

Aktuelle Entwicklungen zur genomischen Selektion

Reinhard Reents

Vereinigte Informationssysteme Tierhaltung w.V. (vit), Verden

Einleitung

Die Analyse von molekulargenetischen Daten erfreut sich eines starken Interesses. Die Aktivitäten in der Humangenetik stimulieren eine sehr intensive Forschungs- und Entwicklungstätigkeit sowohl bei öffentlich geförderten Forschungseinrichtungen als auch privaten Anbietern von Labortechnik zur Genotypisierung. In der Humangenetik werden zunehmend dichtere ‚Chips‘ verwendet, die zur Aufklärung des genetischen Hintergrundes vieler Krankheiten verwendet werden. Komplementär dazu werden auch für Nutztiere (inkl. Haustiere) immer mehr Informationen über das Genom verfügbar. Anders als in der Humangenetik gibt es bei Nutztierpopulationen bereits über Generationen erhobene phänotypische Daten. Diese Daten wurden bisher in konventionellen Zuchtprogrammen für Selektionszwecke verwendet. Mit Hilfe einer Verknüpfung dieser phänotypischen Daten und aktuell erhobener Daten zum Genom kann die Selektion sehr viel effizienter gestaltet werden (Schaeffer 2006). Zusätzlich können Merkmalskomplexe, die bisher über die konventionelle Selektion schwer zu bearbeiten waren, mit der genomischen Selektion verbessert werden.

In diesem Beitrag sollen die aktuellen Entwicklungen zur genomischen Selektion dargestellt werden. Die Methodik funktioniert theoretisch bei allen Nutztieren. Da die Implementierung im Bereich der Milchrinderzucht am weitesten fortgeschritten ist, werden aus der praktischen Anwendung einige Beispiele beleuchtet, um mögliche Einsatzgebiete in der Pferdezucht abzuleiten.

Methodik der genomischen Selektion

Stark vereinfacht sind folgende Entwicklungsschritte zur Einführung einer Genomischen Selektion nötig:

1. Entwicklung eines Standardpanels mit geeigneten SNPs (Single Nucleotide Polymorphisms)

2. Typisierung einer Lernstichprobe, d.h. Feststellung der SNP-Genotypen für eine ausreichende Anzahl von Tieren mit phänotypischen Beobachtungswerten
3. Erarbeitung einer Auswertungsmethodik, die den Effekt der Genotypen auf die Merkmalsausprägung zuverlässig schätzt (= genomische ZWS)
4. Validierung dieser genomischen ZWS an einer geeigneten Validierungsstichprobe
5. Schätzung von genomischen Zuchtwerten für potentielle Selektionskandidaten (kann bereits sehr früh im Leben des Tieres passieren, aktuell bereits an Embryonen) und Paarung der besten Zuchttiere

Der Status quo für diese 5 Schritte soll im Folgenden näher beschrieben werden.

1. Genotypisierung

Inzwischen liegen Informationen über das Genom von fast allen wichtigen Nutztierarten vor (Übersicht 1 im Anhang). Obwohl die Kosten für eine komplette Sequenzierung eines Genoms drastisch gefallen sind (siehe Abbildung 1), ist auf absehbare Zeit die Entwicklung eines sogenannten SNP Panels (oft auch als SNP Chip bezeichnet) nötig. In diesen Panels werden über alle Chromosomen hinweg möglichst gleichmäßig Einzelbasenpaare identifiziert, die im Sinne der Mendelschen Vererbungsregeln Heterozygotie aufweisen. Übersicht 2 im Anhang zeigt, dass inzwischen für alle Nutztierarten kommerziell verfügbare SNP Panels vorliegen, die auch eine ausreichende Abdeckung des Genoms gewährleisten (ca. 1500 – 2000 SNPs pro Chromosom). Weiterhin zeigt sich, dass bei intensiver Nachfrage von den Anbietern auch maßgeschneiderte SNP Panels erstellt werden, wie z.B. der Euro10K Chip des EuroGenomics Consortiums (Abbildung 2). Hier haben sich mehrere mitteleuropäische Länder zusammengetan, um die für das Milchrind aussagekräftigsten SNPs auf einen Standardchip zu bringen, der kommerziell so interessant ist, dass ein noch breiterer Einsatz in der Milchrinderzucht ermöglicht wird.

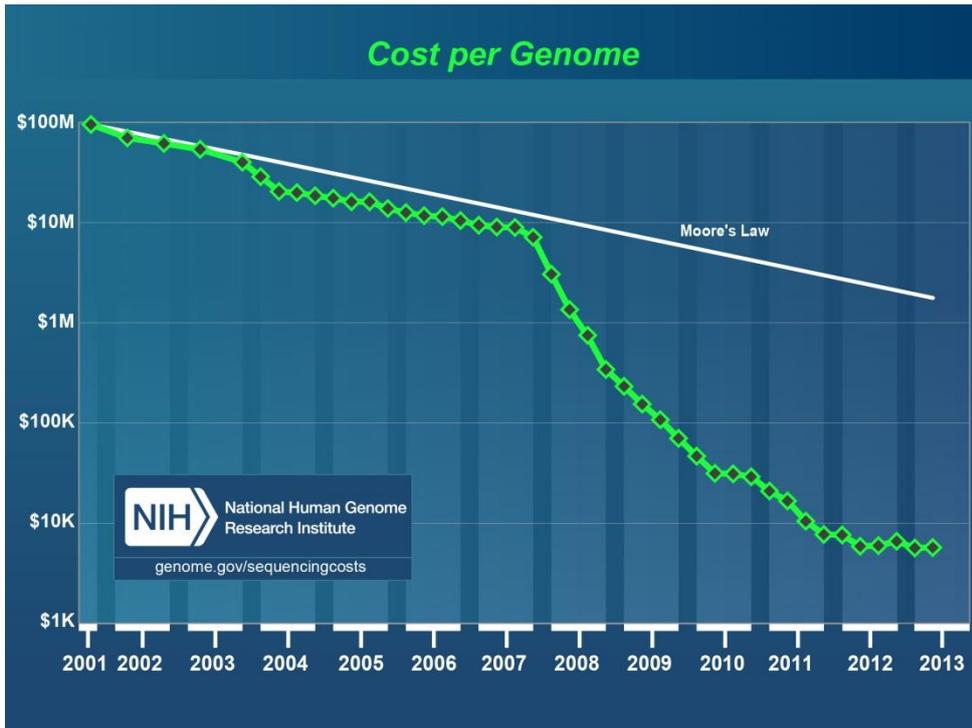


Abb. 1. Analysekosten einer Komplettsequenzierung im Zeitablauf für ein ca. 3000 Mb Paare großes Genom, Quelle: Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) online verfügbar unter <http://www.genome.gov/sequencingcosts>. Zugriff am 9. Januar 2014

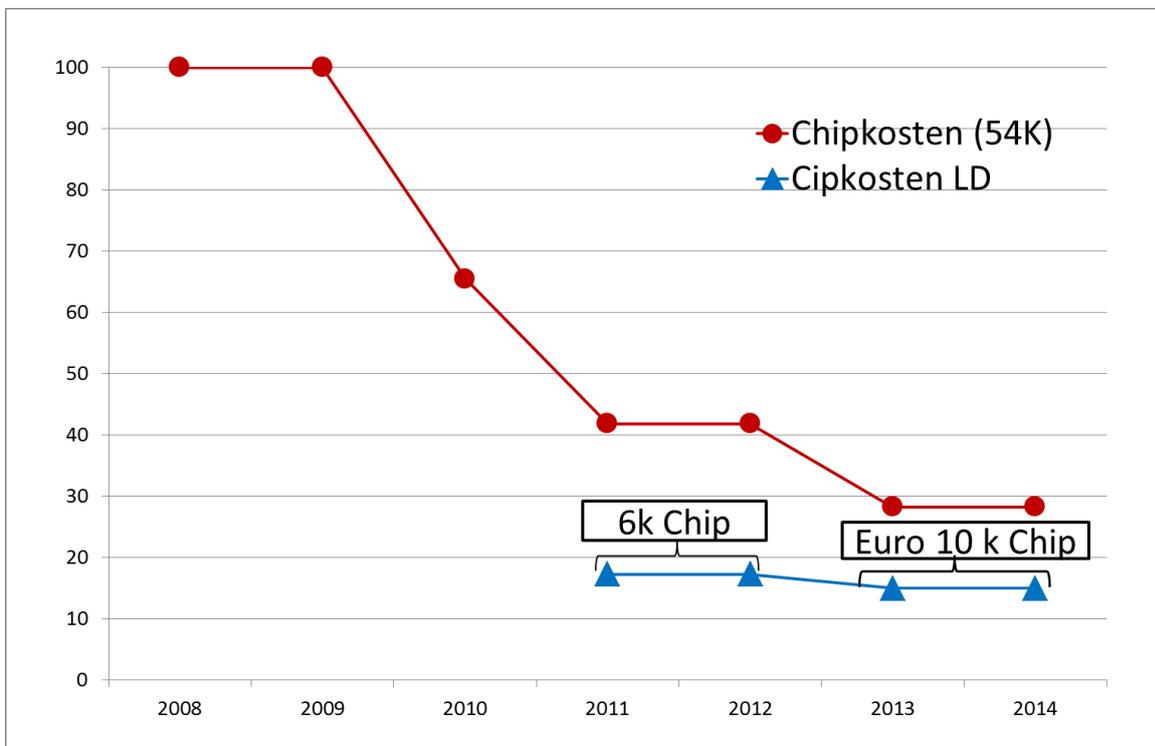


Abb. 2. Kosten für SNP Chips für Milchrinder, ausgedrückt in Relation zum Preis des Illumina Standard 54K Chip im Jahr 2008/2009

2. Typisierung von Lernstichproben

Für ‚etablierte‘ Merkmale, für die bereits eine historische Datensammlung vorliegt, hat es sich bewährt, genau die Tiere zu genotypisieren, die bereits genau geschätzte konventionelle Zuchtwerte haben. In der Milchrinderzucht liegen so für das etablierte Merkmalspektrum Lernstichproben in einer Größenordnung von über 20.000 sehr genau geschätzten Besamungsbullen vor (Liu et al., 2013a). Für neu zu erhebende Merkmale versucht man zunehmend, eine direkte Genotypisierung der Tiere vorzunehmen, die auch Phänotypen aufweisen. Das erfordert allerdings komplett andere Auswertungssoftware (auch als One-Step-Approach bezeichnet), die derzeit Rechenzentren noch vor modell- und rechentechnische Herausforderungen stellt (Liu et al, 2013b).

Die Reitferdezucht hat anderen Spezies gegenüber einen Vorteil darin, dass seit einigen Jahren routinemäßig von fast allen Fohlen Gewebematerial (meist in Form von Haarwurzeln) vorliegt, und somit die erhobenen Phänotypen mit Genominformationen verknüpft werden können. Es setzt sich immer mehr die Erkenntnis durch, dass aus DNA, die aus Haarwurzeln gewonnen wurde, eine korrekte Genotypisierung durchgeführt werden kann. So sind zum Beispiel in den USA bei fast 200.000 Milchrindern in 60% der Fälle die SNP Daten aus Haarwurzeln ermittelt worden, in 23% aus Gewebeproben (z.B. Ohrstanzen), zu 14% aus Blut und der Rest aus Nasenschleim und Sperma (Wiggans, 2013).

3. Auswertungsmethodik

In der wissenschaftlichen Literatur sind inzwischen mehrere Verfahren beschrieben, mit denen sichere genomische Zuchtwerte geschätzt werden können. Es hat sich allerdings gezeigt, dass auch eine ausgereifte Auswertungsmethodik eine intensive Anpassung an die speziellen Gegebenheiten der zu analysierenden Population erfordert (Liu et al, 2010). Abbildung 3 zeigt die Sicherheit einer genomischen ZWS im Vergleich zu anderen ZWS Verfahren auf, die bei Verwendung einer ausreichend großen Lernstichprobe und einer ausgereiften Auswertungsmethodik erzielt werden kann. In der praktischen Anwendung wird inzwischen bei der Genotypisierung von Selektionskandidaten auch mit einem kostengünstigeren ‚Low Density‘ (LD) Chip gearbeitet (siehe Übersicht 2 im Anhang). Diese LD Daten werden in einem sogenannten Imputing auf ein dichteres Standardpanel hochgerechnet (bei Milchrindern 54K), so dass die daran ermittelten Schätzformeln des dichteren Standardchips angewendet werden

können. Auswertungen haben ergeben, dass der Verlust an Sicherheit zu vernachlässigen ist und somit eine weitere Effizienzverbesserung des Gesamtsystems erzielt werden kann (Segelke et al, 2012).

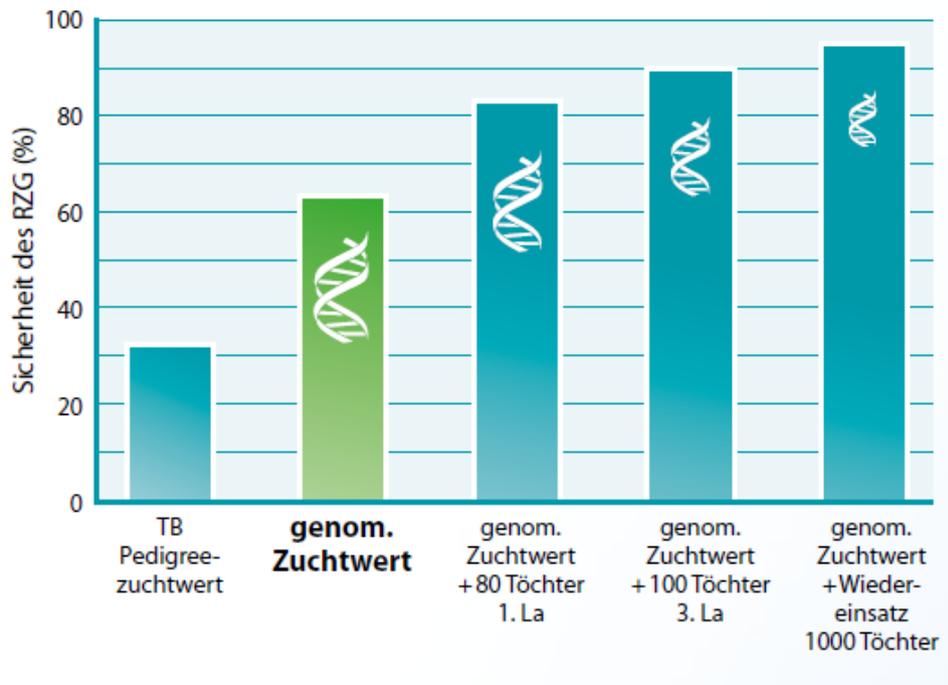


Abb. 3. zeigt die Sicherheiten einer genomischen ZWS im Vergleich mit einer ZWS, die nur auf Pedigreeinformationen beruht sowie die Sicherheiten aus einem umfangreichen Nachkommentest.

4. Validierung

Die Erfahrung aus der Anwendung der genomischen ZWS bei unterschiedlichen Milchrinder-Populationen, aber auch unterschiedlichen Spezies hat gezeigt, dass nur eine Validierung der geschätzten genomischen Zuchtwerte anhand von tatsächlich erbrachten Selektionsergebnissen eine realistische Einschätzung der Güte des genomischen ZWS Verfahrens zulässt. Deshalb wurden in der Milchrinderzucht anhand der inzwischen vorliegenden Töchterleistungen konventionelle Zuchtwerte geschätzt und diese in Relation gesetzt zu den genomischen Zuchtwerten dieser Bullen OHNE Töchterinformationen. Abbildung 4 zeigt, dass die genomische ZWS sehr gut frühzeitig die überlegenen Zuchttiere identifiziert hat. Inzwischen gibt es wissenschaftlich anerkannte Methoden, mit deren Hilfe man die Güte von genomischen ZWS Systemen objektiv einschätzen kann (Mäntysaari et al, 2010). Interbull bietet im Rahmen der Qualitätssicherung der genomischen Zuchtwerte von Besamungsbullen dieses Verfahren auf EU Ebene an, d.h., in breitem Maßstab darf Samen von jungen,

genomisch geprüften Bullen in der EU nur vertrieben werden, wenn eine Validierung des entsprechenden ZWS Systems durch Interbull erfolgt ist.

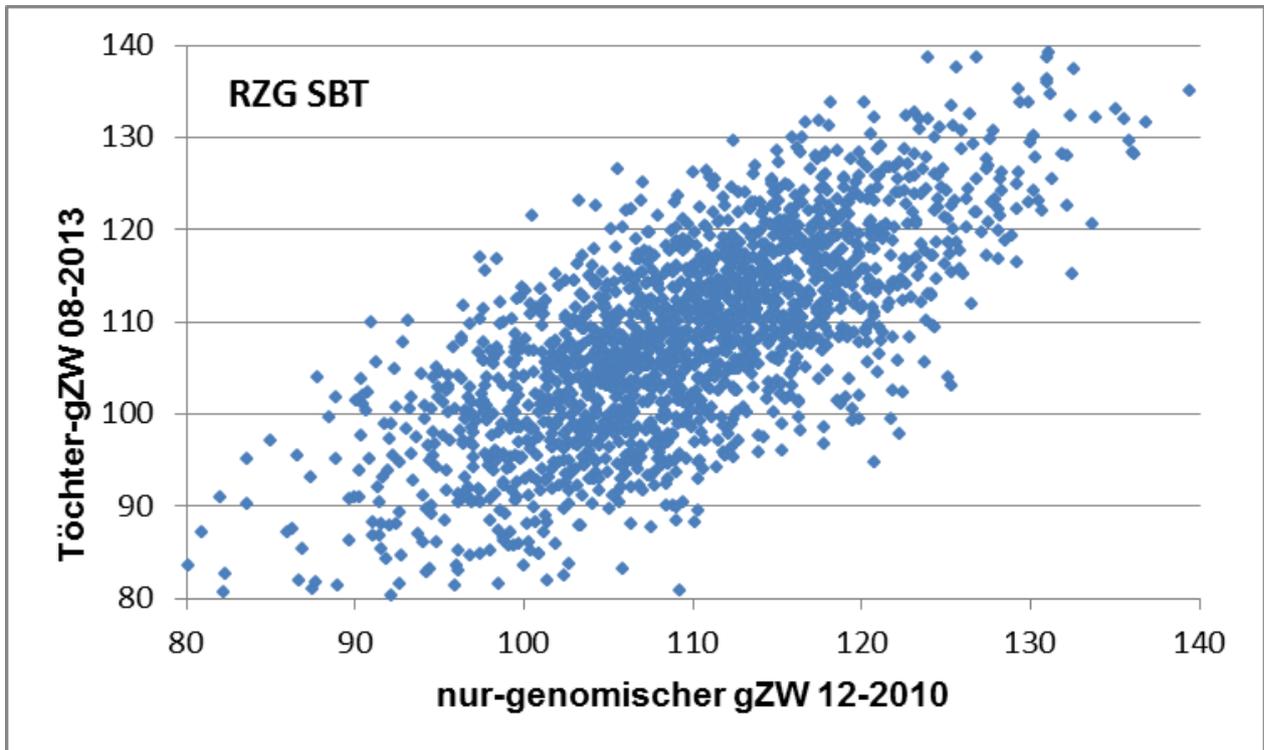


Abb. 4.: Validierung von genomisch geschätzten Zuchtwerten für 1785 SBT KB Bullen aus dem Dezember 2010 in Relation zu konventionellen Töchterzuchtwerten aus dem August 2013 (Geburtsjahre der Bullen 2006-2008), Merkmal: Gesamtzuchtwert RZG

5. Durchführung der Genomischen ZWS / Anwendung der genomischen Selektion

Im herkömmlichen System folgt die Durchführung der konventionellen ZWS konsequenterweise direkt dem Vorhandensein von neuen phänotypischen Informationen. Deshalb wird z.B. die ZWS in der Pferdezucht einmal jährlich im Herbst durchgeführt, weil die saisonal anfallenden Daten bis dahin aufbereitet sind und für eine ZWS zusammengeführt wurden. In der Milchrinderzucht hat sich ein Intervall von vier Monaten zwischen den konventionellen ZWS Terminen eingebürgert. Die genomische ZWS als Routineanwendung ist vom Anfall von neuen phänotypischen Daten entkoppelt. Deshalb hat sich international ein monatlicher Rhythmus etabliert, in welchem die in den vier vorangegangenen Wochen genotypisierten Tiere genomische Zuchtwerte erhalten. Viele Zuchtverbände wünschen ein noch kürzeres Intervall, weil die Ankaufentscheidungen für die wenigen selektierten Bullenkälber (einer aus 20-25 Kandidaten) so früher getätigt und somit

Große Perspektiven bietet die genomische Selektion demnach bei Tierarten, die folgende Bedingungen erfüllen:

- langes Generationsintervall
- hohe Vermehrungsrate einzelner Zuchttiere
- teure Leistungsprüfung
- hochentwickelte Datenerfassung, insbesondere, wenn diese standardisiert über Ländergrenzen hinweg vorgenommen wird (internationale Referenzpopulationen)

Allgemein kann man schlussfolgern, dass zukünftig das Vorhandensein von Genomdaten kein beschränkender Faktor mehr darstellen wird, sondern vielmehr die notwendigen Phänotypen begrenzend sind.

Ausblick für die Pferdezucht

Die Kenndaten eines Pferdezuchtprogrammes sind dadurch gekennzeichnet, dass ein relativ langes Generationsintervall vorliegt, Hengste durch die Besamung eine hohe Vermehrungsrate haben und die züchterisch bearbeiteten Merkmale relativ aufwändig erhoben werden müssen. Aus diesem Grunde würde man ein gutes Kosten-Nutzen Verhältnis der genomischen Selektion erwarten. Allerdings besteht ein Nachholbedarf in der Erfassung der benötigten Phänotypen, um eine ausreichend große Lernstichprobe zu generieren. Deshalb sind aktuelle Entwicklungen vielversprechend, die die in vielfältiger Weise erhobenen Daten zu Gesundheitsmerkmalen des Pferdes zusammenführen möchten (Sarnowski et al., 2014). Könnte doch hier ein Merkmalskomplex züchterisch bearbeitet werden, der die Pferdezucht derzeit mit einem sehr hohen Kostenfaktor belastet.

Zusammenfassung

Die genomische Selektion hat sich in der Milchrinderzucht etabliert und hat zu Ergebnissen geführt, die von der Zuchtpraxis inzwischen voll anerkannt und zum eigenen Vorteil genutzt werden. Andere Nutztierarten wie Schwein und Geflügel arbeiten ebenfalls konkret an der Nutzung dieser zusätzlichen Informationsquelle. Die Vorleistungen, die in diesen Bereichen erbracht wurden und werden gilt es auch in der Pferdezucht sinnvoll einzusetzen. Im Hinblick auf eine zukünftig starke Stellung im internationalen Wettbewerb sind verstärkt Anstrengungen zu unternehmen, um den vergleichsweise hohen Organisationsgrad der deutschen Pferdezüchter für konzertierte Aktionen zur Einführung einer genomischen Selektion zu nutzen.

Literatur

- Liu Z., F. R. Seefried, F. Reinhardt, S. Rensing, G. Thaller, R. Reents (2010) Dairy Cattle Genetic Evaluation Enhanced with Genomic Information (ID009), 9th WCGALP, 1.-6. August 2010, Leipzig.
- Liu Z., G. P. Aamand, S. Fritz, C. Schrooten (2013a) Comparison of national genomic predictions of EuroGenomics exchanged young bulls. *Interbull Bulletin* 47:38-42.
- Liu Z., M. E. Goddard, F. Reinhardt, R. Reents. (2013b) Computing strategies for a single step SNP model with an across country reference population. 64. Jahrestagung der EAAP, 26.-30. August 2013, Nantes / Frankreich.
- Mäntysaari E., Z. Liu, P. M. VanRaden (2010) Interbull validation test for genomic evaluations. *Interbull Bulletin* 41:10-14.
- Sarnowski S., K. F. Stock, E. Kalm, R. Reents (2014) Aufbau einer Pferde-Gesundheitsdatenbank. 7. Pferde-Workshop Uelzen am 18.-19. Februar 2014..
- Segelke D, J. Chen, Z. Liu, F. Reinhardt, G. Thaller, R. Reents (2012) Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips, *J. Dairy Sci.* 95 :5403–5411.
- Schaeffer L. R. (2006) Strategy for applying genomewide-selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218-223.
- Stock K. F., R. Reents (2013) Genomic selection: status in different species and challenges for breeding. *Reprod Dom Anim* 48 (Suppl. 1), 2-10.
- Wiggans G. (2013) Current status of genomic evaluation for U.S. dairy cattle, <http://aipl.arsusda.gov/publish/present.htm>

Übersicht 1: Verfügbare Informationen über das Genom von wichtigen Nutztierarten (Zusammenstellung aus öffentlichen Datenbanken des National Center for Biotechnology Information (NCBI) und des Ensembl project), Größe des Genoms (total sequence length) und Anzahl der Single Nucleotide Polymorphism (SNP; $N_{\text{SNP_NCBI}}$, $N_{\text{SNP_Ens}}$).

Spezies, Jahr der Sequenzierung	Genomsequenz-Information (aktueller Eintrag mit NCBI GenBank assembly ID) ¹	Genomgröße ¹ , Gesamtzahl SNPs ^{2,3*}
Rind (<i>Bos taurus</i>), 2009	Bos_taurus_UMD_3.1 (GCA_000003055.3), Center for Bioinformatics and Computational Biology (CBCB) at University of Maryland, Nov. 2009	2.670.422.299; $N_{\text{SNP_NCBI}}$ 22.055.952, $N_{\text{SNP_Ens}}$ 21.759.918
Schwein (<i>Sus scrofa</i>), 2009	Sscrofa10.2 (GCA_000003025.4), The Swine Genome Sequencing Consortium, Sept. 2011	2.808.525.991; $N_{\text{SNP_NCBI}}$ 28.665.189, $N_{\text{SNP_Ens}}$ 28.650.266
Schaf (<i>Ovis aries</i>), 2008	Oar_v3.1 (GCA_000298735.1), International Sheep Genome Consortium, Sept. 2012	2.619.054.388; $N_{\text{SNP_NCBI}}$ 2.721.261, $N_{\text{SNP_Ens}}$ 694.992
Ziege (<i>Capra hircus</i>), 2012	CHIR_1.0 (GCA_000317765.1), International Goat Genome Consortium, Jan. 2013	2.635.848.900; $N_{\text{SNP_NCBI}}$ 60.094
Pferd (<i>Equus caballus</i>), 2007	EquCab2.0 (GCA_000002305.1), The Genome Assembly Team, Okt. 2007	2.474.929.062; $N_{\text{SNP_NCBI}}$ 4.973.981, $N_{\text{SNP_Ens}}$ 1.154.177
Hund (<i>Canis lupus familiaris</i>), 2003	CanFam3.1 (GCA_000002285.2), Dog Genome Sequencing Consortium, Nov. 2011	2.410.976.875; $N_{\text{SNP_NCBI}}$ 3.298.379, $N_{\text{SNP_Ens}}$ 3.225.735
Huhn (<i>Gallus gallus</i>), 2004	Gallus_gallus-4.0 (GCA_000002315.2), International Chicken Genome Consortium, Nov 2011	1.046.932.099; $N_{\text{SNP_NCBI}}$ 9.415.942, $N_{\text{SNP_Ens}}$ 9.281.439
Pute (<i>Meleagris gallopavo</i>), 2009	Turkey_2.01 (GCA_000146605.2), Turkey Genome Consortium, Feb. 2011	1.061.817.103; $N_{\text{SNP_NCBI}}$ 6.964
Atlantischer Lachs (<i>Salmo salar</i>), 2011	ASM23337v1 (GCA_000233375.1), International Cooperation to Sequence the Atlantic Salmon Genome, Okt. 2011	2.435.040.521 **; $N_{\text{SNP_NCBI}}$ 8.635
Kabeljau (<i>Gadus morhua</i>), 2010	GadMor_May2010 (GCA_000231765.1), Genofisk, Aug. 2011	824.311.139 **; $N_{\text{SNP_NCBI}}$ 2.141

Quellen (Zugriff am 13. Januar 2014): ¹ <http://www.ncbi.nlm.nih.gov/genome/>; ² <http://www.ncbi.nlm.nih.gov/snp/>; ³ <http://www.ensembl.org/>;

* für Ensembl: Gesamtzahl kurzer Genomvarianten, d.h. SNPs, Indels (Insertionen/Deletionen), somatische Mutationen;

** noch keine chromosomale Anordnung, sondern Vorstadien (*Salmo salar*: Contig; *Gadus morhua*: Scaffold)

Übersicht 2: Verfügbare genomweite Single Nucleotide Polymorphism (SNP) Chips für einige wichtige Tierarten, modifiziert von Eggen (2012) unter Berücksichtigung der Anbieter-Informationen (Zugriff am 13. Januar 2014: <http://www.illumina.com/>, <http://www.affymetrix.com/>).

Spezies	Klassifikation und Consortium (für private Chips)	Identifikation (Anbieter) und Anzahl SNPs
Rind (<i>Bos taurus</i>),	kommerziell	BovineHD (Illumina), N _{SNP} 777.962
	kommerziell	BOS 1 (Affymetrix), N _{SNP} 648.855
	kommerziell	BovineSNP50v2 (Illumina), N _{SNP} 54.609
	kommerziell	BovineLD (Illumina), N _{SNP} 6.912
	privat (Consortium-Chip)	Euro10k (Illumina), Deutschland: N _{SNP} 9.072
Schwein (<i>Sus scrofa</i>),	kommerziell	PorcineSNP60v2 (Illumina), N _{SNP} 64.232
Schaf (<i>Ovis aries</i>),	kommerziell	OvineSNP50 (Illumina), N _{SNP} 54.241
	privat (freier Verkauf), AgResearch	Ovine (Illumina), N _{SNP} 5.409
Ziege (<i>Capra hircus</i>)	privat (Consortium-Chip)	Goat (Illumina), N _{SNP} 53.347
Pferd (<i>Equus caballus</i>)	privat (freier Verkauf), Neogen (GeneSeek)	Equine (Illumina) N _{SNP} 65.157
Hund (<i>Canis lupus familiaris</i>)	kommerziell	CanineHD (Illumina) N _{SNP} 173.662
	privat (freier Verkauf ausschließlich zu Forschungs- zwecken), Broad Institute	DogSNPs520431 (Affymetrix), N _{SNP} 127.132
Chicken (<i>Gallus gallus</i>)	privat (freier Verkauf), Cobb Vantress-Hendrix-USDA	Chicken 60k (Illumina) N _{SNP} 57.636
	kommerziell	Chicken (Affymetrix) N _{SNP} 580.961