

## Genomische Selektion beim Schwein

Prof. Dr. Jörn Bennewitz, Institut für Tierhaltung und Tierzucht, Uni Hohenheim

### Einleitung

Die genomische Selektion beschreibt den Gebrauch von genomweiten und massiven genetischen Markern zur Schätzung von Zuchtwerten und die Selektion der Tiere anhand dieser genomisch geschätzten Zuchtwerte. Sie gilt in der Rinderzucht als die Technik, welche seit Einführung der künstlichen Besamung die größte Innovationskraft hat und wird in nahezu allen großen Milchrinderpopulationen zurzeit implementiert. In der Schweinezucht hingegen sind die Erwartungen an die genomische Selektion geringer. Im ersten Teil dieses Beitrags wird die Idee der genomischen Selektion dem Leser näher gebracht. Anschließend wird die Bedeutung der Referenzpopulation bei der Schätzung der Markereffekte beschrieben und in diesem Zusammenhang auch die Gründe für die reduzierten Erwartungen an die genomische Selektion in der Schweinezucht erläutert. Darauf folgend wird die Projektinitiative zur Implementierung der genomischen Selektion bei der Vaterrasse Piétrain im Land Baden-Württemberg beschrieben und deren Potentiale diskutiert. Der Beitrag endet mit einem kurzen Fazit.

### Wie funktioniert die genomische Selektion?

Ein QTL (Quantitative Trait Locus) beschreibt einen chromosomalen Abschnitt, der ein oder mehrere segregierende Gene mit einem Einfluss auf ein bestimmtes Merkmal beherbergt. Mit Hilfe genetischer Marker ist es in den vergangenen ~10 Jahren gelungen, einige dieser QTL zu kartieren. Man vermutet, dass es mehrere hundert bis einige tausend solcher QTL für ein Merkmal gibt, in Abhängigkeit der effektiven Populationsgröße und der Erbllichkeit des betrachteten Merkmals. Jedoch hat nicht jeder QTL den gleichen Effekt auf ein Merkmal. Die genomische Selektion berücksichtigt alle QTL, die für ein Merkmal in der Population vorkommen. Sie erfordert, dass genomweite Markerkarten zur Verfügung stehen, die mehrere tausend Marker enthalten.

Die genomische Selektion funktioniert wie folgt. Es wird eine Referenzpopulation (auch Trainings- oder Kalibrierungsdatensatz genannt, siehe auch nächster Absatz) aufgestellt, in der die Tiere sowohl genotypisiert als auch in irgendeiner Form phänotypisiert sind. In dieser Population werden dann die Effekte aller Marker (egal, ob sie einen QTL flankieren oder auch nicht) mit entsprechenden statistischen Modellen geschätzt. Das Problem dabei ist, dass sehr viele Markereffekte aus einer begrenzten Anzahl an Beobachtungen geschätzt werden müssen. Der Statistiker spricht in diesem Zusammenhang von einem Mangel an Freiheitsgraden. Es existieren jedoch Modelle, die mit diesem Problem gut fertig werden (z. B. G-BLUP oder Bayes-Modelle, Meuwissen et al. 2001). Die genetischen Marker werden dann an den Selektionskandidaten genotypisiert. Der genomische Zuchtwert eines Selektionskandidaten ist dann die Summe der Markereffekte (geschätzt in der Referenzpopulation), welche das Tier trägt.

### Die Bedeutung der Referenzpopulation

Wie oben beschrieben erfolgt die Schätzung der Markereffekte in der sogenannten Referenzpopulation. In ihr sind die Tiere sowohl genotypisiert als auch in irgendeiner Form phänotypisiert. Die Größe dieser Referenzpopulation entscheidet maßgeblich mit über die Genauigkeit der genomischen Zuchtwerte. Die Sicherheit der genomischen Zuchtwertschätzung kann approximativ wie folgt ausgedrückt werden (Meuwissen 2009):

$$r^2 = \frac{Nh^2}{Nh^2 + 4N_eLv}$$

Dabei ist  $L$  die Länge des Genoms in Morgan (beim Schwein ca. 18 Morgan),  $N_e$  die effektive Populationsgröße (ein Maß für die genetische Diversität der Population),  $N$  die Größe der Referenzpopulation,  $h^2$  die Heritabilität

des Merkmals und der Term  $4*N_eLv$  bezeichnet die Anzahl effektiver QTL im Genom ( $v$  ist das Verhältnis der effektiven und aktuellen Anzahl an Segmenten (siehe Meuwissen (2009) für weitere Details). Werden in der Referenzpopulation BLUP-Zuchtwerte als Phänotypen verwendet, so muss  $h^2$  durch die Sicherheit der BLUP-Zuchtwerte ersetzt werden. Folgendes kann aus dieser Gleichung abgeleitet werden. Je diverser die Population ist (höheres  $N_e$ ), desto größer muss die Referenzpopulation sein. Gleiches gilt für die Länge des Genoms. Je zuverlässiger die Phänotypen, desto genauer die genomischen Zuchtwerte. Es kann festgehalten werden, dass eine sichere genomische Zuchtwertschätzung eine große Referenzstichprobe voraussetzt. Wichtig ist noch hervorzuheben, dass es von großem Vorteil sein kann, wenn die Referenzpopulation und die Tiere, in denen der genomische Zuchtwert geschätzt werden soll, miteinander verwandt sind, da dann zusätzlich verwandtschaftliche Beziehungen substantiell zur Genauigkeit der genomischen Zuchtwertschätzung beitragen (Meuwissen 2009).

### **Genomische Selektion bei den baden-württembergischen Piétrains**

Wie eingangs erwähnt, wird die genomische Selektion zurzeit in vielen Milchrinderpopulationen implementiert. Das Hauptziel ist die Verkürzung des Generationsintervalls auf der männlichen Seite von ca. 6 Jahren auf vielleicht 2 Jahre und zudem die Optimierung der zeit- und kostenaufwendigen Nachkommenschaftsprüfung. In der Schweinezucht sind die Potentiale weniger offensichtlich, da dort die Generationsintervalle bereits deutlich geringer sind und es schwierig ist, eine Referenzstichprobe von substantieller Größe aufzustellen. Zudem ist eine Refinanzierung der hohen Anfangskosten, die durch die Genotypisierung der Referenzstichprobe entstehen, deutlich schwieriger, da ein Zuchttier im Vergleich zur Rinderzucht deutlich weniger wert ist, die Typisierungskosten jedoch ungefähr gleich sind.

Aktuell wird ein Projekt zur alternativen Implementierung der genomischen Selektion bei der baden-württembergischen Piétrainpopulation mit dem Ziel einer routinemäßigen genomischen Vorselektion von Ebern gestartet. Beteiligt sind das Bildungs- und Wissenszentrum Boxberg, der Schweinezuchtverband Baden-Württemberg e.V. sowie das Hohenheimer Tierzuchtinstitut. Finanziert wird dieses Projekt im Wesentlichen vom Land Baden-Württemberg mit Beteiligung des Schweinezuchtverbandes B.-W.. Gegenwärtig werden jeweils 8 Nachkommen von ca. 80-100 Ebern pro Jahr in der Leistungsprüfungsanstalt Boxberg geprüft. Es soll herausgearbeitet werden, ob durch eine genomische Vorselektion von Ebern die Prüfkapazität reduziert werden kann, ohne dass signifikante Einbußen im Zuchtfortschritt zu erwarten sind. Ein ähnlicher Ansatz wurde von den Rinderzüchtern bei der markergestützten Vorselektion von Testbullen verfolgt (Bennewitz et al., 2004). Im Projekt werden zunächst ca. 500 nachkommengeprüfte Eber mit dem porcinen SNP-Chip mit ca. 50.000 SNPs genotypisiert; sie bilden die erste Referenzpopulation. Aus den Ergebnissen der SNP-Chip-Typisierungen und den BLUP-Zuchtwerten der Eber soll ein optimaler Set von mehreren hundert bis tausend SNP-Markern für die Population zusammengestellt werden, der ein Großteil der genetischen Variation der Merkmale erklärt und mit dem es möglich ist, genomische Zuchtwerte mit einer definierten Genauigkeit bei gegebener Größe der Lernstichprobe zu schätzen. Diese SNPs sollen dann in der Routine an den Selektionskandidaten genotypisiert werden. Naturgegeben werden die Genauigkeiten der geschätzten genomischen Zuchtwerte geringer sein, als die mit dem vollen Satz an SNPs. Diese reduzierten Genauigkeiten scheinen jedoch vor dem Hintergrund des Einsparpotentials bei den Typisierungskosten gerechtfertigt. Dies gilt umso mehr, als dass die Eber anschließend noch nachkommengeprüft und nicht allein aufgrund ihres genomischen Zuchtwertes eingesetzt werden. Ob dies tatsächlich so ist, wird im Laufe des Projektes herausgearbeitet.

Die Größe der Referenzpopulation, welche zurzeit die ca. 500 nachkommengeprüften Eber mit der höchsten Genauigkeit der Zuchtwertschätzung umfasst, sollte weiter ausgebaut werden. Es wäre aus Sicht der Wissenschaft sehr wünschenswert, wenn weitere Tiere die Referenzstichprobe ergänzen.

## **Fazit**

Die genomische Selektion wird zurzeit weltweit in die Milchrinderzucht implementiert. In der Schweinezucht sind mit der Verfügbarkeit des porcinen SNP-Chips die Voraussetzungen mittlerweile auch geschaffen. Die Sicherheit der genomischen Zuchtwertschätzung hängt maßgeblich von der Größe und Zusammensetzung der Referenzpopulation ab. Sie wird naturgegeben in der Schweinezucht geringer sein, als in der Rinderzucht. Eine genomische Vorselektion von Ebern, die anschließend noch die Nachkommenschaftsprüfung durchlaufen, ist ein guter Startpunkt mit überschaubaren Risiken und doch guten Potentialen für die Implementierung der Technik in die Zucht von Vaterrassen. Dabei wird versucht, eine möglichst große Referenzpopulation von Tieren mit sehr genau geschätzten Zuchtwerten aufzustellen.

## **Literatur**

- Bennewitz J., Reinsch N., Reinhardt F., Liu Z., Kalm, E. (2004). Top down marker assisted selection in dairy cattle using marker assisted estimates of breeding values. *Journal of Animal Breeding and Genetics* 121: 307-318.
- Meuwissen T.H.E., Hayes B.J., Goddard M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.
- Meuwissen T.H.E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genetics Selection Evolution*, 41, 35.