

# Bruk av genomisk avlsverdi ved innkjøp av seminokseemner i NRF avlen

MORTEN SVENDSEN<sup>1</sup>, BJØRG HERINGSTAD<sup>1,2</sup> OG TRYGVE R. SOLBERG<sup>1</sup>

<sup>1</sup>Geno SA, <sup>2</sup>Institutt for husdyr- og akvakulturvitenskap, UMB

## Plan

Genombaserte avlsverdier (GS verdier) ble fra høsten 2012 tatt i bruk som en viktig tilleggsplysningsopplysning ved innkjøp av seminokseemner. Genomisk seleksjon (GS) kan brukes for å redusere generasjonsintervallet og derved øke avlsframgangen. Da må imidlertid sikkerheten være så god at eliteoksene kan tas ut direkte basert på GS alene. Avlsmålet i NRF har så stort innslag av egenskaper med lav arvbarhet hvor GS virker dårlig, at avkomsgranskningen må fortsette. Likevel kan bruk av GS som preseleksjon ved innkjøpet gi 13 % økt avlsframgang (Lillehammer et al., 2011). Den økte sikkerheten ved GS skal også utnyttes til å redusere antall innkjøpte kalver til fenotypetesting fra 300 til 230 per år.

## Implementering

Som tidligere listes potensielle seminokseemner ut, basert på tradisjonelle BLUP avlsverdier til foreldrene. Hvis kalven blir okse og bonden vil tilby den som seminokseemne, kommer rådgiveren og tar en biologisk prøve av oksemor og kalv ved besiktigelsen. Hårprøve ble tidligere tatt ut for PCR analyse av rekesmaksgenet. SNP genotyping krever imidlertid mer DNA og av høyere kvalitet enn PCR analyse av én punktmutasjon. En har ikke behersket DNA ekstraksjon fra hårprøve godt nok til å få vedvarende høyt utbytte. Nasale svabre (bomullspinner for å skrape epitelceller fra neselimplimhinnene) har gitt mer konsistente resultater, både i prøveuttak og DNA ekstraksjon, og brukes nå til å ta den biologiske prøven fra seminokseemnene.

Oksekalvene genotypes rutinemessig med BovineSNP50® fra Illumina, som analyserer 54.000 SNP loci (punktmutasjoner) samtidig. Disse dataene brukes også til å prediktere hva resultatet ville blitt (imputere) på en eldre 25.000 SNP-chip fra Affymetrix. Den eldste delen av referansepopulasjonen er bare genotypet med chipen fra Affymetrix, så for disse oksene er det resultatet fra Illumina chipen som blir imputert. Begge chipene er laget for å passe en rekke ulike raser. Basert på relevansen for bruk i NRF, blir 48.204 loci fra den imputerte kombinasjonen av de to brukt videre i analysen.

GS verdier beregnes, fra september 2012, annenhver uke eller oftere. Ved færre enn 48 kalver blir brettet fylt opp med prøver fra ennå ikke testede venteokser. Første steg er å analysere noen utvalgte SNP'er med MassARRAY® for å verifisere farskap og å bestemme genotype for rekesmak, kollethet og kaseinvarianter. Per i dag sorteres de som har allelet for rekesmak bort og en sparer kostnaden til full SNP analyse av disse. I fremtiden kan en eventuelt selektere for andre enkeltgener i dette steget. SNP genotypene fra 48 prøver om gangen blir så bestemt gjennom en tre dagers automatisert prosess.

GS avlsverdier for alle de 50 egenskapene i avkomsgranskningen beregnes fra prediksjonsligninger der allelene i hvert locus har fått en verdi for egenskapen. Nye individer trenger derfor bare SNP genotype for å få GS avlsverdi. Prediksjonsligningene baseres på gBLUP analyse (Meuwissen et al., 2001) av avlsverdiene i referansepopulasjonen. Den består av alle NRF seminokser (litt over 3.000 okser) som har både SNP genotype og BLUP avlsverdi. Dette gjøres slik at summen av løsningsverdiene for alle SNP-loci i hver referanseokse

kommer så nær avlsverdien til oksen som mulig. GS verdiene blir skalert i forhold til den samme basepopulasjonen som de vanlige indeksene skaleres i forhold til. En samla GS verdi blir også laget fra GS verdiene til avlsmåleegenskapene ved å bruke de vanlige vektene.

Tabell 1. Arvbarhet (%) brukt i beregningene av BLUP avlsverdier; optimal arvbarhet (%) brukt i gBLUP; prediksjonsevne ved validering av gBLUP ved optimal arvbarhet.

Egenskap	Arvbarhet (BLUP)	Arvbarhet (gBLUP)	Korrelasjon gBLUP - BLUP
Samla avlsverdi		80	0,40
Mjølkk, indeks		95	0,63
Protein,kg	20,5	90	0,61
Mjølkk, kg	26,3	99	0,54
Fett, kg	19,4	80	0,66
Protein, %	53,7	95	0,55
Fett, %	40,9	90	0,61
Celletal	13,6	80	0,65
Kjøtt, indeks		99	0,62
Slaktevekt	15,4	95	0,57
Slakteklasse	22,1	90	0,77
Fettgruppe	20,8	95	0,71
Utmjølkingshastighet	19,1	70	0,65
Lekkasje	10,8	90	0,50
Lynne	8,1	80	0,62
Fruktbarhet, indeks		30	0,40
Ikkeomløp ved 56d, kviger	2,9	20	0,41
Ikkeomløp ved 56d, kyr	5,5	40	0,40
Kalving 1.ins.	7,7	70	0,31
Kalvingsvansker, far kalv	3,0	99	0,47
Kalvingsvansker, far ku	2,0	99	0,40
Dødfødsler, far kalv	0,9	90	0,33
Dødfødsler, far ku	1,2	20	0,16
Mastitt	3,5	80	0,46
Andre sjukdommer	2,4	30	0,35
Bein, indeks	7,0	30	0,66
Hasevinkel	8,6	50	0,68
Beinstilling	8,8	70	0,60
Kodeledd	8,8	50	0,71
Klauver	4,7	80	0,65
Jur og spener, indeks	19,3	90	0,60
Jurbotn-golv	35,8	70	0,66
Jurdybde	24,5	70	0,63
Spene plassering framme	18,3	95	0,64
Spene plassering bak	21,7	95	0,56
Ekstraspenner	20,9	70	0,54
Jurbalanse	10,6	50	0,57
Jurfeste framme	11,5	90	0,49
Jurfeste bak, bredde	12,9	95	0,64
Jurfeste bak, høyde	11,3	90	0,49
Midtband	10,4	90	0,48
Jurtype	8,2	70	0,40
Spenelengde	27,2	80	0,61
Spenetykkelse	20,2	50	0,54
Jurpoeng	8,8	95	0,33
Kroppsdybde	13,7	99	0,63
Krysshøyde	38,9	60	0,60
Kryssretning	19,9	99	0,74
Poeng kropp bein	7,4	90	0,56

gBLUP er den enkleste og overlegent raskeste av mange konkurrerende algoritmer for å beregne GS verdier. Ved å kjøre 12 parallelle serier tar hele beregningsjobben for 50 egenskaper kun 4 timer totalt.

## Validering

For å få et bilde av sikkerheten på GS verdiene kan en validere modellen ved å ta vekk en del av datasettet og prediktere disse verdiene fra den resterende delen. Korrelasjonen mellom de predikterte og de originale verdiene gir prediksjonsevnen til den statistiske prosedyren man har benyttet (Luan et al., 2011).

Okseårgangen som var gransket for første gang i 2011, dvs født i 2006, ble brukt som test-datasett og deres avlsverdier fra gransking 2012-1 skulle predikteres. Referansepopulasjonen besto av okser født i 2005 eller tidligere, totalt 2.866. Prediksjonslikningene ble beregnet med gBLUP på deres avlsverdier fra gransking 2012-1.

På samme måte som i BLUP forutsettes det at feilvariansen og den genetiske variansen er kjent, eller i det minste forholdet mellom den. En må derfor ha et estimat på "arvbarheten" til avlsverdiene som blir brukt som input i analysene. Med 48.000 forklaringsvariable og 2.866 datapunkter er det innlysende at en for høy arvbarhet vil overtilpasse løsningene og forklare feilvariasjon, mens en for lav arvbarhet ikke vil ekstrahere all informasjonen i dataene. En kjørte derfor 12 analyser med varierende arvbarhet fra 20 % til 99,9 %, og gikk videre med den arvbarheten som gav best prediksjonsevne. Det viste seg at alle egenskapene var robuste mot feil i arvbarheten, fordi et vidt spenn av arvbarheter ga tilnærmet like god prediksjonsevne. Den optimale arvbarheten med tilhørende prediksjonsevne er gitt i tabell 1. Korrelasjonene varierte fra 0,16 til 0,77, med overveiende høye korrelasjoner for mjølk, kjøtt og eksteriør; mens korrelasjonene var mye lavere for kalving, helse og fruktbarhet.

Oppsettet i denne valideringen gir et litt for optimistisk bilde av prediksjonsevnen til GS verdiene ved innkjøp av seminokseemne. Oksene i testdatasettet ble født og innkjøpt i 2006. En skulle derfor brukt indeksene fra 2006 og ikke 2012-1 i referansedatasettet. Referansepopulasjonen skulle også vært redusert fra okser født i 2005 og tidligere til okser granska første gang i 2005 eller tidligere. På den annen side ville referansepopulasjonen da blitt vesentlig mindre og av den grunn vist for lav prediksjonsevne.

## High density SNP-chip

Illumina markedsfører nå BovineHD® med 777.000 SNP-loci. En del NRF okser er genotypet med begge chipene slik at det kan imputeres fra 54.000 til 777.000 SNP-loci. Kryssvalidering på dette materialet ble prøvd, men ga kun marginal økning i prediksjonsevnen (Solberg et al., 2011). Tidsforbruket til den statistiske analysen ble dessuten tidoblet.

## Felles nordisk referansepopulasjon

Felles nordisk referansepopulasjon på 7.000 okser er prøvd (Heringstad et al., 2011). Tross mye større referansepopulasjon ble ikke resultatene vesentlig forbedret, antageligvis fordi populasjonene har for få felles fedre.

## Blending

Sammenveining (blending) av GS verdien og foreldremiddelet fra regulær BLUP ble prøvd for noen egenskaper. En prøvde ut en rekke blandingsforhold, men fant ingen miks som ga særlig

bedre prediksjon enn GS verdien alene. Foreldremiddelet var imidlertid ikke optimalt beregnet for egenskaper med farmodell. Overgang til dyremodell og inkludering av dummydyr for venta kalvinger et kvartal framover kan imidlertid gi foreldremiddel av bedre kvalitet for de fleste egenskapene.

## Referanser

- Heringstad, B., G. Su, T.R. Solberg, B. Guldbrandtsen, M. Svendsen, and M.S. Lund. 2011. Genomic predictions based on a joint reference population for Scandinavian red breeds. Book of abstracts of the 62nd Annual meeting of the EAAP, 29.*
- Lillehammer, M., T.H.E. Meuwissen and A.K. Sonesson. 2011. A comparison of dairy cattle breeding designs that use genomic selection. J. Dairy Sci. 94, 493-500.*
- Luan, T., J.A. Woolliams, S. Lien, M. Kent, M. Svendsen, and T.H.E. Meuwissen. 2009. The accuracy of genomic selection in Norwegian Red cattle assessed by cross validation. Genetics 183: 1119-1126.*
- Meuwissen, T. H. E., B. J. Hayes and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.*
- Solberg, T.R., B. Heringstad, M. Svendsen, H. Grove, and T. H. E. Meuwissen. 2011. Genomic predictions for production- and functional traits in Norwegian Red from BLUP analyses of imputed 54K and 777K SNP data. Interbull Bulletin no 44:240-243.*