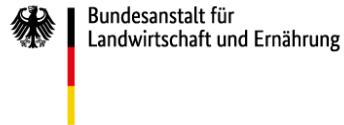


Gefördert durch



Projektträger



aufgrund eines Beschlusses
des Deutschen Bundestages

Abschlussbericht

zu dem Projekt

DSN-SNP:

Entwicklung eines SNP-Chips für die Zuchtwertschätzung und zum Diversitätsmanagement für das Deutsche Schwarzbunte Niederungsring

Förderkennzeichen: 2818BM090 und 2818BM091

Die Förderung des Vorhabens erfolgte aus Mitteln des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) aufgrund eines Beschlusses des deutschen Bundestages. Die Projektträgerschaft erfolgte über die Bundesanstalt für Landwirtschaft und Ernährung (BLE).



HUMBOLDT-
UNIVERSITÄT
ZU BERLIN



Abschlussbericht

Zuwendungsempfänger:

Justus-Liebig-Universität Gießen (Förderkennzeichen: 2818BM090)

Humboldt-Universität zu Berlin (Förderkennzeichen: 2818BM091)

Thema: „Entwicklung eines SNP-Chips zur Zuchtwertschätzung und zum Diversitätsmanagement für das Deutsche Schwarzbunte Niederungsrind“

Laufzeit: 01.04.2019 – 31.01.2023 (46 Monate)

Berichtszeitraum: 01.07.2019 – 31.01.2023 (43 Monate)

Projektkoordination:

Justus-Liebig-Universität Gießen, Institut für Tierzucht und Haustiergenetik, Professur für Tierzüchtung, Prof. Dr. Sven König, Ludwigstraße 21b, 35390 Gießen (**TZ-GI**)

Verbundpartner:

Humboldt-Universität zu Berlin, Lebensmittelwissenschaftliche Fakultät, Albrecht Daniel Thaer-Institut für Agrar- und Gartenbauwissenschaften, Prof. Dr. Gudrun A. Brockmann, Invalidenstraße 42, 10115 Berlin (**TZ-BER**)

Inhaltsverzeichnis

Abkürzungsverzeichnis	III
Tabellenverzeichnis	IV
Abbildungsverzeichnis	VI
1. Ziele und Aufgabenstellung des Vorhabens	1
1.1 Planung und Ablauf des Vorhabens.....	3
1.2 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde	10
Das Deutsche Schwarzbunte Niederungsgrind.....	10
Genomweite Assoziationsstudien beim DSN	11
Genomische Zuchtwertschätzung in kleinen Populationen	12
Lernstichprobe für die genomische Zuchtwertschätzung.....	13
2. Material und Methoden	15
2.1 Aufbau einer informativen Lernstichprobe (APIa).....	15
2.2 SNP-Chip Design (APII).....	19
Technische Konstruktion des SNP-Chips	19
Detaillierte Erläuterung der zehn Kategorien zur Auswahl von Sequenzvarianten für das Design des DSN200K-SNP-Chips:.....	20
2.3 Evaluierung des neuen DSN200K-SNP-Chip (APIII).....	23
Genotypisierung	23
Vereinigung von Genotypen, die mit verschiedenen SNP-Chips erzeugt wurden, durch Imputation.....	24
Vergleich verschiedener SNP-Datensätze in genomweiten Assoziationsstudien	25
Kreuzvalidierung im Rahmen der genomischer Zuchtwertschätzung	26
Kreuzvalidierung und Berechnung der gZW-Genauigkeiten.....	30
2.4 Ermittlung von Maßzahlen genetischer Diversität (APIV)	31
Phylogenetische Untersuchungen und Nukleotid-Diversitätsanalysen	31
Anwendung der „Optimum genetic contribution“-Methode	32
2.5 Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen (APV).....	33
Sequenzierungsanalyse von Tieren der Rassen Angler und Rotbunt-Doppelnutzung	33
2.6 Transfer der Ergebnisse in die Praxis (APVI).....	34
3. Ergebnisse.....	35
3.1 Aufbau einer Lernstichprobe: Sequenzierung, Aufbereitung und Imputation (API)	35
3.2 Genotypisierung mit dem neuen DSN200K-SNP-Chip und Performance-Analysen der Genotypdaten (APII)	36
3.3 Evaluierung des DSN200K-SNP-Chip (APIII)	38
Imputation von SNPs in Genotypen-Datensätze geringerer Dichte	38
Vergleichende genomweite Assoziationsstudien und Heritabilitäten für ausgewählte Merkmale auf Basis von drei verschiedenen Markerdichten (50K, 200K, VGS)	38
Genomische Zuchtwertschätzung und Genauigkeiten mit verschiedenen Markerdichten.....	41

3.4	Untersuchungen zur Sicherung der genetischen Diversität in der DSN-Population (APIV).....	53
	Phylogenetische Untersuchungen und Nukleotid-Diversitätsanalysen beim DSN und 68 weiteren taurinen Rinderrassen	53
	„Optimum genetic contribution“-Anwendung innerhalb der DSN-Population	58
3.5	Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen (APV).....	60
	Verwendung des DSN200K-SNP-Chips in anderen Rassen.....	60
	Sequenzierungsanalyse von Tieren der Rassen Angler und Rotbunt-Doppelnutzung	62
3.6	Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse.....	62
	Übersicht über erfolgte oder geplante Veröffentlichungen.....	64
4.	Zusammenfassung.....	68
5.	Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen	70
6.	Literaturverzeichnis	73

Abkürzungsverzeichnis

AAF	Alternative Allelfrequenz
ALCAM	„Activated leukocyte cell adhesion molecule“ - Gen
ANG	Angler Rind
AP	Arbeitspaket
BLE	Bundesanstalt für Landwirtschaft und Ernährung
BMEL	Bundesministeriums für Ernährung und Landwirtschaft
bp	Basenpaare
BTA	<i>Bos taurus</i> Autosom
DDR	Deutsche Demokratische Republik Deutschland
DGfZ	Deutsche Gesellschaft für Züchtungskunde
DIM	Tage in Milch (engl.: „days in milk“)
DNA	Desoxyribonukleinsäure
DSN	Deutsches Schwarzbuntes Niederungsring
E%	305-d-Milcheiweißanteil in Prozent
E305	305-d-Milcheiweißleistung in Kilogramm
EBV	Pedigree-Zuchtwert
EKA	Erstkalbealter
EKM	Energie-korrigierte Milchleistung
F%	305-d-Milchfettanteil in Prozent
F305	305-d-Milchfettleistung in Kilogramm
FEQ	Fett-Eiweiß-Quotient
F _{ST}	Fixationsindex
FV	Fleckvieh
GI	Gießen
GIN	Gastrointestinale Nematoden
GWAS	Genomweite Assoziationsstudie
gZW	genomischer Zuchtwert
gZWS	Genomische Zuchtwertschätzung
h^2	Heritabilität
HB	Haplotypblock (HB)
HF	Holstein-Friesian
HOL	Deutsch Holstein
HWG	Hardy-Weinberg-Gleichgewicht
ICAR	International Committee for Animal Recording
IFN	Institut für Fortpflanzung landwirtschaftlicher Nutztiere

INDEL	Mutation durch eine Insertion oder Deletion
ISAG	International Society for Animal Genetics
kb	Kilobasenpaare
Kg	Kilogramm
KV	Kalbeverlauf
M305	305-d-Milchleistung in Kilogramm
MAF	Frequenz des selteneren Allels (engl.: „minor allele frequency“)
Mb	Megabasenpaare
MHC	Haupthistokompatibilitäts-Komplex (engl.: „Major Histocompatibility Complex“)
N_e	Effektive Populationsgröße
OGC	„optimum genetic contribution“
r	Korrelation
RBB	Rinderunion Berlin-Brandenburg GmbH
RDN	Rotbunt Doppelnutzung
RoHs	Runs of Homozygosity
RSH	Rinderzuchtverbands Schleswig-Holstein
RZG	Relativzuchtwert Gesamt
RZM	Relativzuchtwert Milch
SCS	transformierte somatische Zellzahl (engl.: „somatic cell score“)
SD	Standardabweichung (engl.: „standard deviation“)
SE	Standardfehler
SNP	Einzelnukleotid-Polymorphismus (engl.: „single nucleotide polymorphism“)
TGR	Totgeburtenrate
TZ-BER	Arbeitsgruppe Tierzucht Berlin
TZ-GI	Arbeitsgruppe Tierzucht Gießen
VEP	„variant effect predictor“
VGS	Vollgenomsequenzen
vit	Vereinigte Informationssysteme Tierhaltung w.V.
XP-EHH	„Cross Population Extended Haplotype Homozygosity“

Tabellenverzeichnis

Tabelle 1: Chronologischer Ablauf der Arbeitspakete (AP) nach Quartalen (Balkenplan).	3
Tabelle 2: Anzahl an Sequenzvarianten in einzelnen Kategorien, welche bei der Konstruktion des DSN200K-SNP-Chips berücksichtigt wurden.....	19
Tabelle 3: Anzahl der Sequenzvarianten, die in genomweiten Assoziationsstudien und Gen-Sequenzierungen mit den jeweiligen Merkmalen identifiziert und für das Design des DSN200K-SNP-Chips benutzt wurden.	22
Tabelle 4: Deskriptive Statistik für 305-d-Produktionsmerkmale und Testtagsmerkmale.....	27
Tabelle 5: Deskriptive Statistik für lineare Exterieurmerkmale und deren Merkmalskomplexe.	28
Tabelle 6: Anzahl assoziierter Marker (SNPs/Sequenzvarianten [SVs]) und Anzahl annotierter Gene in den GWAS und Heritabilitäten (h^2) mit Standardfehlern (SE) für die drei unterschiedlichen Markerdichten.	40
Tabelle 7: Vergleich der GWAS-Ergebnisse unter Verwendung verschiedener SNP-Chips...	41
Tabelle 8: Heritabilitäten (h^2), additiv-genetische Varianzen (σ_g^2) und Restvarianzen (σ_e^2) mit den dazugehörigen Standardfehlern (SE) der Schätzungen für die 305-d Produktions- und Testtagsmerkmale basierend auf verschiedenen Markerdichten.	43
Tabelle 9: Heritabilitäten (h^2), additiv-genetische Varianzen (σ_g^2) und Restvarianzen (σ_e^2) mit den dazugehörigen Standardfehlern (SE) der Schätzungen für die Exterieurmerkmale, basierend auf verschiedenen Markerdichten.	45
Tabelle 10: Genauigkeit der genomischen Vorhersagen für 305-d Produktionsmerkmale und Testtagsmerkmale unter Verwendung verschiedener Markerdichten.....	50
Tabelle 11: Genauigkeit der genomischen Vorhersagen für Exterieurmerkmale und Testtagsmerkmale unter Verwendung verschiedener Markerdichten.....	51
Tabelle 12: Vergleich der Werte für genomische Inzucht und genetischen Diversität zwischen DSN und Holstein.....	57
Tabelle 13: Ergebnisse aus der Anwendung der „optimum genetic contribution“-Methode mit dem Programm GENCONT (Meuwissen, 2002) für das Merkmal 305-d-Milchleistung in kg. ...	60

Abbildungsverzeichnis

Abbildung 1: Deutsches Schwarzbuntes Niederungsrind (Foto: L. Aufmhof)	11
Abbildung 2: Hauptkomponenten-Analyse für die Merkmale durchschnittliche Eiausscheidung gastrointestinaler Nematoden (mean FEC-GIN) und durchschnittliche Eiausscheidung des großen Leberegels (<i>Fasciola hepatica</i>) (mean FEC-FH) bei Deutschen Schwarzbunten Niederungsrindern (DSN) aus drei Betrieben mit Weidehaltung.	16
Abbildung 3: Hauptkomponenten-Analyse für Kühe der Rasse Deutsches Schwarzbuntes Niederungsrind (DSN), basierend auf der Milchleistung der ersten Laktation für selektierte DSN-Kühe (=Selektion, schwarze Punkte) sowie die DSN-Gesamtpopulation (=Population, hellblaue Punkte).....	16
Abbildung 4: Überblick der Pipeline zur Prozessierung der Rohsequenzdaten. Die Pipeline wurde in Anlehnung an die Richtlinien vom 1000-Bullen-Genom-Projekt entwickelt. Die Namen der verwendeten Softwares sind rot markiert.....	18
Abbildung 5: Exemplarische Darstellung der Vorgehensweise bei dem Vergleich von 50K Genotypen, 200K Genotypen und Vollgenomsequenzdaten im Rahmen einer genomischen Zuchtwertschätzung innerhalb der Rasse Deutsches Schwarzbuntes Niederungsrind (DSN).30	
Abbildung 6: Pipeline für das DSN200K-SNP-Chip Design. A) Filterung des initialen Datensatzes. B) Selektion der Varianten auf Basis mehrerer Kategorien.	36
Abbildung 7: Manhattan-Plots für die 305-d Milchleistung in der ersten Laktation; A) 50K Illumina BovineSNP50 BeadChip, B) neu entwickelter DSN200K-SNP-Chip und C) Vollgenomsequenzvarianten.	39
Abbildung 8: Unterschiede in der Vorhersagegenauigkeit (Mittelwert aller 10 Läufe der Kreuzvalidierung) für 305-d-Milchproduktionsmerkmale, Testtagsmerkmale und Exterieurmerkmale.	52
Abbildung 9: Phylogenetischer Baum von 69 Rinderrassen.....	54
Abbildung 10: (a) Die durchschnittliche Nukleotid-Diversität in der Rasse DSN pro Autosom (BTA), (b) Nukleotid-Diversität in der Rasse DSN pro Fenster auf Autosom 23.	56
Abbildung 11: (a) Überblick über Diversitätsmessungen und Inzucht bei verschiedenen Rassen.....	57
Abbildung 12: Darstellung der Ergebnisse aus der Anwendung der „optimum genetic contribution“ Methode mit dem Programm GENCONT (Meuwissen 2002) für das Merkmal 305-d Milchleistung in kg.	59
Abbildung 13: Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen.	61

1. Ziele und Aufgabenstellung des Vorhabens

In der Richtlinie des Bundesministeriums für Ernährung und Landwirtschaft (BMEL) vom 28. Januar 2015 zum Modell- und Demonstrationsvorhaben definiert Abschnitt 2 übergeordnet die Erhaltung und nachhaltige Nutzung der Agrobiodiversität, also die Erhaltung der Vielfalt der landwirtschaftlichen Nutztiere, als Gegenstand der Förderung. Erhaltung kann nur dann nachhaltig realisiert werden, wenn bedrohte Rassen hinsichtlich ihrer Leistung und/oder Alleinstellungsmerkmale im Vergleich zu „größeren“ Rassen konkurrenzfähig bleiben. Während insbesondere in den großen Milchviehpopulationen wie Deutsch Holstein (HOL) oder Fleckvieh (FV) die Einführung der genomischen Selektion zu einer weitergehenden Optimierung der Zuchtprogramme mit rasantem Zuchtfortschritt für Merkmale der konventionellen Leistungsprüfung beigetragen hat, profitieren kleine, vom Aussterben bedrohte Rassen wie das Deutsche Schwarzbunte Niederungsrind (DSN) von diesen hocheffizienten Zuchtverfahren und vom aktuellen Entwicklungsfortschritt auf dem Gebiet der genomischen Zuchtwertschätzung (gZWS) trotz des offensichtlichen Potenzials bisher nicht. Basis für die gZWS sind kommerzielle Genotypisierungs-Chips auf Basis von Einzelnukleotid-Polymorphismen (SNPs; engl.: „single nucleotide polymorphisms“), die speziell auf die populären Rassen „zugeschnitten“ sind. Bei dieser kommerziellen SNP-Chip Entwicklung spielen die genomische Architektur der Hochleistungsrassen und die konventionellen Zuchtzielmerkmale (insbesondere Produktionsmerkmale) eine entscheidende Rolle. Während in großen Populationen wie HOL und FV durch genomische Selektionsschemata ein beschleunigter Zuchtfortschritt beobachtet werden kann, bleiben kleinere, bedrohte Populationen im Hinblick auf ihre Wettbewerbsfähigkeit im Vergleich zu den größeren Rassen züchterisch zunehmend weiter zurück.

Darüber hinaus sind ein gutes Inzuchtmanagement und die Erhaltung der genetischen Diversität für Rassen mit kleiner Populationsgröße von besonderer Bedeutung. Züchterische Strategien müssen somit zwei zentrale Bedingungen gleichzeitig berücksichtigen: Die Maximierung von Zuchtfortschritt bei gleichzeitiger Restriktion für zukünftige verwandtschaftliche Beziehungen bzw. Maximierung von Kennzahlen der genetischen Diversität. Die genaueste Beurteilung der Diversität ist möglich (insbesondere bei unzureichender Tiefe der Abstammungsdateien), wenn hierzu genomische SNP-Markerdaten verwendet werden. Dabei erscheint der Heterozygotiegrad basierend auf SNP-Markerdaten die aktuell beste Maßzahl, um genetische Variabilität innerhalb von Populationen zu beschreiben. Allerdings kann auch dieser Heterozygotiegrad nur dann genau abgeleitet werden, wenn die Vollgenomsequenzen (VGS) der Tiere bekannt sind (Malomane et al., 2018). VGS-Daten haben den Vorteil, dass die für die Ausprägung eines Merkmals verantwortlichen Genvarianten mit sehr hoher Wahrscheinlichkeit enthalten sind. Im Rahmen der Sequenzierung werden Informationen von über 3 Milliarden Basen des gesamten Genoms generiert und daher alle genomischen Unterschiede zwischen den sequenzierten Tieren erfasst und abgebildet. Sequenzierungen von VGS in großen

Stichproben ist aber aufgrund der derzeit noch hohen Kosten der Sequenzierungsverfahren nicht finanzierbar. Es muss das Ziel sein, im Modellvorhaben auf der Basis von VGS-Daten weniger Tiere einer Population diejenigen SNP-Marker zu identifizieren, die notwendig sind, um zukünftig einen kostengünstigen „low density“ SNP-Chip im Sinne effizienter Diversitätsstudien unter Berücksichtigung von vielen genotypisierten Tieren anwenden zu können. Genotypisierung meint hierbei (und auch in allen weiteren diesbezüglichen Textstellen) die Nutzung kommerzieller SNP-Chips. Hier gibt es z.B. den sogenannten Illumina Bovine Eurogenomics 10K low-density chip (ca. 10.000 SNPs = „low-density chip“) oder den Illumina BovineSNP50 BeadChip (ca. 50.000 SNPs = „medium-density chip“), mit denen die Tiere für routinemäßige gZWS genotypisiert werden. Insgesamt umfasst das Rindergenom aber 25-30 Millionen SNPs. Es besteht somit die Herausforderung, die wesentlichen SNPs aus dem VGS-Datensatz zu identifizieren, die sowohl für genetische Diversitätsstudien als auch für eine kommerziell zu entwickelnde gZWS beim DSN am besten geeignet sind.

Das zentrale Ziel des Vorhabens bestand in der Entwicklung und Erprobung eines spezifischen, auf die Rasse DSN angepassten custom-SNP-Chip-Arrays für den Einsatz in der gZWS sowie für weitergehende Anwendungen zum Management der genetischen Diversität. Das Vorhaben diente als Modell für die züchterische Weiterentwicklung und Sicherung der genetischen Diversität in bedrohten Rassen bei Nutzung modernster Verfahren der Molekulargenetik, Bioinformatik und Zuchtwertschätzung am Beispiel DSN. Innovativ war insbesondere die Vorgehensweise, auf Basis engmaschiger genomischer Informationen kombiniert mit modernsten genetisch-statistischen Verfahren der Bioinformatik, Zielvorstellungen zur züchterischen Rasseverbesserung zu realisieren, ohne auf Fremdgensequenzen bzw. genmanipulative Maßnahmen zurückgreifen oder ausweichen zu müssen.

Der Modellcharakter des Vorhabens manifestierte sich in folgenden **Arbeitszielen**:

1. Auf- und Ausbau einer informativen Lernstichprobe in der Rasse DSN (**Arbeitspaket (AP) I**)
2. Entwicklung und Weiterentwicklung einer Pipeline zur Imputation der DSN-Sequenzdaten in SNP-Chip Daten (**API**)
3. Konstruktion eines custom-made SNP-Chips für die Rasse DSN unter Berücksichtigung von
 - i) „neuen Merkmalen“ der Adaptation und des Tierverhaltens und
 - ii) von Diversitätskennzahlen (u.a. Heterozygotiegrad über SNP-Marker) (**APII**)
4. Erprobung des neu entwickelten DSN-spezifischen SNP-Chips mittels genetisch-statistischer Anwendungen (**APIII**)
5. Entwicklung und Umsetzung von Monitoringverfahren zur Sicherung der genetischen Diversität in der DSN-Population (**APIV**)
6. Evaluierung der Übertragbarkeit des DSN SNP-Chips auf andere Rassen (**APV**)
7. Transfer der Ergebnisse in die Praxis (**APVI**)

1.1 Planung und Ablauf des Vorhabens

Der geplante Projektstart am 01.04.2019 verzögerte sich um drei Monate, da die/der wissenschaftliche Mitarbeiter/in der Arbeitsgruppe Tierzucht-Gießen (TZ-GI) erst zum 01.07.2019 und der Arbeitsgruppe Tierzucht-Berlin (TZ-BER) mit vier Monaten Verzögerung nach Projektstart, also zum 01.08.2019 eingestellt werden konnten. Somit ergab sich ein verspäteter Projektstart ab dem 01.07.2019 und damit eine Laufzeit des Projektes bis 31.07.2022. Neben den Verzögerungen des Projektstartes in 2019 hat die Corona-Pandemie in 2020 und 2021, sowie ein Cyber-Angriff Ende 2019 auf die Justus-Liebig-Universität Gießen mit der Folge eines mehrmonatigen Ausfalls des Internets und einem beschränkten Zugriff auf Datenbanken und Rechenserver zu Verzögerungen im Projekt geführt. Der Projektzeitraum verlängerte sich somit um nochmals sechs Monate bis zum 31.01.2023. Der angepasste chronologische Ablauf der Arbeiten in dem Projektvorhaben ist in Tabelle 1 aufgeführt. Die Planung und Umsetzung bzw. Abläufe der Arbeitspakete werden nachfolgend beschrieben.

Tabelle 1: Chronologischer Ablauf der Arbeitspakete (AP) nach Quartalen (Balkenplan). APIa - Auswahl der Tiere für die Sequenzierung und Logistik der Sequenzierung; APIb - Datenbank mit aufbereiteten Sequenzdaten inklusive der imputierten Genotypen erstellen; APII - Identifikation der relevanten SNPs für den DSN spezifischen SNP-Chip; APIII - Entscheidungskriterien zum Erfolg der genomischen Zuchtwertschätzung (u.a. Genauigkeiten der genomischen Zuchtwerte) herausarbeiten; APIV - Maßzahlen der genetischen Diversität auf Basis der identifizierten SNPs bestimmen; APV - Evaluierung des entwickelten SNP-Chips in anderen Rassen; APVI - Abschlussworkshop.

II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	Jan.	
2019			2020				2021				2022				2023	
APIa																
		APIb														
			APII													
						APIII										
											APIV					
										APV						
																APVI

Laut **API** sollten in gemeinsamer Zusammenarbeit der beiden Projektpartner eine informative Lernstichprobe auf- bzw. ausgebaut werden. Das Arbeitspaket unterteilte sich in zwei Teile: **a)** Sequenzierung von 100 DSN Tieren, sowie die Identifizierung von Sequenzvarianten (TZ-GI und TZ-BER) und **b)** Entwicklung bzw. Weiterentwicklung einer Pipeline zur Imputation von 50K SNP-Chip Genotypen von DSN Tieren auf VGS-Niveau (TZ-BER).

Die Anzahl der zu sequenzierenden Tiere ließ sich von den ursprünglich geplanten 100 Tieren auf 214 Tiere (107 pro Arbeitsgruppe) erweitern, da für die Sequenzierung ein günstigeres Angebot als ursprünglich geplant von der Firma Novogene eingeholt werden konnte. Durch die deutlich höhere Anzahl an Tieren für die Sequenzierung konnten DSN-Kühe aus einer sehr weiten Bandbreite innovativer Phänotypen selektiert werden, was einen positiven Einfluss auf

die Weiterentwicklung der Imputations-Pipeline (**APIb**) sowie auf die zu entwickelnde Zuchtwertschätzung hatte (**APIII**). Die Vorgehensweise und die notwendigen Voranalysen zur Selektion der Tiere wurden zu Projektstart zwischen beiden Projektpartnern festgelegt. Auch für die Auswahl der Tiere, die sequenziert werden sollten, wurde in Absprache der beiden Arbeitsgruppen eine Strategie festgelegt. Es wurden 30 DSN-Bullen für die Sequenzierung verwendet; das entspricht der Anzahl aller Bullen, von denen Blut- und Spermaproben in den Gewebebanken von TZ-BER und TZ-GI vorlagen. Die Auswahl der restlichen 184 DSN-Kühe erfolgte in jeder Arbeitsgruppe individuell. Die Sequenzierung der Proben wurde bei der Firma Novogene auf der Illumina NovaSeq Plattform mit 150 gepaarten Enden und einer 15-fachen Abdeckung des Genoms durchgeführt. Die von der Firma zurückgesandten Rohdaten werden sowohl bei TZ-BER als auch bei TZ-GI gespeichert und aufbewahrt. Die Aufbereitung der Sequenzdaten erfolgte durch TZ-BER unter Verwendung einer eigens erarbeiteten Pipeline. Des Weiteren wurden weitere 90, bereits sequenzierte DSN Tiere von TZ-BER in die Pipeline integriert. Somit lagen insgesamt Sequenzdaten für 304 DSN für die Identifizierung von Sequenzvarianten vor. Die einzelnen Arbeitsschritte inkludierten das Sequenz-Alignment („mapping“), das Detektieren der Varianten („variant calling“) und die Rekalibrierung in Anlehnung an die Richtlinien des 1000-Bullen-Genom-Projekts (<http://www.1000bullgenomes.com>).

Die unter **APIa** detektierten Sequenzvarianten aller DSN Tiere dienten als Referenzpanel für die Imputation auf Sequenzniveau von ca. 1.800 DSN Genotypen, die mit dem Illumina BovineSNP50 BeadChip (50K SNP-Chip) genotypisiert wurden. Die hier verwendeten 50K DSN Genotypen lagen bereits aus vorangegangenen Projekten vor. Die Imputation erfolgte mit der Software Beagle v.5.1 (Browning et al., 2018) nach dem beschriebenen Vorgehen von Korcuć et al. (2019). Im Anschluss wurden die 1.800 auf Sequenzlevel imputierten DSN Genotypen mit den 304 sequenzierten DSN Genotypen zusammengeführt, sodass ein imputierter Datensatz mit etwa 2.100 Genotypen auf Sequenzniveau erstellt wurde. Um die Genauigkeit der Imputation zu untersuchen, wurden die Varianten von 30 DSN für die sowohl Sequenzdaten als auch 50K Genotypen vorlagen mit dem Referenzpanel der 304 sequenzierten DSN Tiere imputiert. Im Anschluss wurden die imputierten Genotypen und die originalen Sequenz-Genotypen der 30 DSN Tiere miteinander verglichen.

In Arbeitspaket II (**APII**) sollte ein 50K custom-made SNP-Chip für die Rasse DSN entwickelt werden. Dabei sollten insbesondere Varianten berücksichtigt werden, die den folgenden Kriterien entsprechen:

- SNPs, die sich in Regionen befinden, die große Unterschiede in den Allelfrequenzen zwischen den Rassen DSN und HOL aufweisen.
- SNPs, die in genomweiten Assoziationsstudien (GWAS) mit den imputierten Genotypen (aus **APIb**) innerhalb der DSN-Population Assoziationen zu Merkmalen wie

Milchleistung, Exterieur und Fruchtbarkeit zeigen. Darüber hinaus sollten neue funktionale Merkmale für Adaptationsfähigkeit (Hitzestress) und Gesundheit (Parasitenresistenz) für die DSN-Population entwickelt und in Assoziationstests genutzt werden.

- SNPs, die im Kopplungsungleichgewicht zu Sequenzvarianten stehen, die signifikante Assoziationen zu Merkmalen in HOL haben.

Nach umfänglichen Videokonferenzen zwischen TZ-BER und TZ-GI und dem Einholen von Vergleichsangeboten (Berücksichtigung von Kosten, Effizienz, Aussagekraft) verschiedener Firmen zur Entwicklung des DSN-SNP-Chips fiel die Wahl auf das Axiom™ myDesign™ TG Array der Firma Thermo Fisher Scientific. Das Array umfasst ca. 200.000 Varianten, wobei neben SNPs auch Mutationen die durch Insertion oder Deletion verursacht wurden (INDELs), berücksichtigt werden können. Der neuentwickelte Chip erhielt das Kürzel „DSN200K-SNP-Chip“. Mit den im Projekt beantragten Kosten für Genotypisierungen war es möglich, 1.488 Tiere mit dem neu entwickelten Chip zu genotypisieren.

Als Vorbereitungsschritt für die GWAS mit den imputierten Genotypdaten für konventionelle und funktionelle Merkmale innerhalb der DSN-Population wurden zunächst die aktuellen Phänotypdaten aller Tiere aus elf DSN-Betrieben, die in das Projekt einbezogen waren, aufbereitet. Die Daten wurden bei den Vereinigten Informationssystemen Tierhaltung w.V. (vit) abgefragt und als Vorbereitung für die GWAS in **APII** und für die gZWS in **APIII** zunächst einer Plausibilitätsprüfung unterzogen. Die aufzubereitenden Daten beinhalteten Stamm- und Pedigreedaten, Testtagsdaten, Daten zur Nutzungsdauer, Abkalbedaten, Besamungsdaten, Leistungsdaten zur Fruchtbarkeit, Exterieurdaten, sowie Kuhzuchtwerte. Um den Datenpool an innovativen Phänotypen zu erweitern, wurden Kotproben von weiteren ca. 70 Tieren in einem hessischen DSN-Betrieb mit Weidehaltung gesammelt und im Labor des Instituts für Tierzucht der JLU Gießen auf Endoparasiten-Infektionen untersucht. Für die Entwicklung des DSN200K-SNP-Chips wurden GWAS und Analysen mit VGS-Daten und SNP-Chip Daten (aus **APIb**) für verschiedenste Merkmale durchgeführt. Betrachtet wurden hierbei Infektionen mit gastrointestinalen Nematoden, dem großen Leberegel (*Fasciola hepatica*) sowie dem bovinen Lungenwurm (*Dictyocaulus viviparus*). Zudem wurden Milchleistungsmerkmale, Fruchtbarkeitsmerkmale und Exterieurmerkmale in Vorbereitung auf die Konstruktion des DSN200K-SNP-Chips durch TZ-GI und TZ-BER mittels GWAS analysiert. Eine Auswahl von SNPs, die mit Indikatormerkmalen für Hitzestress (z.B. Milchfettsäuren, physiologische Merkmale) beim DSN assoziiert werden konnten, wurden für das Chip-Design ebenfalls berücksichtigt.

Neben den merkmalsassoziierten Varianten wurden auf Basis der in **APII** identifizierten 20,5 Millionen Sequenzvarianten in der untersuchten DSN-Population gezielt Varianten nach bestimmten Kriterien ausgewählt, welche auf dem DSN200K-SNP-Chip berücksichtigt werden

sollten. Nach einigen Filterungsschritten, sowie der Überprüfung der technischen Eignung der Varianten verblieb ein Pool von 3,1 Millionen Varianten, die sich für den Chip empfahlen. Auf Basis der 3,1 Millionen empfohlenen Sequenzvarianten (SNPs und INDELS) wurden die besten Varianten für den neuen DSN200K-SNP-Chip ausgewählt. Auf den 200.000 vorhandenen Positionen auf dem Chip konnten ausgewählte 182.154 Sequenzvarianten (173.569 SNPs und 8.585 INDELS) untergebracht werden. Die restlichen Positionen dienten technischen Zwecken.

Mit dem DSN200K-SNP-Chip wurden in diesem Projekt 1.488 Tiere genotypisiert (1.201 DSN, 123 Rotes Höhenvieh, 42 Rotbunt Doppelnutzung (RDN), 39 Angler (ANG), 24 Pinzgauer, 20 Holstein, 14 Original Braunvieh, 10 Gelbvieh, 8 White Park, 2 Hinterwälder und 5 Butana). Weitere 432 Genotypen, die mit dem DSN200K-SNP-Chip erzeugt wurden, stammten von DSN-Rindern aus einem anderen Projekt. Die Genotypen wurden mit der Axiom™ Analysis Suite Software v5 gemäß dem „Best-Practice“-Workflow aufbereitet und analysiert.

In Arbeitspaket 3 (**APIII**) sollten von TZ-GI eine Erprobung des DSN-spezifischen SNP-Chips mittels statistischer Verfahren im Kontext einer rasseübergreifenden gZWS durchgeführt werden. Dafür sollten zunächst alle Tiere, für die bis dato keine Genotypen vorlagen, mit dem neuen DSN200K-SNP-Chip genotypisiert und dann auf Sequenzniveau imputiert werden. Anschließend sollten vergleichende GWAS durchgeführt werden a) mit den SNPs vom herkömmlichen 50K SNP-Chip und b) mit dem neuen DSN200K-SNP-Chip. Zudem sollte ein Vergleich der beiden Arrays im Rahmen der gZWS umgesetzt werden.

Die Auswahl der Tiere für die Genotypisierung mit dem DSN200K-SNP-Chip erfolgte durch beide Projektpartner im Rahmen von **APIII**. Zusätzlich zu den bereits vorliegenden DNA-Proben nicht genotypisierter DSN wurden durch TZ-GI in einem hessischen DSN-Betrieb im November 2020 von 61 weiteren (nicht genotypisierten) DSN-Kühen Haarproben gezogen, welche ebenfalls in **APIII** mit dem DSN200K-SNP-Chip genotypisiert wurden.

Die Sequenzvarianten von 289 DSN Tieren dienten als Referenzpanel für die Imputation auf VGS-Level von DSN-Rindern, die zuvor mit dem DSN200K-SNP-Chip oder dem Illumina BovineSNP50 BeadChip genotypisiert wurden. Um die Genauigkeit der Imputation zu überprüfen, wurden die Varianten von 15 sequenzierten DSN Tieren, die ebenfalls mit dem 50K Illumina BeadChip und dem DSN200K-SNP-Chip genotypisiert wurden, zur Überprüfung der Konkordanz herangezogen. Aus diesem Grund wurden diese 15 DSN nicht in das Referenzpanel aufgenommen (VGS-Daten lagen zu 304 Tieren vor). Die Illumina BovineSNP50 BeadChip Daten, die aus anderen Projekten zur Verfügung standen, wurden ebenfalls auf das Niveau des DSN200K-SNP-Chips imputiert. Dabei wurden alle DSN Tiere, die mit dem DSN200K-SNP-Chip genotypisiert wurden, als Referenzpanel herangezogen. Ausgeschlossen wurden lediglich die zuvor erwähnten 15 DSN, die bereits für die Genauigkeitsprüfung verwendet wurden. Die Imputation wurde mit Beagle v.5.1 (Browning et al., 2018) mittels „One-phase“ Methodik durchgeführt (Korkuć et al., 2019).

Ein wesentliches Ziel von **APIII** war es, vergleichende GWAS basierend auf drei verschiedenen Genotypdatensätzen mit unterschiedlichen Markerdichten durchzuführen. Um die Unterschiede in der Detektion von SNP-Effekten auf verschiedene Merkmale aufzuzeigen, wurden folgende Markerinformationen verwendet: Illumina BovineSNP50 BeadChip, DSN200K-SNPChip, sowie imputierte VGS-Daten. Bei den Phänotypen handelte es sich um Beobachtungen erstlaktierender Kühe von sechs Betrieben. Dabei standen 2.020 Beobachtungen für den 305-Tage (305-d) Milch-, Fett- und Eiweißertrag in kg (M305, F305, E305) mit moderaten bis hohen Heritabilitäten zur Verfügung. Als Gesundheitsindikator-Merkmale mit niedrigen bis mittleren Erblichkeiten dienten jeweils 2.402 Aufzeichnungen des ersten Testtags für den Fett-Eiweiß-Quotienten (FEQ), sowie die logarithmierte (bzw. transformierte) somatische Zellzahl (SCS; engl.: „somatic cell score“). Für die Abkalbmerkmale Kalbeverlauf (KV) und Totgeburtenrate (TGR) mit niedrigen Erblichkeiten lagen 2.606 Beobachtungen vor. Anschließend wurden für die oben genannten Merkmale GWAS mit dem Programm GCTA (Yang et al., 2011; Yang et al., 2014) mit der „leaving one chromosome out“ (--mlma-loco) Option durchgeführt.

Ein weiteres Ziel von **APIII** war es, im Rahmen einer zu entwickelnden gZWS eine Validierung des neu entwickelten DSN200K-SNP-Chips durchzuführen. Dazu wurden genetische Parameter und die Genauigkeiten der genomisch geschätzten Zuchtwerte (gZW) für die 305-d-Produktionsmerkmale, den FEQ und den SCS des ersten Testtags nach dem ersten Abkalben, sowie für die Exterieurmerkmale der Rasse DSN unter Verwendung von vier verschiedenen Markerpanels miteinander verglichen: (1) der kommerzielle 50K Illumina BovineSNP50 BeadChip, (2) DSN200K-SNP-Chip, (3) zufällig generierte 200K Chips auf der Grundlage von VGS-Daten und (4) ein VGS-Panel. Die gleiche Anzahl von Tieren wurde für alle Markerpanel-Analysen berücksichtigt, d.h. 1.811 genotypisierte bzw. sequenzierte Kühe für Exterieurmerkmale, 2.383 Kühe für 305-d-Produktionsmerkmale und 2.420 Kühe für FEQ und SCS. Die gZWs wurden anhand des BLUP-Schätzverfahrens ermittelt. Für die Berechnungen der gZW-Genauigkeiten wurde eine wiederholte zufällige Unterstichprobenvalidierung durchgeführt mit einem Validierungsset mit 20 % der Kühe mit maskierten Phänotypen und ein Trainingsset mit 80 % der Kühe. Die Genauigkeit wurde definiert als die Korrelation zwischen den direkten gZWs und den Phänotypen, wobei die entsprechenden festen Effekte für die Kühe im Validierungssatz subtrahiert wurden.

In **APIV** sollte ein Monitoringverfahren zur Sicherung der genetischen Diversität in der DSN-Population entwickelt und umgesetzt werden. Durch TZ-BER sollten auf Basis der Sequenzdaten die SNP-Marker bestimmt werden, die notwendig sind, um den Heterozygotiegrad als Kenngröße der genetischen Diversität möglichst genau zu bestimmen. Ziel war es, mit dem neuen SNP-Chip ein Verfahren umzusetzen, das zwei Bedingungen erfüllt: Eine möglichst genaue Zuchtwertschätzung bei gleichzeitigem Monitoring der genetischen

Diversität. Dafür kam die „optimum genetic contribution (OGC)“-Methode durch TZ-GI zur Anwendung (siehe hierzu u.a. König und Simianer, 2006). Hierbei kann eine Gewichtung vorgenommen werden, wieviel Verlust an genetischer Diversität in Kauf genommen wird, um eine Steigerung der Zuchtwerte in der nachfolgenden Generation zu erzielen. Voraussetzung für die Anwendung ist eine Berechnung der Zuchtwerte für Zielmerkmale. Um kleine Rassen wie das DSN konkurrenzfähiger zu machen, ist es sinnvoll, eine Steigerung in den Zuchtwerten von Leistungsmerkmalen zu erzielen. Um dabei, wie bei dem DSN den Doppelnutzungscharakter Milch/Fleisch zu erhalten, sollte nur ein begrenzter Verlust an genetischer Diversität in Kauf genommen werden. Daher wurden Zuchtwerte für das 305-d-Produktionsmerkmal Milch-kg nach der in **APIII** etablierten gZWS, sowie genetische Parameter basierend auf dem konventionellen 50K Chip und dem neu entwickelten DSN200K-SNP-Chip geschätzt. In dem Zuchtwertschätzverfahren wurden sowohl Pedigree-Informationen als auch genetische Beziehungen auf Basis von SNP-Markerdaten in einer sogenannten H-Verwandtschaftsmatrix berücksichtigt, weshalb auch gZW für DSN Kühe und Bullen geschätzt werden konnten, von denen keine genomischen Informationen vorlagen.

In **APV** sollte die Übertragbarkeit des entwickelten DSN200K-SNP-Chips auf andere Rassen evaluiert werden. Dabei galt es zu testen, ob der für das DSN entwickelte SNP-Chip auch die wesentlichen kausalen Varianten in anderen Milchrinder- oder Zweinutzungspopulationen abbildet und für züchterische Entscheidungen und Diversitätsstudien genutzt werden kann. Durch die TZ-BER wurden mit DSN nahverwandte Rinderrassen, andere einheimische Rassen und weitere taurine Rassen mit dem neu für DSN entwickelten SNP-Chip genotypisiert und die Genotypen ausgewertet. Hierbei wurden konkret alle Rassen, deren Sequenzen in der Datenbank des „1000-Bullen-Genom-Projekts“ hinterlegt sind, einbezogen. Basierend auf den ersten Daten mit dem 1000-Bullen-Genom-Projekt waren die zu DSN am nächsten verwandten Rassen Holstein, Dutch Friesian Red aber auch die Rasse Normande aus Frankreich sowie skandinavischen Rotviehrassen Swedish Red, Norwegian Red und Danish Red. Die Arbeitsgruppe von TZ-GI sollte dazu zwei weitere bedrohte einheimische Rassen in das Modellvorhaben einbinden. Vorüberlegungen deuteten auf die Rassen Rotbunt Doppelnutzung (RDN) und Angler (ANG) als zu berücksichtigende Populationen hin. Die Untersuchungen zur Anwendbarkeit des Chips auf diese beiden bedrohten einheimischen Rassen wurden durch TZ-BER vorgenommen. Ferner war vorgesehen am Ende des Projekts je 15 Tiere der Rassen RDN und ANG zu sequenzieren. Dafür wurde durch TZ-GI aus Besamungsportionen von 45 Bullen der Rassen RDN und ANG von der Station des Rinderzuchtverbands Schleswig-Holstein (RSH eG) DNA extrahiert und durch die Firma Novogene sequenziert. Die Auswertung dieser Sequenzdaten wurde begonnen, konnte aber nicht abgeschlossen werden, da dafür am Projektende keine Zeit mehr zur Verfügung stand.

Ein Abschlussworkshop gemäß **APVI** war vorgesehen, um die Projektergebnisse in die Praxis zu transferieren. Dazu sollten zwei Aktivitäten umgesetzt werden: Zum einen sollte ein Abschlussworkshop an der Humboldt-Universität Berlin mit den partizipierenden Zuchtverbänden, Arbeitsgruppen anderer Universitäten und Vertretern des Ministeriums durchgeführt werden. Dieser Workshop sollte von der TZ-BER organisiert werden. Zum anderen war die Einbindung der Ergebnisse in die Frankenhäuser DSN-Züchertagung vorgesehen, die federführend von TZ-GI organisiert werden sollte. Da in den Jahren 2021 und 2022 die Projektergebnisse bereits sehr umfangreich der Öffentlichkeit im wissenschaftlichen und praktischen Rahmen präsentiert wurden (siehe „Übersicht über erfolgte oder geplante Veröffentlichungen“), wurde in Rücksprache mit dem Projektgeber vereinbart, dass ein zusätzlicher Workshop an der Humboldt-Universität Berlin zum Ende des Projektes keinen Mehrwert bringt, so dass auf diesen im allgegenwärtigen Einvernehmen verzichtet wurde. Eine Vorstellung der Ergebnisse auf der Frankenhäuser DSN-Züchertagung konnte nicht erfolgen, da diese aufgrund der Corona-Pandemie zwischen den Jahren 2020 bis 2022 nicht stattfinden konnte. Es ist vorgesehen, die Projektergebnisse nach Projektabschluss auf der nächsten Frankenhäuser DSN-Züchertagung (für November 2023 geplant) vorzustellen.

1.2 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Das Deutsche Schwarzbunte Niederungsring

Bei der Rasse DSN handelt es sich um eine robuste Zweinutzungsrasse, welche zur Milch- und Fleischproduktion genutzt wird. Im Jahr 2021 umfasste die lebende DSN-Population 2.514 weibliche und 27 männliche Zuchttiere. Das entspricht einer effektiven Populationsgröße (N_e) von 106,85. Somit handelt es sich bei der DSN-Population nach der Definition der Bundesanstalt für Landwirtschaft und Ernährung (BLE) um eine Erhaltungspopulation ($N_e < 200$).

Charakteristisch für DSN Tiere sind die erhöhten Milchfettanteile, hohe Dauerleistungen, gute Anpassungsfähigkeit und Fruchtbarkeit. Auch bei reiner Grundfutterzufuhr weisen DSN gute Mastleistungen auf und sind daher besonders gut für Grünlandstandorte geeignet (Jaeger, 2018).

Die Tiere weisen einen mittleren Rahmen und eine mittlere Bemuskelung auf (Abbildung 1). Weibliche Tiere werden im Schnitt 1,30 bis 1,42 m groß bei einem durchschnittlichen Gewicht von 550 bis 650 kg. Mit einer durchschnittlichen Größe von 1,50 bis 1,62 m und 1.000 bis 1.150 kg sind die männlichen Tiere etwas größer und deutlich schwerer (Rahmann, 2011). Laut Rassedefinition sind DSN regelmäßig schwarz-weiß gescheckt und genetisch horntragend. Schwarze Zeichnungen finden sich an Kopf, Hals, Brust und Becken. Weiße Scheckung haben die Tiere an Bauch, Euter und an den Beinen bis zum Karpal- bzw. Sprunggelenk. Als Zuchtziel soll die Doppelnutzung aufrechterhalten werden, jedoch soll das genetische Milchleistungspotential auf über 7.000 kg mit einem Fettgehalt von 4,30 % und Eiweißgehalt 3,70 % gesteigert werden (ZBH, 2019). Außerdem sollen Zuchttiere aufgrund der gemeinsamen Zuchtgeschichte mit HOL maximal einen HOL-Anteil von 10 % aufweisen.

Ihren Ursprung hat die Rasse in den Marsch- und Niederungsgebieten der Nordseeregion. Im 18. Jahrhundert gab es in den dänischen, holländischen und deutschen Küstengebieten viele schwarzbunte Landschläge mit guten Milchleistungen, welche den Ursprung der heutigen Rasse DSN bildeten (Mügge, 1999). Ab Mitte des 19. Jahrhunderts wurden viele Rinder mit englischen Shorthornrindern gekreuzt, wodurch die Milchleistung der schwarzbunten Kühe zurückging. Das hatte zur Folge, dass wieder zur Reinzucht der schwarzbunten Rinder zurückgefunden wurde. So entstand das erste Zuchtbuch der Rasse im Jahr 1876 (Haller, 2015). Im 19. Jahrhundert wurden schwarzbunte Tiere der Nordseeregion in die USA exportiert, wo sie auf eine hohe Milchleistung selektiert wurden und schon bald die Rasse Holstein-Friesian (HF) bildeten. Die heutige kleine Populationsgröße der Rasse DSN ist auf eine Verdrängungskreuzung durch HF in den 1960er Jahren zurückzuführen. Damals wurden stark HF-Bullen aus Amerika eingesetzt, um hohe Milchleistungen zu erreichen. Die ursprüngliche Zuchtrichtung der Rasse DSN wurde in der damaligen Deutschen Demokratischen Republik (DDR) als Genreserve gehalten. Dort wurden DSN-Kühe mit dänischen Jersey Bullen gekreuzt.

Mit deren F1-Generation wurde durch Kreuzung mit der Rasse HF das Schwarzbunte Milchrind gezüchtet, das die dominierende Milchrasse in der DDR wurde (Jaeger, 2018).

Um den Bestand der Rasse DSN zu sichern, wurde 1989 der „Verein zum Erhalt und Förderung des alten Deutschen Schwarzbunten Niederungsrind“ gegründet. Bei dem Friedrich-Loeffler-Institut in Mariensee gibt es eine Kryoreserve von 300 Embryonen der Rasse DSN. Diese dient dazu, ein Inzuchtmanagement zu ermöglichen und die genetische Diversität zu bewahren (Jaeger, 2018).



Abbildung 1: Deutsches Schwarzbuntes Niederungsrind (Foto: L. Aufmhof)

Genomweite Assoziationsstudien beim DSN

In den Arbeitsgruppen der TZ-GI und TZ-BER wurden bereits vor dem Start des Projektes GWAS für verschiedene Merkmale innerhalb der DSN-Population durchgeführt. May et al. (2019) identifizierten 44 signifikant assoziierte SNPs und 23 Kandidatengene für eine Resistenz gegen Endoparasiteninfektionen (gastrointestinale Nematoden, Lungenwürmer, großer Leberegel) anhand von imputierten 700K Genotypdaten. Ziel der Studie war es, herauszufinden, ob es SNP-Marker gibt, die mit einer Endoparasiteninfektion in Verbindung stehen, ob diese SNP-Marker mit Kandidatengenen assoziiert werden können und ob die SNP-Marker die Korrelation zwischen der Endoparasiteninfektion und der Milchleistung beeinflussen. Diese identifizierten SNPs lagen auf den Chromosomen 4, 5, 18, 21, 22, 24, und 26, wobei fünf der potentiellen Kandidatengene bei der Aktivierung von Immunantworten beteiligt sind. Außerdem wurde eine antagonistische Assoziation zwischen Endoparasiten- und Euterinfektionen, welche Einfluss auf die Leistung hatte, beschrieben. Yin et al. (2019) fokussierten im Rahmen einer ‚multi-breed‘ GWAS auf neue Phänotypen zum Futteraufnahmeverhalten und zur Wiederkauaktivität aus Sensordaten, wobei über 300 Tiere der Rasse DSN inkludiert waren. In dieser Analyse wurden 22 potentielle Kandidatengene mit Merkmalen zum Weideverhalten

assoziiert. Eine ‚multi-breed‘ GWAS, wie von Yin et al. (2019) durchgeführt, kann in Rassen mit kleiner Populationsgröße und wenigen Phänotypen für neue innovative Merkmale, wie es der Fall in DSN ist, zur Identifizierung genetischer Marker sehr zielführend sein. Ein Nachteil der ‚multi-breed‘ GWAS in Rassen mit kleiner Populationsgröße wie DSN ist aber, dass rassespezifische Varianten mit einer zu geringen Power detektiert werden. Das Gen *DGAT1* auf Chromosom 14, welches in HOL hochsignifikant mit dem Milchfettgehalt assoziiert ist, auch in der ‚multi-breed‘ GWAS unter Inkludierung von über knapp 300 DSN von Yin et al. (2019) als potentieller Kandidatengen für den Milchfettgehalt identifiziert. Dies führte zu einer Fehlinterpretation für DSN. In rassespezifischen GWAS innerhalb der Rasse DSN zeigt sich hingegen keine Assoziation zwischen *DGAT1* und Milchleistungsmerkmalen, da diese Region in DSN fast komplett fixiert ist (Korkuć et al., 2021). Dies ist ein signifikantes Beispiel für die Notwendigkeit zur Entwicklung eines rassespezifischen SNP-Chips und die sorgfältige Analyse kleiner Rassen.

Meier et al. (2020) führten eine GWAS mit 1.062 DSN-Kühen für klinische Mastitis durch, um die Mastitis-Anfälligkeit beim DSN besser zu verstehen. 489 der untersuchten Kühe hatten mindestens einen Fall der klinischen Mastitis und 573 Kühe erkrankten in den ersten drei Laktationen nie an klinischer Mastitis. Es wurden fünf assoziierte SNP-Marker beschrieben, drei davon signifikant assoziiert mit einer klinischen Mastitis. Ein SNP lag im Gen *NEURL1* auf Chromosom 26 und die Autoren hoben ebenfalls das Kandidatengen *BMP1B* auf Chromosom 6 hervor, welches schon in Studien zu anderen Rinderrassen mit der klinischen Mastitis in Verbindung gebracht werden konnte. Die Autoren konnten 1-3 % der Varianz der Krankheit in der untersuchten Population mit den gefundenen SNPs erklären.

In GWAS mit Milchleistungsdaten in 1.490 DSN-Kühen, die von TZ-BER durchgeführt wurde (Korkuć et al., 2021), sind 41 signifikante und 20 suggestive SNPs für Merkmale der Milchleistung und 15 SNPs für den Eiweißgehalt in der Milch identifiziert worden. Die signifikantesten Auswirkungen auf die Milchleistung wurden auf den Chromosomen 1, 6 und 20 festgestellt. Die Region auf Chromosom 6 befand sich in der Nähe des Kasein-Genclusters; der signifikante SNP in einem Haplotyp, der sich mit dem *CSN3*-Gen (Kappa-Kasein) überschneidet. Assoziationen für den Fett- und Eiweißanteil und -gehalt wurden ebenfalls identifiziert. Wie schon oben erwähnt, wurde trotz der engen Beziehung zwischen DSN und HOL kein Zusammenhang zwischen *DGAT1* und anderen bekannten Genen, die Einfluss auf Milchproduktionsmerkmale bei HOL-Kühen aufweisen (z.B. *GHR*, *ABCG2*, *GPAT4*), festgestellt.

Genomische Zuchtwertschätzung in kleinen Populationen

Kleine, vom Aussterben bedrohte Rassen wie das DSN profitieren derzeit nur gering vom Entwicklungsfortschritt in der gZWS, die in den zahlenmäßig großen Populationen wie HOL und FV bereits in der Routine angewandt wird. Basis dafür sind kommerzielle SNP-Chips, die

speziell auf die großen Populationen „zugeschnitten“ sind. Bei dieser kommerziellen SNP-Chip Entwicklung spielen die genomische Architektur der Hochleistungsrassen und die konventionellen Zuchtzielmerkmale (insbesondere Produktionsmerkmale) eine entscheidende Rolle. Während in den großen Populationen wie HOL und FV durch genomische Selektionsschemata ein beschleunigter Zuchtfortschritt beobachtet werden kann, bleiben kleinere, bedrohte Populationen im Hinblick auf ihre Wettbewerbsfähigkeit im Vergleich zu den größeren Rassen züchterisch zunehmend weiter zurück. Darüber hinaus sind für kleine Populationen das Inzuchtmanagement bzw. die Erhaltung der genetischen Diversität von großer Bedeutung. Züchterische Strategien müssen somit zwei zentrale Bedingungen gleichzeitig berücksichtigen: Die Maximierung von Zuchtfortschritt bei gleichzeitiger Restriktion für zukünftige verwandtschaftliche Beziehungen bzw. Maximierung von Kennzahlen der genetischen Diversität. Die genaueste Beurteilung der Diversität ist möglich (insbesondere bei unzureichender Tiefe der Abstammungsdateien), wenn hierzu genomische Markerdaten verwendet werden. Dabei erscheint der Heterozygotiegrad basierend auf SNP-Markerdaten die aktuell beste Maßzahl, um genetische Variabilität innerhalb von Populationen zu beschreiben. Allerdings kann auch dieser Heterozygotiegrad nur dann genau abgeleitet werden, wenn VGS Informationen der Tiere zur Verfügung stehen (siehe u.a. Malomane et al., 2018)

Lernstichprobe für die genomische Zuchtwertschätzung

Für eine funktionierende gZWS, d.h. geschätzte Zuchtwerte auf genomischer Basis mit einer hohen Sicherheit, bedarf es einer ausreichend großen (Lern-)Stichprobe an genotypisierten Tieren innerhalb einer Rasse. Als kleinste Größe werden hier in der Literatur mindestens 2.000 mit dem 50K SNP-Chip genotypisierte Tiere genannt (u.a. Meszaros et al., 2015). Die aktuelle DSN-Population an Tieren mit geringem Fremdblutanteil (< 10 %) in Deutschland beträgt mehr als 2.000 Tiere und erfüllt daher dieses Kriterium. Die derzeit publizierten Forschungsergebnisse und Praxiserfahrungen zeigen, dass die genetischen SNP-Markereffekte, die den gZW eines Tieres bestimmen, spezifisch in der jeweiligen Population geschätzt werden sollten. Das heißt, dass der für HOL oder FV entwickelte kommerzielle SNP-Chip nicht genutzt werden kann. Die genetische Architektur der wichtigsten züchterisch bearbeiteten Merkmale, d.h. die für die Merkmalsausprägungen kausalen Genvarianten, unterscheidet sich in der Regel deutlich zwischen Populationen und Rassen. Das gilt insbesondere für Rassen mit verschiedenen Zuchtzielen (hier: Hochleistungs-Milchviehrassen *versus* kleine lokale Doppelnutzungsrassen DSN). Allerdings vertreten wir die Hypothese, dass ein neuentwickelter SNP-Chip für DSN auf andere lokale Rassen mit ähnlichem Zuchtziel und ähnlicher Selektionssignatur teilweise übertragbar ist.

Bei lokalen Rassen, die vom Aussterben bedroht sind, stehen Merkmale der Adaptation, der Tiergesundheit oder des Tierverhaltens im Fokus, die in den großen Populationen bisher kaum

züchterisch bearbeitet wurden. In diesem Zusammenhang ist es wichtig festzuhalten, dass die derzeit in der Routine der gZWS eingesetzten SNP-Marker, die die kausalen Genvarianten abbilden sollen, ausschließlich auf der Grundlage von Erkenntnissen in den großen Rassen ausgewählt wurden. Die spezifischen kausalen Genvarianten, wie zum Beispiel der Rasse DSN, die auf genetischer und genomischer Ebene maßgeblich den Beitrag der Rasse DSN zur genetischen (Bio-)Diversität ausmachen, sind dementsprechend bisher nur unzureichend über die aktuell eingesetzten SNP-Marker erfassbar. Ein spezifisches Marker-Genotypisierungs-Panel (sog. custom-SNP-Arrays oder custom-SNP-Chips) für eine bedrohte Rasse wie DSN könnte helfen, die beschriebenen Grenzen zu überwinden. Auf solchen Chips könnten auch bisherige in GWAS identifizierte Marker berücksichtigt werden, welche beim DSN mit Milchleistungsmerkmalen oder insbesondere auch mit funktionalen Merkmalen assoziiert sind, um wichtige Alleinstellungsmerkmale der Rasse züchterisch weiterzuentwickeln. Die mit solchen rassespezifischen Chips erzeugten Genotypen repräsentieren den Genotyp der Tiere dieser Rasse genauer. Damit sind sie besser für den Aufbau einer gZWS innerhalb kleiner Populationen geeignet. Wichtige Voraussetzung für die gZWS ist aber eine sichere, valide Schätzung der Allelfrequenzen für möglichst alle rassetypischen Allelvarianten auf der Basis von VGS. Für die Erprobung eines custom SNP-Chips für die gZWS in kleinen Populationen gibt es derzeit international keine Erfahrungswerte aus Grundlagenforschung oder Praxis, was die Modellhaftigkeit des Projektes deutlich unterstreicht und die Erprobung eines rassespezifischen SNP-Chips rechtfertigt.

2. Material und Methoden

2.1 Aufbau einer informativen Lernstichprobe (APIa)

Für **APIa** wurden zunächst 214 DSN für die Sequenzierung ausgewählt, darunter 30 DSN-Bullen und 184 DSN-Kühe, die Schlüsseltiere in der Population sind. Entsprechend der Definition der Rassezugehörigkeit lag der Anteil von Fremd-DNA bei allen Tieren unter 10 %. Von TZ-GI wurden DSN Tiere aus einem hessischen Lehr- und Forschungsbetrieb ausgewählt, zu welchen bereits Blut,- Haar, - oder Spermaproben aus früheren Projekten in den DNA-Banken vorlagen. Für einige dieser Tiere wurden DNA-Proben in der DNA-Bank von TZ-BER beherbergt. Die Tiere wurden zunächst anhand ihres DSN-Anteils ausgewählt. Die Berechnung des DSN-Anteils erfolgte nach einem von TZ-GI entwickelten Algorithmus (Jaeger, 2018). In Zusammenarbeit mit dem Verein zur Erhaltung des Deutschen Schwarzbunten Niederungsrindes und der Rinderproduktion Berlin-Brandenburg GmbH (RBB) erfolgte im Anschluss daran eine händische Prüfung des Pedigrees aller Tiere, sodass final ausschließlich Tiere mit einem DSN-Anteil von > 90 % mit einer sehr hohen Genauigkeit ausgewählt werden konnten. Im nächsten Schritt erfolgte ein Abgleich mit vorliegenden innovativen Phänotypen und Leistungsmerkmalen. Ziel war es, die Tiere in einem Case-Control Design für das funktionale Merkmal „Endoparasitenresistenz“ auszuwählen, welches beim DSN aufgrund der intensiven Nutzung in Weideproduktionsbedingungen eine besondere Rolle spielt. Dafür wurden die Tiere basierend auf wiederholten parasitologischen Untersuchungen zum Nachweis von Infektionen mit gastrointestinalen Nematoden (GIN) und dem großen Leberegel (*Fasciola hepatica*) als „resistent“ (Controls) und „anfällig“ (Cases) klassifiziert. Ziel war es, möglichst die Hälfte „anfällige“ Tiere sowohl für GIN- als auch *Fasciola hepatica*-Infektionen unter den selektierten Tieren zu realisieren. Im Anschluss daran wurden Hauptkomponenten-Analysen durchgeführt, um auf Stratifikation im Datensatz zu prüfen und die Verwandtschaft zwischen Cases und Controls abzubilden (Abbildung 2). Weiterhin erfolgten Auswertungen zu den vorhandenen Testtagsdaten (u.a. Milchleistung, Exterieur, Fruchtbarkeit, Nutzungsdauer) und Zuchtwerten der Tiere. Da einige der selektierten DSN-Kühe bereits abgegangen waren, war das vollständige Vorhandensein von Testtagsdaten über mehrere Jahre eine Voraussetzung, damit diese Tiere für eine Sequenzierung berücksichtigt wurden. Zudem erfolgte ein Abgleich mit weiteren Gesundheitsmerkmalen wie Mastitis und Merkmalen der Adaptationsfähigkeit (Merkmale für Hitzestress).

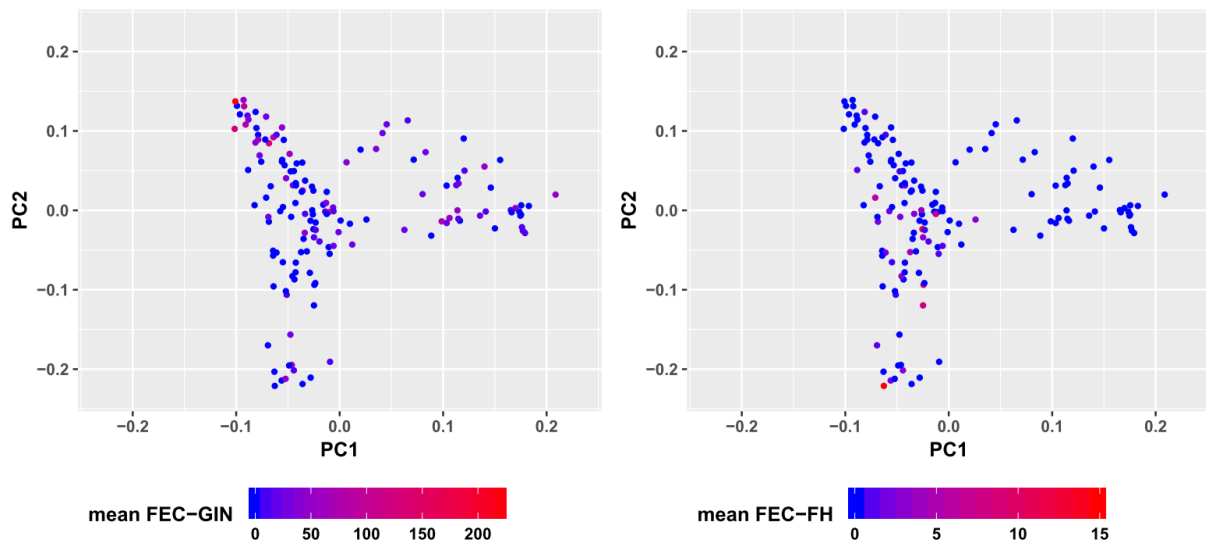


Abbildung 2: Hauptkomponenten-Analyse für die Merkmale durchschnittliche Eiausscheidung gastrointestinaler Nematoden (mean FEC-GIN) und durchschnittliche Eiausscheidung des großen Leberegels (*Fasciola hepatica*) (mean FEC-FH) bei Deutschen Schwarzbunten Niederungsrindern (DSN) aus drei Betrieben mit Weidehaltung.

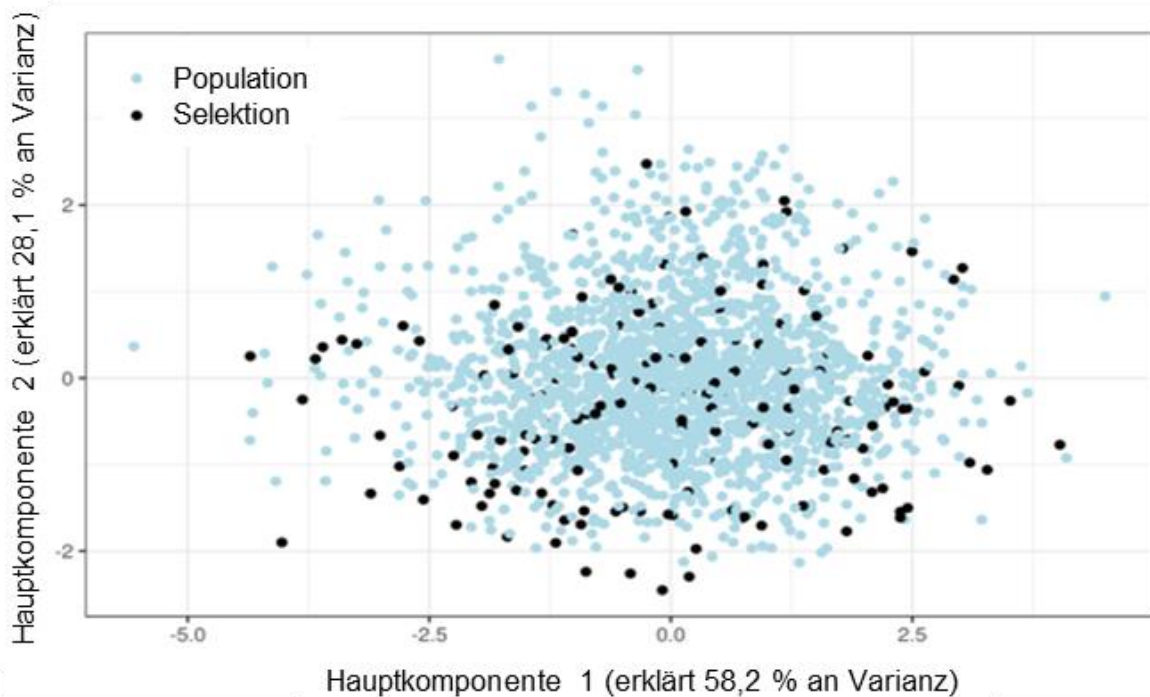


Abbildung 3: Hauptkomponenten-Analyse für Kühe der Rasse Deutsches Schwarzbuntes Niederungsrind (DSN), basierend auf der Milchleistung der ersten Laktation für selektierte DSN-Kühe (=Selektion, schwarze Punkte) sowie die DSN-Gesamtpopulation (=Population, hellblaue Punkte).

TZ-BER konnte bei der Auswahl der Tiere für die Sequenzierung auf die von ihnen aufgebaute umfangreiche DNA-Datenbank für ca. 2.500 Tiere der DSN-Population zurückgreifen. Erstes Auswahlkriterium war der DSN-Anteil der Tiere, welcher > 90 % betragen sollte. Unter Verwendung der Pedigreedaten wurden Abstammung und Verwandtschaftsgrad aller Tiere analysiert, um möglichst Individuen verschiedener Bullen und Kühe aus verschiedenen Geburtsjahren einzubeziehen. Ziel war es, solche Tiere auszuwählen, die die gesamte DSN-Population möglichst gleichmäßig repräsentieren. Von TZ-BER wurden alle in der DNA-Bank vorliegenden Bullen in die Sequenzierung inkludiert. Die Auswahl der DSN-Kühe erfolgte aus neun verschiedenen Betrieben. Für die weiteren Tiere wurden die Testtagsdaten (u.a. Milchleistung, Exterieur, Fruchtbarkeit, Nutzungsdauer) und die 305-d Produktionsmerkmale (Milch-kg, Milchlaktose-%, Milcheiweiß-%) und Zuchtwerte (Relativzuchtwert Milch (RZM), Relativzuchtwert Gesamt (RZG)) analysiert. Die Phänotypen der ausgewählten Tiere sollten der Normalverteilung der DSN-Population entsprechen, d.h. dass sowohl Tiere mit hoher, also auch mittlerer und niedriger Leistung inkludiert wurden (Abbildung 3). Als Merkmale der Krankheitsresistenz lagen zudem Mastitis-Diagnosedaten zu einem Großteil der Kühe vor, welche als weiteres Kriterium für die Auswahl der Tiere berücksichtigt wurden.

Die Sequenzierung der 214 Proben erfolgte bei der Firma Novogene in der Illumina PE150, auf der Illumina NovaSeq Plattform mit 150 gepaarten Enden und einer 15-fachen Abdeckung des Genoms. Für die Identifizierung der Sequenzvarianten aus den VGS-Daten hat TZ-BER eine Pipeline basierend auf den Richtlinien des 1000-Bullen-Genom-Projekts (<http://www.1000bullgenomes.com>) erarbeitet (Abbildung 4). Die VGS-Daten von 57 DSN Tieren wurden zudem in die Datenbank des 1000-Bullen-Genom-Projekts eingespeist, wo ebenfalls Sequenzvarianten identifiziert wurden. Die Übereinstimmung der Genotypen für diese 57 DSN Tiere aus der erarbeiteten Pipeline und aus der Datenbank des 1000-Bullen-Genom-Projekts war 99,996 %. Dies zeigt, dass die erarbeitete Pipeline sehr gut funktionierte und sehr gut für die Detektion der Sequenzvarianten geeignet ist.

Die unter **APIa** detektierten Sequenzvarianten aller DSN Tiere wurden in **APIb** als Referenzpanel für die Imputation von Sequenzvarianten in Genotypen der ca. 1.800 DSN, die mit dem Illumina BovineSNP50 BeadChip (50K SNP-Chip) genotypisiert wurden, auf Sequenzniveau verwendet. Die hier verwendeten 50K Genotypen lagen bereits aus vorangegangenen Projekten vor. Die Imputation erfolgte mit der Software Beagle v.5.1 (Browning et al., 2018) nach dem beschriebenen Vorgehen von Korcuć et al. (2019). Im Anschluss wurden die 1.800 auf Sequenzlevel imputierten DSN Genotypen mit den 304 sequenzierten DSN Genotypen zusammengeführt, so dass ein imputierter Datensatz mit etwa 2.100 Genotypen auf Sequenzniveau erstellt wurde.

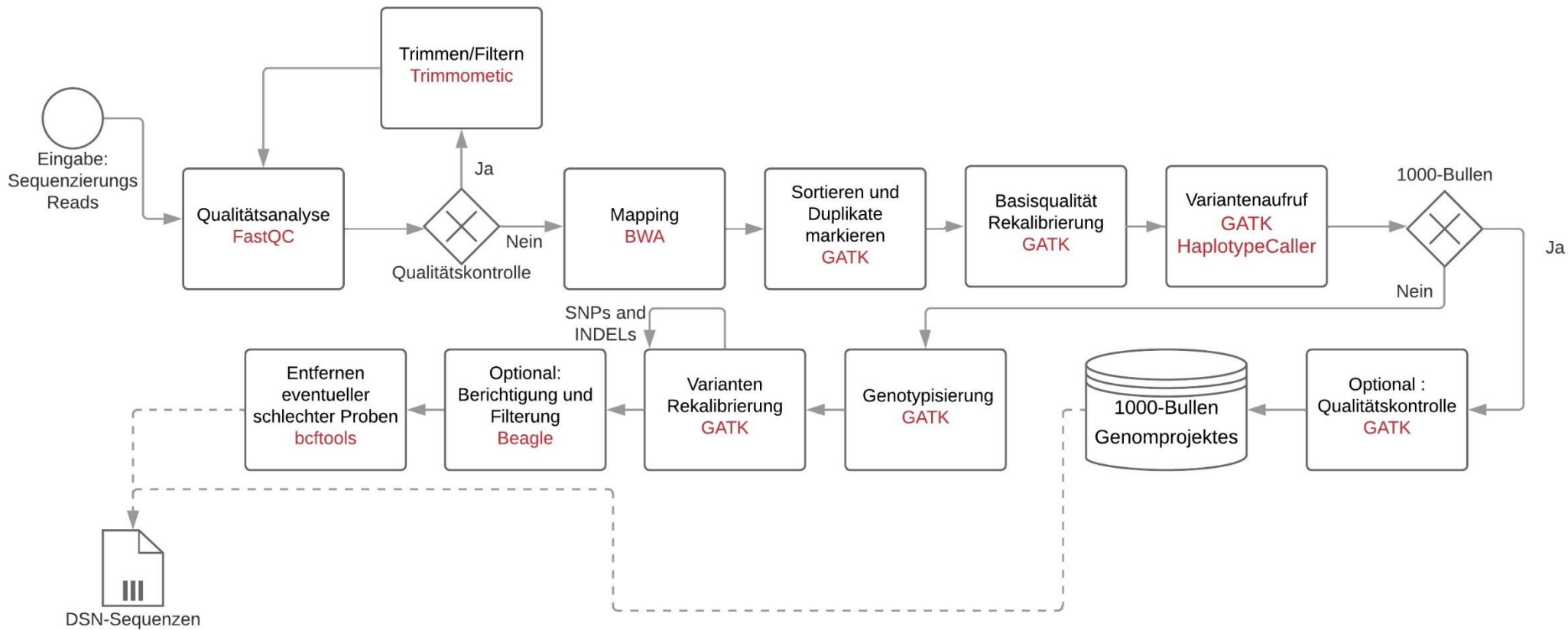


Abbildung 4: Überblick der Pipeline zur Prozessierung der Rohsequenzdaten. Die Pipeline wurde in Anlehnung an die Richtlinien vom 1000-Bullen-Genom-Projekt entwickelt. Die Namen der verwendeten Softwares sind rot markiert.

2.2 SNP-Chip Design (APII)

Für die Entwicklung eines custom-made SNP-Chips für die Rasse DSN haben wir uns nach Vergleich der Kosten, Leistungsfähigkeit und technischer Umsetzungsmöglichkeiten in bestehender Infrastruktur für das Axiom™ myDesign™ TG Array mit ca. 200.000 Varianten (SNPs und INDELs) der Firma Thermo Fisher Scientific entschieden. Der für die Rasse DSN neuentwickelte Chip (DSN200K-SNP-Chip) umfasst 182.154 Sequenzvarianten, die nach zehn Kategorien ausgewählt wurden (Tabelle 2) (Neumann et al., 2021). Eine detaillierte Erläuterung der Kategorien erfolgt weiter unten.

Tabelle 2: Anzahl an Sequenzvarianten in einzelnen Kategorien, welche bei der Konstruktion des DSN200K-SNP-Chips berücksichtigt wurden. Sequenzvarianten können zu mehreren Kategorien zugeordnet werden, daher muss die absolute Anzahl von Varianten ohne doppelte Zuordnungen berechnet werden.

Kategorie	Anzahl Sequenzvarianten
1) Varianten vom Illumina BovineSNP50 Chip	34.039
2) Varianten aus GWAS in DSN	2.071
3) Varianten mit funktionellem Einfluss auf Gene	50.611
4) DSN-spezifische Varianten	38.198
5) Varianten mit hohem Unterschied in AAF zwischen DSN und HOL	55
6) Varianten auf dem Y-Chromosom	321
7) Varianten auf den mitochondrialen Chromosomen	278
8) Varianten zur Abstammungsbestimmung (ICAR ¹ & ISAG ²)	554
9) Varianten in Haplotypblöcken	103.801
10) Varianten zum Füllen von Lücken	16
Absolute Anzahl (ohne doppelte Zuordnung zu Kategorien)	182.154

¹ International Committee for Animal Recording

² International Society for Animal Genetics

Technische Konstruktion des SNP-Chips

Das Design des DSN200K-SNP-Chips erfolgte nach den Qualitätsparametern der Affymetrix Pipeline (Kranis et al., 2013). Die in **API** identifizierten 20,5 Millionen Sequenzvarianten in der gegebenen DSN Population wurden zunächst gefiltert, um für das Design des SNP-Chips einen Datensatz mit technisch geeigneten Varianten zu erhalten. Es wurden nur hoch-qualitative biallelische Varianten benutzt, die keine weiteren Varianten in einer Entfernung von 35 Nukleotiden zueinander besitzen. Diese Entfernung von Varianten zueinander musste eingehalten werden, um sicherzustellen, dass die für den SNP-Chip benötigten Oligonukleotide (71 Nukleotid-lange Sequenzstücke, die zur SNP-Detektion auf dem SNP-Chip benötigt werden) einzigartig sind und es zur eindeutigen Detektion der gewünschten SNPs kommt. Durch diese ersten Filterungsschritte wurde die Zahl der 20,5 Millionen Varianten auf 8,9 Millionen Varianten reduziert. Weitere Filterung umfasste die Abdeckung der SNPs mit Reads, die

Allelfrequenzen in DSN und anderen Rassen aus dem 1000-Bullen-Genom-Projekt hatten. Nach der Filterung blieben 5,5 Millionen Varianten übrig, die einer Überprüfung auf technische Tauglichkeit durch Thermo Fisher Scientific unterzogen wurden. Dazu wurde für jeden SNP ein sogenannter „*in-silico* score“ berechnet, der angibt mit welcher Wahrscheinlichkeit ein SNP auf dem Chip funktionieren würde. Von 5,5 Millionen überprüften Varianten wurden 3,1 Millionen Varianten als technisch geeignet deklariert und somit für das weitere Chipdesign empfohlen. Im Rahmen des Filterungsprozesses wurden neben den genomischen Informationen der DSN Tiere an dieser Stelle bereits genomische Informationen weiterer Rassen wie HOL, Braunvieh und Gelbvieh aus dem 1000-Bullen-Genom-Projekt mitberücksichtigt. Grund dafür ist, dass der DSN200K-SNP-Chip auch für weitere bedrohte Rassen evaluiert werden sollte bzw. sich später in der Praxis optimaler Weise auch für andere Rassen eignet (siehe **APV**). Auf Basis der 3,1 Millionen empfohlenen Sequenzvarianten (SNPs und INDELS) wurden die Varianten für den neuen DSN200K-SNP-Chip nach den oben genannten zehn Kategorien ausgewählt.

Detaillierte Erläuterung der zehn Kategorien zur Auswahl von Sequenzvarianten für das Design des DSN200K-SNP-Chips:

1) Varianten vom Illumina BovineSNP50 Chip

33,4K Varianten des Illumina BovineSNP50 v3 BeadChip (Illumina Inc., San Diego, CA) mit einer SNP Call Rate > 90 % und einer Frequenz des selteneren Allels (MAF, engl.: „minor allele frequency“) < 5 % in der Rasse DSN wurden ausgewählt. Die Auswahl von Varianten des Illumina Chips erfolgte aus mehreren Gründen: Bei den auf dem Illumina Chip berücksichtigten SNPs handelt es sich bereits um eine optimierte Auswahl von SNPs, welche das *Bos taurus* Genom sehr gut abdecken und eine hohe genetische Relevanz und Merkmalsassoziationen beim Rind besitzen. Zudem werden die Genotypisierungs-Chips der Firma Illumina häufig in Forschung und Praxis verwendet. Daher werden diese SNPs benötigt, um zukünftige Ergebnisse auf Basis unseres DSN200K-SNP-Chips mit den Ergebnissen anderer genetischer Studien beim Rind vergleichen zu können. Weiterhin wird es dadurch ermöglicht, Genotypen die von dem Illumina BovineSNP50 v3 BeadChip und dem DSN200K-SNP-Chip stammen zu kombinieren.

2) Varianten aus GWAS und Gen-Sequenzierungen in DSN

Insgesamt wurden 2.071 Sequenzvarianten aus GWAS-Analysen, die in DSN durchgeführt wurden, ausgewählt (Tabelle 3). Zuerst wurden alle Varianten zusammengetragen, welche in vorangegangenen GWAS mit verschiedensten Merkmalen beim DSN assoziiert wurden. Dabei handelte es sich um Varianten für funktionale Merkmale wie Mastitisresistenz (Meier et al., 2020), Endoparasitenresistenz (May et al., 2019), aber auch Produktionsmerkmale wie Milchleistung (Korkuč et al., 2021). Des Weiteren wurden aber auch neue GWAS mit der in **API** generierten imputierten Genotypen durchgeführt. Für die GWAS wurde entweder die

Programmiersprache R oder die Software GCTA (Yang et al., 2011) verwendet. Es wurden Phänotypdaten und Zuchtwerte zu Milchleistung, Fruchtbarkeit, Exterieur, die beim vit abgefragt wurden, analysiert. Zusätzlich wurden weitere phänotypische Daten zu Endoparasitenresistenz gesammelt und analysiert. Des Weiteren wurden auch Phänotypen zu Hitzestress (z.B. Milchfettsäuren, physiologische Merkmale) und Futtereffizienz in DSN analysiert (Bohlouli et al., 2022; Yin et al., 2019). Eine Auswahl von assoziierten Varianten aus den verschiedenen GWAS wurde für das Design des DSN200K-SNP-Chips berücksichtigt.

Für das funktionale Merkmal Endoparasitenresistenz war es zudem möglich, Sequenzvarianten für das Activated leukocyte cell adhesion molecule (*ALCAM*) Gen im DSN-Genom zu identifizieren, welches vorab in der GWAS mit Leberegel-Infektionen assoziiert war (May et al., 2019). Als Datengrundlage dienten hier 94 DSN-Kühe mit vorliegenden Sequenzdaten für das *ALCAM* Gen. Das Gen codiert für Proteine, welche eine bedeutende Rolle in immunologischen Prozessen spielen. Bei diesen Analysen wurden drei Polymorphismen in Exon 9 (*c.1017T>C*) und in Intron 9 (*c.1104+10T>A*, *c.1104+85T>C*) des Gens identifiziert. Für diese drei SNPs wurden Genotyp- und Allelfrequenzen sowie Unterschiede in den Genotyp- und Allelfrequenzen zwischen *Fasciola hepatica* bzw. GIN infizierten und nicht-infizierten Tieren berechnet. Ziel war es, Allelsubstitutions- und Dominanzeffekte für die identifizierten SNPs für Merkmale der Endoparasitenresistenz sowie für weitere Testtagsmerkmale (Milch-kg, Milcheiweiß-%, Milchfett-%, FEQ, SCS, Harnstoff, Laktationspersistenz) zu schätzen. Die geschätzten Effekte sollten Aufschluss darüber geben, ob die identifizierten Sequenzvarianten günstig für eine Endoparasitenresistenz als auch für weitere Leistungs- sowie Krankheitsresistenzmerkmale (z.B. SCS) sind, sodass diese SNPs auf dem neu konstruierten custom-made Chip berücksichtigt werden können. Die Schätzung der Markereffekte in den drei identifizierten SNPs erfolgte mittels logistischer Regressionsanalysen in der Statistik-Software SAS. Die Analysen ergaben signifikante Unterschiede in den Allel- und Genotypfrequenzen zwischen mit Leberegeln (*Fasciola hepatica*) infizierten und nicht-infizierten Tieren für alle drei SNPs. Zudem wurden signifikante Allelsubstitutionen in den drei SNPs für eine Endoparasitenresistenz (sowohl *Fasciola hepatica*-Infektionen als auch Infektionen mit gastrointestinalen Nematoden) als auch für Milcheiweiß-%, Milchfett-% sowie FEQ geschätzt. Genotyp TC der SNPs *c.1017T>C* and *c.1104+85T>C* zeichnete sich als günstiger Genotyp hinsichtlich einer Verbesserung der *Fasciola hepatica*-Resistenz beim DSN aus, hatte aber gleichzeitig einen negativen Effekt auf die Merkmale Milch-kg und SCS. Für den SNP *c.1104+10T>A* zeigten sich günstige Allelsubstitutions-Effekte sowohl für eine *Fasciola hepatica*-Resistenz als auch für die Merkmale Milchfett-%, Milcheiweiß-%, FEQ und Laktationspersistenz beim DSN. Die Ergebnisse der Analysen wurden im Rahmen des Projektes in dem Journal „*Mammalian Genome*“ publiziert (May et al., 2019) und die identifizierten SNPs auf den neuen DSN200K-SNP-Chip inkludiert.

Tabelle 3: Anzahl der Sequenzvarianten, die in genomweiten Assoziationsstudien und Gen-Sequenzierungen mit den jeweiligen Merkmalen identifiziert und für das Design des DSN200K-SNP-Chips benutzt wurden.

Untersuchtes Merkmal	Anzahl Sequenzvarianten
Milch	568
Mastitisresistenz	67
Endoparasitenresistenz	163
Fruchtbarkeit	895
Exterieur	326
Hitzestress	35
Futtereffizienz	7

3) Varianten mit funktionellem Einfluss auf Gene

Die Funktion der Sequenzvarianten wurde mit der Software VEP („Variant Effect Predictor“) vorausgesagt (McLaren et al., 2016). Für insgesamt 50.611 Varianten wurde ein hoher, mittlerer oder niedriger Einfluss geschätzt, weswegen sie für den DSN200K-SNP-Chip berücksichtigt wurden. Ein hoher Einfluss aufs Genom hat z.B. eine „Missense“-Mutation, die Änderungen beim Ablesen des genetischen Codes und des translatierten Proteins mit sich bringt.

4) DSN-spezifische Varianten

Es wurden 38.198 Varianten als „einzigartig für die Rasse DSN“ eingestuft. Dies war der Fall für Varianten, die eine alternative Allelfrequenz (AAF) $> 1\%$ in DSN aufwiesen und in DSN verwandten Rassen mit vorhandenen Sequenzdaten aus dem 1000-Bullen-Genom-Projekt (HOL, Modern Danish Red, Original Braunvieh, Gelbvieh, Finnish Ayrshire, Normande, Norwegian Red, Swedish Red, und Jersey) vollkommen fehlten. Dafür wurden die Allelfrequenzen für alle Rassen mit der Software vcfTools v0.1.16 (Danecek et al., 2011) kalkuliert. Zudem wurden Varianten mit einer MAF $> 1\%$ in DSN in diesem Schritt inkludiert. Die Einzigartigkeit dieser Varianten ist der Schlüssel zur Sicherung der genetischen Diversität und rassespezifischer Merkmale.

5) Varianten mit hohem Unterschied in AAF zwischen DSN und HOL

Es wurden 55 Varianten mit einem hohen Unterschied in den alternativen Allelfrequenzen (AAF $> 0,7$) zwischen Holstein und DSN ausgewählt. Diese Varianten können helfen, gemeinsame und unterschiedliche Mutationen zu erkennen, die wahrscheinlich die größten Unterschiede zwischen Holstein und DSN in Bezug auf die Merkmalsleistungen verursachen.

6) Varianten auf dem Y-Chromosom

Es wurden alle verfügbaren 321 Varianten auf dem Y-Chromosom für das Design des DSN200K-Chips ausgewählt. Diese Varianten sind von besonderem Interesse zur Analyse männlicher Individuen.

7) Varianten auf den mitochondrialen Chromosomen

Es wurden alle 278 verfügbaren Varianten auf den mitochondrialen Chromosomen (278 Varianten) ausgewählt. Diese Varianten sind von besonderem Interesse zur Ableitung maternaler Effekte, da mitochondriale Vererbung nur über die mütterliche Linie stattfindet.

8) Varianten zur Abstammungsbestimmung (ICAR & ISAG)

Es wurden 554 Varianten auf Basis des Abstammungspanels der Organisationen ICAR (International Committee for Animal Recording) und ISAG (International Society for Animal Genetics) ausgewählt. Es handelte sich bei diesen um besonders diverse Varianten zwischen Tieren innerhalb einer Rasse als auch zwischen Rassen.

9) Varianten in Haplotypblöcken

Insgesamt wurden 103.801 Varianten ausgewählt, die sich in Haplotypblöcken (HBs) befinden. Ein HB ist eine Region, in der alle enthaltenden Varianten im Kopplungsungleichgewicht zueinanderstehen, was bedeutet, dass es wenig Hinweise auf genetische Rekombinationsereignisse gibt. Innerhalb eines solchen HBs treten unterschiedliche Haplotypen auf. Die HBs wurden in Anlehnung an die Standarddefinition eines HBs aus dem Human Genome Project (Gabriel et al., 2002) mit PLINK v1.9 berechnet. Die wurden als Regionen, die sich innerhalb von Fenstern mit 200 Kilobasenpaaren (kb) und dem unteren Ende des 90 % D-Prime-Konfidenzintervalls $\geq 0,60$ und dem oberen Ende des Konfidenzintervalls $\geq 0,85$ definiert (Purcell et al., 2007). Es wurde je eine Variante innerhalb eines HBs mit einer Länge von > 1 kb ausgewählt. Zudem wurden Varianten, die sich bereits in der Thermo Fisher Scientific Axiom™ Datenbank befinden und die höchste Frequenzdifferenz zwischen Holstein und DSN aufwiesen, priorisiert. Mit dieser Vorgehensweise wurde eine breite Auswahl von Varianten über das gesamte Genom garantiert, sodass noch vorhandene Lücken von Regionen ohne Marker minimiert werden konnten.

10) Varianten zum Füllen von Lücken

Abschließend wurden noch weitere 16 Varianten selektiert um Lücken zwischen bereits ausgewählten Varianten abzudecken, die größer als 250 kb sind.

2.3 Evaluierung des neuen DSN200K-SNP-Chip (APIII)

Genotypisierung

Die Evaluierung der Funktionalität des DSN200K-SNP-Chips erfolgte 1. durch die Überprüfung der technischen Zuverlässigkeit hinsichtlich Genauigkeit und Wiederholbarkeit der Genotypisierung und 2. durch die Anwendbarkeit in der Rasse DSN. Insgesamt wurden 1.201 DSN Tiere mit dem DSN200K-SNP-Chip genotypisiert. Die Genotypisierung der Rinderproben mit dem DSN200K-SNP-Chip wurde vom Institut für Fortpflanzung landwirtschaftlicher Nutztiere (IFN) Schönow e.V. durchgeführt. Aus den Rohdateien wurden die Genotypen von TZ-BER mit der Axiom Analysis Suite Software v5 von Thermo Fisher Scientific unter Befolgung der „Best

Practice“-Richtlinien ausgelesen. Die Genotypen wurden basierend auf der Konfidenz und einer Call Rate von mindestens 95 % gefiltert, sodass nur hoch-qualitative Genotypen in die Analysen aufgenommen und als erfolgreich ausgelesenen definiert wurden (Neumann et al., 2021). Zusätzlich standen noch DSN200K-Genotypen von weiteren 432 DSN Tieren aus einem anderen Projekt zur Verfügung.

Für die Bestimmung der technischen Zuverlässigkeit des DSN200K-SNP-Chips wurde der Grad der Übereinstimmung (=Konkordanz) zwischen den tatsächlichen Sequenzierungsdaten und den aus Genotypisierung mit dem DSN200K-SNP-Chip erhaltenen Daten berechnet. Zur Bestimmung der Wiederholbarkeit wurde der Grad der Übereinstimmung zwischen den Genotypen von Tieren, die mehrfach mit dem DSN200K-SNP-Chip genotypisiert wurden, bestimmt. Die Anwendbarkeit in der Rasse DSN wurde dadurch bestimmt, indem ermittelt wurde, wie viele Varianten vom DSN200K-SNP-Chip informativ waren und mit einer MAF von mehr als 0 % oder 1 % segregieren.

Vereinigung von Genotypen, die mit verschiedenen SNP-Chips erzeugt wurden, durch Imputation

Durch die Verwendung von originalen Genotypisierungsdaten und Imputation von SNPs wurden drei Datensätze mit unterschiedlicher Dichte an Sequenzvarianten geschaffen: 50K, 200K und VGS. Es lagen Genotypen von 1.797 DSN Tieren vom Illumina BovineSNP50 BeadChip, 1.183 DSN Tieren vom DSN200K-SNP-Chip und 304 DSN Tiere aus Sequenzierungen vor. Bei der Imputation werden Varianten eines Referenzpanels mit höherer Dichte an Varianten in Genotypen niedrigerer Dichte statistisch eingefügt.

Vollgenomsequenzinformationen sind die umfänglichsten Genomdaten eines Tieres. Als Referenzpanel für die Imputation zu VGS-Niveau dienten die zuvor in den sequenzierten DSN Tieren gefundenen 16.175.216 Sequenzvarianten. Dieses Referenzpanel wurde zuvor gefiltert, um die Genauigkeit der Imputation zu erhöhen. Es wurden nur Varianten mit einer MAF > 0,01 und einer Call Rate von > 95 % berücksichtigt. Varianten mit einer Mendelian Error Rate von mehr als 5 % (basierend auf gegensätzlichen Homozygoten für die verfügbaren 156 Eltern- und 192 Nachkommenpaare) wurden entfernt. Darüber hinaus wurden Varianten mit einer Quality by depth < 10 und < 3000 wie mit dem Ensembl VEP (McLaren et al., 2016) vorhergesagt aufgrund eines niedrigen Übergangs/Transversions-Verhältnisses verworfen. Phasing und Imputation von Varianten wurde mit Beagle v.5.1 durchgeführt (Browning et al., 2018).

Um die Genauigkeit der Imputation zu überprüfen, wurde der Grad der Übereinstimmung von Sequenzvarianten in 15 DSN Tieren geprüft, die sowohl vollgenomisch sequenziert als auch mit dem DSN200K-SNP-Chip oder dem 50K Illumina BeadChip genotypisiert wurden. Die Sequenzvarianten dieser 15 Tiere waren nicht im SNP-Referenzpanel für die Imputation enthalten (VGS-Daten lagen zu 304 Tieren vor).

Die Daten vom Illumina BovineSNP50 BeadChip wurden ebenfalls auf das Niveau des DSN200K-SNPChips imputiert. Dabei wurden die erfolgreich genotypisierten Varianten vom DSN200K-SNPChip als Referenzpanel (167.271 Varianten mit $MAF > 0,1$) herangezogen. Für die Berechnung der Imputationsgenauigkeit wurden die Varianten von 15 sequenzierten DSN Tieren herangezogen, von denen auch Genotypen sowohl mit dem 50K-Chip als auch vom DSN200K-SNPChip vorlagen.

Vergleich verschiedener SNP-Datensätze in genomweiten Assoziationsstudien

GWAS für verschiedene Merkmale wurden mit allen drei verfügbaren (imputierten) SNP-Datensätzen durchgeführt: 50K (Illumina BovineSNP50 BeadChip), 200K (DSN200K-SNP-Chip) und VGS. GWAS, die mit denselben Tieren, Merkmalen und Modellen, jedoch unterschiedlichen Markerdichten durchgeführt werden, erlauben Aussagen zum Informationsgehalt der verwendeten SNP-Datensätze.

Für die Testung verschiedener SNP-Datensätze wurden Phänotypen erstlaktierender Kühe von sechs Betrieben verwendet. Dabei standen 2.020 Beobachtungen für die 305-d-Leistungen für Milch-, Fett- und Eiweißmenge in kg (M305, F305, E305) zur Verfügung. Diese Merkmale weisen eine moderate bis hohe Erblichkeit auf. Für Gesundheitsindikator-Merkmale mit einer niedrigen bis moderaten Erblichkeit wurden 2.402 Tiere mit Daten für FEQ und SCS-Werten in der ersten Laktation untersucht. Die Anzahl der somatischen Zellen wurde nach der Formel von Ali und Shook (1980) in den SCS umgewandelt: $\log_2(\text{Anzahl der somatischen Zellen}/100.000) + 3$. Für die Abkalbmerkmale KV und TGR mit niedrigen Erblichkeiten lagen 2.606 Beobachtungen vor. Nach Filterung der Genotypen verblieben abhängig vom Merkmal ~37K SNPs für den 50K-Chip, ~125K Varianten für den DSN200K-SNPChip und ~11.6 Millionen Varianten für die VGS-Daten. Die GWAS wurden mit dem Programm GCTA (Yang et al., 2011; Yang et al., 2014) mit der „leaving one chromosome out“ (--mlma-loco) Option durchgeführt. Folgendes statistisches Modell in Matrixschreibweise wurde angewandt:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{S}\mathbf{s} + \mathbf{e}, \quad (1)$$

dabei ist \mathbf{y} der Vektor mit der jeweiligen Beobachtung für M305, F305, E305, FEQ, SCS, KV und TGR; $\boldsymbol{\beta}$ = Vektor der fixen Effekte (Betrieb, Jahr und Saison der ersten Kalbung und lineare Regressionen der Tage in Milch (DIM) und dem Erstkalbealter (EKA) in Monaten für M305, F305 und E305; Betrieb, Jahr und Saison des ersten Testtags, linearer Regressionen der DIM und dem Fettanteil für SCS; Betrieb, Jahr und Saison des ersten Testtags, lineare Regression des EKA für FEQ; Betrieb, Jahr und Saison der Kalbung, sowie Geschlecht des Kalbes für KV und TGR), \mathbf{u} = Vektor der polygenen Effekte mit $\mathbf{u} \sim N(0, \mathbf{G}\sigma_u^2)$, wobei \mathbf{G} die genomische Verwandtschaftsmatrix (VanRaden, 2008) und σ_u^2 die polygene Varianz bezeichnet; \mathbf{s} = Vektor

für Markereffekte; e = Vektor der zufälligen Residuen; und X , Z , S sind die Inzidenzmatrizen für β , u und s . Für die Ermittlung geeigneter Schwellen wurde zuvor die effektive Anzahl unabhängiger Varianten mit der „--indep-pairwise“ Option in PLINK v1.9 (Chang et al., 2015) ermittelt. Dazu wurde in einem Fenster von 5.000 Basenpaaren (bp), welches in jedem Schritt um 500 bp verschoben wurde, von zwei Varianten, die eine Kopplung von $r^2 > 0.5$ aufwiesen, je eine Variante eines Paares entfernt. Im nächsten Schritt wurde eine nach Bonferroni angepasste Signifikanzschwelle (p_{Bonf}) folgendermaßen berechnet: $p_{Bonf} = 0,05/n_{eff}$, mit n_{eff} = effektive Anzahl unabhängiger Varianten. Des Weiteren wurde eine Suggestivschwelle bestimmt: $p_{Sug} = 1/n_{eff}$. Die in der GWAS identifizierten SNPs und Varianten wurden anschließend mit potentiellen Kandidatengen unter Verwendung der Ensembl (release 104) Datenbank assoziiert. Ein Gen wurde als Kandidatengen angesehen, wenn mindestens ein SNP/eine Variante über p_{Sug} in dem Gen oder innerhalb 150 kb auf- oder abwärts des Gens lokalisiert war. Weiterhin wurden genomweite und chromosomenweite SNP-Effekt Korrelationen innerhalb jedes Merkmals für alle zwischen den drei Markerdichten überlappenden SNPs bzw. strukturellen Varianten kalkuliert.

Das Modell (1) wurde zudem genutzt, um für alle Merkmale Erblichkeiten mit den drei unterschiedlichen Markerdichten zu schätzen. Dazu wurde der restricted maximum likelihood Algorithmus (--reml) in GCTA (Yang et al., 2011) verwendet.

Kreuzvalidierung im Rahmen der genomischer Zuchtwertschätzung

Basierend auf den drei verfügbaren Datensätzen (50K, 200K, VGS) für die DSN-Population wurden jeweils Kreuzvalidierungen im Rahmen der gZWS durchgeführt. Insgesamt waren Informationen von 2.420 erstkalbenden DSN-Kühen von zehn verschiedenen Betrieben verfügbar. Die Kühe wurden zwischen 2006 und 2018 geboren. Das EKA lag zwischen 22 bis 46 Monaten (Mittelwert: 27,93 Monate). Für die 305-d-Merkmale wurden nur Kühe mit einer Laktationsdauer von mehr als 250 Tagen berücksichtigt. Für 2.383 Kühe lagen Aufzeichnungen für M305, F305, E305, Milchfett-% (F%), Eiweiß-% (E%) und energiekorrigierte Milchleistung (EKM) vor. Die EKM wurde nach der folgenden Formel von Zumwald et al. (2018) berechnet: $M305/3.14*(0.38*F\% + 146.024*E\% + 0,816)$. Testtags-Aufzeichnungen vom ersten Testtag nach dem Kalben umfassten FEQ und SCS von 2.420 Kühen. Für FEQ und SCS wurde sich auf die schwierige frühe Laktationsperiode konzentriert, da die Messungen in der frühen Laktation zuverlässige Indikatoren auf Herdenebene sind, insbesondere in den großen Herden in Ostdeutschland (Bergk und Swalve, 2011). Daher wurde nur der erste offizielle Testtag nach dem Kalben zwischen 5 und 40 Laktationstagen berücksichtigt. Deskriptive Statistiken für 305-d-Milchleistungsmerkmale, FEQ und SCS sind in **Tabelle 4** aufgeführt. Die Merkmale des Exterieurs wurden in den fünf größten Herden, welche sich in Ostdeutschland befinden, erfasst. Von 1.811 Kühen lagen Exterieurbeurteilungen für 18 lineare Merkmale und vier

Merkmalskomplexe vor. Zu den linearen Merkmalen gehörten: Milchcharakter, Größe, Körpertiefe, Stärke, Beckenneigung, Beckenbreite, Body-Condition-Score (BCS), Hinterbeinwinkel, Klauenwinkel, Sprunggelenk, Hinterbeinstellung, Hintereuterhöhe, Zentralband, Strichplatzierung vorne, Strichplatzierung hinten, Vordereuteraufhängung, Eutertiefe und die Strichlänge. Bis auf die in Zentimetern gemessene Größe wurden alle linearen Merkmale auf einer Skala von 1 bis 9 bewertet. Die Merkmalskomplexe Milchtyp, Körper, Fundament und Euter wurden auf einer Skala von 65 bis 88 Punkten erfasst. Die Beurteilung der Exterieurmerkmale wurde von geschulten Mitarbeitern der deutschen Zuchtverbände vorgenommen. Die deskriptiven Statistiken für die Exterieurmerkmale sind in Tabelle 5 dargestellt.

Tabelle 4: Deskriptive Statistik für 305-d-Produktionsmerkmale und Testtagsmerkmale.

Merkmal ¹	Beobachtungen	Mittelwert	SD ²	Min.	Max.
305-d Produktionsmerkmale					
M305	2.383	6.162,11	1.246,86	2.248,00	10.574,00
EKM	2.383	6.398,44	1.222,40	2.093,07	9.728,75
F305	2.383	256,71	48,61	76,00	412,00
F%	2.383	4,19	0,37	2,16	5,60
E305	2.383	221,16	44,72	77,00	346,00
E%	2.383	3,59	0,22	2,85	4,31
Testtagsmerkmale					
FEQ	2.420	1,18	0,19	0,45	2,52
SCS	2.420	2,61	1,51	-1,06	8,88

¹ M305 = 305-d Milchleistung in kg; E305 = 305-d Eiweißertrag in kg; F305 = 305-d Fettertrag in kg; F% = 305-d Fettprozent; E% = 305-d Eiweißprozent; EKM = 305-d energie-korrigierte Milchleistung; FEQ = Fetteiweißquotient am ersten Testtag; SCS = Somatischer Zellscore am ersten Testtag. ² SD = Standardabweichung (engl.: „standard deviation“)

Tabelle 5: Deskriptive Statistik für lineare Exterieurmerkmale und deren Merkmalskomplexe.

Merkmal	Beobachtungen	Mittelwert	SD ²	Min.	Max.
<u>Merkmalskomplexe</u>					
Milchtyp	1.811	79,01	3,02	68,0	86,0
Körper	1.811	79,21	3,42	65,0	87,0
Fundament	1.811	81,82	3,08	65,0	87,0
Euter	1.811	78,24	3,66	65,0	86,0
<u>Lineare Merkmale</u>					
Milchcharakter	1.811	3,38	1,35	1,0	9,0
Größe (in cm)	1.811	136,83	2,99	126,0	152,0
Körpertiefe	1.811	5,34	1,29	1,0	9,0
Stärke	1.811	6,28	1,13	2,0	9,0
Beckenneigung	1.811	4,79	0,97	1,0	9,0
Beckenbreite	1.811	5,74	1,23	2,0	9,0
BCS ¹	1.811	6,06	1,16	2,0	9,0
Hinterbeinwinkelung	1.811	5,23	0,98	1,0	9,0
Klauenwinkel	1.811	5,04	1,00	1,0	9,0
Sprunggelenk	1.811	5,95	1,38	1,0	9,0
Hinterbeinstellung	1.811	5,86	1,52	1,0	9,0
Hintereuter	1.811	4,25	1,36	1,0	9,0
Zentralband	1.811	4,34	1,37	1,0	8,0
Strichplatzierung vorne	1.811	4,00	1,16	1,0	8,0
Strichplatzierung hinten	1.811	4,96	1,06	1,0	8,0
Vordereuteraufhängung	1.811	5,32	1,32	1,0	9,0
Eutertiefe	1.811	4,39	1,23	1,0	9,0
Strichlänge	1.811	4,95	1,00	2,0	9,0

¹ BCS = Body-Condition-Score; ² SD =Standardabweichung (engl.: „standard deviation“)

Um dieselben Tiere in allen drei Genotypen-Panels (50K, DSN200K und VGS) vorliegen zu haben, wurden Datensätze durch Imputation und Downscaling angepasst. Die Qualitätskontrolle der drei Genotyp-Panels (50K, 200K und VGS) wurde in PLINK (Purcell et al., 2007) in zwei Schritten durchgeführt. In einem ersten Schritt wurden die Genotypen und imputierten VGS-Daten gefiltert, indem nur Varianten und Individuen mit einer Genotypisierungsrate von > 90 %, Varianten mit MAF > 0,05 und Varianten ohne signifikante Abweichung ($p < 10^{-6}$) vom Hardy-Weinberg-Gleichgewicht (HWG) beibehalten wurden. Es wurden nur SNPs auf den Autosomen berücksichtigt. Dieser Schritt führte zu einer Anzahl von 31.539 Varianten für 50K, 125.678 Varianten für 200K und 11.827.995 Varianten für VGS. Anschließend wurden diese Datensätze auf die Tiere mit verfügbaren Phänotypen für die

untersuchten Merkmale (305-d-Produktion, FEQ, SCS, Exterieur) reduziert, und es wurden die gleichen Filterschritte für MAF, Call-Raten und HWG angewendet. Die endgültige Anzahl der Varianten nach der Filterung in Abhängigkeit von den jeweiligen Merkmalen betrug im Durchschnitt ~31.400 für 50K, ~124.300 für DSN200K und ~11,6 Millionen für VGS-Daten. Die 304 sequenzierten Kühe wurden in genomischen Studien für die Entwicklung des DSN200K-SNP-Chip und in der vorliegenden Studie für genomische Vorhersagen berücksichtigt, was mögliche verzerrte Schätzungen oder ein gewisses Maß an Überanpassung implizieren könnte. Zu den drei Genotypen-Panels wurden zufällige 20 x 200K-SNP-Chip Datensätze aus den vorliegenden VGS Daten generiert, um die Auswirkungen des Designs des neuen DSN200K-SNP-Chip auf die Ergebnisse der gZWS zu prüfen. Von den 3.069.815 von Thermo Fisher Scientific als „technisch geeignet“ eingestuft Varianten waren 2.833.148 Varianten aus VGS-Daten nach vorherigen Filterschritten verfügbar. Die `sample()`-Funktion in R (R Core Team, 2022) wurde verwendet, um die zuvor genannten 20 zufälligen Chips aus VGS-Daten zu erzeugen, die dieselbe Anzahl von 125.678 Varianten enthielten, die nach dem ersten Filterungsschritt für den DSN200K übrig geblieben waren. Die gleiche Anzahl von Markern garantierte den Vergleich der genetischen Parameter und der gZW-Genauigkeit zwischen den rassespezifischen DSN200K mit der gleichen Markerdichte, aber ohne Berücksichtigung der rassespezifischen Varianten (im Folgenden als "zufällige 200K-Chips" bezeichnet).

Zur Berechnung genetischer Parameter und für die Durchführung der gZWS wurden vorab genomische Verwandtschaftsmatrizen (**G**) in GCTA (Yang et al., 2011) unter Berücksichtigung aller Kühe mit Genotypen für jedes Merkmal erstellt. Zur Integration der **G**-Matrizen in die BLUPF90-Anwendungen (Aguilar, 2018) wurde die Inverse von $\mathbf{G} + 0,02 \mathbf{I}$ berechnet, wobei **I** eine Identitätsmatrix ist. Zur Schätzung der Varianzkomponenten und Erblichkeiten wurden Einmerkmals-Tiermodelle aufgestellt und der AI-REML-Algorithmus durch Ausführung des AIREMLF90-Programmes aus dem BLUPF90-Softwarepaket (Misztal et al., 2014) angewendet. Das statistische Modell, das den gesamten Datensatz für jedes Merkmal berücksichtigte, lautete:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{g} + \mathbf{e}, \quad (2)$$

wobei \mathbf{y} = Vektor der Phänotypen für die 305-d-Produktionsmerkmale, FEQ, SCS und Exterieurmerkmale; \mathbf{b} = Vektor der festen Effekte, einschließlich der folgenden festen Effekte für die 305-d-Produktionsmerkmale, FEQ und SCS: Herde, Abkalbejahr, Abkalbesaison, EKA und DIM als lineare Regression auf die Laktationslänge für die unvollständigen 305-d-Beobachtungen und auf die Tage in der Laktation nach dem Abkalben für die Testtagsdatensätze; und für die Exterieurmerkmale: Betrieb, Klassifizierer, Jahr der Bewertung des Exterieurmerkmals, Laktationsstadium und Erstkalbealter; \mathbf{g} = Vektor der additiven genetischen Effekte mit $\sim N(0, \mathbf{G}\sigma_g^2)$, wobei **G** die genomische Beziehungsmatrix und σ_g^2 die

genetische Varianz bezeichnet; e = Residualeffekte mit $\sim N(0, I\sigma_e^2)$, und σ_e^2 die Restvarianz; \mathbf{X} und \mathbf{Z} = Design-Matrizen für b und g , und \mathbf{I} = Identitätsmatrix.

Kreuzvalidierung und Berechnung der gZW-Genauigkeiten

Für die Berechnung der gZW-Genauigkeiten wurde eine wiederholte Random-Sub-Sampling-Validierung (RSV), d. h. eine Monte-Carlo-Kreuzvalidierung für jedes Merkmal separat mit jedem Genotypdatensatz, durchgeführt (Picard und Cook, 1984). Zu diesem Zweck wurden die Datensätze geteilt in jeweils einen Validierungssatz (20 % der Kühe) mit maskierten Phänotypen und einen Trainingsatz (80 % der Kühe). Die Anzahl der Tiere im Validierungssatz betrug 477 für 305-d-Produktionsmerkmale, 484 für FEQ und SCS und 363 für Exterieurmerkmale. Die Kühe wurden nach dem Zufallsprinzip mit zehn Wiederholungen ausgewählt. In jedem Validierungsschritt wurden die Zuchtwerte der Kühe in dem Validierungssatz, auf der Grundlage der genomischen Beziehungsinformationen und der Phänotypen für die Kühe im Trainingsatz geschätzt. Die Genauigkeit wurde definiert als die Korrelation zwischen den gZWs und den bekannten Phänotypen, abzüglich der entsprechenden festen Effekten (d.h. $y - \mathbf{X}b$), für die Kühe im Validierungsset. Anschließend wurde die mittlere Genauigkeit unter Berücksichtigung der 10 Wiederholungen für jedes Merkmal für 50K-, DSN200K-, zufällige 200K- (20 zufällige Datensätze) und VGS-Datensätze berechnet.

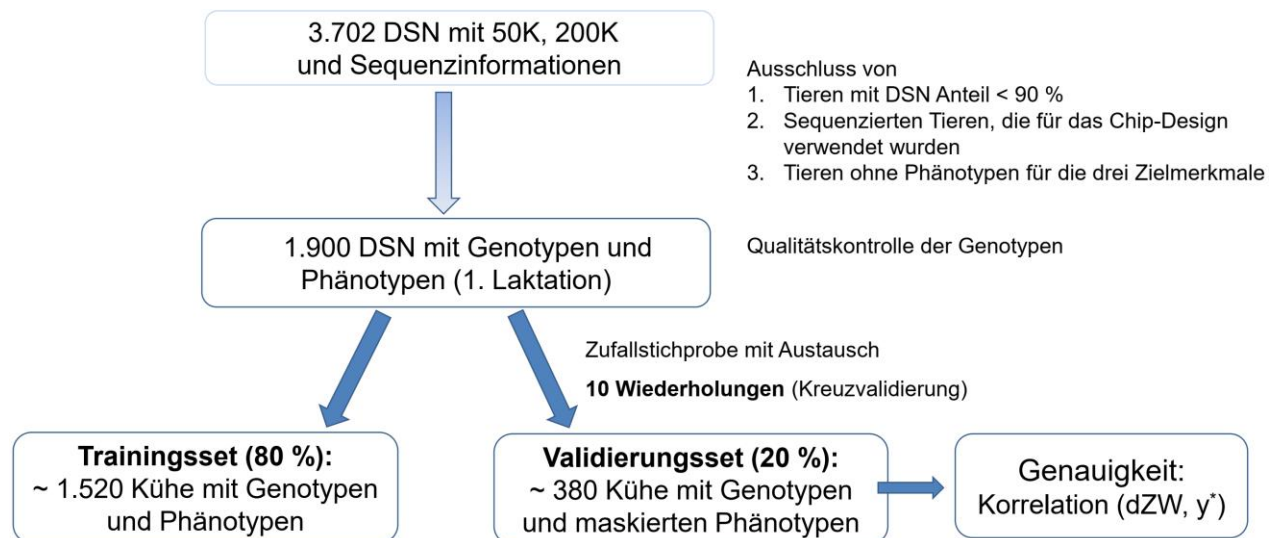


Abbildung 5: Exemplarische Darstellung der Vorgehensweise bei dem Vergleich von 50K Genotypen, 200K Genotypen und Vollgenomsequenzdaten im Rahmen einer genomischen Zuchtwertschätzung innerhalb der Rasse Deutsches Schwarzbuntes Niederungsriind (DSN). dZW = direkte Zuchtwerte; y^* = Phänotypen, abzüglich der festen Effekte ($y^* = y - \mathbf{X}b$).

2.4 Ermittlung von Maßzahlen genetischer Diversität (APIV)

Durch die TZ-BER sollten auf Basis der Sequenzdaten die SNP-Marker bestimmt werden, die notwendig sind um den Heterozygotiegrad als Kenngröße der genetischen Diversität möglichst genau zu bestimmen. Durch die TZ-GI sollte ein Monitoringverfahren getestet werden, dass eine Steigerung der Zuchtwerte vorsieht, doch dabei den Erhalt der Diversität innerhalb der Rasse DSN berücksichtigt.

Phylogenetische Untersuchungen und Nukleotid-Diversitätsanalysen

Durch die vorhergegangenen Arbeitsschritte standen für die folgenden Arbeiten Varianten auf VGS-Level von 304 DSN, 15 ANG und 15 RDN (generiert in **APV**) zur Verfügung. Weitere Sequenzvarianten von 1.394 Rindern aus 66 weiteren *Bos taurus* Rassen, sowie eines Auerochsen (*Bos primigenius*) konnten durch Zugriff auf die Daten des 1000-Bullen-Genom-Projekts (Run 9) (Hayes and Daetwyler, 2018) herangezogen werden. Zur Sicherung der Datenqualität sollte mindestens eine 8-fache Sequenzabdeckung und mindestens fünf Tiere je Rasse vorliegen. Für Rassen mit mehr als 30 Tieren wurden 30 Tiere nach dem Zufallsprinzip ausgewählt. Von Tieren mit der Rassekennzeichnung „Holstein Friesian“ wurden 30 Tiere je Land (Deutschland, Dänemark, Niederlande und U.S.A) randomisiert ausgewählt. Die Sequenzvarianten der DSN Tiere aus diesem Projekt und der DSN Tiere aus dem 1000-Bullen-Genom-Projekt wurden mit dem Programm BCFtools v1.9 (Danecek et al., 2021) zusammengetragen. Insgesamt wurden 74.834.092 biallelische Varianten (68.486.380 SNPs und 6.347.712 INDELS), die in der Tranche 99 % (aus der vom 1000-Bullen-Genom-Projekt durchgeführten Varianten-Rekalibrierung) mit einer Genotypisierungsrate ≥ 90 % vorkommen, für die Analyse berücksichtigt.

Im nächsten Schritt wurde ein genomweiter phylogenetischer Baum für alle *Bos taurus* Autosomen (BTA) erstellt. Die 74.834.092 Sequenzvarianten wurden zunächst mit PLINK v1.9 (Chang et al., 2015) unter Verwendung einer Fenstergröße von 50 bp und einer Schrittweite von 5 SNPs von allen Varianten mit hoher Kopplung ($r^2 > 0,6$) bereinigt. Mit den daraus resultierenden 21.426.106 Varianten wurde der Manhattan-Abstand zwischen den Tieren berechnet und der in der Biotit-Bibliothek in Python implementierte UPGMA-Algorithmus zum Clustering verwendet. Der Baum wurde mit den verfügbaren Daten eines Auerochsen, der die Outgroup bildete, verwurzelt. Der phylogenetische Baum wurde mit iTOL v6 (Letunic and Bork, 2019) visualisiert. Die Zweige des Baums stellen die Position der Mehrheit der Tiere jeder Rasse dar. Die Nukleotid-Diversität (π) wurde pro 10 kb Fenster (π_{window}) für alle BTAs unter Verwendung von 74.834.092 Sequenzvarianten und dem scikit-allele Paket in Python wie folgt berechnet:

$$\pi_{\text{window}} = \frac{\frac{n}{n-1} \sum_{ij} p_i p_j \pi_{ij}}{10 \text{ kb}}, \quad (3)$$

mit der Häufigkeit des i -ten und j -ten Haplotypen p_i ; π_{ij} ist die Anzahl der Nukleotidunterschiede pro Stelle zwischen dem i -ten und j -ten Haplotypen und n die Anzahl der Haplotypen. Die durchschnittliche Nukleotid-Diversität pro Chromosom (π_{chrom}) wurde als Mittelwert über alle Fenster der jeweiligen Chromosomen und die gesamte Nukleotid-Diversität (π_{total}) als Mittelwert über alle BTAs berechnet.

Zusätzlich wurden Verwandtschaftsanalysen unter Verwendung des Fixationsindex (F_{ST}) zwischen DSN und 68 anderen Rassen aus dem 1000-Bullen-Genom-Projekt durchgeführt. Auch wurde die Heterozygotie als Metrik für die genomische Diversität berechnet. Inzucht wurde als Überschuss an Homozygotie (F_{Hom}) und genomische Inzucht (F_{RoH}) durch Runs of Homozygosity (RoHs) gemessen.

Anwendung der „Optimum genetic contribution“-Methode

In **APIV** lag der Fokus von TZ-GI auf der Anwendung der zuvor entwickelten Zuchtwertschätzung auf Basis aller DSN Bullen und Kühe, für die Pedigree-, als auch Genotypeninformationen und Phänotypdaten für das Merkmal 305-d Produktionsleistung vorlagen. Die daraus resultierenden Zuchtwerte wurden für verschiedene Szenarien in eine „optimum genetic contribution“ (OGC) Anwendung implementiert. Die OGC maximiert den genetischen Fortschritt bei gleichzeitiger Beschränkung der Inzuchtwerte bzw. den Beziehungen zwischen den Selektionskandidaten.

Insgesamt lagen Phänotypen zur 305-d-Milchleistung (in kg) der ersten Laktation von 8.373 DSN-Kühen vor. Zudem konnte auf Genotypeninformationen von dem Illumina 50K und dem DSN200K-SNP-Chip für 2.797 DSN, sowie Pedigreeinformationen von etwa 70.000 Tieren zurückgegriffen werden. Anhand dieser Informationen wurde mit AIREMLF90 aus der BLUPF90 Programmfamilie (Misztal et al., 2014) eine modifizierte Verwandtschaftsmatrix **H** erstellt, in der genomische errechnete Verwandtschaftsverhältnisse (**G**-Matrix) und Pedigree-basierte Beziehungsinformationen (**A**-Matrix) vereint werden können. Die Schätzungen der Varianzkomponenten und Erblichkeiten wurden abermals mit einem Einmerkmals-Tiermodell in AIREMLF90 durchgeführt. Das statistische Modell für das Merkmal 305-d-Produktionsleistung entspricht dem oben beschriebenen Modell (2).

Für die folgende Analyse wurden nur Kühe berücksichtigt, die auf Basis vorliegender Stammdaten als lebend deklariert sind ($n = 1.189$). Bullen, die zur künstlichen Besamung zur Verfügung stehen ($n = 89$) verblieben in dem Datensatz. Das OGC-Konzept, das in dem Programm GENCONT (Meuwissen, 2002) implementiert ist, wurde angewandt um Anpaarungsempfehlungen zu erhalten, bei denen unter vorgegebenen Bedingungen ein maximaler Zuchtfortschritt in dem Leistungsmerkmal 305-d Milchleistung in kg, bei gleichzeitigem Erhalt der Diversität in der Rasse DSN realisiert wird.

Die Methode impliziert eine Maximierung des durchschnittlichen genetischen Verdienstes der Selektionskandidaten, $c'u$, durch Einschränkung der durchschnittlichen Beziehung $c'Ac$ innerhalb der Population, wobei c ein Vektor der Anteile in der nächsten Generation jedes ausgewählten Tieres ist, u ein Vektor der geschätzten Zuchtwerte der Selektionskandidaten und A die numerische Beziehungsmatrix zwischen den Selektionskandidaten (Meuwissen, 2002) ist. Die Pedigreedatei, welche für das für das GENCONT-Programm obligatorisch ist, umfasste 67.258 Tiere. In einer Parameterdatei für GENCONT können Bedingungen für ein Zuchtprogramm vorgegeben werden. Hierbei wurde die Anzahl an Kühe auf 30 beschränkt, um möglichst geeignete Bullenmütter für die kommende Generation zu identifizieren. Der maximale genetische Beitrag der einzelnen Bullen wurde auf 20 % begrenzt. In verschiedenen Szenarien wurden die Beschränkungen für die durchschnittliche Verwandtschaft zwischen den künftigen Nachkommen zwischen 0,02 und 0,10 variiert. In der Auswertung wurde die durchschnittliche Verwandtschaft zwischen den Nachkommen und der genetische Fortschritt in Beziehung gesetzt.

2.5 Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen (APV)

In **APV** sollte evaluiert werden, ob der für DSN entwickelte SNP-Chip auch die wesentlichen kausalen Varianten in anderen Milchrinder- oder Zweinutzungspopulationen abbildet und für züchterische Entscheidungen und Diversitätsstudien genutzt werden kann. Hierzu wurden 287 Rinder aus acht verschiedenen einheimischen Rinderrassen (123 Rotes Höhenvieh, 42 RDN, 39 ANG, 24 Pinzgauer, 20 HOL, 14 Original Braunvieh, 10 Gelbvieh, 2 Hinterwälder), einer bedrohten englischen Rinderrasse (8 White Park) und einer *Bos indicus* Rasse (5 Butana) mit dem DSN200K-SNP-Chip genotypisiert. Die Genotypisierung sowie das Auslesen und Filtern der Rohdaten für diese Tiere erfolgte ebenfalls im IFN Schönau e.V.

Sequenzierungsanalyse von Tieren der Rassen Angler und Rotbunt-Doppelnutzung

Für die Sequenzierung von 15 ANG und 15 RDN Tieren wurden pro Bulle zwei Besamungsportionen von der RSH eG geordert, die für die DNA-Gewinnung verwendet wurden. Zur Sequenzierung wurden Bullen ausgewählt, welche eine möglichst geringe Verwandtschaft zueinander aufwiesen. Dazu wurde auf Basis von Internetrecherchen (Bullenkataloge, bulliweb) ein Pedigree erstellt und Tiere mit hohem Verwandtschaftsgrad aussortiert. Damit ist es möglich, sogenannte Schlüsselerber zu identifizieren, auf welche ursächliche Mutationen in der Population zurückzuführen sind. Die Sequenzierung wurde wie bei den DSN Tieren durch die Firma Novogene durchgeführt und auch die gesamte Pipeline bis zur Identifizierung der Sequenzvarianten war äquivalent zu der Pipeline, die für die DSN Tiere benutzt und in **API** erläutert wurde.

2.6 Transfer der Ergebnisse in die Praxis (APVI)

Ein Transfer der Ergebnisse in die Praxis erfolgte in 2020/2021 aufgrund der Corona-Pandemie vor allem über Online-Veranstaltungen und ab 2022 auf verschiedensten nationalen und internationalen Veranstaltungen in Präsenz. Eine detaillierte Auflistung von Beiträgen auf wissenschaftlichen Tagungen findet sich im Kapitel 3.6. Im Rahmen des Fachgesprächs „Erhaltung/Biodiversität in kleinen Nutztierassen“ am 26.10.2021, welches von TZ-BER organisiert und ausgetragen wurde, waren Teilnehmer aus der Wissenschaft und Praxis (u.a. Landwirte, Züchter) beteiligt, mit welchen die ersten Ergebnisse des Projektes in großer Runde diskutiert werden konnten. Am 27.10.2021 hat TZ-BER die Ergebnisse DSN-Züchtern und beteiligten Mitarbeitern der RBB in den Räumlichkeiten der RBB vorgestellt. Ein besonderer Erfolg für das Projekt war auch die Vorstellung durch TZ-GI auf der Plenartagung im Rahmen der Jahrestagung der Deutschen Gesellschaft für Züchtungskunde (DGfZ) 2022, sowie die Präsentation der Ergebnisse im Rahmen von zwei Vorträgen von TZ-BER und einem Vortrag von TZ-GI auf der anschließenden Fachtagung. Wie zuvor aufgeführt, wurde der Transfer der Ergebnisse in die Praxis im Rahmen zahlreicher Veranstaltungen zwischen 2020 und 2022 sehr umfangreich betrieben, so dass von einem zusätzlichen Workshop zum Ende des Projektes abgesehen wurde.

3. Ergebnisse

In den folgenden Absätzen werden die wesentlichen Ergebnisse der Arbeitspakete **API** bis **APVI** aufgeführt.

3.1 Aufbau einer Lernstichprobe: Sequenzierung, Aufbereitung und Imputation (API)

Wie die Angaben in dem „Material und Methoden“-Teil für **API** bereits zeigen, konnte für die Rasse DSN erfolgreich eine informative Lernstichprobe aufgebaut werden.

Insgesamt 214 DSN Tiere, davon 30 Bullen, wurden genomweit sequenziert. Um die Diversität der Rasse möglichst genau abbilden zu können, wurden Schlüsseltiere ausgewählt, die das gesamte Pedigree repräsentieren. Die durchschnittliche Genomabdeckung lag über dem Faktor 18. Die Qualitätskontrolle der Firma Novogene nach der Sequenzierung zeigte, dass die Sequenzdaten eine Qualität von 99,9 % aufweisen. Es konnte gezeigt werden, dass alle sequenzierten Tiere einen DSN-Anteil > 90 % aufwiesen. Unter Verwendung unserer eigenen Analysepipeline, die sich an das 1000-Bullen-Genom-Projekt orientiert (Abbildung 6), wurden 25.938.563 Sequenzvarianten für die Rasse DSN detektiert. Nach Filterung wies der Datensatz insgesamt 20.567.619 Varianten auf (18,5 Millionen SNPs und 2,0 Millionen INDELS). Die Übereinstimmung zwischen den von uns ermittelten Sequenzvarianten und denen, die vom 1000-Bullen-Genom-Projekt für die Rasse DSN bestimmt wurden, war mit durchschnittlich 99,996 % außerordentlich hoch. Dies zeigte, dass die erarbeitete Pipeline zur Aufarbeitung der Rohsequenzdaten und der anschließenden Identifikation von Sequenzvarianten sehr gut geeignet ist.

Die insgesamt 304 sequenzierten DSN dienten als Referenzpanel zur Imputation von ca. 1.800 DSN, zu denen Genotypen vom Illumina BovineSNP50 BeadChip vorlagen. Die Genauigkeit der Imputation lag bei 96,69 % und ist damit höher als bis dato in anderen publizierten Studien beim Rind (z.B. 94 % (Iso-Touru et al., 2016), 95 % (Pausch et al., 2017)).

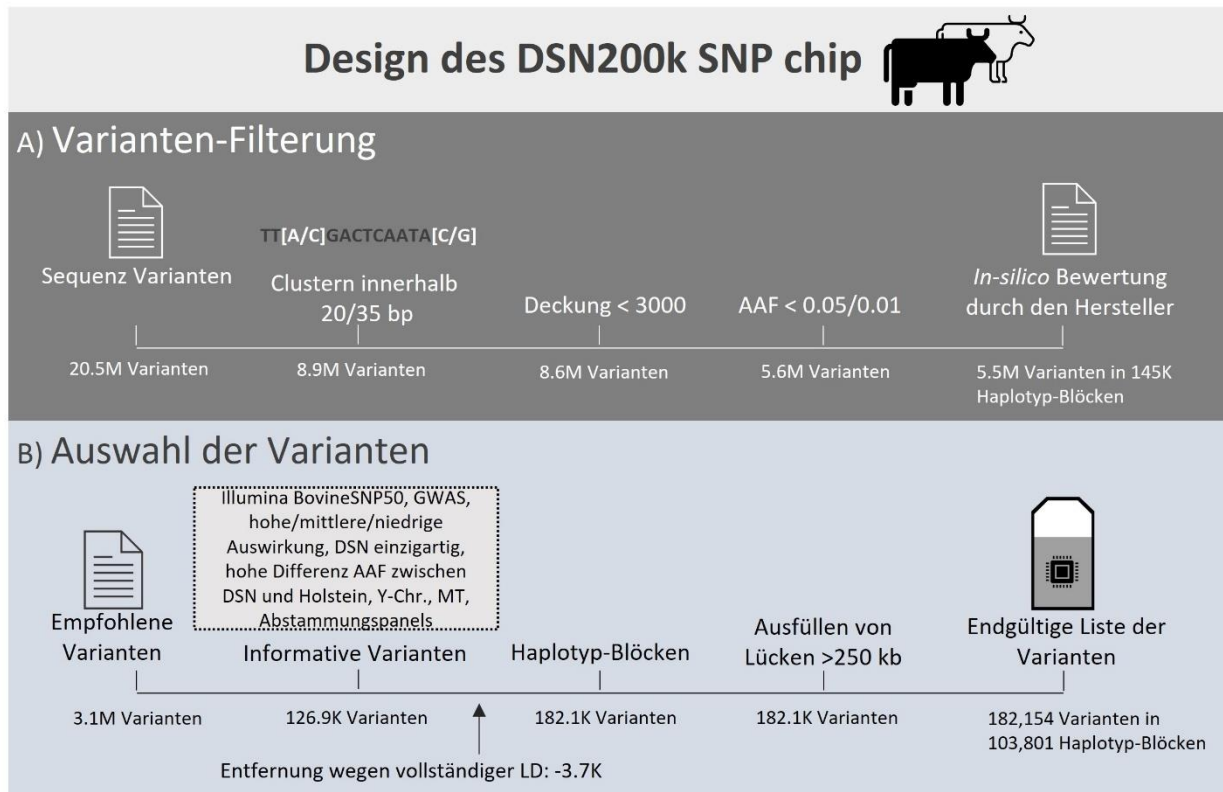


Abbildung 6: Pipeline für das DSN200K-SNP-Chip Design. A) Filterung des initialen Datensatzes mit 20,5 Millionen Varianten auf eine Scoring-Liste von 5,5 Millionen Varianten, von denen 3,1 Millionen Varianten von Thermo Fisher Scientific empfohlen wurden. B) Nach Selektion der Varianten auf Basis mehrerer Kategorien verblieben 182.154 Varianten.

3.2 Genotypisierung mit dem neuen DSN200K-SNP-Chip und Performance-Analysen der Genotypdaten (APII)

Unter Berücksichtigung der beschriebenen Kriterien für die Auswahl von Sequenzvarianten wurde der DSN-spezifische SNP-Chip „DSN200K-SNP-Chip“ bei der Firma Thermo Fisher Scientific entwickelt und produziert. Insgesamt wurden 182.154 (173.569 SNPs und 8.585 INDELS) für den DSN200K-SNP-Chip ausgewählt. Aufgrund technischer Grenzen der SNP-Chip-Technologie von Thermo Fisher Scientific beanspruchten G/C- oder A/T-Varianten doppelten Platz auf dem Chip, was die Kapazität des Chips für abzubildende Sequenzvarianten reduzierte. Die verbleibenden Positionen auf dem Chip wurden für technische Replikate von wichtigen Varianten oder Standard- Kontrollvarianten verwendet, die für Qualitätsberechnungen, Geschlechtsbestimmung und Hintergrundsignaldetektion verwendet werden.

In diesem Projekt wurden 1.201 DSN Tiere mit dem DSN200K-SNP-Chip genotypisiert. Nach Qualitätskontrolle der Genotypen, standen 1.162 DSN Tiere für die Analysen zur Verfügung (39 DSN Tiere sind durch die Qualitätskontrolle gefallen). Aus einem anderen Projekt standen weitere 432 DSN Tiere zur Verfügung, die ebenfalls mit dem DSN200K-SNP-Chip genotypisiert wurden, so dass die Gesamtzahl der mit dem DSN200K-SNP-Chip genotypisierten DSN Tiere 1.594 betrug. Die Genotypisierungen zeigten, dass von den 182.154 (173.569 SNPs und 8.585

INDELs) auf dem DSN200K-SNP-Chip vorhandenen Sequenzvarianten 167.271 Varianten erfolgreich in der Rasse DSN abgebildet wurden. Das sind 91,8 % aller Sequenzvarianten auf dem Chip. Die 167.271 funktionierenden Sequenzvarianten setzen sich aus 160.290 SNPs (92,4 % aller SNPs) und 6.981 INDELs (81,3 % aller INDELs) zusammen. 158.382 funktionierende Sequenzvarianten besaßen eine $MAF > 0,01$. Mit dieser Anzahl abgebildeter Varianten kann der DSN200K-SNP-Chip erfolgreich in der Rasse DSN eingesetzt werden.

Die Übereinstimmung der Genotypen zwischen den technischen Replikaten auf dem DSN200K-SNP-Chip war $99,83 \% \pm 0,19 \%$ Standardabweichung (SD) und die Übereinstimmung von erhaltenen Genotypen und vorliegenden Sequenzinformationen von 57 Tieren betrug $99,60 \% \pm 0,41 \%$ SD. Dies sind beides sehr hohe Übereinstimmungen, die das technische Funktionieren und die sehr hohe Genauigkeit des Auslesens der Varianten vom DSN200K-SNP-Chip bestätigen.

Die insgesamt 15.591 Varianten, die bei den DSN Tieren nicht funktionierten, verursachen keine großen Lücken im Genom, auch keine Lücken im Sinne von genetischer Distanz, die durch das Kopplungsungleichgewicht definiert ist. Ein durchschnittlicher D-Prime von $0,97 \pm 0,07$ und ein r^2 von $0,73 \pm 0,26$, wie von PLINK v1.9 (Chang et al., 2015) berechnet, wurden zwischen ausgefallenen und funktionierenden Varianten für DSN beobachtet. Für die erfolgreich ausgelesenen Sequenzvarianten in den einzelnen Selektionskategorien, die zur Entwicklung des DSN200K-SNP-Chips beitrugen, wurden folgende Beobachtungen gemacht: Durch die Varianten werden 97.187 HB (von 103.801 auf dem Chip) markiert. In den meisten HBs liegen Varianten, die einen hohen, mittleren oder niedrigen prognostizierten funktionellen Einfluss auf das Genprodukt haben. Von den in den DSN Tieren identifizierten 131.341 Varianten mit hohem oder mittlerem funktionellem Einfluss auf das Genprodukt befinden sich 112.389 Varianten innerhalb eines HBs. Von den in DSN Tieren identifizierten 455.486 Varianten mit geringem funktionellem Einfluss wurden 400.561 Varianten durch gekoppelte SNPs, die einer anderen Auswahlkategorie angehören, auf dem Chip abgedeckt. Nur wenige Varianten mit hohem, mittlerem oder niedrigem funktionellem Einfluss wurden nicht abgerufen. Zum Beispiel fielen von 38.198 DSN-spezifischen Varianten nur 35 mit hohem oder mittlerem Einfluss und 78 mit niedrigem Einfluss raus. Von 34.039 informativen Varianten für die Rasse DSN, die von dem Illumina BovineSNP50-Chip übernommen wurden, wurden elf mit einem hohen oder mittleren Einfluss und 45 mit einem geringen Einfluss nicht detektiert. Von 2.071 Varianten, die aus GWAS mit Gesundheits- und Resistenzmerkmalen (z.B. Endoparasitenresistenz) stammen, fiel nur eine Variante mit niedrigem Einfluss aus. Von den 554 Varianten aus Abstammungspanelen fielen zwei Varianten mit niedrigem Einfluss durch. Auf Basis dieser Analysen ist anzunehmen, dass die 167.271 Varianten, die in der Rasse DSN mit dem DSN200K-SNP-Chip identifiziert wurden, den Großteil der DSN Diversität durch Kopplung zwischen den Varianten repräsentieren können.

3.3 Evaluierung des DSN200K-SNP-Chip (APIII)

Imputation von SNPs in Genotypen-Datensätze geringerer Dichte

Bevor eine Erprobung des neuen DSN200K-SNP-Chip anhand von GWAS und Zuchtwertschätzverfahren umgesetzt werden konnte, wurden alle vorliegenden Genotypdatensätze (50K, DSN200K, VGS) durch Imputation und Downscaling angeglichen. Um die Imputationsverfahren zu bewerten wurden 15 sequenzierte DSN sowohl mit dem DSN200K, als auch mit dem konventionellen 50K Chip genotypisiert. Zur Bildung der entsprechenden Referenzpanele wurden diese 15 Tiere ausgelassen, um Verzerrungen bei den Ermittlungen der Imputationsgenauigkeiten vorzubeugen. Die Imputationsgenauigkeiten auf Sequenzebene wurden im Durchschnitt auf 96,50 % (0,58 % SD) für die Illumina BovineSNP50 BeadChip Daten und 98,58 % (0,21 % SD) für den DSN200K-SNP-Chip geschätzt. Ein Vergleich mit anderen Imputationsgenauigkeiten aus der Literatur zeigt, dass die Qualität der Imputation generell hoch war (z. B. 94 % bei Iso-Touru et al., 2016, 95 % bei Pausch et al., 2017). Der im Rahmen des Projekts entwickelte DSN200K-SNP-Chip erzielte höhere Imputationsgenauigkeiten im Vergleich zu dem kommerziellen Illumina BovineSNP50 Chip, was aufgrund der höheren Markerdichte zu erwarten war. Bei der Imputation von 50K Illumina auf DSN200K konnte eine Genauigkeit von 99,99 % (0,07 % SD) erzielt werden. Die Daten von 3.703 imputierten DSN-Genotypen wurden in GWAS für Milchproduktionsmerkmale (TZ-BER), Gesundheitsindikatoren (TZ-GI) sowie in der gZWS (TZ-GI) verwendet.

Vergleichende genomweite Assoziationsstudien und Heritabilitäten für ausgewählte Merkmale auf Basis von drei verschiedenen Markerdichten (50K, 200K, VGS)

In den GWAS für verschiedene Merkmale, in denen der Effekt der drei Markerdichten geprüft wurde, wurden unterschiedlich viele Sequenzvarianten gefunden, die mit dem untersuchten Merkmal signifikant assoziiert waren. Die Anzahl assoziierter Varianten und annotierter Gene für die drei unterschiedlichen Markerdichten je Merkmal sind in Tabelle 6 dargestellt. Für M305, F305 und E305 wurden 23 SNPs für 50K, 84 SNPs für 200K und 950 Varianten für VGS mit einem p -Wert $< p_{\text{Bonf}}/p_{\text{Sug}}$ identifiziert. Bei den GWAS Ergebnissen zum Merkmal M305 (Abbildung 7) stimmten 14,8 % (4/27) der Gene (*KLF3*, *LIMCH1*, *NSUN7*, *RALGAPA2*) in den Analysen mit 50K und 200K Sequenzvarianten überein. Die Gene mit der größten Anzahl von signifikanten Varianten pro Gen für M305 waren *MGST1* (BTA 5), *ADGRL3* (BTA 6) und *GNAL* (BTA 24).

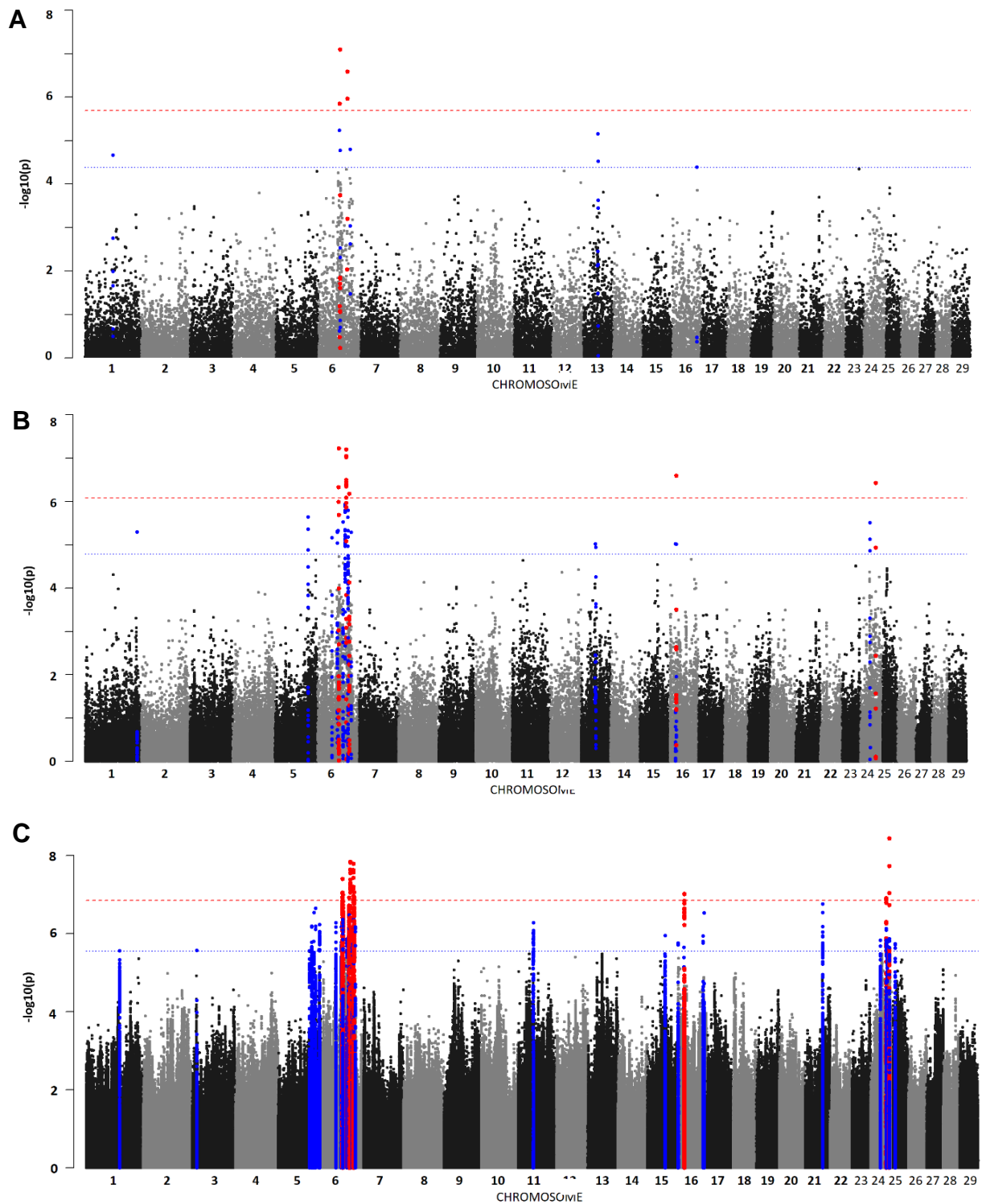


Abbildung 7: Manhattan-Plots für die 305-d Milchleistung in der ersten Laktation; A) 50K Illumina BovineSNP50 BeadChip, B) neu entwickelter DSN200K-SNP-Chip und C) Vollgenomsequenzvarianten; Marker oberhalb von p_{Bonf} (rote Linie) sind rot hervorgehoben; Marker oberhalb von p_{Sug} (blaue Linie) sind blau hervorgehoben; Marker in einem Abstand von 125 Kilobasen vor und nach den assoziierten Markern sind ebenfalls mit den entsprechenden Farben markiert.

Für SCS wurden 62 Varianten mit VGS identifiziert, aber kein SNP erreichte p_{Sig} für die 50K und 200K Chips. Für die Kalbeeigenschaften KV und TGR wurden 11 Gene mit 50K und 23 Gene mit 200K detektiert. Darunter waren sieben gemeinsame Gene, die mit beiden Genotypisierungs-Chips annotiert wurden. Die genomweiten und chromosomenweiten SNP-Effekt-Korrelationen zwischen 50K und 200K betragen $\geq 0,99$ ($p \leq 0,001$) für alle Merkmale und reichten von 0,62 bis 0,81 ($p \leq 0,001$) zwischen VGS mit 50K und 200K.

Die geschätzten Heritabilitäten waren für die drei Markerdichten recht ähnlich (Tabelle 6). Bis auf die geschätzte Erblichkeit für das Merkmal FEQ wurden die Heritabilitäten mit den VGS-Daten für alle anderen Merkmale am höchsten geschätzt. Sowohl bei den Merkmalen mit moderaten bis hohen, als auch bei den mittleren bis geringen Erblichkeiten sind mal die Schätzungen mit 50K Markerdichte am höchsten, mal die mit dem DSN200K-SNPChip.

Tabelle 6: Anzahl assoziierter Marker (SNPs/Sequenzvarianten [SVs]) und Anzahl annotierter Gene in den GWAS und Heritabilitäten (h^2) mit Standardfehlern (SE) für die drei unterschiedlichen Markerdichten.

Merkmal	50K		200K		VGS	
	Anzahl SNPs/Gene	h^2	Anzahl SVs/Gene	h^2	Anzahl SVs/Gene	h^2
M305	11/11	0,40 (0,04)	64/20	0,39 (0,04)	724/48	0,41 (0,04)
F305	8/10	0,33 (0,04)	11/8	0,33 (0,04)	115/20	0,34 (0,04)
E305	4/6	0,34 (0,04)	9/7	0,33 (0,04)	111/11	0,35 (0,04)
FEQ	4/4	0,13 (0,04)	10/6	0,14 (0,04)	52/12	0,13 (0,04)
SCS	0/0	0,11 (0,04)	0/0	0,12 (0,03)	62/2	0,13 (0,04)
KV	0/0	0,03 (0,02)	1/1	0,03 (0,02)	5/2	0,04 (0,02)
TGR	1/1	0,04 (0,02)	2/1	0,05 (0,02)	10/8	0,05 (0,02)

M305/F305/E305 = 305-d Laktationsertrag für Milch, Milchfett und Milcheiweiß; FEQ = Fett-Eiweiß Quotient; SCS = somatischer Zellscore; KV = Kalbeverlauf; TGR = Totgeburtenrate.

In Übereinstimmung mit den Ergebnissen von Wu et al. (2015) wurden mehr potentielle Kandidatengene durch die Verwendung des DSN200K-SNP-Chips und der VGS-Daten im Vergleich zu 50K erkannt. Dies weist darauf hin, dass funktionale Marker, die in der Rasse DSN eine Rolle spielen, auf den kommerziellen Chips fehlen und auch nicht durch gekoppelte Marker repräsentiert werden. Dieses „Fehlen“ wird auch als Ascertainment Bias bezeichnet, der dadurch zustande kommt, dass die Rasse DSN bei der Auswahl der SNPs auf dem kommerziellen Chip nicht berücksichtigt wurde. Außerdem besteht mit hoher Wahrscheinlichkeit ein stärkeres Kopplungsungleichgewicht zwischen Markern und den Regionen eines quantitativen Merkmals (engl.: „Quantitative Trait Loci“) bei höherer Markerdichte. Die SNP-Effekt-Korrelationen zwischen den GWAS mit 50K und 200K lagen nahe bei eins, waren aber zwischen VGS und den beiden Genotypisierungs-Chips nur mäßig bis hoch. Daher sind VGS besser geeignet, rassespezifische Varianten zu erkennen, die sich auf Merkmale auswirken, die

bei kleinen, gefährdeten Rinderrassen von Interesse sind, und die in genomische Vorhersagemodelle integriert werden können, um die Vorhersagegenauigkeit zu verbessern. In einer weiteren Studie wurde die Güte des DSN200K-SNP-Chips im Vergleich zu randomisierten Chips derselben Dichte (200K) am Beispiel von GWAS mit Milchleistungsdaten verglichen (Korkuć et al., 2021). Damit sollte untersucht werden, ob eine zufällige Auswahl der Varianten für einen 200K Chip genauso gut gewesen wäre wie die Auswahl nach den 10 definierten Kriterien, die in APII aufgelistet wurden. Hierfür wurde GWAS mit 2.099 DSN-Kühen und den 305-d Produktionsmerkmalen mit dem Statistikprogramm R durchgeführt. Die randomisierten Datensätze, die randomisierte SNP-Chips darstellten, wurden aus denselben 3,3 Millionen Sequenzvarianten zusammengestellt, die schon für den DSN200K zur Verfügung standen. Für die GWAS wurden aber nur 131,273 Varianten mit einer MAF $\geq 5\%$ verwendet. Des Weiteren wurde bestimmt, welche der mit Milchleistungsdaten assoziierten Varianten ($p > 0,1$) aus vorheriger GWAS stammen, einen funktionellen Einfluss haben oder DSN-spezifisch sind. Insgesamt konnte beobachtet werden, dass mit dem DSN200K-SNP-Chip 2,7-mal mehr assoziierte Varianten identifiziert wurden als mit einem randomisierten 200K Chips, obwohl die gleiche Anzahl von Varianten für die Analysen benutzt wurde (Tabelle 7). Wegen der Zufallsauswahl ist auch die Zahl der Varianten mit MAF $\geq 5\%$, die aus vorheriger GWAS stammen, funktionell oder DSN-spezifisch sind, auf den randomisierten 200K Chips um ein Vielfaches geringer als auf dem originalen DSN200K-SNP-Chip. Dies spiegelt sich dementsprechend auch in der Zahl der identifizierten Varianten ($p < 0,1$) wieder. Da die funktionellen Varianten ein hohes Potenzial haben, für wirtschaftlich wichtige Phänotypen kausal zu sein, sollten sie generell in rassenspezifische SNP-Chips aufgenommen werden.

Tabelle 7: Vergleich der GWAS-Ergebnisse unter Verwendung verschiedener SNP-Chips.

Kategorie von Sequenzvarianten	DSN200K (10 Kriterien)		Random 200K (Zufällig)	
	MAF $\geq 5\%$	p < 0.1	MAF $\geq 5\%$	p < 0.1
# Total (Vergleich zu DSN200K)	131,273	163	131,273 \pm 0 (1 x)	60 \pm 8 (-2.7 x)
# signifikant in vorheriger GWAS (Vergleich zu DSN200K)	1,898	99	197 \pm 8 (-9.6 x)	12 \pm 2 (-8.3 x)
# Funktionelle Varianten (Vergleich zu DSN200K)	40,120	16	2,491 \pm 36 (-16.1 x)	1 \pm 1 (-16.0 x)
# DSN-spezifisch (Vergleich zu DSN200K)	7,295	8	294 \pm 17 (-24.8 x)	0 \pm 1 (-)

Genomische Zuchtwertschätzung und Genauigkeiten mit verschiedenen Markerdichten

Die Heritabilitäten und Varianzkomponenten mit den entsprechenden Standardfehlern (SE, engl.: „standard error“) für 305-d-Produktionsmerkmale, FEQ und SCS auf der Grundlage

verschiedener Markerdichten sind in Tabelle 8 dargestellt. Bei den 305-d-Produktionsmerkmalen lagen die Heritabilitäten zwischen 0,355 (VGS) für F305 und 0,646 (VGS) für E%. Für alle 305-d-Produktionsmerkmale wurden die höchsten Heritabilitäten auf der Grundlage der VGS-Daten geschätzt. Die Heritabilitäten stiegen jedoch nur um 0,011 bis 0,028, wenn VGS-Daten verwendet wurden, verglichen mit 50K- oder 200K-Marker-Dichten. Für SCS und FEQ war die additiv-genetische Varianz basierend auf den VGS-Daten recht gering. In den meisten Fällen war die Erhöhung der Markerdichten mit steigenden Heritabilitäten verbunden, aber die Unterschiede in den Heritabilitäten zwischen den vier Chips lagen nah beieinander. Die entsprechenden SE von 0 bis 0,036 weisen auf nicht-signifikante Effekte hin. Die auf den verschiedenen Chips basierenden Heritabilitäten für FEQ betrugen 0,145 (50K), 0,152 (DSN200K), 0,143 (random 200K) und 0,148 (VGS). Für SCS waren die Heritabilitäten der verschiedenen Genotyp-Datensätze sehr ähnlich, d. h. 0,097 basierend auf den VGS-Daten, 0,095 basierend auf den zufälligen 200K-Chipdaten, 0,088 basierend auf dem 50K-Chip und 0,085 basierend auf dem DSN200K. Auch die genetische Varianz für SCS auf der Grundlage von VGS-Daten war größer (0,202) im Vergleich zu den anderen drei Datensätzen, d. h. 0,198 für die zufälligen 200K-Chips, 0,183 für den 50K-Chip und 0,179 für den DSN200K-SNP-Chip. Die Heritabilitäten und Varianzen mit den entsprechenden SE für Exterieurmerkmale basierend auf verschiedenen Markerdichten sind in Tabelle 9 dargestellt. Für die vier Merkmalskomplexe lagen die Heritabilitäten zwischen 0,276 (50K) und 0,292 (VGS) für den Milchtyp, 0,317 (50K) und 0,353 (VGS) für den Körper, 0,080 (50K) und 0,089 (DSN200K) für das Fundament und zwischen 0,215 (50K) und 0,230 (VGS) für Euter. Die Heritabilitäten für die 18 linearen Merkmale reichten von 0,010 für den Klauenwinkel mit den 200K-Chips bis zu 0,450 für die Größe mit den VGS-Daten. Für die Strichplatzierung vorne und das Zentralband waren die Heritabilitäten mit dem 50K-Chip am größten. Im Allgemeinen waren die Unterschiede in den Heritabilitäten für ein und dasselbe Merkmal auf der Grundlage der verschiedenen Markerpanels recht gering. Jedoch wurden mit den VGS-Daten bei 15 der 22 Exterieurmerkmale die höchsten Heritabilitäten geschätzt. Dieses Ergebnis deutet erneut darauf hin, dass dichtere Markerdaten, d. h. VGS im Vergleich zur 50K- oder 200K-Dichte, die additiv-genetische Varianz für Merkmale des Körperbaus bei DSN genauer erfassen. Für das zusammengesetzte Merkmal Fundament und für die linearen Merkmale Milchcharakter, Hintereuterhöhe und Strichlänge wurden mit dem DSN200K-SNP-Chip die höchsten Heritabilitäten geschätzt. Die Heritabilitäten stiegen mit dem DSN200K-SNP-Chip um bis zu 2,2 %-Punkte im Vergleich zum 50K-Chip für die meisten der 22 Exterieurmerkmale, mit Ausnahme der Stärke, der Beckenneigung, des BCS, des Klauenwinkels, des Zentralbands und der Strichplatzierung vorne. Nur für die Beckenneigung war die Erblichkeit der zufälligen 200K-Chips die gleiche wie die anhand der VGS-Daten geschätzt.

Tabelle 8: Heritabilitäten (h^2), additiv-genetische Varianzen (σ_g^2) und Restvarianzen (σ_e^2) mit den dazugehörigen Standardfehlern (SE) der Schätzungen für die 305-d Produktions- und Testtagsmerkmale basierend auf verschiedenen Markerdichten. Die höchste Heritabilität innerhalb eines jeden Merkmals ist fett hervorgehoben.

Merkmal	Markerdichte							
	50K		DSN200K		Zufällige 200K (1-20) ²		VGS	
	Mittelwert	SE	Mittelwert	SE	Mittelwert	SE	Mittelwert	SE
M305								
h^2	0,410	0,036	0,412	0,036	0,414	0,037	0,433	0,037
σ_g^2	354020,000	38857,000	354370,000	38593,000	357023,500	39674,600	374540,000	40714,000
σ_e^2	507870,000	25385,000	504540,000	25263,000	503745,000	26368,600	488960,000	26632,000
EKM								
h^2	0,348	0,036	0,349	0,036	0,350	0,037	0,367	0,038
σ_g^2	260210,000	32101,000	260680,000	31968,000	262169,000	32824,250	275160,000	33784,000
σ_e^2	486380,000	22974,000	484230,000	22909,000	483819,500	23785,100	472840,000	24075,000
F305								
h^2	0,338	0,036	0,342	0,036	0,341	0,037	0,355	0,038
σ_g^2	423,420	52,902	428,030	53,093	426,869	54,161	445,460	55,673
σ_e^2	827,410	38,545	821,870	38,527	822,737	39,882	806,860	40,405
F%								
h^2	0,543	0,033	0,558	0,033	0,546	0,034	0,562	0,035
σ_g^2	0,060	0,005	0,062	0,005	0,060	0,005	0,062	0,005
σ_e^2	0,050	0,003	0,049	0,003	0,050	0,003	0,048	0,003
E305								
h^2	0,349	0,036	0,347	0,036	0,350	0,037	0,367	0,038
σ_g^2	317,710	39,305	315,500	38,927	318,838	40,020	334,740	41,189
σ_e^2	591,060	28,037	590,000	27,930	588,329	28,975	575,030	29,330

Ergebnisse

E%								
h^2	0,618	0,030	0,640	0,030	0,635	0,032	0,646	0,032
σ_g^2	0,021	0,002	0,022	0,002	0,021	0,002	0,022	0,002
σ_e^2	0,013	0,001	0,012	0,001	0,012	0,001	0,012	0,001
FEQ								
h^2	0,145	0,031	0,152	0,032	0,143	0,032	0,148	0,033
σ_g^2	0,005	0,001	0,005	0,001	0,005	0,001	0,005	0,001
σ_e^2	0,028	0,001	0,028	0,001	0,028	0,001	0,028	0,001
SCS								
h^2	0,088	0,025	0,086	0,025	0,095	0,026	0,097	0,027
σ_g^2	0,183	0,052	0,179	0,052	0,198	0,056	0,202	0,057
σ_e^2	1,893	0,068	1,895	0,067	1,877	0,069	1,874	0,069

¹ M305/F305/E305 = 305-d Laktationsertrag für Milch, Milcfett und Milcheiweiß; F% = 305-d Fettanteil in %; E% = 305-d Eiweißanteil in %; EKM = 305-d energiekorrigierte Milchleistung; FEQ = Fett-Eiweiß-Quotient am ersten Testtag; SCS = Somatischer Zellscore am ersten Testtag; ²Mittlere Heritabilitäten, basierend auf den 20 zufällig zusammengestellten 200K-Chips.

Tabelle 9: Heritabilitäten (h^2), additiv-genetische Varianzen (σ_g^2) und Restvarianzen (σ_e^2) mit den dazugehörigen Standardfehlern (SE) der Schätzungen für die Exterieurmerkmale, basierend auf verschiedenen Markerdichten. Die höchste Heritabilität innerhalb eines jeden Merkmals ist fett hervorgehoben.

Merkmal	Markerdichte							
	50K		DSN 200K		Zufällige 200K (1-20) ¹		VGS	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Milchtyp								
h^2	0,276	0,039	0,291	0,040	0,286	0,041	0,292	0,041
σ_g^2	2,358	0,373	2,490	0,386	2,439	0,387	2,499	0,396
σ_e^2	6,141	0,312	6,040	0,314	6,070	0,322	6,018	0,326
Körper								
h^2	0,317	0,042	0,328	0,043	0,346	0,044	0,353	0,045
σ_g^2	3,327	0,506	3,451	0,518	3,641	0,538	3,712	0,548
σ_e^2	7,124	0,389	7,035	0,392	6,840	0,402	6,772	0,408
Fundament								
h^2	0,080	0,030	0,089	0,031	0,086	0,031	0,088	0,032
σ_g^2	0,694	0,259	0,779	0,273	0,750	0,272	0,763	0,276
σ_e^2	7,970	0,336	7,894	0,339	7,913	0,342	7,901	0,344
Euter								
h^2	0,215	0,039	0,221	0,040	0,228	0,040	0,230	0,040
σ_g^2	2,817	0,551	2,905	0,559	2,980	0,564	3,016	0,569
σ_e^2	10,260	0,504	10,179	0,505	10,069	0,510	10,022	0,514
Größe								
h^2	0,416	0,043	0,428	0,043	0,437	0,045	0,450	0,046
σ_g^2	3,404	0,439	3,512	0,445	3,583	0,460	3,704	0,469
σ_e^2	4,757	0,292	4,664	0,292	4,597	0,305	4,492	0,309

Ergebnisse

Milchcharakter

h^2	0,227	0,039	0,242	0,041	0,229	0,041	0,234	0,042
σ_g^2	0,360	0,068	0,387	0,071	0,363	0,069	0,372	0,071
σ_e^2	1,221	0,061	1,202	0,061	1,218	0,062	1,211	0,063

Körpertiefe

h^2	0,332	0,043	0,351	0,044	0,354	0,045	0,360	0,046
σ_g^2	0,521	0,079	0,553	0,081	0,555	0,082	0,565	0,083
σ_e^2	1,041	0,059	1,017	0,059	1,009	0,061	0,999	0,061

Stärke

h^2	0,180	0,037	0,183	0,038	0,195	0,039	0,199	0,040
σ_g^2	0,212	0,046	0,216	0,047	0,231	0,049	0,235	0,050
σ_e^2	0,963	0,045	0,961	0,046	0,945	0,047	0,941	0,047

Beckenneigung

h^2	0,315	0,043	0,305	0,043	0,331	0,045	0,331	0,045
σ_g^2	0,294	0,046	0,284	0,045	0,309	0,048	0,308	0,048
σ_e^2	0,635	0,035	0,642	0,035	0,620	0,036	0,620	0,037

Beckenbreite

h^2	0,303	0,043	0,309	0,043	0,316	0,044	0,325	0,045
σ_g^2	0,419	0,067	0,427	0,068	0,436	0,070	0,449	0,071
σ_e^2	0,957	0,052	0,950	0,053	0,940	0,054	0,928	0,055

Vordereuteraufhängung

h^2	0,177	0,037	0,191	0,038	0,189	0,038	0,196	0,039
σ_g^2	0,296	0,065	0,320	0,068	0,317	0,067	0,327	0,069
σ_e^2	1,371	0,065	1,350	0,065	1,349	0,065	1,339	0,066

Hintereuterhöhe

h^2	0,221	0,042	0,243	0,043	0,237	0,043	0,246	0,044
σ_g^2	0,364	0,074	0,402	0,077	0,391	0,077	0,406	0,079
σ_e^2	1,277	0,065	1,245	0,065	1,250	0,067	1,238	0,068

Ergebnisse

Eutertiefe

h^2	0,268	0,042	0,270	0,042	0,285	0,044	0,287	0,044
σ_g^2	0,387	0,067	0,389	0,067	0,411	0,070	0,414	0,071
σ_e^2	1,049	0,055	1,046	0,055	1,025	0,057	1,022	0,057

Strichlänge

h^2	0,335	0,043	0,343	0,044	0,337	0,045	0,344	0,045
σ_g^2	0,336	0,051	0,346	0,051	0,338	0,051	0,345	0,053
σ_e^2	0,664	0,038	0,658	0,038	0,660	0,039	0,654	0,039

Strichplatzierung vorne

h^2	0,312	0,041	0,306	0,041	0,307	0,041	0,307	0,042
σ_g^2	0,380	0,057	0,371	0,056	0,372	0,057	0,372	0,057
σ_e^2	0,834	0,045	0,839	0,045	0,837	0,046	0,836	0,046

Strichplatzierung hinten

h^2	0,199	0,038	0,206	0,038	0,199	0,039	0,207	0,040
σ_g^2	0,213	0,043	0,220	0,044	0,213	0,044	0,221	0,045
σ_e^2	0,854	0,041	0,846	0,041	0,852	0,042	0,844	0,042

Hinterbeinwinkelung

h^2	0,122	0,033	0,140	0,036	0,148	0,037	0,153	0,038
σ_g^2	0,112	0,031	0,128	0,034	0,136	0,035	0,140	0,036
σ_e^2	0,800	0,035	0,786	0,036	0,778	0,037	0,774	0,037

Hinterbeinstellung

h^2	0,097	0,032	0,111	0,034	0,112	0,035	0,116	0,036
σ_g^2	0,195	0,066	0,223	0,070	0,226	0,072	0,233	0,073
σ_e^2	1,807	0,079	1,782	0,080	1,777	0,081	1,770	0,082

Ergebnisse

Zentralband

h^2	0,166	0,039	0,157	0,038	0,162	0,039	0,159	0,039
σ_g^2	0,298	0,073	0,282	0,072	0,290	0,073	0,284	0,073
σ_e^2	1,490	0,072	1,502	0,072	1,493	0,073	1,497	0,073

Sprunggelenk

h^2	0,150	0,036	0,158	0,037	0,159	0,038	0,161	0,038
σ_g^2	0,279	0,069	0,293	0,071	0,295	0,073	0,299	0,074
σ_e^2	1,569	0,072	1,558	0,073	1,554	0,074	1,550	0,075

Klauenwinkel

h^2	0,011	0,017	0,010	0,017	0,011	0,018	0,012	0,019
σ_g^2	0,009	0,015	0,009	0,015	0,010	0,016	0,010	0,016
σ_e^2	0,869	0,032	0,870	0,032	0,869	0,033	0,868	0,033

BCS

h^2	0,260	0,042	0,260	0,042	0,265	0,043	0,266	0,044
σ_g^2	0,307	0,054	0,308	0,055	0,313	0,056	0,315	0,057
σ_e^2	0,871	0,045	0,871	0,046	0,864	0,047	0,862	0,047

¹Mittlere Heritabilitäten, basierend auf den 20 zufällig zusammengestellten 200K-Chips.

Die mittleren gZW-Genauigkeiten der Validierungsläufe für 305-d-Produktionsmerkmale, FEQ und SCS sind in Tabelle 10 dargestellt. Bei den mäßig bis stark vererbaren 305-d-Produktionsmerkmalen wurden die höchsten gZW-Genauigkeiten auf der Grundlage des neu entwickelten DSN200K-SNP-Chips erzielt. Die Genauigkeiten lagen um 0,002 bis 0,008 höher im Vergleich zum 50K SNP-Chip und um 0,0002 bis 0,011 im Vergleich zu VGS-Daten (Abbildung 8). Bei den 305-d-Produktionsmerkmalen waren die Genauigkeiten der 20 zufälligen 200K-Chips im Vergleich zum DSN200K durchweg niedriger. Der Anstieg der gZW-Genauigkeit ist also nicht nur auf die erhöhte Markerdichte von 50K auf 200K zurückzuführen, sondern auch auf die Aufnahme von DSN-spezifischen Varianten auf dem DSN200K-SNP-Chip. Für SCS erreichten wir die höchste gZW-Genauigkeit mit VGS-Daten. Wie oben erläutert, wurde bei der Entwicklung des neuen DSN200K-SNP-Chips Marker aus vorangegangenen GWAS berücksichtigt, die signifikant mit M305, P305, F305, E% und F% in DSN assoziiert sind. Varianten, die mit SCS in einem assoziativen Zusammenhang standen wurden nicht berücksichtigt. Im Allgemeinen waren die gZW-Genauigkeiten für die Testtagsmerkmale FEQ und SCS mit geringer Heritabilität im Vergleich zu den 305-d-Produktionsmerkmalen viel geringer und reichten von 0,172 bis 0,185 für FEQ und von 0,156 bis 0,161 für SCS. Für beide Testtagsmerkmale wurde die erste Testtagsaufzeichnung nach dem ersten Kalben verwendet, die eine recht große phänotypische und residuale Varianz aufweist, was möglicherweise die niedrigen gZW-Genauigkeiten erklärt.

Tabelle 11 zeigt die mittleren gZW-Genauigkeiten von den Validierungsläufen für die Exterieurmerkmale. Die Anzahl der Datensätze für Exterieurmerkmale in unseren genomischen Vorhersagemodellen war auf 1.811 Kühe begrenzt. Dennoch erreichten wir moderate gZW-Genauigkeiten für die Merkmale mit moderater Heritabilität von bis zu 0,41 für die Größe, basierend auf VGS und DSN200K. Für 10 der 22 Exterieurmerkmale waren die gZW-Genauigkeiten mit VGS-Daten am höchsten. Für das Merkmalskomplexe Fundament waren die gZW-Genauigkeiten mit dem DSN200K-SNP-Chip, dem zufälligen 200K-Chip und VGS identisch. Auch bei der genetischen Bewertung der Körpertiefe war die durchschnittliche Genauigkeit mit den Schätzungen mit den zufälligen 200K-Chips und den VGS-Daten identisch. Die gZW-Genauigkeit wurde durch die Verwendung des neuentwickelten DSN200K-SNP-Chips im Vergleich zu den VGS-Daten für 11 der 22 Konformitätsmerkmale um bis zu 0,012 verbessert (Abbildung 8). Verbesserungen der gZW-Genauigkeit vom 50K-Chip zum DSN200K-SNP-Chip wurden für die drei Merkmalskategorien 305-d Produktionsmerkmale, FEQ und SCS am Testtag und Exterieurmerkmalskomplexe erzielt. Bei den linearen Typmerkmalen Beckenneigung, Eutertiefe, Strichplatzierung vorne, Zentralband und BCS ergab der DSN200K-SNP-Chip eine etwas geringere Genauigkeit von 0,001 bis 0,003 im Vergleich zum 50K-Chip (Abbildung 8). Bei den Exterieurmerkmalen war sogar der maximale Unterschied in der Genauigkeit (0,012) kleiner

als die minimale SD über alle Wiederholungen von 0,03, was wiederum auf nicht-signifikante Auswirkungen des gewählten Markerpanels hinweist.

Tabelle 10: Genauigkeit der genomischen Vorhersagen für 305-d Produktionsmerkmale und Testtagsmerkmale unter Verwendung verschiedener Markerdichten. Die Genauigkeit stellt den Mittelwert aller zehn Läufe der Kreuzvalidierung pro Merkmal und Markerdichte dar. Die höchste Genauigkeit pro Merkmal ist durch fett markierte Zahlen hervorgehoben. Die Standardabweichungen der mittleren Genauigkeiten lagen zwischen 0,02 und 0,05.

Merkmal ¹	Markerdichte			
	50K	DSN200K	Zufällige 200K (1-20) ²	VGS
305-d Produktionsmerkmale				
M305	0,453	0,455	0,449	0,452
EKM	0,391	0,392	0,385	0,387
F305	0,375	0,378	0,370	0,371
F%	0,558	0,567	0,557	0,558
E305	0,394	0,393	0,387	0,390
E%	0,605	0,614	0,605	0,605
Testtagsmerkmale				
FEQ	0,178	0,185	0,172	0,173
SCS	0,156	0,157	0,160	0,161

¹ M305/F305/E305 = 305-d Laktationsertrag für Milch, Milchfett und Milcheiweiß; F% = 305-d Fettanteil in %; E% = 305-d Eiweißanteil in %; EKM = 305-d energiekorrigierte Milchleistung; FEQ = Fett-Eiweiß-Quotient am ersten Testtag; SCS = Somatischer Zellscore am ersten Testtag; ²Mittlere Genauigkeiten, basierend auf den 200 ermittelten Genauigkeiten der zufällig zusammengestellten 200K-Chips (20 zufällige Chip x 10 Validierungsläufe)

Tabelle 11: Genauigkeit der genomischen Vorhersagen für Exterieurmerkmale und Testtagsmerkmale unter Verwendung verschiedener Markerdichten. Die Genauigkeit stellt den Mittelwert aller zehn Läufe der Kreuzvalidierung pro Merkmal und Markerdichte dar. Die höchste Genauigkeit pro Merkmal ist durch fett markierte Zahlen hervorgehoben. Die Standardabweichungen der mittleren Genauigkeiten lagen zwischen 0,02 und 0,05.

Merkmale	Markerdichte			
	50K	DSN 200K	Zufällig 200K (1-20) ¹	VGS
Exterieurkomplexe				
Milchtyp	0,339	0,348	0,338	0,339
Körper	0,323	0,328	0,331	0,332
Fundament	0,129	0,134	0,134	0,134
Euter	0,262	0,265	0,270	0,272
Lineare Merkmale				
Größe	0,398	0,408	0,398	0,401
Milchcharakter	0,264	0,271	0,260	0,259
Körpertiefe	0,319	0,324	0,325	0,326
Stärke	0,232	0,234	0,239	0,238
Beckeneneigung	0,316	0,315	0,317	0,314
Beckenbreite	0,274	0,274	0,274	0,275
Vordereuteraufhängung	0,225	0,237	0,236	0,238
Hintereuter	0,225	0,235	0,234	0,236
Eutertiefe	0,277	0,275	0,276	0,278
Strichlänge	0,340	0,347	0,335	0,336
Strichplatzierung vorne	0,385	0,382	0,377	0,376
Strichplatzierung hinten	0,292	0,295	0,285	0,288
Hinterbeinwinkelung	0,197	0,206	0,205	0,206
Hinterbeinstellung	0,148	0,157	0,156	0,157
Zentralband	0,193	0,190	0,190	0,188
Sprunggelenk	0,213	0,216	0,214	0,213
Klauenwinkel	0,039	0,039	0,042	0,044
BCS ²	0,299	0,299	0,299	0,299

¹Mittlere Genauigkeiten, basierend auf den 200 ermittelten Genauigkeiten der zufällig zusammengestellten 200K-Chips (20 zufällige Chip x 10 Validierungsläufe). ²Body-Condition-Score

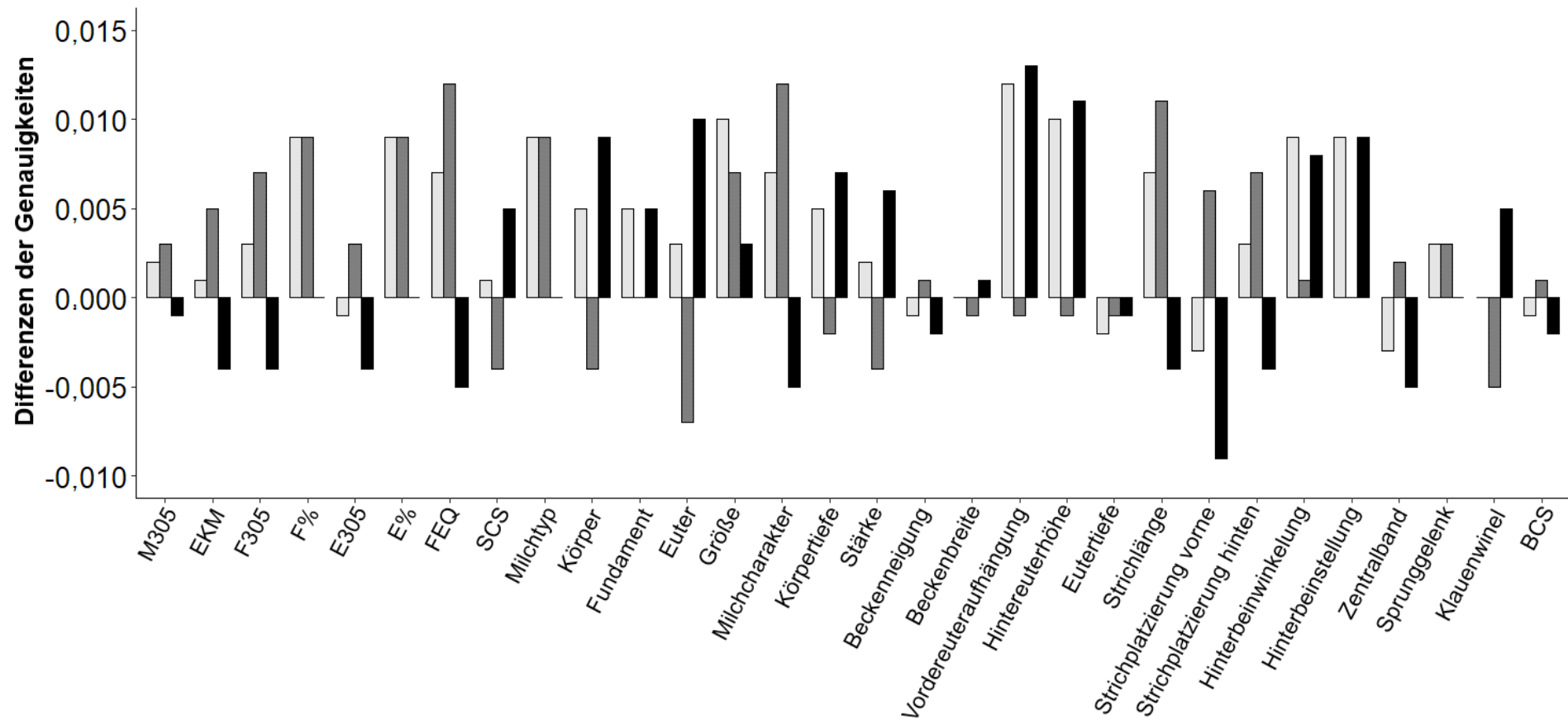


Abbildung 8: Unterschiede in der Vorhersagegenauigkeit (Mittelwert aller 10 Läufe der Kreuzvalidierung) für 305-d-Milchproduktionsmerkmale, Testtagsmerkmale und Exterieurmerkmale, Hellgraue Balken = erzielte Genauigkeiten mit dem DSN200K-SNP-Chip im Vergleich zum 50K-Illumina-Chip; dunkelgraue Balken = erzielte Genauigkeiten mit dem DSN200K-SNP-Chip im Vergleich zu den imputierten Vollgenomsequenzdaten (VGS); schwarze Balken = erzielte Genauigkeit mit den VGS-Daten im Vergleich zum 50K-Illumina-Chip, M305 = 305-d Milchleistung in kg; E305 = 305-d Eiweißertrag in kg; F305 = 305-d Fettertrag in kg; F% = 305-d Fettanteil in %; E% = 305-d Eiweißanteil in %; EKM = 305-d energiekorrigierte Milchleistung; FEQ = Fett-Eiweiß-Quotient am ersten Testtag; SCS = Somatischer Zellscore am ersten Testtag; BCS = Body-Condition-Score

3.4 Untersuchungen zur Sicherung der genetischen Diversität in der DSN-Population (APIV)

Phylogenetische Untersuchungen und Nukleotid-Diversitätsanalysen beim DSN und 68 weiteren taurinen Rinderrassen

Bei der phylogenetischen Analyse von 69 Rinderrassen wurden vier Hauptcluster identifiziert, die geografischen Ursprüngen zugeordnet werden können (Abbildung 9): Nordeuropa (grün), Mitteleuropa (violett), die Inseln Jersey und Guernsey (rot) sowie Osteuropa, Mittelitalien und Asien (blau). Die Gruppe Mitteleuropa umfasste Rassen aus Österreich, der Schweiz, Süddeutschland und Frankreich. Dabei handelt es sich hauptsächlich um Zweinutzungsrasen, die in Gebirgsregionen gehalten werden. Die meisten Rassen, einschließlich DSN und Holstein, bildeten jedoch eine Gruppe, die auf ihren Ursprung in den mittleren bis nördlichen Teilen Europas gründet, Jersey- und Guernsey-Rinder bildeten eine eigene Gruppe. Das vierte Cluster umfasste Rassen aus asiatischen, mittelitalienischen und osteuropäischen Ländern. Innerhalb der Rassen wurden die meisten Tiere ohne, oder mit nur wenigen Ausnahmen, demselben Cluster zugeordnet. Bei den Rassen Modern Danish Red, Swedish Red, Norwegian Red, Ayrshire Finish und Modern Angler wurden die Tiere der jeweiligen Rasse vermehrt zu verschiedenen Clustern zugeordnet. Bei der Rasse DSN konnte festgestellt werden, dass sie am engsten mit den Rassen Dutch Friesian Red, Dutch Belted, HOL, Eastern Flanders White Red und Kholmogory in Beziehung steht. Dies sind alle Rassen, die ihren Ursprung in Deutschland, den Niederlanden und Belgien haben. Kholmogory ist eine russische Rasse und stammt von Tieren ab, die aus den Niederlanden eingeführt wurden. Weitere Rassen aus der Nordseeregion, aus den Gebieten Deutschlands, der Niederlande, Dänemarks, Belgiens und Skandinaviens zeigten ebenfalls eine enge Verwandtschaft zur Rasse DSN.

Dieselben Rassen wie in der phylogenetischen Baumanalyse zeigten die engste Verwandtschaft mit DSN in Bezug auf die niedrigsten F_{ST} -Werte. Die Rassen sortiert nach einem aufsteigenden F_{ST} -Wert von 0,032 bis 0,075 sind Dutch Friesian Red, Dutch Improved Belted, Eastern Belgian Red White, Belgian Red White Campine, Deep Red Cattle, Eastern Flanders White Red, Holstein Red, Kholmogory, Red and White Dual Purpose, Holstein, und Dutch Belted. Rassen, die ebenfalls niedrige F_{ST} -Werte zwischen 0,053 und 0,078 aufweisen, aber in der phylogenetischen Analyse etwas weiter entfernt sind, sind Modern Angler, Modern Danish Red, German Red Angler, Polish Red, Norwegian Red, Eastern Finncattle, West Vlaams Rood und Swedish Red Polled (in aufsteigender Reihenfolge). Charolais ist die einzige Rasse, die im phylogenetischen Baum außerhalb des nordeuropäischen Clusters liegt, aber einen niedrigen F_{ST} -Wert von 0,070 zu DSN aufweist.

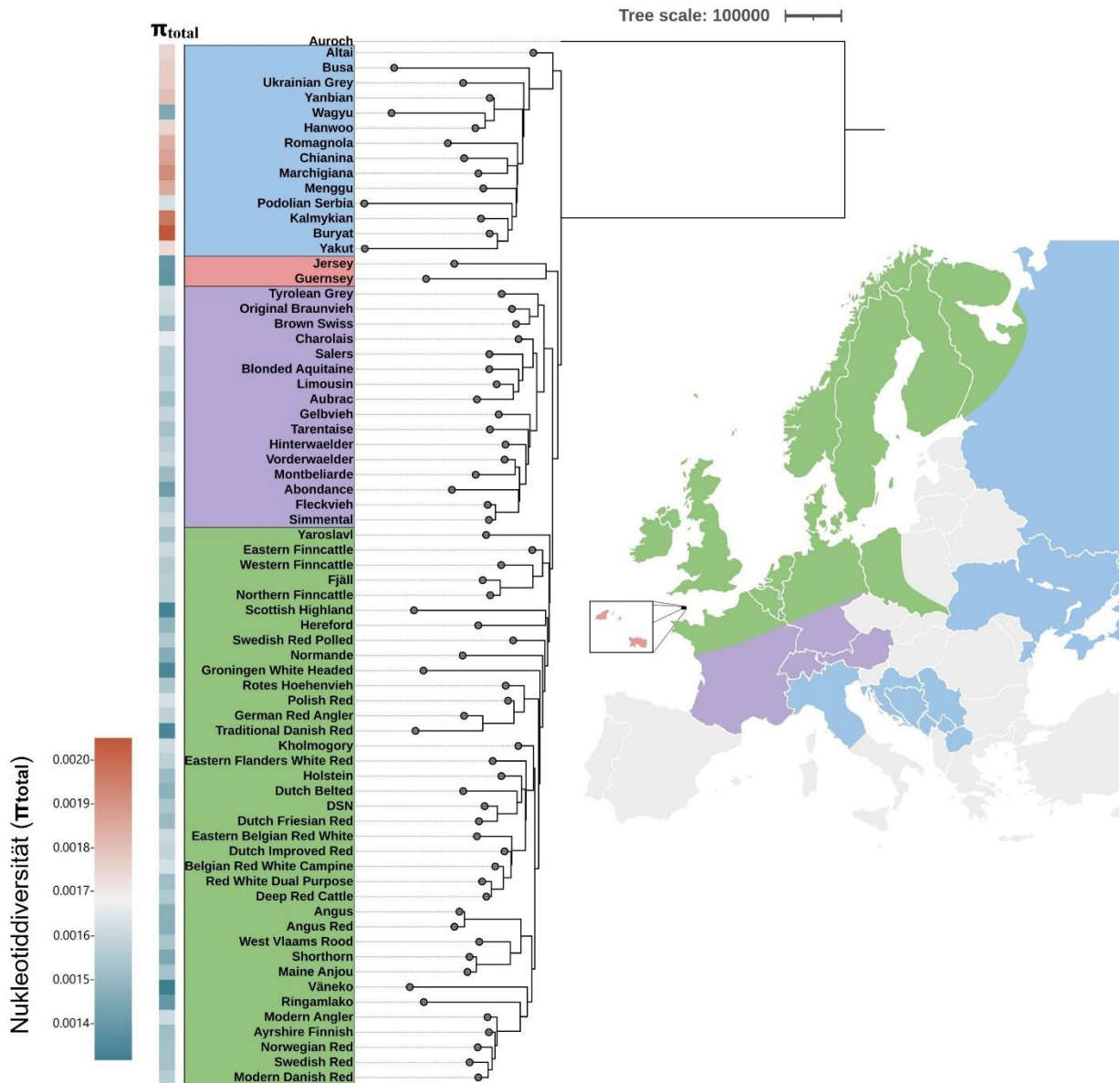


Abbildung 9: Phylogenetischer Baum von 69 Rinderrassen. Die Cluster sind nach ihrer geographischen Lage gefärbt: Mittelitalien, osteuropäische und asiatische Länder (blau), Jersey und Guernsey (rot), Mitteleuropa (violett), nordeuropäische Länder (grün). Die Individuen jeder Rasse wurden zu einzelnen Clustern zusammengefasst (graue Kugeln).

Die Nukleotid-Diversität (π_{total}) für die 69 untersuchten Rassen lag zwischen 0,001316 bei Väneko und 0,002048 bei Buryat. Bei der Rasse DSN war π_{total} mit einem Wert von 0,001541 höher als bei der Rasse HOL (0,001505) und bei 28 weiteren Rassen (20 aus Nordeuropa), jedoch geringer als bei anderen mitteleuropäischen und asiatischen Rassen. Die höchsten π_{total} -Werte wurden bei asiatischen und osteuropäischen Rassen festgestellt. Die anderen europäischen Rassen wiesen ähnliche π_{total} -Werte auf, mit einigen wenigen Ausnahmen ($< 0,0014$): Ringamålako (0,001396), Jersey (0,001396), Guernsey (0,001385), Groningen White Headed (0,001345), Traditional Danish Red (0,001339), Scottish Highland (0,001328) und Väneko (0,001316), Regionen mit besonders hoher Nukleotiddiversität (π_{window}) wurden beim DSN, wie auch bei allen anderen Rassen auf BTA 4, 10, 12, 15, 29 (Abbildung 10a) und insbesondere auf BTA 23 zwischen 19 und 37 Megabasenpaaren (Mb) (Abbildung 10b) beobachtet, wo sich der Haupthistokompatibilitäts-Komplex (MHC, engl.: „Major Histocompatibility Complex“) und das Bovine Leukozyten-Antigen befinden.

Trotz der geringen Populationsgröße zeigte DSN eine ähnliche Nukleotid-Diversität wie die Rassen desselben Clusters, was auf ein gutes Management der gefährdeten Population hinweist. Wie erwartet, enthält BTA 23 beim DSN und allen anderen Rassen eine hoch polymorphe Region, in der sich der MHC befindet. Diese hohe Diversität um den MHC ist auch bei anderen Arten zu beobachten (Shiina et al., 2017). Asiatische und osteuropäische Rassen weisen eine höhere Nukleotid-Diversität (π_{total}) auf, was wahrscheinlich auf weniger intensive Zuchtprogramme, genetische Engpässe oder Gründereffekte im Vergleich zu mittel- und nordeuropäischen Rassen zurückzuführen ist (Felius et al., 2014). Die weitere Migration dieser Rinder nach Süd-, Mittel- und schließlich Nordeuropa steht im Einklang mit den Clustern in unserem phylogenetischen Baum (Felius et al., 2014). Die geringste Nukleotid-Diversität der nordeuropäischen Rotviehpopulationen wurde bei den Rassen Traditional Danish Red und Groningen White Headed festgestellt, was auf jüngste Inzuchtereignisse in den kleinen Populationen zurückzuführen ist (Schmidtman et al., 2021). Kürzlich haben Schmidtman et al. (2021) ein hohes Maß an Vermischung zwischen Modern Angler, Ayrshire Finnish und Norwegian Red festgestellt, was erklärt, warum unsere Clusterbildung zwischen diesen Rassen viele Fehlzusammenordnungen aufwies.

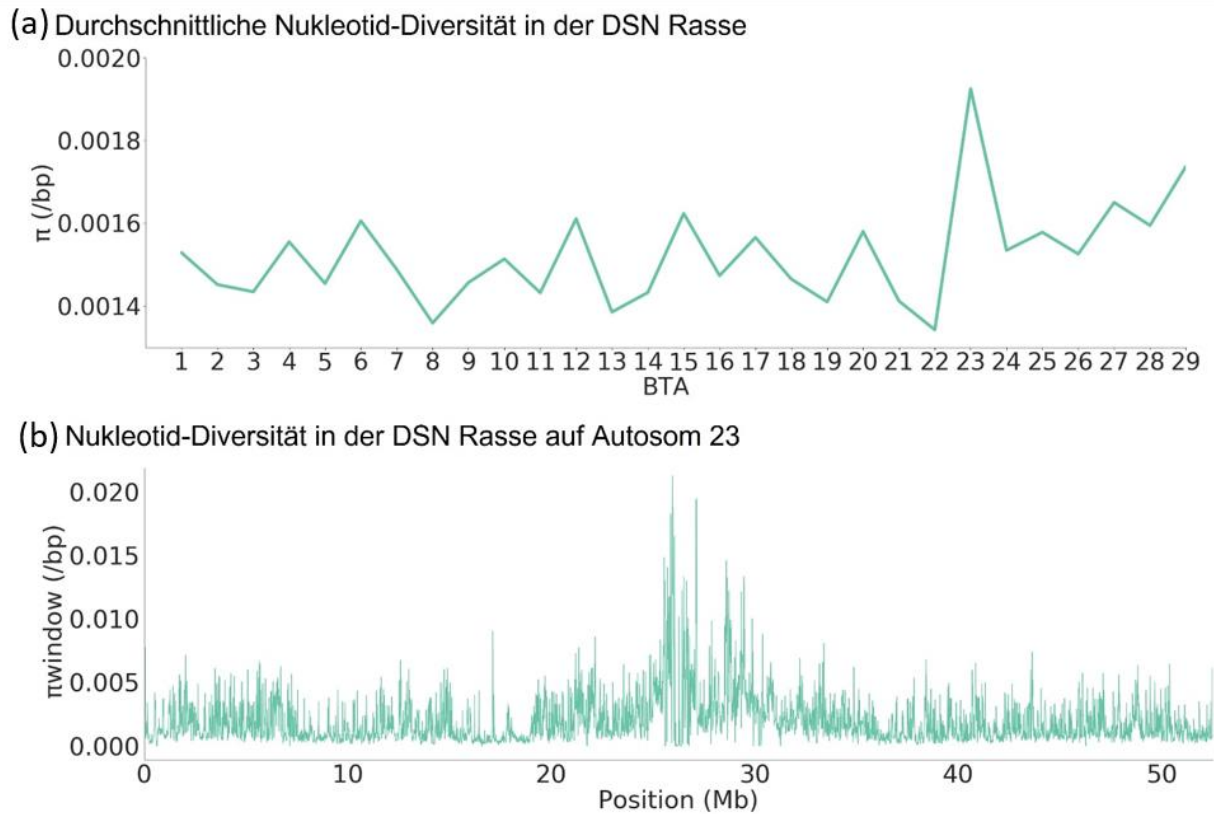


Abbildung 10: (a) Die durchschnittliche Nukleotid-Diversität in der Rasse DSN pro Autosom (BTA), (b) Nukleotid-Diversität in der Rasse DSN pro Fenster auf Autosom 23.

Zusätzlich wurde die Diversität als Durchschnittswert der erwarteten und beobachteten Heterozygotie pro Individuum berechnet. Je geringer diese Werte sind, desto geringer ist die genetische Variabilität. Die erwartete Heterozygotie (H_e) lag zwischen 9,4 % in DSN und 11,9 % in Yakut und die durchschnittliche beobachtete Heterozygotie (H_o) zwischen 9,2 % in Jersey und 11,3 % in Yakut (Abbildung 11a). Die genomische Inzucht (F) wurde basierend auf dem Überschuss an Homozygotie (F_{Hom}) und auch als Inzuchtrate F_{RoH} , die auf RoHs unterschiedlicher Länge (50kb, 100kb, 1Mb, 2Mb, 4Mb) basiert. Je größer die Werte für die genomische Inzucht sind, desto höher ist die Inzucht. Je länger ein homozygotes Segment, desto kürzer liegt das Inzucht-Ereignis zurück. Die Inzuchtrate, die als Überschuss an Homozygotie (F_{Hom}) geschätzt wurde, reichte von 2,7 % bei DSN bis 9,8 % bei Hereford (Abbildung 11b). Damit war F_{Hom} bei DSN niedriger als bei HOL (6,4 %) und allen anderen Rassen. Wie erwartet, wiesen die kommerziellen Rassen im Allgemeinen höhere Inzuchtraten auf, wie z. B. Charolais (7,0 %), Shorthorn (8,5 %), Brown Swiss (9,4 %) und Hereford (9,8 %). Die durch F_{RoH} geschätzte Inzuchtrate unter Berücksichtigung aller RoHs, die länger als 50 kb sind, betrug im Durchschnitt 16,0 % bei DSN und 19,7 % bei HOL. Von $F_{RoH > 50 \text{ kb}}$ bis $F_{RoH > 2 \text{ Mb}}$ wurde die niedrigste bzw. höchste Inzuchtrate immer bei Modern Danish Red und Jersey beobachtet, während $F_{RoH > 4 \text{ Mb}}$ von 0,7 % bei Limousin bis 4,0 % bei Brown Swiss reichte. $F_{RoH > 1 \text{ Mb}}$ und $F_{RoH > 2 \text{ Mb}}$ zeigten ähnliche Ergebnisse wie bei F_{Hom} .

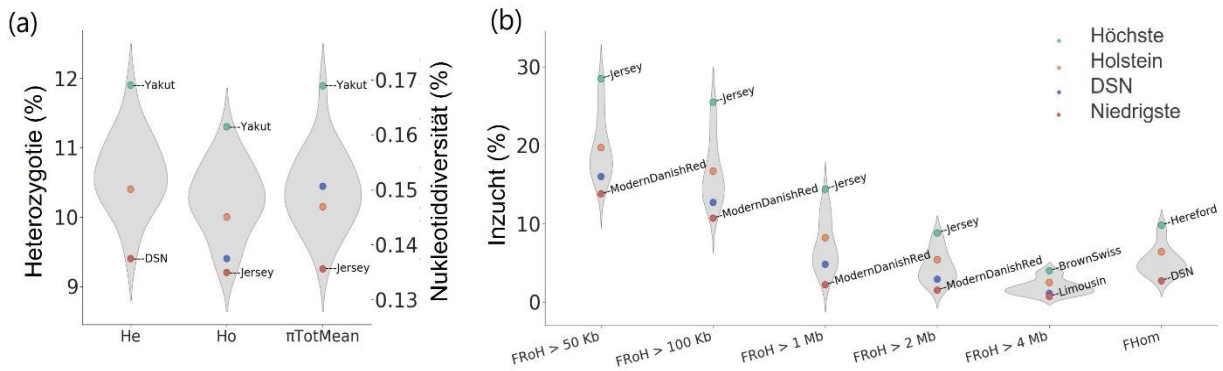


Abbildung 11: (a) Überblick über Diversitätsmessungen und Inzucht bei verschiedenen Rassen. Links: Durchschnittliche Nukleotiddiversität ($\pi_{TotMean}$), beobachtete Heterozygotie (Ho) und erwartete Heterozygotie (He), (b) Inzucht, basierend auf Runs of Homozygotie (RoHs) verschiedener Länge (F_{RoH}) und Überschuss an Homozygotie (F_{Hom}).

Trotz der geringen Populationsgröße von 2.500 Tieren zeigt die Rasse DSN im Vergleich zu anderen Rinderrassen keine Anzeichen von Diversitätsverlust oder erhöhter Inzucht. Im Gegenteil, der Inzuchtgrad bei DSN ist sogar niedriger und die Diversität höher als bei HOL (Tabelle 12). Dies ist ein bemerkenswertes Ergebnis der Zuchtstrategie, die zur Erhaltung von DSN als genetische Ressource angewandt wird, und zeigt das Potenzial der Erhaltung kleiner lokaler Populationen bei gleichzeitiger Wahrung der Vielfalt und Kontrolle der Inzucht. Darüber hinaus könnten Rinderrassen, die nachweislich eng mit DSN verwandt sind, als externer Genpool dienen, um die Vielfalt in DSN zu erhalten oder sogar zu erhöhen. Eine hohe genomische Vielfalt wurde auch bei Rassen wie Yakut, Charolais, Kholmogory und Modern Danish Red festgestellt, während die Inzucht bei Jersey, Wagyu, Hereford und Shorthorn höher war, was darauf hindeutet, dass für diese Rassen besondere Sorgfalt erforderlich ist.

Tabelle 12: Vergleich der Werte für genomische Inzucht und genetischen Diversität zwischen DSN und Holstein.

		DSN	Holstein
Genomische	Überschuss der Homozygotie (F_{Hom})	2,7 (2,7)	6,4 (7,0)
Inzucht (%)	Inzuchtkoeffizient (F_{RoH})		
	> 1 Mb	4,8 (2,5)	8,2 (3,5)
	> 2 Mb	2,9 (1,9)	5,4 (2,8)
	> 4 Mb	1,1 (1,0)	2,5 (1,7)
Genetische	Nukleotiddiversität (π_{total})	0,152 (0,012)	0,148 (0,013)
Diversität (%)	Beobachtete Heterozygotie (Ho)	22,4 (0,9)	21,2 (1,9)

„Optimum genetic contribution“-Anwendung innerhalb der DSN-Population

Die Ergebnisse aus der Anwendung der „optimum genetic contribution“ Methode mit dem Programm GENCONT (Meuwissen, 2002) sind der Tabelle 13 zu entnehmen. Eine graphische Übersicht der Ergebnisse zeigt Abbildung 12. Die angestrebten durchschnittlichen Verwandtschaften von 0,02 bis 0,10 konnten nicht immer erzielt werden. Verwandtschaften zwischen 0,03 und 0,07 wurden sowohl auf Basis der gZWs basierend auf 50K als auch mit den 200K Genotypen rechnerisch realisiert. Umso geringer die Verwandtschaft zwischen den Nachkommen ist, desto mehr Bullen kommen zum Einsatz. So wären es bei 30 Kühen 19 eingesetzte Bullen um eine Verwandtschaft von 0,03 und 7 Bullen um eine angestrebte Verwandtschaft von 0,07 zu erzielen. Umso mehr Verwandtschaft toleriert wurde, desto höher war die Steigerung der durchschnittlichen gZWs der selektierten Elterntiere von 354 kg bei einer Verwandtschaft von 0,21 (50K) bis 1.206 kg bei einer Verwandtschaft von 0,07 (50K). Die Unterschiede zwischen den gZWs, geschätzt anhand der 50K und 200K Genotypen waren marginal. Die Anzahl an ausgewählten Bullen, bzw. Kühen unterschieden sich maximal um ein Tier.

Die Untersuchungen zeigen, dass es möglich ist eine Nachkommengeneration und somit auch Bullen zu erzeugen, die Potenzial für genetischen Fortschritt für das Leistungsmerkmal Milch in kg (305-d) vorweist, ohne eine Steigerung der Verwandtschaft in Kauf zu nehmen. Würde die Anpaarungsstrategie gewählt werden, bei der eine durchschnittliche Verwandtschaft von 0,04 als Grenze genommen wurde, würden 15 Bullen (200K), bzw. 16 Bullen (50K) mit 30 Kühen angepaart werden und man könnte eine Steigerung des Zuchtfortschritts von 1.003 kg (200K) bis 1.013 kg (50K) erzielen. Die dabei gewählte durchschnittliche Verwandtschaft von 0,04 liegt 3 %-Punkte unter dem Durchschnitt der gesamten DSN-Population (DSN-Anteil > 90 %) (Jaeger, 2018).

Generell lässt sich feststellen, dass trotz der gefährdeten Populationsgröße der Rasse DSN eine Steigerung in den Zuchtwerten mit Verzicht auf Einkreuzung von anderer Genetik, ohne dabei die Inzucht innerhalb der Population zu erhöhen, durchaus möglich wäre. Es ist denkbar, dass der langfristige Einsatz solcher Monitoringverfahren, basierend auf dem OGC Ansatz in der Praxis die Attraktivität der Zweinutzungsrasse DSN in Form von steigenden Zuchtwerten in den Leistungsmerkmalen, bei gleichzeitiger Limitierung der innerpopulären Verwandtschaft, steigert und somit zum Erhalt der Rasse beiträgt.

Es ist eine Publikation Anfang des Jahres 2023 mit weitergehenden Analysen zur Anwendung der OGC-Theorie in der Rasse DSN durch TZ-GI geplant.

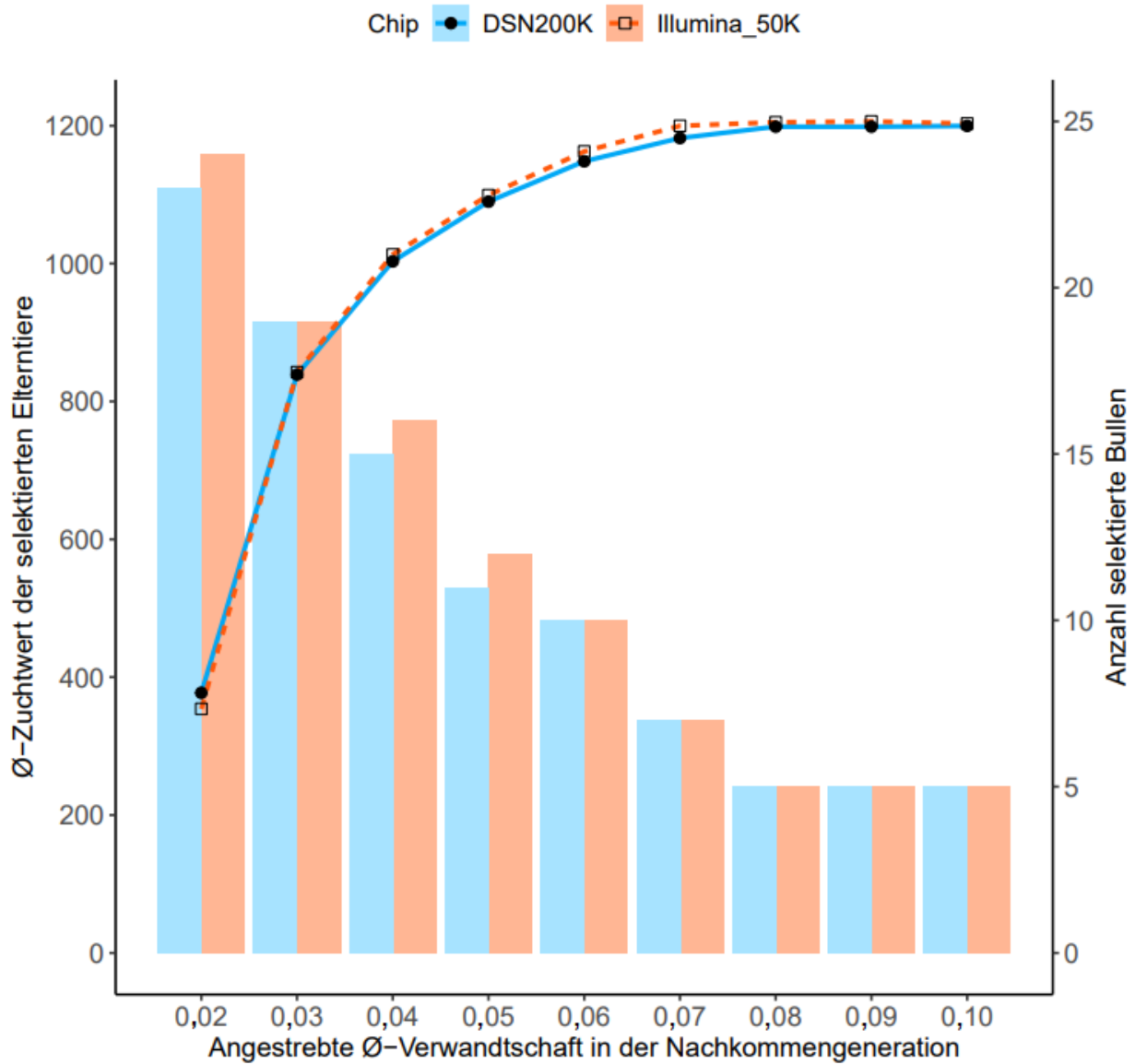


Abbildung 12: Darstellung der Ergebnisse aus der Anwendung der „optimum genetic contribution“ Methode mit dem Programm GENCONT (Meuwissen 2002) für das Merkmal 305-d Milchleistung in kg. Die Linien zeigen die Ø-Zuchtwerte der Elterntiere an. Die Balken zeigen an, wie viele Bullen in die Berechnungen einbezogen wurden, um bei der Anpaarung mit insgesamt 30 Kühen, die angestrebten Verwandtschaften in den Nachkommengenerationen zu erzielen. Dabei beziehen sich die blaue Linie und Balken auf die Ergebnisse basierend auf dem DSN200K-SNP-Chip, die rote gestrichelte Linie und Balken auf den Illumina 50K Chip.

Tabelle 13: Ergebnisse aus der Anwendung der „optimum genetic contribution“-Methode mit dem Programm GENCONT (Meuwissen, 2002) für das Merkmal 305-d-Milchleistung in kg. Es wurden neun Szenarien mit jeweils unterschiedlicher, angestrebter durchschnittlicher Verwandtschaft der Tiere in der Nachkommengeneration berechnet. Die erzielte Steigerung der durchschnittlichen Zuchtwerte in der Nachkommengeneration im Vergleich zur Gesamtpopulation sind in der Spalte „Ø Zuchtwert der selektierten Eltern (in kg)“ aufgelistet.

Angestrebte Ø- Verwandtschaft	Erzielte Ø- Verwandtschaft		Ø Zuchtwert der selektierten Eltern (in kg)		Selektierte Bullen		Selektierte Kühe	
	50K	200K	50K	200K	50K	200K	50K	200K
0,02	0,021	0,022	354	377	24	23	30	30
0,03	0,030	0,030	843	839	19	19	30	30
0,04	0,040	0,040	1013	1003	16	15	30	30
0,05	0,050	0,050	1099	1090	12	11	30	30
0,06	0,060	0,060	1163	1148	10	10	30	30
0,07	0,070	0,070	1200	1182	7	7	30	30
0,08	0,077	0,082	1205	1199	5	5	31	30
0,09	0,077	0,082	1206	1199	5	5	31	31
0,10	0,076	0,082	1203	1200	5	5	31	31

3.5 Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen (APV)

Verwendung des DSN200K-SNP-Chips in anderen Rassen

Der DSN200K-SNP-Chip wurde in 9 verschiedenen *Bos taurus* und einer *Bos indicus* Rinderrasse getestet. Von den 287 Rindern dieser Rassen haben 269 die Qualitätskontrolle bestanden und konnten zusammen mit den 1.594 DSN Tieren (1.162 aus diesem Projekt und 432 aus einem anderen Projekt), die ebenfalls die Qualitätskontrolle bestanden haben, analysiert werden. Insgesamt konnten von den 182.154 Varianten auf dem DSN200K-SNP-Chip 171.979 Varianten (94,4 %) in mindestens einer Rasse ausgelesen werden, d.h. sie haben technisch funktioniert. Diese setzten sich aus 164.535 SNPs und 7.444 INDELS zusammen. Im Rückschluss heißt das, dass 10.174 (5,6 %) Varianten trotz der durch Thermo Fisher Scientific ermittelten Eignung nicht funktioniert haben und folglich nicht ausgelesen werden konnten. Von den 171.979 funktionierenden Sequenzvarianten segregieren 167.271 Varianten mit einer MAF > 0 % in DSN (Abbildung 13). Überraschenderweise ist die Anzahl segregierender Varianten mit MAF > 0% in der Rasse Pinzgauer am höchsten (169.182; 92,9 %). Danach folgten HOL (167.118; 91,8 %), RDN (165.678; 90,9 %), Hinterwälder (165.274; 90,7 %), Rotes Höhenvieh (161.804; 88,9 %), White Park (159.957; 87,8 %), ANG (158.963; 87,3 %), Original Braunvieh (158.224; 86,9 %), Butana (158.273; 86,9 %), und Gelbvieh (154.921; 85,1 %). Allerdings überrascht es nicht, dass die Zahl der segregierenden Varianten mit MAF > 1 % und damit informativen Varianten bei den DSN Tieren am höchsten war (158.382; 87,0 %), denn

schließlich beinhaltet der DSN200K-SNP-Chip über 38K DSN-spezifische Varianten. Diese DSN-spezifische Varianten sind in den anderen Rassen monomorph. Die Häufigkeiten informativer Varianten mit $MAF > 1\%$ bei den anderen Rassen lagen zwischen 48,3 % (Butana) und 68,5 % (RDN). Nur die Rasse Hinterwälder hatte eine sehr niedrige Anzahl segregierender Varianten von 21,9 % (39.910). Dies lag höchstwahrscheinlich daran, dass nur zwei Tiere dieser Rasse genotypisiert wurden und auch der Verwandtschaftsgrad zwischen diesen beiden Tieren nicht bekannt ist.

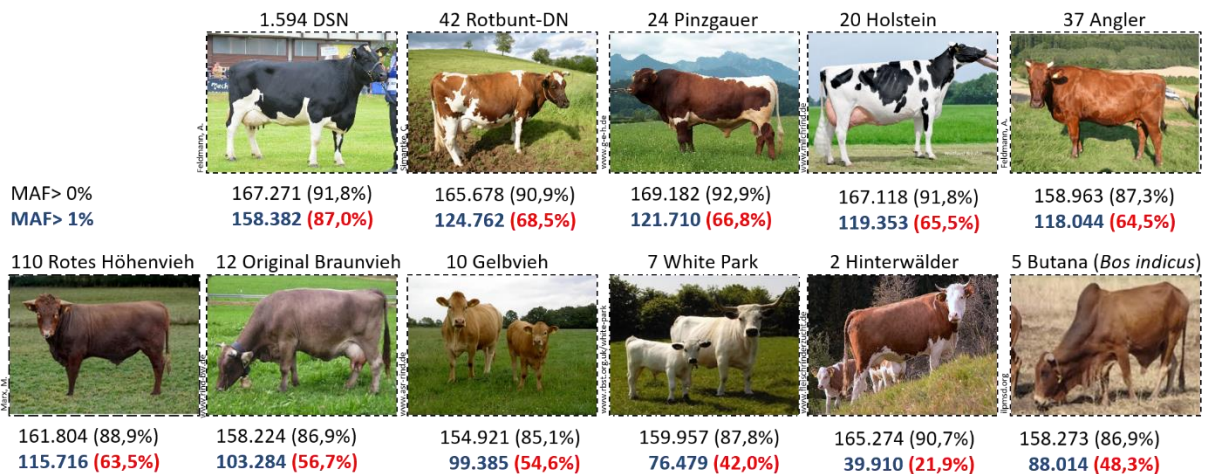


Abbildung 13: Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen. Hierfür wurden Rinder aus neun *Bos taurus* und einer *Bos indicus* Rassen mit dem DSN200K-SNP-Chip genotypisiert. Rotbunt-DN = RDN; Angler = ANG; Holstein = HOL.

Obwohl der DSN200K-SNP-Chip für DSN entwickelt wurde, hat sich seine Verwendung für andere Rassen als erfolgreich erwiesen und liefert mehr als 154K Varianten, die je Rasse mit einer $MAF > 0\%$ ausgelesen werden konnten.

Die Anzahl segregierender Varianten ($MAF > 1\%$) unterscheidet sich stark je nach Rasse und Anzahl genotypisierter Tiere. Der Einsatz dieses neuen SNP-Chips könnte vor allem für Rassen, die nah mit DSN verwandt sind von Interesse sein, da die Anzahl der segregierenden Varianten höher ist als bei entfernt verwandten Rassen. Zu den eng verwandten Rassen zählen z.B. Dutch Friesian Red, Dutch Belted, Dutch Improved Red, Rotbunt-DN, Angler, Rotes Höhenvieh und natürlich HOL. Folglich ist die Verwendung des gesamten DSN200K-SNP-Chips oder einzelner polymorpher Varianten des Chips für andere Rassen möglich. Die in anderen Rassen segregierenden Varianten können zum Beispiel als Grundlage zur Erstellung weiterer rassespezifischer Chips für andere lokale Rassen dienen. Das ist sogar für die Rinderrasse Butana denkbar, die eine Kurzhorn-Zeburasse aus Ostafrika repräsentiert. Das Chipdesign, sowie die Leistungsfähigkeit des Chips wurde im Jahr 2021 in der Zeitschrift BMC Genomics veröffentlicht (Neumann et al., 2021).

Sequenzierungsanalyse von Tieren der Rassen Angler und Rotbunt-Doppelnutzung

Mit Hilfe der Sequenzdaten wurden insgesamt 16.148.426 Varianten in der Rasse Angler (ANG) und 14.615.330 Varianten in der Rasse Rotbunt-Doppelnutzung (RDN) identifiziert. Diese SNPs wurden indirekt in den Diversitätsuntersuchungen, in denen Daten aus dem 1000-Bullen-Genom-Projekt verwendet wurden, berücksichtigt. Eine weitere Auswertung der Daten war in diesem Projekt nicht vorgesehen. Auf Basis der vorliegenden Vollgenomsequenzdaten der Rassen Angler und RDN können weiterhin rassespezifische und funktionelle Sequenzvarianten in diesen beiden Rassen identifiziert und in den bestehenden 200K DSN-SNP Chip inkludiert werden.

3.6 Voraussichtlicher Nutzen und Verwertbarkeit der Ergebnisse

In dem Verbundprojekt mit Modellcharakter ist es gelungen, einen rassespezifischen SNP-Chip für das DSN zu konstruieren, welcher als erster Chip rassespezifische Varianten von DSN abdeckt. Ein Vergleich des neuen DSN200K-SNP-Chips mit dem Illumina 50K Chip in genetischen Parameterschätzungen und in der gZWS hat gezeigt, dass die Erblichkeiten und Genauigkeiten der gZW mit dem neuen 200K Chip nur zu einer geringfügigen Verbesserung führen. Für die Detektion rassespezifischer Varianten in GWAS für die Rasse DSN ist der neue DSN200K-SNP-Chip hingegen besser geeignet als der Illumina 50K Chip oder ein randomisierter 200K Chip. Es hat sich zudem gezeigt, dass der neu konstruierte DSN200K-SNP-Chip auch für andere, vom Aussterben bedrohte Rinderrassen geeignet ist. Somit kann der DSN200K-SNP-Chip weiterentwickelt und bisherige Marker auf dem Chip durch „besser geeignete“ Varianten ersetzt werden, auch könnten neu detektierte Varianten mitberücksichtigt werden. Aufgrund der kleinen Populationsgröße der Rasse DSN und der derzeit noch hohen Kosten für den neuen Chip (ca. 102,90 € pro Probe exkl. MwSt. und Servicekosten; Stand: 21.11.2022), könnten besonders bedeutende rassespezifische Varianten der Rasse DSN auch auf die derzeit kostengünstigen SNP-Chips für größere Rinderpopulationen (z.B. EuroG MD Chip) inkludiert werden. Dies würde zu dem Erhalt der Rasse sehr wesentlich beitragen. Der Ausbau der Lernstichprobe der Rasse DSN, welcher sehr intensiv im Rahmen des Projektes betrieben wurde, führt zu einer Verbesserung der gZW in der Rasse DSN und damit zu einer besseren Akzeptanz der genomischen Selektion bei den DSN-Züchtern. Dies wiederum trägt zum Inzucht- und Diversitätsmanagement der Rasse bei. Es konnte gezeigt werden, dass mit der erzielten Lernstichprobe von über 3.000 DSN eine rassespezifische gZWS möglich ist, wenn auch die Genauigkeiten recht niedrig sind. Die Lernstichprobe kann zukünftig zudem von dem vit für eine rassespezifische gZWS oder für die gZWS im gemeinsamen Datenpool mit HOL genutzt werden. Die in **APIII** entwickelten Modelle für die gZWS können dabei als Basis fungieren. Wichtig ist, dass die Genotypisierung von Nachkommen und Jungtieren der Rasse DSN weiter vorangetrieben wird.

Mit dem Aufbau einer Pipeline zur Imputation von 50K und 200K Daten auf VGS-Niveau (**API**) wurde ein wichtiges Instrument geschaffen, welches für zukünftige wissenschaftliche Fragestellungen genutzt werden kann und soll, beispielsweise um Marker für weitere neue und innovative Merkmale der Adaptation, Futtereffizienz, Ressourceneffizienz und Gesundheit mit Hilfe von VGS-Daten zu identifizieren. Mögliche Merkmale in diesem Zusammenhang könnten Merkmale der Stoffwechselgesundheit (z.B. Ketose), Methanausscheidung, Wiederkauaktivität oder auch Merkmale der Schlachtkörperqualität sein, welche für das DSN als Zweinutzungsrasse von besonderer Bedeutung sind. Die entwickelte Pipeline zur Imputation der Genotypdaten könnte zudem Anwendung in anderen vom Aussterben bedrohten Rinderrassen mit kleiner Populationsgröße finden. Die in **APII** und **APIII** durchgeführten Markervergleiche (50K vs. 200K vs. VGS) in GWAS und in der gZWS zeigen zudem, dass VGS-Daten wichtige Informationen zu rassespezifischen Varianten in Rassen mit kleiner Populationsgröße liefern und daher zukünftig verstärkt genutzt werden sollten. Mittels „Optimum Genetic Contribution“-Methodik wurden Strategien für ein Diversitätsmonitoring und ein Inzuchtmanagement in der Rasse DSN erarbeitet. Die Ergebnisse dieser Analysen sind insbesondere für Zuchtverbände und DSN-Züchter von Interesse und könnten in gemeinsamer Kooperation weiterentwickelt werden. Auch zeigte sich in den Diversitätsanalysen, dass das DSN zwar eng verwandt ist mit HOL, dass es sich beim DSN aber um eine eigenständige Rasse handelt, sodass eine Einkreuzung von HOL nicht erforderlich und auch nicht anzuraten ist.

Datenmanagement:

Alle im Projekt generierten Daten werden lokal bei beiden Projektpartnern TZ-GI und TZ-BER gespeichert, sowie in Teilen zentral in Cloud-Speichersystemen der Rechenzentren beider Universitäten über einen begrenzten Zeitraum gespeichert. Die Genotypendaten und Sequenzdaten von DSN sind darüber hinaus bei vit (Verden) gespeichert. Die Daten und Ergebnisse des Projektes sind zudem im Rahmen der wissenschaftlichen Publikationen online publiziert und einsehbar. Die zugehörigen Links sind für jede einzelne wissenschaftliche Publikation im Folgenden separat gelistet. Das Design und die Auswahl der SNP-Marker für den DSN-SNP Chip ist unter folgendem Link veröffentlicht: <https://doi.org/10.1186/s12864-021-08237-2>.

Da es sich bei dem entwickelten DSN200K-SNP-Chip um ein Custom-Design handelt, werden Informationen zu dem Chip (z.B. Marker-Inhalt, Preis) durch Thermo Fisher Scientific nicht an Dritte weitergegeben. Bei Interesse werden detaillierte Informationen zum DSN200K-SNP-Chip diskriminierungsfrei durch die Projektkoordination zur Verfügung gestellt.

Übersicht über erfolgte oder geplante Veröffentlichungen

Wissenschaftliche Publikationen

Wolf, M.J., Neumann, G.B., Korcuć, P., Yin, T., Brockmann, G.A., König, S., May, K. (in press). Genetic evaluations for endangered dual-purpose German Black Pied cattle using 50K SNPs, a breed-specific 200K chip, and whole-genome sequencing. *Journal of Dairy Science* 106, 3345-3358. <https://doi.org/10.3168/jds.2022-22665>.

Neumann, G.B., Korcuć, P., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A. (2023). Genomic diversity and relationship analyses of endangered German Black Pied cattle (DSN) to 68 other taurine breeds based on whole-genome sequencing. *Frontiers in Genetics* 13. <https://doi.org/10.3389/fgene.2022.993959>

Neumann, G.B., Korcuć, P., Arends, D., Wolf, M.J., May, K., Reißmann, M., Elzaki, S., König, S., Brockmann, G.A (2021). Design and performance of a bovine 200 k SNP chip developed for endangered German Black Pied cattle (DSN). *BMC Genomics* 22, 905. <https://doi.org/10.1186/s12864-021-08237-2>

Wolf, M.J., Yin, T., Neumann, G.B., Korcuć, P., Brockmann, G.A., König, S., May, K. (2021). Genome-Wide Association Study Using Whole-Genome Sequence Data for Fertility, Health Indicator, and Endoparasite Infection Traits in German Black Pied Cattle. *Genes* 12(8), 1163. <https://doi.org/10.3390/GENES12081163>

Naderi, S., Moradi, M.H., Farhadian, M., Yin, T., Jaeger, M., Scheper, C., Korcuć, P., Brockmann, G.A., König, S., May, K. (2020). Assessing selection signatures within and between selected lines of dual-purpose black and white and German Holstein cattle. *Animal Genetics* 51, 391-408. <https://doi.org/10.1111/age.12925>

May, K., Weimann, C., Scheper, C., Strube, C., König, S. (2019). Allele substitution and dominance effects of *CD166/ALCAM* gene polymorphisms for endoparasite resistance and test-day traits in a small cattle population using logistic regression analyses. *Mammalian Genome* 30(9), 301-317. <https://doi.org/10.1007/s00335-019-09818-z>

Beiträge auf wissenschaftlichen Tagungen:

Internationale Tagungen

Korkuć, P., Neumann, G.B., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A. (2022) Improved genome-wide associations using a breed-specific 200K SNP chip for German Black Pied (DSN) cattle. Proceedings of the World Congress on Genetics Applied to Livestock Production, Rotterdam, 03.-08.07.2022. https://www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/66_008.pdf

May, K., Wolf, M.J., Neumann, G.B., Korkuć, P., Brockmann, G.A., König, S. (2022) Genetic evaluations and genomic characteristics for local cattle using genome sequences, 50K and a specific SNP chip. Proceedings of the World Congress on Genetics Applied to Livestock Production, Rotterdam, 03.-08.07.2022. https://www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/54_011.pdf

Neumann, G.N., Korkuć, P., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A. (2022). Phylogenetic analysis and nucleotide diversity of 69 cattle breeds including German Black Pied cattle using VGS. Proceedings of the World Congress on Genetics Applied to Livestock Production, Rotterdam, 03.-08.07.2022. https://www.wageningenacademic.com/pb-assets/wagen/WCGALP2022/67_017.pdf

Neumann, G.B., Korkuć, P., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A. (2022). Genomic diversity and relatedness between endangered German Black Pied cattle and 68 cattle breeds. 73rd Annual Meeting of European Federation of Animal Science, 05.-09.09.2022, Porto.

Wolf, M.J., Neumann, G.B., Korkuć, P., May, K., Brockmann, G.A., König, S., (2022). Genetic evaluations using a novel custom-made SNP chip for local cattle breeds. 73rd Annual Meeting of European Federation of Animal Science, 05.-09.09.2022, Porto.

Neumann, G.B., Korkuć, P., Arends, D., Wolf, M., May, K., König, S., Brockmann, G.A. (2021): A 200K SNP-Chip for the endangered German Black Pied cattle (DSN). 72nd Annual Meeting of European Federation of Animal Science, 30.08.-03.09.2021, Davos.

Neumann, G.B., Korkuć, P., Arends, D., Wolf, M., May, K., König, S., Brockmann, G.A. (2021): Content and design of the DSN200K SNP chip: a new tool to assist genetic diversity management in an endangered German cattle population. Thermo Fisher Scientific Virtual Workshop. Inn: 72nd Annual Meeting of European Federation of Animal Science, 30.08.-03.09.2021, Davos.

May, K., Scheper, C., Weimann, C., Strube, C., König, S. (2020). Strategy to infer candidate genes, allele substitutions and dominance for endoparasite resistance. 71th Annual Meeting of the European Federation of Animal Science, 01.-04.12.2020, Virtual Meeting.

Neumann, G.B., Korcuć, P., Arends, D., Brockmann, G.A. (2020). Functional Annotation of Gene Variants in Endangered Breeds: A Study in German Black Pied Cattle (DSN). 2nd FAANG-Europe Workshop on Functional Annotation of Animal Genomes, 11.-13.02.2020, Prag.

Nationale Tagungen

Brockmann, G.A. (2022) DSN - Eine alte und bedrohte Rinderrasse im Fokus der Wissenschaft. Messe EuroTier, Hannover, 15.-18.11.2022.

Brockmann, G.A., König, S., Buchholz, C., (2022) DSN: Informationen der Genanalyse, neue Merkmale & Zuchtwertschätzung, praktische Zuchtarbeit. BMEL Kongress zur Zucht und Erhaltung alter und bedrohter einheimischer Nutztierassen, Bonn, 26.-28.09.2022.

Brockmann, G.A. (2022). Was lesen wir aus dem Genom der DSN? 50 Jahre Genreservezucht Brandenburg, DSN-Fachtagung, Gräfendorf, 16.06.2022.

Korcuć, P., Neumann, G.B., Arends, D., Wolf, M.J., May, K., Koenig, S., Brockmann, G.A. (2022) Ein 200K SNP-Chip für das gefährdete Deutsche Schwarzbunte Niederungsring (DSN). DVG-Vet-Congress 2022, Berlin, 13.-15.10.2022.

Korcuć, P., Neumann, G.B., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A., (2022). Phylogenetische Analyse und Nukleotiddiversität von 69 Rinderrassen, einschließlich des Deutschen Schwarzbunten Niederungsringes. Verbesserte genomweite Assoziationen bei Verwendung eines rassespezifischen 200K SNP-Chips für Deutsche Schwarzbunte Niederungsring (DSN). Vortragstagung der Deutschengesellschaft für Züchtungskunde e.V. und der Gesellschaft für Tierzuchtwissenschaften e.V., 21.-22.09.2022, Kiel.

Neumann, G.B., Korcuć, P., Arends, D., Wolf, M.J., May, K., König, S., Brockmann, G.A., (2022). Phylogenetische Analyse und Nukleotiddiversität von 69 Rinderrassen, einschließlich des Deutschen Schwarzbunten Niederungsringes. Vortragstagung der Deutschengesellschaft für Züchtungskunde e.V. und der Gesellschaft für Tierzuchtwissenschaften e.V., 21.-22.09.2022, Kiel.

Wolf, M.J., May, K., Yin, T., Neumann, G.B., Korcuć, P., Brockmann, G.A., König, S., (2022). Genomische Zuchtwertschätzung beim Deutschen Schwarzbunten Niederungsring mit Sequenzdaten. Vortragstagung der Deutschengesellschaft für Züchtungskunde e.V. und der Gesellschaft für Tierzuchtwissenschaften e.V., 21.-22.09.2022, Kiel.

Korcuć, P. (2021) Forschungsergebnisse zu Milchleistung in DSN und Entwicklung eines DSN-spezifischen SNP-Chips. DSN-Kolloquium: Ergebnisse zur Forschung in DSN, Groß Kreutz, 27.10.2021, organisiert von TZ-BER.

Korkuć, P. (2021) DSN-spezifischer SNP-Chip und Ergebnisse zu Milchleistung in DSN. Fachgespräch „Erhaltung/ Biodiversität in kleinen Nutztierassen“, online, 26.10.2021, organisiert von TZ-BER.

May, K. (2021) Genomische Analysen für funktionale Merkmale und Aufbau einer genomischen Zuchtwertschätzung beim Deutschen Schwarzbunten Niederungsrind (DSN). Fachgespräch „Erhaltung/ Biodiversität in kleinen Nutztierassen“, online, 26.10.2021, organisiert von TZ-BER.

May, K., Wolf, M.J., Neumann, G.B., Yin, T., Korkuć, P., Brockmann, G.A., König, S. (2021): Nutzung von Vollgenomsequenzdaten zur Aufklärung genetischer Mechanismen und zur Entwicklung eines rassespezifischen SNP-Chips für die genomische Zuchtwertschätzung beim Deutschen Schwarzbunten Niederungsrind. Treffen des Genetisch-Statistischen Ausschusses in Kassel, 20.09.2021.

Qualifizierungsarbeiten:

Rühl, Alena. Genomweite Assoziationsstudie für Exterieurmerkmale beim Deutschen Schwarzbunten Niederungsrind. Bachelorthesis, Justus-Liebig-Universität Gießen, Januar 2022.

Lutz, Eliana. Vergleich von zwei SNP-Genotypisierungschips in genomweiten Assoziationsstudien für Milchleistungs- und Abkalbmerkmale beim Deutschen Schwarzbunten Niederungsrind. Masterthesis, Justus-Liebig-Universität Gießen, Dezember 2021.

Weitere geplante Veröffentlichungen

Es ist bei beiden Projektpartnern je eine Promotion geplant. Die Dissertation von Guilherme B. Neumann wird im 1. Quartal 2023 eingereicht. Die Dissertation von Manuel J. Wolf wird voraussichtlich zum 2. Quartal 2023 eingereicht.

Im Rahmen der Identifizierung von DSN-spezifischen Sequenzvarianten sollen Sequenzvarianten und deren potentielle Funktion publiziert werden, die beim Alignment mit dem Referenzgenom herausfallen. Die Ergebnisse liegen vor. Die Publikation wird durch TZ-BER vorbereitet.

Zu der im Rahmen von Arbeitspaket IV angewandten „optimum genetic contribution“- Methode ist eine zeitnahe Publikation durch TZ-GI in Bearbeitung.

4. Zusammenfassung

Bedrohte Rinderrassen, wie das DSN, profitieren bisher nicht von hocheffizienten Schätzverfahren auf genomischer Ebene, die in den großen Rassen wie etwa HOL und FV mittlerweile zur Routine gehören und für einen rasanten Zuchtfortschritt verantwortlich sind. Dadurch bleiben kleine Rassen hinsichtlich ihrer Wettbewerbsfähigkeit zunehmend zurück. Neben einer Steigerung des Zuchtfortschritts könnten genomische Informationen helfen die genomische Diversität innerhalb bedrohter Rassen im Blick zu halten und das Risiko von Inzuchtdepressionen zu vermindern. Standardmäßig werden Genotypisierungs-Chips auf Basis von SNPs mit Markerdichten von 10K, 50K oder auch 700K genutzt. Dabei sind die Marker auf den Chips so angelegt, dass vor allem Regionen mit Leistungsbezug (Milch, Fleisch) innerhalb der großen konventionellen Rinderrassen abgedeckt sind. Da jedoch die genetischen Architekturen zwischen Rassen durch die unterschiedlichen Entwicklungsgeschichten stark divergieren können, sind Genotypisierungs-Arrays, die auf den leistungsstarken Rassen basieren, gegebenenfalls für kleine, bedrohte Rassen ungeeignet.

Ein wesentliches Ziel des Vorhabens war daher der Aufbau einer aussagekräftigen Lernstichprobe mit genomischen Daten für die Rasse DSN. Anhand von VGS-Daten einer repräsentativen Stichprobe der DSN-Population sollte ein Chip designet werden, der geeignete Marker beinhaltet, die neben Produktionsmerkmalen auch neuartige Merkmale wie Hitzetoleranz und Endoparasitenresistenz berücksichtigen. Anhand des neuen Chips sollten weitere Analysen durchgeführt werden, um einen Mehrgewinn des rassespezifischen SNP-Chips für das DSN zu evaluieren. Im Rahmen dieser Untersuchungen sollten auch erstmalig genomische Zuchtwerte für die Rasse DSN geschätzt werden. Um das gesamte Potenzial der neugewonnenen Informationen auszuschöpfen sollte abschließend ein Monitoringverfahren getestet werden, um neben einer effektiven Steigerung der Zuchtwerte und somit einen Zugewinn an Konkurrenzfähigkeit auch die Diversität innerhalb der Rassen weitestgehend zu erhalten.

Für die Konstruktion des neuen SNP-Chips wurden VGS-Daten von 304 DSN erfolgreich mit einer selbst entwickelten Pipeline, in Anlehnung an das 1000-Bullen-Genom-Projekt, aufbereitet. Darüber hinaus wurde ein Imputationsansatz entwickelt, um DSN, die zuvor mit dem konventionellen Illumina 50K BeadChip genotypisiert wurden, auf VGS-Level zu imputieren. Dieser Datensatz mit Genotypeninformationen von ca. 2.100 DSN diente dazu, GWAS für Produktionsmerkmale und neuartige Merkmale durchzuführen. Die dabei detektierten assoziierten Marker wurden anschließend bei der Konstruktion des DSN200K-SNP-Chips berücksichtigt. Final beinhaltete der DSN200K-SNP-Chip 182.154 Varianten (SNPs und INDELS). Dabei wurden Varianten aus zehn verschiedenen Kategorien berücksichtigt, wie z.B. identifizierte Varianten aus GWAS mit DSN (2.071), DSN-spezifische Varianten (38.198), Varianten um möglichst viele HBs abzubilden (58.886), SNPs entsprechend eines

Abstammungspanel (554), funktionierende SNPs von dem Illumina BovineSNP50 BeadChip (34.039) und weitere. Insgesamt wurden im Rahmen des Projekts 1.488 Rinder mit dem neuen Chip genotypisiert. Darunter waren 1.201 von der Rasse DSN. Zusätzliche Rinderrassen wurden genotypisiert um zu testen, wie viele Varianten des neuen Chips in anderen Rassen, wie z.B. Rotes Höhenvieh, ANG, HOL oder RDN abgerufen werden können. Die Weiterentwicklungen des gegebenen Imputationsverfahrens für die DSN-Population ermöglichten es schließlich von 50K auf 200K (99,99 % (0,07 % SD)), 50K auf VGS (96,50 % (0,58 % SD)) und 200K auf VGS (98,58 % (0,21 % SD)) imputieren zu können. Final waren Datensätze in allen drei Markerdichten für 3.703 DSN verfügbar. Damit wurden erfolgreich vergleichende GWAS mit dem DSN200K-SNP-Chip und dem 50K Chip durchgeführt. Die GWAS mit den VGS Daten deckten eine sehr große Anzahl an merkmalsassoziierten Varianten auf, beispielsweise für das Merkmal SCS, für welches auf Basis des 50K und 200K Chips keine Assoziationen ermittelt werden konnten. Darüber hinaus wurde eine Zuchtwertschätzung entwickelt und für alle drei Markerdichten (50K, 200K, VGS) angewendet. Dabei wurden nur geringfügig höhere Erbllichkeiten und Genauigkeiten mit dem 200K Chip im Vergleich zum 50K Chip erzielt. Insgesamt muss die angewandte Zuchtwertschätzung jedoch als erfolgreich bewertet werden, da sie einen Modellcharakter hat und eine sehr gute Grundlage für weitere Studien bildet. Durch eine Weiterentwicklung in Form von Vergrößerung der Tierzahlen, auch durch das Einbeziehen anderer Rassen und der Anwendung komplexerer Modelle sind höhere Sicherheiten zu erwarten. Untersuchungen zur Nukleotid-Diversität innerhalb der Rasse zeigten, dass π_{total} in DSN geringfügig höher war als in HOL sowie 28 weiteren Nordeuropäischen Rassen. Das deutet darauf hin, dass trotz des Gefährdungsstatus der Rasse bereits ein gutes Management vorliegt. Der OGC-Ansatz macht es möglich eine Steigerung von bestimmten Zuchtwerten zu erreichen, aber dabei stets die Verwandtschaft innerhalb der Nachkommengeneration zu limitieren. Für das Leistungsmerkmal M305 konnte gezeigt werden, dass auch bei starker Limitierung der durchschnittlichen Verwandtschaft auf 0,03 Eltern zur Anpaarung ausgewählt werden können, deren Zuchtwerte im Durchschnitt >800 kg höher sind als in der restlichen Population. Im Rahmen des Projekts wurde durch die Sequenzierung von 15 ANG und 15 RDN zudem der Grundstein gelegt, um die Erkenntnisse des erfolgreich abgeschlossenen Projektes auf weitere heimische bedrohte Rassen zu übertragen.

Zusammenfassend lässt sich festhalten, dass die **technischen Voraussetzungen** für die Erprobung und Anwendung einer gZWS in der Rasse DSN erfüllt sind. Die Förderung des hier vorgestellten Vorhabens könnte daher einen wichtigen Baustein zur Überwindung bestehender ökonomischer Hürden darstellen, bei gleichzeitig hoher Wahrscheinlichkeit der direkten Umsetzung der Vorhabenziele in die konkrete Zuchtpraxis. Die Entwicklung einer gZWS in der bedrohten Rasse DSN kann somit einen wesentlichen Beitrag zur besseren Erhaltung und

nachhaltigen Nutzung der Rasse und somit der biologischen Vielfalt im Bereich der Rinderrassen leisten, was eindeutig die Modellhaftigkeit des Vorhabens widerspiegelt.

5. Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erreichten Zielen

Die Ziele wurden in allen Punkten erreicht.

Die im Verlauf des Projektes zu erreichenden Arbeitsziele waren:

- a) Auf- und Ausbau einer informativen Lernstichprobe in der Rasse DSN
- b) Entwicklung bzw. Weiterentwicklung einer Pipeline zur Imputation von SNP-Chip Daten (50K) auf Sequenzniveau
- c) Konstruktion eines custom-made SNP-Chips für die Rasse DSN unter Berücksichtigung von
 - i) „neuen Merkmalen“ der Adaptation und des Tierverhaltens und
 - ii) von Diversitätskennzahlen (u.a. Heterozygotiegrad über SNP-Marker)
- d) Erprobung des neu entwickelten DSN-spezifischen SNP-Chips mittels genetisch-statistischer Anwendungen
- e) Entwicklung und Umsetzung von Monitoringverfahren zur Sicherung der genetischen Diversität in der DSN-Population
- f) Evaluierung der Übertragbarkeit des DSN SNP-Chips auf andere Rassen
- g) Transfer der Ergebnisse in die Praxis

Die im Verlauf des Projektes erreichten Ziele sind:

Zu a) Für den Auf- und Ausbau der Lernstichprobe der Rasse DSN sollten 100 DSN vollgenomsequenziert werden, um darauf aufbauend DSN-spezifische Allele, Allelfrequenzen und Haplotypen ableiten zu können, welche später bei der Konstruktion des rassespezifischen SNP-Chips berücksichtigt werden sollten. Weiterhin sollten die Sequenzdaten der Imputation vorhandener DSN-Genotypen dienen, um die Lernstichprobe für die gZWS auszubauen. Entgegen der ursprünglich geplanten 100 DSN war es möglich, 214 DSN zu sequenzieren, wobei die Bandbreite der DSN-Population sowohl genetisch als auch phänotypisch sehr genau abgebildet wurde. Berücksichtigt wurden bei der Auswahl Zuchtwerte für die Milchleistung sowie Gesundheitsmerkmale, d.h. es wurden sowohl resistente als auch anfällige Tiere für Erkrankungen wie Mastitis und Endoparasiten-Infektionen inkludiert. Die Sequenzierung einer höheren Anzahl von 214 DSN entgegen der ursprünglich geplanten 100 DSN begründete sich durch ein günstiges Sequenzierungs-Angebot von der Firma Novogene.

Zu b) Es wurde erfolgreich eine Pipeline zur Imputation der vorhandenen DSN 50K Genotypen auf VGS-Niveau erarbeitet werden. Dafür wurden bereits vorliegende DSN-Sequenzdaten von TZ-BER mit integriert, sodass das Referenzpanel für die Imputation bei 304 DSN Tieren lag. Final konnte eine sehr hohe Imputationsgenauigkeit von über 97 % erzielt werden.

Zu c) Ursprünglich war geplant, einen 50K SNP-Chip für die Rasse DSN zu entwickeln. Aufgrund sinkender Kosten für die Konstruktion eines custom-made SNP-Chips ist es gelungen,

einen 200K SNP-Chip mit deutlich höherer Markerdichte zu konstruieren. Dadurch konnte eine deutlich höhere Anzahl rassespezifischer Varianten sowie solche des bestehenden 50K Illumina Chips, die in der Rasse DSN informativ sind, auf dem neuen Chip berücksichtigt werden. Das Ziel war es insbesondere, SNP-Marker für Alleinstellungsmerkmale und „neue Merkmale“ zur Adaptation und des Tierverhaltens, sowie zur Krankheitsresistenz (u.a. Mastitis, Endoparasitenresistenz) auf dem Chip zu berücksichtigen. Hierbei konnten 2.071 Varianten aus GWAS mit Genotyp- und Sequenzdaten, die mit neuen Merkmalen zur Adaptation und Krankheitsresistenz assoziiert waren, erfolgreich auf dem neuen 200K SNP-Chip berücksichtigt und abgerufen werden.

Zu d) Es wurden erfolgreich vergleichende GWAS und eine gZWS mit Kreuzvalidierung mit dem herkömmlichen 50K Chip und dem neu konstruierten DSN200K-SNP-Chip durchgeführt. Die Ergebnisse zeigten, dass insbesondere für moderat bis hoch erbliche Merkmale (z.B. Milchleistung, Exterieur) SNPs mit dem neuen 200K Chip in den GWAS detektiert werden, welche unter Verwendung des 50K Chips nicht assoziiert sind. Die Ergebnisse der gZWS zeigten niedrige bis moderate Genauigkeiten, welche mit dem neuen 200K Chip nur geringfügig besser waren. Dieses Ergebnis deutet darauf hin, dass neben der Markerdichte auch weitere Strategien zur Verbesserung der gZWS in kleinen vom Aussterben bedrohten Populationen wie dem DSN zu berücksichtigen sind (z.B. weitere Vergrößerung der Lernstichprobe mit DSN/DSN-Verwandten Tieren aus anderen Ländern). Wie geplant wurden im Rahmen der gZWS verschiedene Verfahren (BLUP, Bayes) getestet. Da diese in ersten Rechenläufen zu sehr ähnlichen Ergebnissen führten, erfolgten die Berechnungen für alle hier vorgestellten Merkmale mit dem BLUP-Verfahren. Neben dem eigentlichen Ziel, nämlich dem Vergleich des 50K und 200K Chips in GWAS und der gZWS, wurden darüber hinaus alle Rechenläufe mit den VGS-Daten durchgeführt, um deren zukünftigen Nutzen in praktischen Anwendungen für die Rasse DSN zu evaluieren.

Zu e) Unter Anwendung des „optimum genetic contribution“ Verfahrens wurden verschiedene Szenarien in Richtung eines optimalen Monitoringverfahrens zur Sicherung der genetischen Diversität in der Rasse DSN erprobt. Die Ergebnisse zeigen, dass es möglich ist eine Nachkommengeneration und somit auch Bullen zu erzeugen, die Potenzial für genetischen Fortschritt für das Leistungsmerkmal M305 vorweist, ohne eine Steigerung der Verwandtschaft in Kauf zu nehmen. Zudem zeigt sich, dass Steigerung in den Zuchtwerten beim DSN mit Verzicht auf Einkreuzung von anderer Genetik, ohne dabei die Inzucht innerhalb der DSN-Population zu erhöhen, möglich ist.

Zu f) Die Übertragbarkeit des DSN200K-SNP-Chips auf andere Rassen wurde anhand der Genotypisierung von 287 Rindern aus neun verschiedenen *Bos taurus* und einer *Bos indicus* Rasse durchgeführt. Dabei zeigte sich, dass der neu konstruierte DSN200K-SNP-Chip sehr gut auf viele weitere vom Aussterben bedrohte Rinderrassen übertragen werden kann.

Zu g) Ein Transfer der Ergebnisse in die Praxis war im Rahmen von zwei finalen Veranstaltungen geplant. Entgegen dieses ursprünglichen Zieles war ein erfolgreicher Transfer der Ergebnisse in Wissenschaft und Praxis im Rahmen zahlreicher nationaler und internationaler Veranstaltungen möglich (siehe Material und Methoden: Transfer der Ergebnisse in die Praxis).

6. Literaturverzeichnis

- Aguilar, I. 2018. Creation of genomic relationship matrices and BLUPF90 "user file". http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=user_file.pdf. Zugriff: 15.06.2022.
- Bergk, N., & Swalve, H. H. (2011). Der Fett-Eiweiß-Quotient in der Früh lactation als Indikator für den Verbleib von Erstkalbinnen in der Milchkuhherde. *Züchtungskunde* 83.2: 89-103.
- Bohlouli, M., Halli, K., Yin, T., Gengler, N., & König, S. (2022). Genome-wide associations for heat stress response suggest potential candidate genes underlying milk fatty acid composition in dairy cattle. *Journal of Dairy Science*, 105(4), 3323-3340.
- Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *American Journal of Human Genetics*, 103(3), 338–348.
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4(1), s13742-015.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., ... & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2), giab008. Felius et al., 2014
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., ... & Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *science*, 296(5576), 2225-2229.
- Haller, M. (2015) *Alte Haus- & Nutztierassen neu entdeckt*, Graz, Stuttgart, Leopold Stocker Verlag.
- Hayes, B. J., & Daetwyler, H. D. (2019). 1000 bull genomes project to map simple and complex genetic traits in cattle: applications and outcomes. *Annual review of animal biosciences*, 7, 89-102.
- Iso-Touru, T., Sahana, G., Guldbandsen, B., Lund, M. S., & Vilkki, J. (2016). Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. *BMC genetics*, 17(1), 1-12.
- Jaeger, M. (2018). Phenotypic, quantitative genetic and genomic characterization of the German black and white dual-purpose cattle breed. (Dissertation, Agrarwissenschaften) Justus-Liebig-Universität Gießen. <http://geb.uni-giessen.de/geb/volltexte/2019/14000>.

- Koenig, S., & Simianer, H. (2006). Approaches to the management of inbreeding and relationship in the German Holstein dairy cattle population. *Livestock Science* 103(1-2), 40-53.
- Korkuć, P., Arends, D., & Brockmann, G. A. (2019). Finding the optimal imputation strategy for small cattle populations. *Frontiers in Genetics*, 10(FEB), 1–10.
- Korkuć, P., Arends, D., May, K., König, S., & Brockmann, G. A. (2021). Genomic loci affecting milk production in German Black Pied cattle (DSN). *Frontiers in Genetics*, 12, 640039.
- Kranis, A., Gheyas, A. A., Boschiero, C., Turner, F., Yu, L., Smith, S., ... & Burt, D. W. (2013). Development of a high density 600K SNP genotyping array for chicken. *BMC genomics*, 14(1), 1-13.
- Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research*, 47(W1), W256-W259.
- Malomane, D. K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A. R., & Simianer, H. (2018). Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC genomics*, 19(1), 1-16.
- May, K., Scheper, C., Brügemann, K., Yin, T., Strube, C., Korkuć, P., Brockmann, G. A., & König, S. (2019). Genome-wide associations and functional gene analyses for endoparasite resistance in an endangered population of native German Black Pied cattle. *BMC genomics*, 20(1).
- May, K., Bohlsen, E., König, S., & Strube, C. (2020). *Fasciola hepatica* seroprevalence in Northern German dairy herds and associations with milk production parameters and milk ketone bodies. *Veterinary parasitology*, 277, 109016.
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., ... & Cunningham, F. (2016). The ensembl variant effect predictor. *Genome biology*, 17(1), 1-14.
- Meier, S., Arends, D., Korkuć, P., Neumann, G. B., & Brockmann, G. A. (2020). A genome-wide association study for clinical mastitis in the dual-purpose German Black Pied cattle breed. *Journal of dairy science*, 103(11), 10289-10298.
- Mészáros, G., Boison, S. A., Pérez O'Brien, A. M., Ferenčaković, M., Curik, I., Da Silva, M. V. B., ... & Sölkner, J. (2015). Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. *Frontiers in Genetics*, 6, 173.
- Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2014). Manual for BLUPF90 family of programs. Zugriff: 21.06.2021. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all7.pdf.

- Meuwissen, T. H. (2002). GENCONT: an operational tool for controlling inbreeding in selection and conservation schemes. In Proceedings of the 7th Congress on Genetics Applied to Livestock Production (pp. 19-23).
- Mügge, B. (1999) Deutsche Holsteins: Die Geschichte einer Zucht, Stuttgart, Ulmer.
- Neumann, G. B., Korkuć, P., Arends, D., Wolf, M. J., May, K., Reißmann, M., ... & Brockmann, G. A. (2021). Design and performance of a bovine 200 k SNP chip developed for endangered German Black Pied cattle (DSN). *BMC genomics*, 22(1), 905.
- Pausch, H., MacLeod, I. M., Fries, R., Emmerling, R., Bowman, P. J., Daetwyler, H. D., & Goddard, M. E. (2017). Evaluation of the accuracy of imputed sequence variant genotypes and their utility for causal variant detection in cattle. *Genetics Selection Evolution*, 49, 1-14.
- Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, 79(387), 575-583.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3), 559-575.
- Rahmann, G. (2011). Biodiversity and organic farming: what do we know?. *vTI Agriculture and Forstery Research*, 3, 189-208.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Schmidtman, C., Schönherz, A., Guldbrandtsen, B., Marjanovic, J., Calus, M., Hinrichs, D., & Thaller, G. (2021). Assessing the genetic background and genomic relatedness of red cattle populations originating from Northern Europe. *Genetics Selection Evolution*, 53(1), 1-18.
- Shiina, T., Blancher, A., Inoko, H., & Kulski, J. K. (2017). Comparative genomics of the human, macaque and mouse major histocompatibility complex. *Immunology*, 150(2), 127-138.
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *American journal of human genetics*, 88(1), 76–82.
- Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., & Price, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nature genetics*, 46(2), 100–106.
- Yin, T., & König, S. (2019). Genome-wide associations and detection of potential candidate genes for direct genetic and maternal genetic effects influencing dairy cattle body weight at different ages. *Genetics Selection Evolution*, 51(1), 4.

- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of dairy science*, 91(11), 4414-4423.
- Wu, X., Lund, M. S., Sahana, G., Guldbandsen, B., Sun, D., Zhang, Q., & Su, G. (2015). Association analysis for udder health based on SNP-panel and sequence data in Danish Holsteins. *Genetics Selection Evolution*, 47(1), 1-14.
- Zucht- und Besamungsunion Hessen eG (ZBH) (2019). Zuchtprogramm der Zucht- und Besamungsunion Hessen eG - Deutsches Schwarzbuntes Niederungsrind. https://www.qnetics.de/fileadmin/downloads/Zuchtprogramme/Hessen_Milch_und_Zweinutzung/Zuchtprogramm_DSN_31.05.2019.pdf. Stand: 31.05.2019.
- Zumwald, J., Braunschweig, M., Nemecek, T., Schüpbach, B., Jeanneret, P., Roesch, A., Hofstetter, P., & Reidy, B. 2018. Life cycle assessment of grassland-based dairy production systems in Switzerland. Sustainable meat and milk production from grasslands: proceedings of the 27th General Meeting of the European Grassland Federation. *Grassland Science in Europe*. 23: 541-543.