

Einfluss der genetischen Architektur auf die empirische Genauigkeit der genomischen Zuchtwertschätzung



M. Kramer¹, F. Biscarini¹, B. Bapst², C. Stricker³, H. Simianer¹

¹ Department für Nutztierwissenschaften, Abteilung Tierzucht und Haustiergenetik, Georg-August-Universität Göttingen, Albrecht-Thaer-Weg 3, 37075 Göttingen

² QUALITAS AG, Chamerstrasse 56, Zug 6300, Schweiz

³ agn Genetics GmbH, Börtjistraße 8b, 7260, Davos, Schweiz



6./7.9.2011

DGfZ/GfT-Gemeinschaftstagung, Weihenstephan

Einleitung



- Für die gZWS haben GBLUP und Bayesverfahren die größte Bedeutung in der Praxis
- Genauigkeit ist abhängig von (Daetwyler et al., 2010):
 - Anzahl Tiere in der Lernstichprobe (N_p)
 - Effektive Populationsgröße (N_e)
 - h^2 des Merkmals
 - Anzahl Gene die ein Merkmal beeinflussen (N_G)
- Abhängigkeiten zwischen diesen Faktoren

Fragestellung



- Welches Verfahren liefert anhand eines realen Datensatzes die höchsten Genauigkeiten der genomischen Zuchtwerte?

Fragestellung



- Welches Verfahren liefert anhand eines realen Datensatzes die höchsten Genauigkeiten der genomischen Zuchtwerte?
- Lassen sich Gesetzmäßigkeiten zur Vorteilhaftigkeit von GBLUP oder Bayes ableiten?

Datenmaterial



- 1.138 Bullen des Schweizer Braunvieh.
- 54.001 SNP Marker (Illumina 54k SNP Chip)
- Konventionelle Zuchtwerte der ZWS April 2010
 - Milchmenge (Mkg)
 - Somatic Cell Score (SCS)
 - Non-Return Rate (NRR)
 - Rastzeit (RZ)

Datenaufbereitung und Imputing



- Ausschluss von SNP Markern:

- Unbekannte Position
- Callrate $< 95\%$
- MAF $< 5\%$

⇒ 34.474 SNP zur Analyse

- Imputing mit BEAGLE 3.2
(Browning and Browning, 2009)

Statistische Modelle



- GBLUP in ASReml,
G Matrix nach Astle and Balding (2009):

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

Statistische Modelle



- GBLUP in ASReml,
G Matrix nach Astle and Balding (2009):

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

- BayesCpi in GenSel:

$$y = \mu + \sum_{j=1}^n X_{ij} \beta_j + e_i$$

80.000 Iterationen

Burnin 8.000

pi Wert 0,90

Kreuzvalidierung



- 5 fach Kreuzvalidierung mit 10 Wiederholungen.
- Zufällige Aufteilung der Bullen in Lernstichprobe und Validierungsstichprobe
 - Alle Genauigkeiten sind ein Durchschnitt aus 50 Wiederholungen
- Berechnung der Genauigkeit:

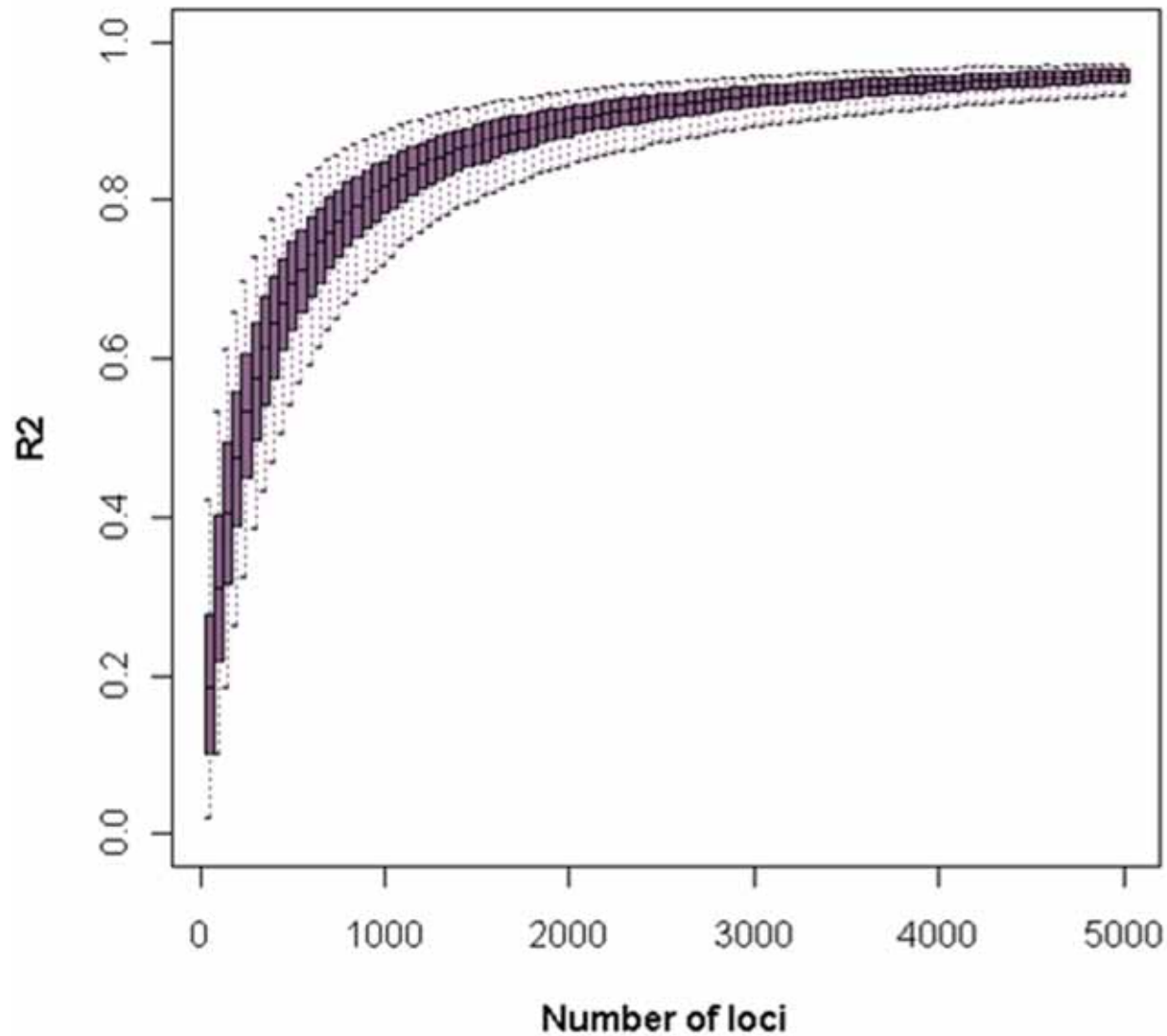
$$r(g, \hat{g}) = \frac{r(y, \hat{g})}{h}$$

Berechnung der N_G



- **Ansatz gem. Pimentel et al. (2011)**
- Es gilt: $\sigma_{Chr}^2 \propto L_{Chr}$
- N_G wird in 100er Schritten von 100 bis 5.000 variiert
- $\sigma_{add.gen}^2$ wird auf die Gene verteilt $\sigma_{Gen}^2 = \frac{\sigma_{add.gen}^2}{N_G}$
- Die Gene werden entsprechend der Größe auf die Chromosomen verteilt, die σ_{Gen}^2 werden je Chromosom aufsummiert zur σ_{Chr}^2
 - 1.000 Wiederholungen
 - R^2 der Regression von σ_{Chr}^2 auf L_{Chr} aus der Simulation wird mit den realen R^2 verglichen

Berechnung der N_G



Ergebnisse



	r_{GBLUP}	r_{Bayes}	Diff	N_G
Mkg	0,766	0,767	-0,001	950
SCS	0,692	0,552	0,140	1.150
NRR	0,749	0,598	0,151	950
RZ	0,673	0,481	0,192	1.450

Ergebnisse



	r_{GBLUP}	r_{Bayes}	Diff	N_G
Mkg	0,766	0,767	-0,001	950
SCS	0,692	0,552	0,140	1.150
NRR	0,749	0,598	0,151	950
RZ	0,673	0,481	0,192	1.450

Ergebnisse



	r_{GBLUP}	r_{Bayes}	Diff	N_G	h^2
Mkg	0,766	0,767	-0,001	950	0,33
SCS	0,692	0,552	0,140	1.150	0,27
NRR	0,749	0,598	0,151	950	0,09
RZ	0,673	0,481	0,192	1.450	0,04

Ergebnisse



	r_{GBLUP}	r_{Bayes}	Diff	N_G	h^2	h^2/N_G
Mkg	0,766	0,767	-0,001	950	0,33	0,000347
SCS	0,692	0,552	0,140	1.150	0,27	0,000234
NRR	0,749	0,598	0,151	950	0,09	0,000095
RZ	0,673	0,481	0,192	1.450	0,04	0,000028

Schlussfolgerungen



- GBLUP ist dem Bayes Verfahren überlegen bei Merkmalen denen viele Gene zugrunde liegen

Schlussfolgerungen



- GBLUP ist dem Bayes Verfahren überlegen bei Merkmalen denen viele Gene zugrunde liegen
- Eine Überlegenheit von BayesCpi konnte für kein Merkmal nachgewiesen werden

Schlussfolgerungen



- GBLUP ist dem Bayes Verfahren überlegen bei Merkmalen denen viele Gene zugrunde liegen
- Eine Überlegenheit von BayesCpi konnte für kein Merkmal nachgewiesen werden
- Die betrachteten Merkmale weisen einen hoch polygenen Erbgang auf

Schlussfolgerungen



- GBLUP ist dem Bayes Verfahren überlegen bei Merkmalen denen viele Gene zugrunde liegen
- Eine Überlegenheit von BayesCpi konnte für kein Merkmal nachgewiesen werden
- Die betrachteten Merkmale weisen einen hoch polygenen Erbgang auf
- Der Faktor h^2/N_G erlaubt eine gute Abschätzung der Vorteilhaftigkeit von GBLUP über BayesC
⇒ Überprüfung in weiteren Studien

Danksagung



Das Projekt wird als Verbundprojekt von der Europäischen Kommission im siebten Rahmenprogramm für Forschung und Entwicklung ko-finanziert (Vertrag Nr. 222623). Trotzdem gibt dieser Artikel nicht zwingend die Meinung der Europäischen Kommission wieder und ebenso wenig nimmt er zukünftige Strategien der Kommission vorweg.



Vielen Dank für Ihre Aufmerksamkeit!



6./7.9.2011

DGfZ/GfT-Gemeinschaftstagung, Weihenstephan