	0	808.341
	G	4.149.920	821.375	956.485	0

Anhang

Tabelle 9.6: Übersicht der verifizierten vorzeitigen Stopp-Codons

Chr	Position	REF	ALT	Gen	NMD	Alle Iso-		Homozygote		Genlänge	AS Pos	AS
						formen	214 Tiere	1000 Bullen	Sanger			
15	75.043.001	C	T	ACCSL	ja	ja	ja	ja	ja	576	561	Q
10	90.126.392	G	T	ADCK1	ja	ja	nein	ja	ja	523	297	E
19	27.333.935	C	T	ALOX15	nein	ja	ja	ja	ja	663	262	Q
6	14.427.121	G	A	APIAR	nein	ja	nein	nein	nein	302	302	R
21	62.743.832	C	T	BDKRB2	nein	ja	nein	nein	ja	388	388	Q
18	63.545.639	C	T	BOSTAUV1R430	nein	ja	nein	nein	ja	302	124	Q
23	25.620.856	G	T	BTNL2	ja	ja	nein	ja	nein	527	90	E
1	264.831	T	A	CBX3	ja	ja	ja	ja	ja	181	5	K
19	57.577.546	G	A	CD300LB	nein	nein	ja	ja	ja	200	139	W
9	93.261.416	G	A	CLDN20	nein	ja	nein	-	nein	237	56	W
24	4.248.577	G	T	CNDP1	nein	ja	nein	nein	nein	510	248	Y
5	42.523.802	C	T	CPNE8	nein	nein	nein	-	ja	564	127	Q
8	11.434.686	C	T	DDX58	ja	nein	ja	nein	ja	942	39	Q
27	5.425.524	G	A	DEFB4	nein	nein	ja	ja	ja	63	57	R
25	2.993.664	C	G	DNASE1	ja	ja	nein	ja	ja	282	85	Y
18	35.576.738	G	C	DPEP3	ja	ja	ja	ja	ja	481	368	Y
24	25.915.403	A	C	DSG2	nein	ja	ja	ja	ja	1120	1092	L
10	65.509.933	G	A	DUOX1	ja	ja	nein	nein	nein	1553	741	Q
2	95.390.113	G	A	DYTN	ja	ja	nein	nein	nein	585	335	R
1	44.652.737	T	A	EFHB	nein	nein	nein	nein	nein	838	777	R
24	35.956.168	G	A	ENOSF1	nein	nein	nein	nein	nein	443	178	R
7	98.612.967	G	A	ERAP1	ja	ja	nein	nein	nein	942	54	R
25	18.658.584	G	T	ERI2	ja	nein	nein	nein	nein	689	290	Y
13	52.510.391	G	T	FASTKD5	nein	ja	nein	-	nein	762	539	E
18	49.743.216	C	T	FCGBP	ja	nein	ja	ja	ja	3047	1672	W
3	12.686.233	C	A	FCRL3	nein	nein	ja	ja	ja	718	98	S
8	29.433.977	C	T	FREM1	ja	nein	nein	nein	nein	2180	1131	R
15	42.515.384	G	A	GIN1	nein	nein	ja	ja	ja	521	405	R
15	42.515.812	C	T	GIN1	nein	ja	nein	nein	nein	521	262	W
19	38.204.969	C	T	GIP	ja	ja	nein	nein	ja	146	42	R
8	6.857.405	G	A	GLRA3	ja	ja	nein	nein	nein	461	251	R
1	45.468.030	G	T	GPR128	ja	ja	nein	-	nein	795	616	G
11	36.414.623	G	A	GPR75	nein	ja	nein	nein	ja	540	406	R
23	24.909.264	G	A	GSTA5	ja	ja	ja	ja	ja	222	13	R
3	33.798.941	A	T	GSTM1	nein	nein	ja	ja	ja	218	87	C
15	76.831.914	C	T	GYLTLIB	nein	nein	nein	nein	nein	726	720	Q
6	87.202.522	C	T	HSTN	nein	ja	nein	nein	nein	58	46	R
11	87.942.810	C	T	IAH1	nein	nein	ja	ja	ja	249	214	W
3	66.286.574	T	A	IFI44	ja	ja	nein	nein	nein	449	279	K
16	4.484.297	C	T	IL19	nein	ja	ja	ja	ja	177	170	R
4	66.003.777	G	A	INMT	ja	ja	ja	ja	ja	263	97	W
3	17.883.354	T	A	IVL	nein	nein	nein	nein	ja	422	406	K
13	78.195.467	G	C	KCNB1	nein	ja	ja	nein	ja	858	815	Y
19	3.878.639	T	A	KIF2B	nein	ja	nein	nein	nein	722	534	L
19	3.879.025	C	T	KIF2B	nein	nein	nein	nein	nein	722	663	Q
5	10.067.7081	T	A	KLRF1	nein	nein	ja	ja	ja	233	113	K
5	27.490.704	C	T	KRT74	ja	ja	nein	ja	nein	536	316	R
1	4.771.715	G	A	KRTAP15-1	nein	nein	ja	ja	ja	136	61	Q
16	49.274.237	G	A	LAD1	nein	ja	nein	-	nein	459	350	R
18	63.149.625	G	A	LILRB3_0	nein	ja	nein	nein	nein	578	106	W
27	14.698.133	G	A	LRP2BP	nein	nein	nein	nein	nein	348	66	R
5	44.555.705	G	A	LYSB	ja	ja	ja	ja	ja	147	46	W
5	44.555.742	C	T	LYSB	ja	ja	ja	-	ja	147	59	R
21	59.203.478	C	T	MGC152353	ja	ja	ja	ja	ja	290	105	Q

Anhang

Chr	Position	REF	ALT	Gen	NMD	Alle Iso-		Homozygote			Genlänge	AS-Pos	AS
						formen	214 Tiere	1000 Bullen	Sanger				
27	14.166.045	G	A	MLF1IP	nein	ja	nein	nein	ja	409	357	Q	
1	71.079.140	G	T	MUC20	ja	nein	nein	nein	nein	738	377	E	
1	71.104.148	C	T	MUC4	ja	ja	ja	ja	ja	1101	690	W	
3	87.961.507	T	A	MYSM1	nein	nein	ja	ja	ja	831	822	C	
8	7.463.112	G	A	NEIL2	nein	nein	ja	ja	ja	329	325	Q	
16	65.368.596	G	T	NPL	nein	nein	nein	nein	nein	320	314	G	
6	23.862.14	G	A	NPY5R	nein	ja	ja	ja	ja	446	364	R	
18	10.300.363	C	T	OSGIN1	nein	ja	nein	nein	ja	530	54	Q	
23	45.276.662	G	A	PAK1IP1	nein	nein	nein	nein	nein	392	389	R	
5	92.218.488	G	A	PIK3C2G	nein	nein	nein	ja	nein	1446	412	Q	
17	72.089.872	C	T	PLA2G3	nein	nein	nein	ja	nein	516	225	W	
26	37.302.490	C	G	PNLIP	ja	ja	nein	nein	ja	465	284	Y	
24	54.170.070	C	G	POLI	ja	ja	nein	-	nein	747	390	Y	
9	87.891.924	G	A	PPIL4	nein	nein	nein	nein	ja	492	447	R	
14	21.121.932	G	A	PRKDC	nein	ja	nein	-	nein	4124	1184	R	
3	117.646.636	C	T	PRLH	nein	nein	ja	ja	ja	98	85	R	
11	66.998.234	C	A	PROKR1	nein	ja	ja	ja	ja	393	293	Y	
23	34.719.499	T	A	PRP9	nein	ja	nein	nein	nein	238	228	K	
23	34.719.645	G	T	PRP9	nein	ja	ja	ja	ja	238	179	S	
6	4.152.699	C	T	QRFPR	nein	ja	ja	ja	ja	432	412	R	
12	32.842.933	G	A	RASL11A	nein	ja	nein	ja	nein	244	182	R	
3	16.818.590	C	T	S100A13	nein	ja	nein	ja	nein	98	45	Q	
9	69.289.773	C	T	SAMD3	ja	nein	nein	nein	ja	510	212	W	
19	27.956.790	C	T	SHBG	nein	nein	ja	nein	ja	401	52	Q	
7	26.329.342	G	A	SLC27A6	ja	ja	ja	ja	ja	635	85	Q	
3	48.855.688	G	A	SLC44A3	ja	ja	nein	nein	nein	649	178	R	
13	63.168.226	G	T	SPLUNC2B	ja	ja	ja	ja	ja	240	66	E	
21	55.891.444	G	A	STRC	nein	nein	nein	ja	nein	1764	1143	R	
11	19.632.148	G	A	SULT6B1	ja	ja	ja	ja	ja	303	216	R	
11	19.638.808	G	A	SULT6B1	ja	ja	nein	ja	nein	303	140	R	
4	103.163.208	C	T	SVOPL	ja	ja	nein	-	ja	492	181	W	
4	106.103.710	G	T	TAS2R38	nein	ja	nein	nein	nein	335	112	C	
1	70.697.630	C	T	TMED2	nein	ja	nein	ja	nein	201	172	W	
10	89.523.800	G	T	TMEM63C	nein	nein	nein	ja	nein	804	796	E	
19	27.689.622	C	A	TMEM95_0	nein	nein	ja	ja	ja	183	161	C	
15	29.219.124	C	T	TMPRSS4	ja	nein	nein	nein	nein	446	118	R	
3	14.299.316	G	A	TTC24	ja	ja	nein	ja	nein	568	391	R	
3	42.748.263	A	C	VCAM1	ja	nein	nein	nein	nein	739	654	Y	
3	26.175.832	C	T	VTCN1	ja	ja	nein	nein	ja	282	172	R	
16	50.686.215	G	T	WDR8	nein	nein	nein	nein	ja	458	445	E	
18	57.909.120	C	A	ZNF175	nein	ja	ja	ja	ja	702	156	S	
23	30.123.089	C	T	ZNF323	nein	ja	ja	ja	ja	435	192	Q	
23	30.995.430	G	A	ZNF391	nein	ja	nein	ja	ja	358	241	R	
18	63.986.957	C	T	ZNF471	nein	ja	ja	ja	ja	759	471	Q	
18	63.987.384	C	G	ZNF471	nein	ja	nein	nein	nein	759	613	S	
16	64.649.983	G	A	ZNF648	nein	nein	nein	ja	nein	564	57	R	

Die verifizierten Stopp-Codons sind alphabetisch nach den Gennamen geordnet. Die Chromosomen- und Basenposition entsprechen der Position des UMD3.1-Assembly (ZIMLN et al. 2009). REF = Referenzallele; ALT = Alternativallele; NMD = Nonsense-vermittelter mRNA-Abbau; Homozygote = Erscheinen von Homozygoten im Datensatz der 214 re-sequenzierten Tiere, im Datensatz des 1000-Bullen-Genom-Projekts und bei der Sanger-Re-Sequenzierung; Genlänge = Länge der Proteinsequenz der längsten Isoform nach der manuellen Annotation; AS-Pos = von der Stoppmutation betroffene Aminosäureposition; AS = betroffene Aminosäure.

Tabelle 9.7: Übersicht über die bei der Sanger-Re-Sequenzierung verwendeten Primer

Gen	Primer-ID	Richtung	Sequenz	Region
ACCSL	10369	forward	GAGGAAGCATTTGGTCAGGA	Exon 14
ACCSL	10370	reverse	AGAGGGCCTAGAAGGACTGG	
ADCK1	10371	forward	CGCCATCCATTAATAACACC	Exon 7
ADCK1	10372	reverse	CCTTGCCATCTGATCCTGTT	
ALOX15	10339	forward	TCACCATCCTTCCTCAAAGG	Exon 6
ALOX15	10340	reverse	ACTGTGGAATTCGGTCTTGC	
APIAR	10127	forward	TCAGCCTCTGATGATCCAA	Exon10
APIAR	10128	reverse	GGAAAAGCAGAAGGCAGATG	
BDKRB2	10264	forward	GCGTAACAACGAGATGCAGA	Exon 3
BDKRB2	10265	reverse	CAGGGCAGGAATTTACTGGA	
BTNL2	10304	forward	GAAGATGCCCTCCTCACTTG	Exon3
BTNL2	10305	reverse	CATGAATTTCTTGCCTGAGCA	
BOSTAUV1R430	10137	forward	GCCTACACAATGGCTGCTTT	Exon 1
BOSTAUV1R430	10138	reverse	TGGTGTCTGTGCAGGAGAAG	
CBX3	10343	forward	AGTAAAGGGGCAGGGAGAAA	Exon 1
CBX3	10344	reverse	GCATCTGTAAACCCCTTCCA	
CD300LB	10069	forward	GTTGAGCGTGCAATGTCACT	Exon 2
CD300LB	10070	reverse	TACCCAGATGTCCCCTCTGT	
CLDN20	10377	forward	TCTTCTTTCCCCAGCCTTTT	Exon 2
CLDN20	10378	reverse	GCCGAAATGATCCAAGGTA	
CNDP1	10379	forward	TGATTTGCCCCATCGTAAAT	Exon 6
CNDP1	10380	reverse	CTCCTCTGGGACCCCTTATC	
CPNE8	10453	forward	AAGGCTGTGCAAACCTAGGG	Exon 6
CPNE8	10454	reverse	CATGGCATATGACAGCTGAA	
DDX58	10049	forward	TTTCTGGTTGCTGTGTCCAC	Exon 2
DDX58	10050	reverse	CTGCTTGTGTAAGGGCATCC	
DEFB4	10381	forward	GCAGGAAGGAGGATGTAGCA	Exon 2
DEFB4	10382	reverse	ATTCCGCTTGGGATCAGACAC	
DNASE1	10023	forward	CAGCAACATTTCTGGCCTTGT	Exon 4
DNASE1	10024	reverse	CCAGTTCTTGGCCCTCAAACC	
DPEP3	10073	forward	TCCTTGTCCCTTCCTGTGGTC	Exon 9
DPEP3	10074	reverse	TGAAGCCACTGAACCCCTCTT	
DSG2	10324	forward	CCTAGGGAACCAGCAGTACG	Exon 15
DSG2	10325	reverse	AGCCAGGTCCCTTTGAGAACA	
DUOX1	10383	forward	TCCCAAGGAGTATGACCTG	Exon 17
DUOX1	10384	reverse	TTAGATAACCGCCTTCCTG	
DYTN	10385	forward	CAGCTCCTTCTCCACAGAG	Exon 13
DYTN	10386	reverse	ATCGTCTGGCTGATGAAAGG	
EFHB	10455	forward	ATCTCAAAGCATTGCCATC	Exon 12
EFHB	10456	reverse	AGGGGCTCCTGATGACTTTT	
ENOSF1	10389	forward	GTGGCGTAGGTACTGTGCTGGT	Exon 7
ENOSF1	10390	reverse	GCATCCATACCCAACCATCT	
ERAP1	10353	forward	TATGGGTGGATGCAATCTGA	Exon 1
ERAP1	10354	reverse	CCTCACTTCTCCTTCCTTGG	
ERI2	10393	forward	AATCGGTGCTGACGTTTTTC	Exon 9
ERI2	10394	reverse	TATTTTGCTCGGGCTGAGAT	
FASTKD5	10395	forward	CGATACCCAGAACACCTGCT	Exon 4
FASTKD5	10396	reverse	CTGCCCATTTGCTGACTCAA	
FCGBP	10266	forward	TTGACAGGGTCTTGGACACA	Exon 11
FCGBP	10267	reverse	GCCTTTGTGGATGGAGATGT	
FCRL3	10133	forward	GGATTCCTTTCTCCAGCTC	Exon 3
FCRL3	10134	reverse	CCTTAGGCAATCCCCATGTA	
FREM1	10322	forward	GCTCCATTTTGGTGTGCTCT	Exon 18
FREM1	10323	reverse	CTGCCAAGCTCTGTGATCTG	
GIN1	10457	forward	TTGGAGAAGGAAATGGCAAC	Exon 1
GIN1	10458	reverse	ACCCGGAGGACTTAGAGGAA	
GIP	10254	forward	GCATGGACAAAAGGGTGACT	Exon 3
GIP	10255	reverse	GAGTCATTTGGCAGCAGTTCA	
GLRA3	10355	forward	TCCTGGCTAGAACTCAAGG	Exon 7
GLRA3	10356	reverse	CATGAGTGTCTGCCCTTTCA	
GPR128	10268	forward	TGGCAGTATCAGTGCATTCA	Exon14
GPR128	10269	reverse	AACATGCCCTGCCAATTTCTC	
GPR75	10270	forward	GGCTGAGAAGGAAAGTGCTG	Exon 2
GPR75	10271	reverse	GCAAACCAAGAGATGCCAAT	
GSTA5	10001	forward	TGAGACTTGGTTTGGTGCAA	Exon 2
GSTA5	10002	reverse	CCGCAGGCAGGATTTATTTAC	
GSTM1	10347	forward	AGGGCAGTCTGGAGTTCAAA	Exon 5
GSTM1	10348	reverse	AGGACCTGGGCTAATGAGGT	
GYLTL1B	10399	forward	CTCCCCAGCCTACACAGT	Exon 13
GYLTL1B	10340	reverse	GCTCATGGTGTGTGGGTTTG	
HSTN	10401	forward	AGCAAAGCCACCTAAACTGA	Exon 5
HSTN	10402	reverse	GGACCTACACTAGGCTGCATTT	

Anhang

Gen	Primer-ID	Richtung	Sequenz	Region
IAH1	10403	forward	CTTGGCTGGACGGTACAAAT	
IAH1	10404	reverse	TGCACCAGTGAAGTT	Exon 6
IFI44	9995	forward	CTCCCCAGCATTTTAGCTT	
IFI44	9996	reverse	CCTGGAGAATCCCATGAAAA	Exon 6
IL19	10208	forward	ATGCCCTGCAGAGATAATGG	
IL19	10209	reverse	CAGCCGATGAGATTTTAGCC	Exon 5
INMT	10139	forward	AGGGGGACCCAGTATGTACC	
INMT	10140	reverse	GGATGGGAGACCTCTTGTGA	Exon 2
IVL	10272	forward	GCTGGAGCAGGAGAAAACAC	
IVL	10273	reverse	TCCAAGCCTTCCAGGATTTA	Exon 2
KCNB1	10276	forward	ATTGATGCGGACACAGATGA	
KCNB1	10277	reverse	GAGGGCTGCAGAGCTGTAC	Exon 2
KIF2B	10175	forward	TGGAAGGAAGCAAGCCAATG	
KIF2B	10176	reverse	AGCATCTCTGGCCTTCTACT	Exon 1
KIF2B	10177	forward	CATGGCAAGTTTCCCTTGT	
KIF2B	10178	reverse	CCCTGAAGGGACATTTCTGA	Exon 1
KLRF1	10179	forward	GCCTGGCACAGATAGATCC	
KLRF1	10180	reverse	AGAGGCCAAAGCTTGAGAAA	Exon 3
KRT74	10316	forward	GTGTGAGCCTTTTCCCTCAGC	
KRT74	10317	reverse	GCTTCGTTCCTCCATCAGAG	Exon5
KRTAP15-1	10027	forward	CAACTGCAGCATGGGAAACT	
KRTAP15-1	10028	reverse	AGCCCTGGTGTGAAGATGG	Exon 1
LAD1	10463	forward	TCTCTTCGCCAGAGTGAGGT	
LAD1	10464	reverse	CAGACAGATGCAGGAAAGCA	Exon 6
LILRB3	10127	forward	AGCTCGCCTCTGATTTCCCTT	
LILRB3	10128	reverse	AGGGAGAGACACTGGAGGAT	Exon 3
LRP2BP	10280	forward	GTTAGTGTGACAGCCCACCA	
LRP2BP	10281	reverse	TCATGATCAAAGCCACAAACA	Exon 3
LYSB	10143	forward	CTTACCTAGGGATGGGAGCA	
LYSB	10144	reverse	TACCACCAAATGGGACCTGT	Exon 2
MGC152353	10145	forward	CTGAATCCTTGCAGCCTTTC	
MGC152353	10146	reverse	CTTCTTCTGGGGTGCAGAGA	Exon 4
MLF1IP	10284	forward	TCGTGGGGCTTCACTGCTAA	
MLF1IP	10285	reverse	GTAGACCTGGCGTCTCCACT	Exon 10
MUC20	10409	forward	AGGCCCTATCCTCAGAGAGC	
MUC20	10410	reverse	GGGTGTGGTGTGTGTAGCAG	Exon 4
MUC4	10411	forward	AAGACCAGGCGTCTCTTCAA	
MUC4	10412	reverse	TCCACTCAC TACCCCACTCC	Exon 16
MYSM1	10413	forward	TCC TCAACCTCCTCAACCTC	
MYSM1	10414	reverse	GCCAGATGGTGTATGTGGTG	Exon 20
NEIL2	10035	forward	CTTTTGGCTTCCCTTGTCCC	
NEIL2	10036	reverse	TAGCATCCCAACCACCCAAT	Exon 5
NPL	10333	forward	AACGTCCCAGCCATGTTTC	
NPL	10334	reverse	CTCCTGAAGGCAAATGGAAT	Exon 12
NPY5R	9997	forward	CCATCCTTCCCAAAGAGTG	
NPY5R	9998	reverse	CAGGACATCATGCC TAGCAA	Exon 2
OSGIN1	10290	forward	TCAGGGGATGACACAGATCA	
OSGIN1	10291	reverse	CCCAGGAGAGCAAACCTGTC	Exon 3
PAK1IP1	10466	forward	CAGTCCAGACGCAACAAAGA	
PAK1IP1	10467	reverse	TCCTCTGTCCATGGGATTTCC	Exon 10
PIK3C2G	10419	forward	ACAGCATTTGAGAGGGCTTGC	
PIK3C2G	10420	reverse	TGTTCGGTTTTTGCCATGAIA	Exon 7
PLA2G3	10335	forward	TCTCGCCCTTCCAGTACAAC	
PLA2G3	10336	reverse	ACAGGACGAGACATGGTTCC	Exon 3
PNLIP	10256	forward	TGTTCTGTGACTGCCAGGAG	
PNLIP	10257	reverse	CATGCCCAATCCTATTCC	Exon 9
POLI	10351	forward	AGTCTGCCAAGATGGAAGGA	
POLI	10352	reverse	ACAACCCAGAATTCATGC	Exon 8
PPIL4	10292	forward	CTGCTTGCTAATGGGAGACA	
PPIL4	10293	reverse	TTTAGCGCCCTACACAATCA	Exon 13
PRLH	9999	forward	AAGGAAGGGGAGTGTCTCT	
PRLH	10000	reverse	GAGGGTCCCTCCAGGCTAAC	Exon 2
PROKR1	10345	forward	GAAACGGTCC TCGTCAATGT	
PROKR1	10346	reverse	CGTGGTTTTGAGGTCAAGGT	Exon 2
PRP9	10421	forward	TTGAAGTTCCACGTGTCTGC	
PRP9	10422	reverse	CCCATTACAGAGCATGCATTA	Exon 5
QRFPR	10047	forward	AACCTTATCCCAGCACGAA	
QRFPR	10048	reverse	ACATGAGTGGCTGGTCAGT	Exon 6
RASL11A	10423	forward	CTTTTCAGGTCCAGGACAGC	
RASL11A	10424	reverse	TAGCAGGAATGCATGGGTTA	Exon 6
S100A13	10429	forward	CCCTCATCCTTGTCCCTTTT	
S100A13	10430	reverse	GCTGGGAAAATGAAACCAGA	Exon 1
SAMD3	10288	forward	CAGTGAACACCCACACACCT	
SAMD3	10289	reverse	TGGAACCCCTTCTTGAGAAA	Exon 7
SLC27A6	10059	forward	AGTCAATGTGTGCTGTGCTG	
SLC27A6	10060	reverse	CAGTATCGCCCTTTTCAGC	Exon 1

Gen	Primer-ID	Richtung	Sequenz	Region
SLC44A3	10433	forward	TTGCAAACAGGGCTTTAACC	Exon 6
SLC44A3	10434	reverse	ACTTTCCATGGCCAGTTCAC	
SPLUNC2B	10032	forward	TCCAGACCCATGAGATGCTC	Exon 3
SPLUNC2B	10104	reverse	TAAGATGAAATAAAATAACCTA	
STRC	10318	forward	GCTCTAGAGTGTGCCCTGCT	Exon 15
STRC	10319	reverse	GCTTCCCAGCAGAGCTAATG	
SULT6B1	10437	forward	CAGCTGGTTTAGACTGGTGTCT	Exon 6
SULT6B1	10438	reverse	CAAGGCAGCTTTACCTTTTGC	
SULT6B1	10435	forward	TCCTGAGCAGCTGTATGTGG	Exon 4
SULT6B1	10436	reverse	CAAGGCAGCTTTACCTTTTGC	
SVOPL	10294	forward	GTCTTTTCCAACGCGTCAGT	Exon 7
SVOPL	10295	reverse	TGCCAAAACACCCTTTTCTC	
TAS2R38	10320	forward	TGCCCTCTGAAGTCAGGAAT	Exon 1
TAS2R38	10321	reverse	AGAACACCAGCATCCAGAG	
TMED2	10258	forward	TGCTCACATGGATGGAACAT	Exon 1
TMED2	10259	reverse	TTGGGAGGGCCCTGTATAATG	
TMEM63C	10439	forward	TCCTGAGCAGCTGTATGTGG	Exon 22
TMEM63C	10440	reverse	GGGGAAAGATCCAGAGAGTC	
TMPRSS4	10441	forward	TGGTCACCCCTAGCTCCTGTC	Exon 5
TMPRSS4	10442	reverse	CCCACCACACACACATCT	
TTC24	10443	forward	CCACCTCTGTCCATGGTTCT	Exon 6
TTC24	10444	reverse	CCCCTGGGCTAGTTACCAA	
VCAM1	10337	forward	CTGCTCACTTCCTGCTTGTG	Exon 8
VCAM1	10338	reverse	CCACCAAGCTTCAGACCTGT	
VTCN1	10260	forward	GCCCCTGGGAAAGTACTGA	Exon 4
VTCN1	10261	reverse	TGGAATTGTGATGGGGTCT	
WDR8	10300	forward	CTCTGGCCACCTGCTTCTC	Exon 12
WDR8	10301	reverse	GCCATGCTTCAGAGTGTCA	
ZNF175	10181	forward	CTGGCGAGCGACAGTATGTA	Exon 5
ZNF175	10182	reverse	CGCTGATGCTCTCTGAACTG	
ZNF323	10013	forward	CAGGTGCTGGTTACCCCTTA	Exon 4
ZNF323	10014	reverse	CCCTGTGTGGATCCTTCTGT	
ZNF391	10447	forward	TTCTGCAAAGGAAGCCTTGT	Exon 3
ZNF391	10448	reverse	ATGTTACAGCAAGGGATGAGC	
ZNF471	10449	forward	TGTGACATCTGTGGGAAAGC	Exon 5
ZNF471	10450	reverse	TAAGCGACTGTCGATGGCTA	
ZNF648	10451	forward	GTTCTGTGGCTGTGAAGGT	Exon 1
ZNF648	10452	reverse	TCCAGATTTGCCCTTGAGTC	

Auflistung der verwendeten Primer mit ihrer internen Identifikationsnummer (Primer-ID). Die Sequenz ist in 5'-3'-Orientierung und die Region der Primer in der genomischen Organisation der Gene angegeben.

Tabelle 9.8: Identifizierte Varianten im FGF23-Gen

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r ²
106.207.894	A	C	0,88	PROM	2,92x10 ⁻⁸	0,05
106.208.395	G	C	0,98	INTRON1	0,98	0,01
106.208.520	A	C	0,81	INTRON1	1,25x10 ⁻⁶	0,48
106.208.532	T	C	0,97	INTRON1	1,95x10 ⁻¹⁰	0
106.208.538	T	C	0,99	INTRON1	1	0
106.208.565	G	T	0,95	INTRON1	0,95	0
106.210.249	C	T	0,94	INTRON1	0	0,03
106.210.270	A	T	0,86	INTRON1	0,35	0,03
106.210.506	T	G	0,97	INTRON1	0,01	0,02
106.211.450	G	A	0,8	INTRON1	4,52x10 ⁻⁷	0,48
106.211.829	C	T	0,77	INTRON1	0,45	0,14
106.211.842	G	A	0,98	INTRON1	0,98	0,01
106.211.954	T	A	0,86	INTRON1	7,33x10 ⁻⁵	0,32
106.214.010	T	C	0,98	INTRON1	3,14x10 ⁻⁷	0
106.214.125	A	C	0,94	INTRON1	0	0,03
106.214.130	A	T	0,98	INTRON1	0,03	0,03
106.214.178	T	G	0,62	INTRON1	0,01	0
106.214.730	C	T	0,61	EXON2	0,01	0
106.215.221	A	G	0,86	INTRON2	6,61x10 ⁻⁵	0,32
106.215.328	G	A	0,86	INTRON2	6,40x10 ⁻⁵	0,32

Anhang

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r^2
106.215.570	T	C	0,61	INTRON2	0,01	0
106.216.089	A	C	0,99	INTRON2	$9,87 \times 10^{-8}$	0
106.216.222	C	A	0,98	INTRON2	0,98	0,01
106.216.293	A	C	0,99	INTRON2	$9,37 \times 10^{-8}$	0
106.216.557	G	A	0,6	EXON3	$2,54 \times 10^{-11}$	0,37
106.216.613	C	T	0,79	EXON3	$1,14 \times 10^{-6}$	0,48

Übersicht der in der genomweiten Assoziationsstudie identifizierten Polymorphismen im Gen FGF23. Die Position entspricht der Lokalisation der Variante auf dem bovinen UMD3.1-Assembly (ZIMIN et al. 2009). Die entsprechenden Allele sind mit A1 und A2 dargestellt. Die Region ist in der genomischen Organisation des Gen angegeben. Der P-Wert resultiert aus den Ergebnissen der genomweiten Assoziationsstudie. Das Kopplungsungleichgewicht zum Top-Marker ist mit LD als r^2 angegeben.

Tabelle 9.9: Identifizierte Varianten im *ARRDC3*-Gen

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r^2
93.253.734	T	G	0,97	PROM	0,92	0
93.253.721	A	C	0,98	PROM	0,53	0
93.253.713	G	A	0,91	PROM	0,13	0,01
93.253.697	G	T	0,79	PROM	0,32	0,04
93.253.575	A	G	0,99	PROM	0,11	0
93.253.434	G	C	0,87	PROM	0,05	0,02
93.253.415	A	G	0,96	PROM	0,21	0
93.253.108	C	G	0,97	PROM	0,14	0
93.253.064	C	T	0,97	5'UTR	0,54	0
93.252.569	C	A	0,98	INTRON1	0,02	0
93.252.561	A	C	0,99	INTRON1	0,04	0
93.251.492	A	C	1	INTRON1	0,54	0
93.251.138	C	T	0,89	INTRON1	$2,43 \times 10^{-7}$	0,85
93.250.685	A	G	0,73	INTRON1	0,51	0,04
93.250.624	G	A	0,96	INTRON1	0,22	0
93.250.503	C	T	0,73	INTRON1	0,51	0,04
93.250.422	A	G	0,73	INTRON1	0,5	0,04
93.250.391	A	G	0,73	INTRON1	0,5	0,04
93.250.023	A	C	1	INTRON1	0,53	0
93.249.112	A	G	0,77	INTRON1	$1,25 \times 10^{-7}$	0,55
93.249.110	A	G	0,77	INTRON1	$1,29 \times 10^{-7}$	0,55
93.248.876	A	C	0,67	INTRON1	0,02	0,14
93.248.872	A	C	0,72	INTRON1	0,01	0,09
93.248.857	C	A	0,77	INTRON1	0	0,43
93.248.840	C	T	0,97	INTRON1	0,91	0
93.248.679	T	C	0,81	INTRON1	$5,11 \times 10^{-7}$	0,46
93.248.586	G	A	0,98	INTRON1	0,89	0
93.248.571	T	A	0,57	INTRON1	0,13	0,1
93.248.503	T	G	0,99	INTRON1	0,67	0
93.248.499	G	A	0,98	INTRON1	0,72	0
93.248.492	A	C	0,5	INTRON1	0	0,11
93.248.451	A	T	0,96	INTRON1	0,23	0,01
93.247.869	C	T	0,86	INTRON2	0,04	0,02
93.247.780	T	G	0,8	INTRON2	$2,05 \times 10^{-6}$	0,44
93.247.504	A	G	0,81	INTRON2	$7,05 \times 10^{-7}$	0,46
93.247.383	A	G	0,86	INTRON2	0,23	0,02
93.246.278	C	G	0,72	INTRON2	$7,75 \times 10^{-5}$	0,33
93.246.272	A	C	0,94	INTRON2	0,36	0
93.246.216	C	T	0,5	INTRON2	0	0,11
93.245.632	T	G	0,96	INTRON3	0,23	0
93.245.478	A	G	0,81	INTRON3	$5,31 \times 10^{-7}$	0,46
93.245.412	T	C	0,73	INTRON3	0,41	0,04
93.245.253	C	A	0,98	INTRON3	0,24	0
93.245.010	G	C	0,86	INTRON3	0,23	0,02
93.244.989	A	G	0,87	INTRON3	0,23	0,02
93.244.933	C	T	0,9	EXON3	$4,33 \times 10^{-9}$	1
93.244.823	T	C	0,5	INTRON4	0	0,11
93.244.714	A	T	0,5	INTRON4	0	0,11
93.244.554	A	G	0,5	INTRON4	0	0,11
93.244.537	A	G	0,96	INTRON4	0,52	0
93.244.532	T	A	0,7	INTRON4	0,03	0,05
93.244.226	C	T	0,73	INTRON5	0,39	0,04
93.244.045	T	G	0,81	INTRON5	$5,10 \times 10^{-7}$	0,46
93.243.714	T	C	0,81	INTRON6	$5,10 \times 10^{-7}$	0,46
93.243.389	T	C	0,8	EXON7	$2,89 \times 10^{-7}$	0,46

Anhang

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r^2
93.243.245	A	G	0,81	INTRON7	$3,30 \times 10^{-7}$	0,46
93.242.849	C	T	0,86	INTRON7	0,23	0,02
93.242.662	C	T	0,96	INTRON7	0,14	0
93.242.658	A	C	0,5	INTRON7	0	0,11
93.242.436	G	C	0,52	INTRON7	0	0,1
93.242.153	T	C	0,54	INTRON7	0,01	0,09
93.242.073	A	C	0,96	INTRON7	0,17	0
93.241.938	A	C	0,81	INTRON7	$3,30 \times 10^{-7}$	0,46
93.241.749	A	C	0,96	INTRON7	0,17	0
93.241.710	G	T	0,96	INTRON7	0,32	0
93.241.604	T	C	0,81	INTRON7	$3,30 \times 10^{-7}$	0,46
93.241.556	G	A	0,99	INTRON7	0,35	0
93.240.856	G	A	1	EXON8	0,36	0

Übersicht der in der genomweiten Assoziationsstudie identifizierten Polymorphismen im Gen *ARRDC3*. Die Position entspricht der Lokalisation der Variante auf dem bovinen UMD3.1-Assembly (ZIMIN et al. 2009). Die entsprechenden Allele sind mit A1 und A2 dargestellt. Die Region ist in der genomischen Organisation des Gen angegeben. Der P-Wert resultiert aus den Ergebnissen der genomweiten Assoziationsstudie. Das Kopplungsungleichgewicht zum Top-Marker ist mit LD als r^2 angegeben.

Tabelle 9.10: Identifizierte Varianten im VPS35-Gen

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r2
15.058.306	G	T	0,75	PROM	0,02	0,23
15.058.275	T	A	0,75	PROM	0,02	0,23
15.058.271	T	G	0,77	PROM	0	0,21
15.058.265	T	C	0,77	PROM	0	0,21
15.058.224	C	G	0,75	PROM	0,02	0,23
15.058.187	C	T	0,75	PROM	0,02	0,23
15.058.167	G	A	0,73	PROM	0,03	0,23
15.058.133	T	C	0,61	PROM	0	0,31
15.058.109	T	C	0,77	PROM	0	0,21
15.058.045	A	G	0,75	PROM	0,02	0,23
15.058.041	T	G	0,75	PROM	0,02	0,23
15.058.022	C	T	0,75	PROM	0,02	0,23
15.058.003	G	A	0,77	PROM	0	0,21
15.057.878	T	C	0,74	PROM	0,01	0,23
15.057.851	A	C	0,98	PROM	0	0
15.057.833	G	A	0,97	PROM	0,27	0,02
15.057.804	T	C	0,75	PROM	0,02	0,23
15.057.771	C	A	0,75	PROM	0,02	0,23
15.057.726	G	A	0,75	PROM	0,02	0,23
15.057.707	C	T	0,75	PROM	0,02	0,23
15.057.681	C	T	0,98	PROM	0,27	0,02
15.057.649	T	A	0,95	PROM	0,01	0
15.057.638	A	T	0,6	PROM	0,01	0,41
15.057.583	T	C	0,75	PROM	0,02	0,23
15.057.552	C	T	0,75	PROM	0,02	0,23
15.057.543	C	A	0,74	PROM	0,02	0,23
15.057.442	A	G	0,75	PROM	0	0,21
15.057.370	G	A	0,93	PROM	0,11	0,06
15.057.361	G	T	0,98	PROM	1,37x10 ⁻⁵	0
15.057.336	A	C	0,59	PROM	0,05	0,07
15.057.327	A	G	0,71	5'UTR	0,02	0,23
15.057.312	G	A	0,76	5'UTR	0	0,19
15.057.304	G	T	0,98	5'UTR	0,27	0,02
15.057.289	C	A	0,73	INTRON1	0,02	0,23
15.057.193	T	C	0,75	INTRON1	0,02	0,23
15.057.144	G	A	0,77	INTRON1	0	0,21
15.057.115	C	A	0,75	INTRON1	0,02	0,23
15.057.106	G	C	0,77	INTRON1	0	0,21
15.057.077	C	T	0,75	INTRON1	0,02	0,23
15.056.979	T	A	0,98	INTRON1	0,27	0,02
15.056.762	G	A	0,99	INTRON2	0,01	0
15.056.740	C	T	0,93	INTRON2	0,11	0,06
15.056.635	C	T	0,8	INTRON2	0,77	0,01
15.056.611	A	G	0,61	INTRON2	0,42	0,18
15.056.537	T	C	0,59	INTRON2	1,32x10 ⁻⁶	0,59
15.056.509	T	A	0,95	INTRON2	0,98	0,03
15.056.341	C	A	0,95	INTRON2	0,98	0,03
15.056.323	G	A	0,58	INTRON2	1,14x10 ⁻⁵	0,64
15.056.297	T	C	0,95	INTRON2	0,98	0,03
15.056.228	G	A	0,8	INTRON2	4,47x10 ⁻⁶	0,35
15.056.199	T	C	0,73	INTRON2	0	0,26
15.056.151	T	C	0,95	INTRON2	0,98	0,03
15.055.935	T	C	0,98	INTRON2	0,18	0,01
15.055.918	T	C	0,96	INTRON2	0,72	0,03
15.055.861	T	C	0,96	INTRON2	0,72	0,03

Anhang

15.055.854	A	C	0,96	INTRON2	0,72	0,03
15.055.837	A	G	0,96	INTRON2	0,72	0,03
15.055.682	T	A	0,89	INTRON2	0	0,09
15.055.678	G	A	0,77	INTRON2	0	0,21
15.055.612	T	C	0,77	INTRON2	0	0,21
15.055.465	C	A	0,93	INTRON3	0,11	0,06
15.055.456	T	C	0,95	INTRON3	0,98	0,03
15.055.387	C	T	0,73	INTRON3	0	0,26
15.055.337	T	G	0,77	INTRON3	0	0,21
15.055.321	G	A	0,77	INTRON3	0	0,21
15.055.303	C	G	0,77	INTRON3	0	0,21
15.055.279	A	G	0,77	INTRON3	0	0,21
15.055.262	C	A	0,77	INTRON3	0	0,21
15.055.258	C	A	0,77	INTRON3	0	0,21
15.055.213	T	A	0,95	INTRON3	0,98	0,03
15.055.160	A	G	0,77	INTRON3	0	0,21
15.055.069	C	A	0,73	INTRON3	0	0,26
15.055.061	C	A	0,77	INTRON3	0	0,21
15.055.008	C	G	0,77	INTRON3	0	0,21
15.054.902	A	G	0,77	INTRON3	0	0,21
15.054.885	A	C	0,55	INTRON3	0,49	0,01
15.054.879	T	A	0,77	INTRON3	0	0,21
15.054.828	C	T	0,77	INTRON3	0	0,21
15.054.585	T	A	0,77	INTRON4	0	0,21
15.054.578	A	G	0,77	INTRON4	0	0,21
15.054.565	T	C	0,77	INTRON4	0	0,21
15.054.550	C	T	0,77	INTRON4	0	0,21
15.054.475	T	C	0,77	INTRON4	0	0,21
15.054.433	C	A	0,77	INTRON4	0	0,21
15.054.410	C	T	0,77	INTRON4	0	0,21
15.054.386	C	A	0,62	INTRON4	0,08	0,15
15.054.270	A	G	0,98	INTRON4	0,27	0,02
15.053.847	A	G	0,73	INTRON5	0	0,26
15.053.821	A	G	0,95	INTRON5	0,99	0,03
15.053.813	G	A	0,73	INTRON5	0	0,26
15.053.788	T	C	0,73	INTRON5	0	0,26
15.053.704	A	G	0,73	INTRON5	0	0,26
15.053.677	A	T	0,72	INTRON5	0	0,26
15.053.640	T	C	0,73	INTRON5	0	0,26
15.053.620	G	A	0,73	INTRON5	0	0,26
15.053.552	T	C	0,77	INTRON5	0	0,21
15.053.450	G	A	0,77	INTRON5	0	0,21
15.053.444	A	G	0,73	INTRON5	0	0,26
15.053.375	T	C	0,77	INTRON5	0	0,21
15.053.206	C	T	0,77	INTRON5	0	0,21
15.053.175	C	T	0,77	INTRON5	0	0,21
15.053.131	T	A	0,97	INTRON5	0,06	0
15.052.955	A	G	0,95	INTRON5	0,99	0,03
15.052.818	T	C	0,77	INTRON5	0	0,21
15.052.697	A	T	0,96	INTRON5	$2,58 \times 10^{-6}$	0
15.052.695	C	G	0,96	INTRON5	$2,58 \times 10^{-6}$	0
15.052.643	C	T	0,77	INTRON5	0	0,21
15.052.576	A	C	0,77	INTRON5	0	0,21
15.052.496	T	C	0,75	INTRON5	0	0,23
15.052.220	G	A	0,77	INTRON6	0	0,21
15.052.021	C	T	0,73	INTRON6	0	0,26
15.051.975	A	G	0,77	INTRON6	0	0,21
15.051.903	T	G	0,73	INTRON6	0	0,26
15.051.883	C	A	0,98	INTRON6	0,97	0

Anhang

15.051.877	G	T	0,97	INTRON6	0,12	0
15.051.823	C	T	0,77	INTRON6	0	0,21
15.051.799	G	A	0,54	INTRON6	0,49	0,01
15.051.724	C	T	0,73	INTRON6	0	0,26
15.051.633	C	G	0,58	INTRON6	0,03	0,08
15.051.465	G	A	0,98	INTRON6	0,27	0,02
15.051.456	C	T	0,98	INTRON6	0,15	0,01
15.051.434	C	A	0,98	INTRON6	0,72	0,01
15.051.342	A	T	0,96	INTRON6	$9,70 \times 10^{-6}$	0
15.051.305	A	T	0,93	INTRON6	0,11	0,06
15.051.284	T	A	0,73	INTRON6	0	0,26
15.051.124	G	A	0,89	INTRON7	0	0,09
15.050.993	C	G	0,95	INTRON7	0,99	0,03
15.050.867	A	C	0,77	INTRON7	0	0,21
15.050.859	T	C	0,77	INTRON7	0	0,21
15.050.843	A	T	0,99	INTRON7	0,02	0
15.050.837	G	A	0,67	INTRON7	0	0,12
15.050.488	C	T	0,77	INTRON8	0	0,21
15.050.455	A	G	0,77	INTRON8	0	0,21
15.050.212	C	T	0,73	INTRON8	0	0,26
15.050.197	C	G	0,77	INTRON8	0	0,21
15.050.082	C	T	0,91	INTRON8	0,32	0,07
15.050.037	G	A	0,98	INTRON8	0,27	0,02
15.049.815	C	T	0,98	INTRON8	0,15	0,01
15.049.772	T	C	0,77	INTRON8	0	0,21
15.049.760	T	C	0,77	INTRON8	0	0,21
15.049.386	A	G	0,73	EXON10	0	0,26
15.049.202	C	T	0,98	INTRON10	$2,21 \times 10^{-5}$	0
15.049.171	A	G	0,77	INTRON10	0	0,21
15.049.073	T	C	0,95	INTRON10	0,86	0,03
15.048.975	C	T	0,95	INTRON10	0,86	0,03
15.048.786	G	A	0,74	INTRON10	0	0,24
15.048.614	C	T	0,95	INTRON10	0,86	0,03
15.048.594	T	C	0,87	INTRON10	0,18	0,09
15.048.581	T	G	0,98	INTRON10	0,15	0,01
15.048.545	G	A	1	INTRON10	0,23	0
15.048.429	G	C	0,93	INTRON10	0,11	0,06
15.048.372	C	T	0,55	INTRON10	0,22	0
15.048.275	T	G	0,75	INTRON10	0	0,24
15.048.132	C	T	0,74	INTRON10	0	0,24
15.048.119	G	T	0,98	INTRON10	$1,55 \times 10^{-5}$	0
15.048.117	C	A	0,73	INTRON10	0	0,24
15.048.109	T	C	0,74	INTRON10	0	0,24
15.048.096	C	G	0,95	INTRON10	0,86	0,03
15.048.087	G	T	0,98	INTRON10	0,16	0,01
15.048.068	C	T	0,98	INTRON10	0,02	0
15.048.043	T	C	0,98	INTRON10	0,16	0,01
15.047.927	A	T	0,79	INTRON10	0	0,19
15.047.921	G	A	0,66	INTRON10	0,13	0,02
15.047.912	C	T	0,74	INTRON10	0	0,24
15.047.892	A	G	0,74	INTRON10	0	0,24
15.047.877	C	G	0,79	INTRON10	0	0,19
15.047.706	T	C	0,98	INTRON10	0,17	0,01
15.047.675	G	A	0,79	INTRON10	0	0,19
15.047.436	G	A	0,79	INTRON11	0	0,19
15.047.412	G	A	0,97	INTRON11	0,41	0,02
15.047.238	G	A	0,98	INTRON11	0,77	0,01
15.047.204	G	A	0,99	INTRON11	0,02	0
15.047.189	T	C	0,98	INTRON11	$6,01 \times 10^{-6}$	0

Anhang

15.047.135	A	C	0,63	INTRON11	0,8	0,02
15.047.126	C	T	0,53	INTRON11	0,87	0,07
15.047.102	C	T	0,94	INTRON11	0,09	0,06
15.046.826	C	G	0,98	INTRON12	0,05	0
15.046.795	G	C	0,65	INTRON12	0,08	0,41
15.046.565	G	A	0,98	INTRON12	0,03	0
15.046.386	T	C	0,65	INTRON12	0,05	0,39
15.046.322	G	C	0,59	INTRON12	0,02	0,49
15.046.094	G	A	0,89	INTRON12	0	0,09
15.046.026	G	C	0,99	INTRON12	0,3	0
15.046.020	A	C	0,96	INTRON12	0,44	0,03
15.045.574	G	T	0,6	INTRON12	0,03	0,49
15.045.249	C	T	0,96	INTRON12	0,45	0,03
15.044.989	C	T	0,9	INTRON12	0,26	0,08
15.044.910	T	G	0,9	INTRON12	0,24	0,08
15.044.885	T	A	0,98	INTRON12	0	0
15.044.876	A	T	0,96	INTRON12	0	0
15.044.831	A	G	0,9	INTRON12	0,24	0,08
15.044.489	A	C	0,93	INTRON13	0,09	0,06
15.044.366	G	C	0,69	INTRON13	0,06	0,33
15.043.857	A	C	0,59	INTRON13	0,02	0,49
15.043.766	A	G	0,96	INTRON13	0,49	0,03
15.043.335	T	A	0,9	INTRON13	0,3	0,07
15.043.299	T	C	0,99	INTRON13	0,92	0
15.042.699	G	T	0,98	INTRON14	0,28	0,02
15.042.453	G	A	0,78	INTRON14	0,04	0,29
15.042.359	T	A	0,6	INTRON14	0,03	0,49
15.042.152	C	T	0,93	EXON15	0,08	0,06
15.041.797	G	C	0,7	INTRON15	0,11	0,33
15.041.760	C	T	0,9	INTRON15	0,31	0,07
15.041.664	A	C	0,98	INTRON15	0,28	0,02
15.040.630	T	C	0,9	INTRON16	0,31	0,07
15.040.255	A	T	0,71	INTRON16	0,12	0,33
15.039.948	T	C	0,9	INTRON16	0,31	0,07
15.039.912	A	C	1	INTRON16	0,61	0
15.039.891	G	A	0,94	INTRON16	0	0,01
15.039.844	C	T	0,7	INTRON16	0,09	0,33
15.039.807	A	G	0,7	INTRON16	0,09	0,33
15.039.709	C	T	0,61	EXON17	$9,85 \times 10^{-9}$	1
15.039.449	C	T	0,94	3'UTR	0,47	0,04

Übersicht der in der genomweiten Assoziationsstudie identifizierten Polymorphismen im Gen VPS35. Die Position entspricht der Lokalisation der Variante auf dem bovinen UMD3.1-Assembly (ZIMIN et al. 2009). Die entsprechenden Allele sind mit A1 und A2 dargestellt. Die Region ist in der genomischen Organisation des Gen angegeben. Der P-Wert resultiert aus den Ergebnissen der genomweiten Assoziationsstudie. Das Kopplungsungleichgewicht zum Top-Marker ist mit LD als r^2 angegeben.

Tabelle 9.11: Identifizierte Varianten im LOC787812-Gen

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r2
59.556.115	T	G	0,99	PROM	4,41x10 ⁻⁵	0
59.555.936	G	A	0,95	PROM	0,37	0
59.555.697	G	A	0,99	PROM	0,33	0
59.555.486	G	C	0,97	PROM	0,33	0
59.555.382	C	T	0,95	5'UTR	4,52x10 ⁻⁵	0,01
59.555.340	A	C	0,99	5'UTR	0,34	0
59.554.972	G	C	0,85	INTRON1	0,07	0,02
59.554.114	A	C	0,99	INTRON1	0,34	0
59.554.041	A	C	0,99	INTRON1	2,83x10 ⁻⁵	0
59.553.985	C	T	0,72	INTRON1	0	0,05
59.553.883	A	G	0,95	INTRON1	4,30x10 ⁻⁵	0,01
59.553.619	T	A	0,89	INTRON1	0,18	0,02
59.553.609	C	A	0,89	INTRON1	0,18	0,02
59.552.638	A	G	0,99	INTRON1	0,33	0
59.552.011	T	C	0,77	INTRON1	0,06	0,04
59.551.608	C	T	0,93	INTRON1	0,38	0,01
59.550.539	A	G	0,67	INTRON1	0	0,05
59.550.333	C	A	0,72	INTRON1	0	0,05
59.550.294	C	T	0,72	INTRON1	0	0,05
59.550.287	C	G	0,89	INTRON1	0,17	0,02
59.549.353	T	C	0,89	INTRON1	0,18	0,02
59.549.349	T	C	0,89	INTRON1	0,18	0,02
59.548.049	G	A	0,97	INTRON2	0,33	0
59.547.821	G	C	0,97	INTRON2	0,33	0
59.547.589	A	C	0,94	INTRON2	3,23x10 ⁻⁵	0,01
59.546.781	G	A	0,97	INTRON2	0,33	0
59.546.222	T	C	0,73	INTRON2	0	0,05
59.545.901	C	T	0,96	INTRON2	1,30x10 ⁻¹¹	0
59.545.314	C	T	0,97	INTRON2	0,33	0
59.545.076	A	G	0,66	INTRON2	0	0,05
59.544.983	T	C	0,99	INTRON2	0,3	0
59.544.705	C	G	0,66	INTRON2	0	0,05
59.544.375	A	C	0,95	INTRON2	0,34	0
59.544.028	C	G	0,95	INTRON2	4,31x10 ⁻⁵	0,01
59.543.779	G	A	0,95	INTRON2	4,31x10 ⁻⁵	0,01
59.543.100	A	C	0,98	INTRON3	1,10x10 ⁻⁷	0
59542965	T	C	0,97	INTRON3	0,32	0
59.542.707	C	A	0,92	INTRON3	0,33	0,01
59.542.467	A	G	0,95	INTRON3	4,31x10 ⁻⁵	0,01
59.542.432	C	A	0,67	INTRON3	0	0,05
59.541.806	A	C	0,77	INTRON3	0,07	0,04
59.541.664	G	C	0,77	INTRON3	0,06	0,04
59.541.380	A	G	0,77	INTRON3	0,07	0,04
59.541.266	G	C	0,98	INTRON3	0	0
59.541.230	A	C	0,96	INTRON3	0	0
59.541.225	A	G	0,96	INTRON3	0	0
59.541.118	A	G	0,78	INTRON3	0,07	0,04
59.541.043	T	C	0,77	INTRON3	0,06	0,04
59.540.709	A	G	0,77	INTRON3	0,07	0,04
59.540.534	T	C	0,77	INTRON3	0,06	0,04
59.540.253	C	T	0,99	INTRON3	0,32	0
59.540.215	T	A	0,95	INTRON3	4,70x10 ⁻⁵	0,01
59.540.150	A	G	0,72	INTRON3	0	0,05
59.538.869	T	A	0,67	INTRON3	0	0,05
59.538.646	A	G	0,72	INTRON3	0	0,05

Anhang

Position (bp)	A1	A2	AF	Region	P-Wert	LD als r²
59.537.487	G	A	0,95	INTRON3	0,29	0
59.537.441	G	C	0,99	INTRON3	1,63x10 ⁻¹⁸	0
59.537.245	A	C	0,95	INTRON3	0,29	0
59.537.049	T	A	0,93	INTRON3	0,36	0,01
59.535.738	G	C	0,72	EXON4	0	0,05
59.535.519	G	C	0,93	EXON4	0,36	0,01
59.534.816	T	C	0,89	EXON4	0,16	0,02
59.534.257	C	T	0,71	EXON5	0	0,05
59.533.556	T	C	0,97	INTRON5	0	0
59.533.523	G	A	0,99	INTRON5	0	0
59.533.428	A	G	0,99	INTRON5	1,79x10 ⁻³⁷	0

Übersicht der in der genomweiten Assoziationsstudie identifizierten Polymorphismen im Gen LOC787812. Die Position entspricht der Lokalisation der Variante auf dem bovinen UMD3.1-Assembly (ZIMIN et al. 2009). Die entsprechenden Allele sind mit A1 und A2 dargestellt. Die Region ist in der genomischen Organisation des Gen angegeben. Der P-Wert resultiert aus den Ergebnissen der genomweiten Assoziationsstudie. Das Kopplungsungleichgewicht zum Top-Marker ist mit LD als r^2 angegeben.

PREDICTED: zinc finger protein 845 [Bos taurus]

Sequence ID: [ref|XP_001255074.5|](#) Length: 807 Number of Matches: 3

Range 1: 1 to 760 [GenPept](#) [Graphics](#)

▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
1580 bits(4090)	0.0	Compositional matrix adjust.	760/760(100%)	760/760(100%)	0/760(0%)
Query 1		MALSQTQLTVEDVDIKFTPEEWECLDPAQRALYQDVMVETYNLLSVDVSRIHMTMKLQA			60
Sbjct 1		MALSQTQLTVEDVDIKFTPEEWECLDPAQRALYQDVMVETYNLLSVDVSRIHMTMKLQA			60
Query 61		KGNNKGGEIFQRVMFGRAKTHEIKDYFLRNIQETVIRGFENLWTDERRDNGIFISYIKNL			120
Sbjct 61		KGNNKGGEIFQRVMFGRAKTHEIKDYFLRNIQETVIRGFENLWTDERRDNGIFISYIKNL			120
Query 121		TDGRDHQSRNDAGDKPVERLGSSFHDELQMLQSEGTFECSPPVVTNINSSPLGLPTQRTR			180
Sbjct 121		TDGRDHQSRNDAGDKPVERLGSSFHDELQMLQSEGTFECSPPVVTNINSSPLGLPTQRTR			180
Query 181		RVRKGICHKRESAPMHPSKQAPDQEGHKKKSYKSNECGITILQDSELTRHQIHTGRKPY			240
Sbjct 181		RVRKGICHKRESAPMHPSKQAPDQEGHKKKSYKSNECGITILQDSELTRHQIHTGRKPY			240
Query 241		KADVYGKAFNENKSLAVHRTNHTGEKPYKCDVCSHSFKRNSALQIHLRVHTGERPYKCDV			300
Sbjct 241		KADVYGKAFNENKSLAVHRTNHTGEKPYKCDVCSHSFKRNSALQIHLRVHTGERPYKCDV			300
Query 301		CGHSFKRNTALQIHLRVHTGERPYKCDVCGHTFKRNAALQIHLRVHTGERPYKCDVCGHS			360
Sbjct 301		CGHSFKRNTALQIHLRVHTGERPYKCDVCGHTFKRNAALQIHLRVHTGERPYKCDVCGHS			360
Query 361		FKHQTHLQNHGRTHHTGEKPYKCDICGRGYPRKAQLGIHQRVHTGEKPYKCDVCGKAFTRK			420
Sbjct 361		FKHQTHLQNHGRTHHTGEKPYKCDICGRGYPRKAQLGIHQRVHTGEKPYKCDVCGKAFTRK			420
Query 421		ESHALHQILHTGEKPYKCDLGGQAFTRKESHTLHQILHTGEKPYKCDVCGRGYTRSRQLA			480
Sbjct 421		ESHALHQILHTGEKPYKCDLGGQAFTRKESHTLHQILHTGEKPYKCDVCGRGYTRSRQLA			480
Query 481		IHQRVHTGEKPYKCDKASKCNLCGKMFTSRCYLAVHQRSHTGEKPYKCDVCGKAFTRKQSQ			540
Sbjct 481		IHQRVHTGEKPYKCDKASKCNLCGKMFTSRCYLAVHQRSHTGEKPYKCDVCGKAFTRKQSQ			540
Query 541		HALHQILHSGEKPYKCDVCGGYTRKSHLEIHCVRVHTGEKPYKCDVCGKAFTRKDSCTLH			600
Sbjct 541		HALHQILHSGEKPYKCDVCGGYTRKSHLEIHCVRVHTGEKPYKCDVCGKAFTRKDSCTLH			600
Query 601		QIIHTGEKPYKCDICGRGYTRKSQLGIHRRIHMRIHTGENSYKCAVCGRGFTKRLGSH			660
Sbjct 601		QIIHTGEKPYKCDICGRGYTRKSQLGIHRRIHMRIHTGENSYKCAVCGRGFTKRLGSH			660
Query 661		WRIHTEEKPFKCNCDKRFLAWSSLNTHQAVHTDEKAHKCTFCGKVFSSRCYLAVHQRIH			720
Sbjct 661		WRIHTEEKPFKCNCDKRFLAWSSLNTHQAVHTDEKAHKCTFCGKVFSSRCYLAVHQRIH			720
Query 721		TGEKPYKCDVCGRGYTRSRQLAIHQRVHTGEKHYKCNVCG	760		
Sbjct 721		TGEKPYKCDVCGRGYTRSRQLAIHQRVHTGEKHYKCNVCG	760		

Abbildung 9.2: Protein-BLAST-Ergebnis für LOC787812

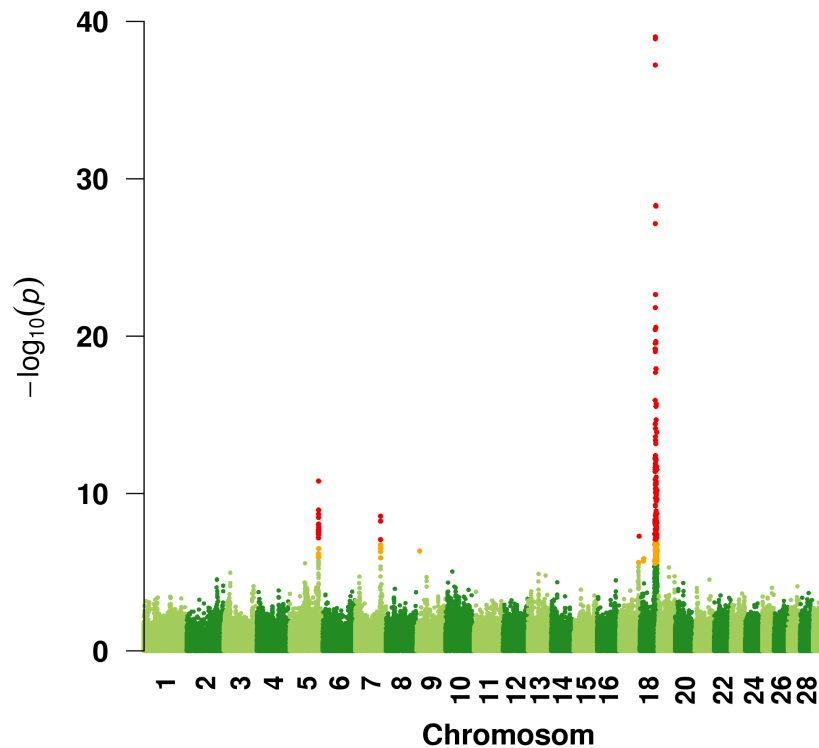


Abbildung 9.3: Manhattan-Plot der Ergebnisse der HD-basierten GWAS bei Holstein für das Merkmal direkte Totgeburten bei Kühen

Die Positionen der einzelnen Varianten, auf der x-Achse dargestellt, basieren auf dem UMD3.1-Assembly (ZIMIN et al. 2009), mit Ausnahme der Geschlechtschromosomen. Auf der y-Achse ist der negativ-dekadische Logarithmus der P-Werte aus der GWAS aufgetragen. Chromosomweit signifikante Polymorphismen sind gelb markiert, genomweit signifikante Varianten sind in rot dargestellt

Danksagung

Bedanken möchte ich mich zuallererst bei Prof. Dr. Ruedi Fries für die Überlassung des Themas meiner Dissertation und die gute Betreuung und Unterstützung, die er mir während der gesamten Arbeit entgegengebracht hat.

Für die finanzielle Unterstützung bedanke ich mich beim Bundesministerium für Bildung und Forschung, da diese Arbeit im Rahmen des Verbundprojekts Synbreed ermöglicht wurde.

Dr. Egbert Feddersen vom Deutschen Holstein Verband danke ich insbesondere für seine Bereitschaft mich im Rahmen des Graduiertenprogramms des WZW als Mentor durch meine Promotion zu begleiten. Durch ihn habe ich außerhalb der Universität viele wichtige Ratschläge erhalten.

Besonderer Dank gilt allen Kollegen des Lehrstuhls für Tierzucht. Bei Birgit Hoffmann bedanke ich mich besonders für die administrative Unterstützung. Den Kollegen und Kolleginnen im Labor möchte ich für ihre Hilfestellung rund um die Laborarbeiten während meiner Promotion bedanken. Meinen Doktorandenkollegen Andrea, Bernhard, Christoph, Gabi, Hubert, Kai, Simone, Tini und Xiaolong danke ich vor allem für die spaßigen Momente und das freundschaftliche Verhältnis, auch außerhalb der Arbeitszeit. Dazu gehören insbesondere Simone und Tini, die mir nicht nur bei Laborproblemen, sondern auch in weiteren fachlichen, als auch anderweitigen, persönlichen Fragen immer mit Rat und Tat zur Seite standen. Auch die oft fruchtbaren abendlichen Diskussionen mit Tini sowie Simonas Versorgung mit Nervennahrung möchte ich nicht missen. Vielen herzlichen Dank dafür. Hubert danke ich für seine Unterstützung in statistischen und bioinformatischen Angelegenheiten. Ohne euch allen wären viele Freitagnachmittage als auch die Mittagspausen nicht so geworden, wie sie waren.

Bei meinen Eltern bedanke ich mich besonders herzlich für ihre ständige Unterstützung. Sie haben mir den Weg hierher ermöglicht und dabei immer an mich geglaubt. Ich danke euch, dass ihr mich immer begleitet.

Zuallerletzt geht mein größter Dank an Basti. Du hast mich während meiner Promotion immer wieder motiviert und mich in allen Situationen unterstützt. In schwierigen Phasen hast du mir besonders viel Kraft durch deine Liebe gegeben. Danke für alles.

LEBENS LAUF

Name	Sandra Kipp, geb. Jansen
Geburtsdatum	13. Juni 1986
Geburtsort	Haselünne
Staatsangehörigkeit	deutsch
Schulabschluss	Abitur

BERUF SERFAHRUNG

seit 10.2014

Vereinigte Informationssysteme Tierhaltung w.V., Verden

- Abteilung Biometrie und Zuchtwertschätzung

11.2010 bis 08.2014

Technische Universität München, Lehrstuhl für Tierzucht

- wissenschaftliche Mitarbeiterin
- Promotion (Dr. agr.) unter der Betreuung von Prof. Ruedi Fries
„Die Analyse von genomweiten Re-Sequenzierungsdaten zur Identifizierung kausaler DNA-Variation beim Rind“

STUDIUM

04.2006 bis 10.2010

Agrarwissenschaften an der Friedrich-Wilhelms-Universität, Bonn

- Schwerpunkte: Tier und Betriebswirtschaftslehre
- Diplomarbeit im Bereich Rinderzucht in Zusammenarbeit mit dem DHV
„Untersuchungen zur Erfassung von Exterieurmängeln im Rahmen der Nachzuchtbewertung von Deutschen Holstein Färsen unter Verwendung des Notenvorschlagsprogramms NoVo“
- Abschluss: Diplom-Agraringenieur (Dipl.-Ing. agr.)

SCHULISCHER WERDEGANG

08.1996 bis 07.2005

Gymnasium Marianum, Meppen

08.1992 bis 07.1996

Grundschule Bokeloh, Meppen

PUBLIKATIONEN UND TAGUNGSBEITRÄGE

Artikel

Assessment of the genomic variation in a cattle population by re-sequencing of key animals at low to medium coverage - **Sandra Jansen**, Bernhard Aigner, Hubert Pausch, Michal Wysocki, Sebastian Eck, Anna Benet-Pagès, Elisabeth Graf, Thomas Wieland, Tim M. Strom, Thomas Meitinger, Ruedi Fries, BMC Genomics 14,1 446 (2013)

A nonsense mutation in TMEM95 encoding a nondescript transmembrane protein causes idopathic male subfertility in cattle - Pausch H, Kölle S, Wurmser C, Schwarzenbacher H, Emmerling R, **Jansen S**, Trottmann M, Fürst C, Götz K-U, Fries R, PLOS Genetics 10 (1) (2014)

Classic Selective Sweeps Revealed by Massive Sequencing in Cattle - Saber Qanbari, Hubert Pausch, **Sandra Jansen**, Mehmet Somel, Tim M. Strom, Ruedi Fries, Rasmus Nielsen, Henner Simianer, PLOS Genetics 10(2) 2014

Evaluation of variant identification methods for whole genome sequencing data in dairy cattle – Christine F Baes, Marlies A Dolezal, James A Koltes, Beat Bapst, Eric Fritz-Waters, **Sandra Jansen**, Christine Flury, Heidi signer-Hasler, Christian Stricker, Rohan Fernando, Ruedi Fries, Juerg Moll, Dorian J Garrick, James M Reecy and Birgit Gredler, BMC Genomics 15:948 (2014)

Konferenzen

Jansen S. et al. (2011): Sichtung der genomischen Variation der Fleckvieh Population durch Resequenzierung bei niedriger bis mittlerer Abdeckung. Tagungsband der DGfZ/GfT-Gemeinschaftstagung 2011

Jansen S. et al.(2012): Extensive variant detection in the Fleckvieh population by low-coverage re-sequencing. Book of Abstracts of the 33rd ISAG conference, Cairns

Pausch H, Wurmser C, **Jansen S**, Aigner B, Fries R (2012): Imputation of whole-genome sequence information for QTL fine-mapping in the Fleckvieh population. Book of Abstracts of the 33rd ISAG conference, Cairns

Fries R, Pausch H, **Jansen S**, Aigner B, Wysocki M (2012): Assessment of the genomic variation in a cattle population based on low-coverage re-sequencing. Book of Abstracts of the 63rd EAAP Annual Meeting, Bratislava

Aigner B, **Jansen S**, Pausch H, Wysocki M, Eck S, Benet-Pagès A, Graf E, Wieland T, Strom T, Meitinger T, Fries R (2012): Erste Ansätze zur Entdeckung von Strukturellen Varianten auf der Grundlage genomweiter Sequenzierung beim Rind. Tagungsband der DgfZ/GfT-Gemeinschaftstagung 2012

Aigner B, **Jansen S**, Pausch H, Wysocki M, Eck S, Wieland T, Benet-Pagès A, Graf E, Strom T, Meitinger T, Fries R (2013): Extensive DNA variant detection in a cattle population by low-coverage re-sequencing. Book of Abstracts of the XXI PAG, San Diego

Baes C, **Jansen S**, Bapst B., Dolezal M., Moll J, Fries R, Gredler B (2013): Analysis of next-generation sequence data of Swiss dairy populations: preliminary studies of variant calling techniques. Book of Abstracts of the 64th EAAP Annual Meeting, Nantes

Baes C, Dolezal M, Bapst B, Flury C, Signer-Hasler H, Stricker C, Fernando R, Garrick D, Schmitz-Hsu F, Fries R, **Jansen S**, Gredler B (2013): Evaluation of a whole genome sequencing data pipeline. Book of Abstracts of the 5th ISAFG, Sao Paulo

Jansen S, Wurmser C, Fries R (2013): Untersuchung von potenziellen Stop-Mutationen im Genom des Rindes. Tagungsband der DgfZ/GfT-Gemeinschaftstagung 2013

Jansen S, Wurmser C, Pausch H, Fries R (2014): Analysis of putatively deleterious alleles in German Fleckvieh by whole genome-resequencing. Book of Abstracts of the 65th EAAP Annual Meeting, Copenhagen