

Genomisessa eläinmallissa käytetään sekä genotyypitettyjen että genotyypittämättömien eläinten tietoja

Minna Koivula, Ismo Strandén ja Esa Mäntysaari

Luonnonvarakeskus Luke, FI-31600 Jokioinen
e-mail. etunimi.sukunimi@luke.fi

TIIVISTELMÄ

Genotyypitettyjen eläinten lukumäärä on noussut nopeasti ja yhä enemmän genotyypitetään myös lehmiä. Tämä luo genomisille arvosteluille laskennallisia haasteita. Lähitulevaisuudessa on siirryttävä genomisen eläinmallin (ssGBLUP) käyttöön, jotta arvostelujen harhattomuus voidaan taata. Yksi kysymys genomiarvostelujen laskennassa on nuorten ja karsittujen eläinten genotyyppien käyttö. Tavallisessa eläinmalli BLUP:ssa eläin, jolla ei ole omaa havaintoa tai jälkeläisiä ei tuo arvosteluun informaatiota ja voidaan siten jättää pois arvostelusta. Teoriassa genotyypitetty nuori kandidaattieläin tai karsittu eläin voi tuoda informaatiota genomiseen eläinmalliin ja vaikuttaa näin muiden eläinten genomisiin jalostusarvoihin (GEBV). Tätä testattiin käyttämällä koelypsyaineistoa ja genomista koelypsymallia. Tässä jalostusarvot lasketaan samanlaisella eläinmalli BLUP menettelyllä kuin aikaisemminkin, mutta nyt genotyypitettyjen eläinten sukulaaisuudet perustuvat genotyypitietoihin ja muiden eläinten sukulaaisuudet sukupuutitietoihin ja genotyypitettyjen sukulaisten genomisiin sukulaisuuksiin. Arvostelussa käytetään suoraan alkuperäisiä havaintoja ja ne laskettiin MiX99-ohjelmalla. Arvosteluissa käytettiin eri määrä genomi-informaatiota joko jättämällä pois karsittujen sonnipoikien genotyypit tai ottamalla ne mukaan. Tulokset osoittivat, että karsittujen genotyypitettyjen sonnipoikien poisjättäminen ei aiheuttanut muutoksia referenssieläinten arvosteluihin. Sen sijaan niiden mukaan ottaminen saattaa vaikuttaa nuorien kandidaattisonnien arvosteluihin siten, että kandidaattisonnit saavat liian korkeat GEBV:t. Näyttäisi siis olevan turvallisempaa jättää karsitut eläimet pois genomisista arvosteluista.

Asiasanat: Genominen valinta, genomiset arvostelut, genominen eläinmalli

Johdanto

Tällä hetkellä lypsykarjalla käytössä olevat genomiset arvostelumallit ovat laskennallisesti kaksi- tai kolmivaiheisia. Näissä eläimille lasketaan ensin perinteiset jalostusarvot. Sonneille, joilla on jopa tuhansia jälkeläisiä, jalostusarvojen luotettavuus on lähes sata prosenttia. Seuraavaksi yhdistetään jälkeläisten perusteella jalostusarvosteltujen sonnien jalostusindeksit ja SNP markkeritiedot. Nämä jälkeläisarvostellut, genotyyppitetyt eläimet muodostavat niin sanotun referenssipopulaation, jota käytetään genomisen arvostelujen ennustemallissa. Genomisen mallin markkeriratkaisujen avulla voidaan ennustaa genomiset arvostelut kaikille genotyyppitetyille eläimille. Yleensä genomiseen arvosteluun yhdistetään eläimen polveutumistiedot, jolloin saadaan genomisen jalostusarvo-indeksi, GEBV.

Kaksivaiheisen genomiarvostelun sijaan ns. single-step menettelyssä eli genomisessa eläinmallissa yhdistyy sekä perinteinen eläinmalli että genomisen arvostelu (Aguilar ym. 2010, Christensen ja Lund 2010). Genomisessa eläinmallissa jalostusarvot lasketaan samanlaisella BLUP menettelyllä kuin aikaisemminkin, mutta nyt genotyyppitettyjen eläinten sukulaisuudet perustuvat genotyyppitietoihin ja muiden eläinten sukulaisuudet sukupuutietoihin sekä genotyyppitettyjen sukulaisten genomisiin sukulaisuuksiin. Arvostelussa voidaan käyttää alkuperäisiä havaintoja, jolloin genomisen informaatio tulee huomioiduksi myös karja- ja muita ympäristövaikutuksia ratkaistaessa.

Genotyyppitettyjen eläinten lukumäärä on noussut nopeasti ja yhä enemmän genotyyppitetään myös lemmiä. Tämä luo genomiselle arvosteluille laskennallisia haasteita. Toisaalta, lähitulevaisuudessa on siirryttävä genomisen eläinmallin (ssGBLUP) käyttöön, koska nykyisin käytettävät monivaiheiset genomiset arvostelumallit käyvät epäluotettavammiksi. Yksi kysymys genomiarvostelujen laskennassa on mitä karsittujen eläinten genotyypeille pitäisi tehdä. Genomisessa valintaohjelmassa saataan genotyyppittää jokaista keinosiemennykseen hankittua sonnia kohti 10, jopa 20 sonnivasikkaa. Tavallisessa eläinmalli BLUP:ssa eläin, jolla ei ole omaa havaintoa tai jälkeläisiä ei tuo arvosteluun informaatiota ja voidaan siten jättää pois arvostelusta. Teoriassa genotyyppitetty nuori kandidaattieläin tai karsittu eläin voi tuoda informaatiota genomiseen eläinmalliin ja vaikuttaa näin muiden eläinten genomisiin jalostusarvoihin. Tämän tutkimuksen tavoitteena oli arvioida miten karsittujen sonnipoikien genotyyppien mukanaolo tai poisjättäminen vaikuttaa referenssieläinten tai nuorten valintakandidaattien genomisiin jalostusarvoihin Pohjoismaisessa ssGBLUP koelypsymallissa.

Aineisto ja menetelmät

Aineisto

Tutkimuksessa käytettiin Pohjoismaista punaisen rodun (RDC) koelypsyaineistoa helmikuulta 2015. Aineistossa oli n. 87 miljoonaa havaintoa 3,9 miljoonalta lehmältä. Sukupuussa oli kaikkiaan noin 5,2 miljoonaa eläintä. Helmikuun genomiaineistossa oli jokaisella genotyyppitetyllä eläimellä 46914 SNP-markkeria. Genotyyppitetyistä RDC eläimistä 20276 oli informatiivisia referenssieläimiä, joilta oli joko omia tai joiden tyttärillä oli havaintoja koelypsyaineistossa. Referenssieläimistä 5696 oli sonneja ja 14580 lemmiä. Lisäksi genotyyppiaineistossa oli 1140 nuorta genotyyppitettyä kandidaattisonnia, sekä 8770 nuorena karsittua genotyyppitettyä sonnia. Kandidaattisonneiksi luokiteltiin vuosina 2009 - 2014 syntyneet genotyyppitetyt, kantakirjanumeron saaneet sonnit, joilla ei ollut jälkeläisiä koelypsyaineistossa. Karsituiksi sonneiksi luokiteltiin nuoret genotyyppitetyt sonnit, joita ei ollut ostettu keinosiemennykseen. Tutkimuksessa analysoitiin NAV:n käyttämällä koelypsymallilla (Lidauer ym. 2014), miten karsittujen sonnien mukana olo tai poisjättäminen vaikuttaa muiden eläinten jalostusarvoennusteisiin (EBV) tai vastaaviin GEBV:hin. Tätä varten muodostettiin kaksi erilaista sukupuutiedostoa ja genomista sukulaisuusmatriisia, joista aina toisessa oli mukana myös karsittujen eläinten sukulaisuustiedot tai genomitiedot. Kaikkiaan siis tehtiin neljä erilaista koelypsymallia. Taulukko 1. esittää eläinten määrät eri koelypsuarvosteluissa.

Taulukko 1. Eläinten määrä sukupuussa ja genomisessa sukulaisuusmatriisissa eri koelypsymalleissa.

	Sukupuus	Genominen sukulaisuusmatriisi
(G)EBV_I	5173381	21416
(G)EBV_II	5182461	30186

Tilastolliset menetelmät

Perinteisessä eläinmallissa tarvitaan sukulaisuusmatriisin käänteismatriisi \mathbf{A}^{-1} . Genomisessa eläinmallissa käytetään \mathbf{A}^{-1} sijasta \mathbf{H}^{-1} -matriisia, joka voidaan kirjoittaa sukulaisuusmatriisin ja genomisen lisäinformaation summana (Aguilar ym. 2010; Christensen ja Lund 2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

jossa \mathbf{A}_{22} on genotyyppitettyjen eläinten sukupuuhun perustuva sukulaisuusmatriisi ja \mathbf{G}_w on genotyyppitettyjen eläinten genominen sukulaisuusmatriisi. Tutkimuksessa oletimme, että kaikkea jalostusarvoissa esiintyvää vaihtelua ei voida kuvata SNP-markkereilla ja siksi $\mathbf{G}_w = (1-w)\mathbf{G}_0 + w\mathbf{A}_{22}$. Tässä w on kerroin, joka määrittää miten suuri osa geneettisestä vaihtelusta johtuu polygeenisistä tekijöistä, joita markerit eivät selitä. Polygeenisen vaihtelun osuutena käytimme arvoa $w=0.10$.

Genomisten sukulaisuusmatriisien laskemista varten on Luke:ssa kehitetty HGinv – ohjelma. Genomisissa arvosteluissa tarvittavat ohjelmistot on liitetty osaksi Luke:n omaa MiX99 jalostusarvosteluohjelmistoa (Strandén ja Lidauer 1999; Vuori ym. 2006). MiX99 – ohjelma lukee erikseen tavallisen sukupuutiedoston ja \mathbf{H}^{-1} -matriisin genomisen osan $\mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1}$.

Tulokset ja tulosten tarkastelu

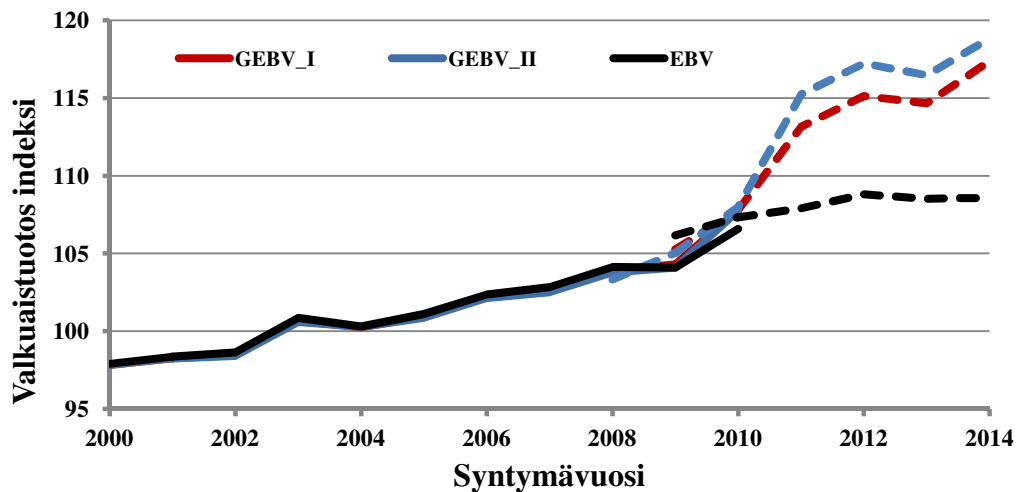
Tulokset osoittivat, että karsittujen sonnien genotyyppien mukanaolo tai poisjättäminen ei vaikuta referenssisonnien jalostusarvoihin (Taulukko 2). Korrelaatio eri sukulaisuus- ja genotyyppitiedoilla laskettujen jalostusarvojen välillä oli yksi. Samoin tarkasteltaessa esimerkiksi valkuaistuotoksen geneettistä trendiä (sonnien keskimääräinen jalostusarvo/syntymävuosi), voidaan todeta, että referenssisonnien trendit ovat yhteneväiset tavallisesta koelypsymallista tai genomisista eläinmalleista, riippumatta siitä käytetäänkö kaikkia genomitietoja vai vain valittujen eläinten genomitietoja (Kuva 1). Myös keskihajonnoista voidaan todeta, että luotettavasti arvosteltujen sonnien (G)EBV hajonnat ovat samansuuruiset. Vuosiluokissa 2009 - 2010 on mukana referenssisonneja, joilla on vielä suhteellisen vähän informaatiota ja arvosteluvarmuus on siten alhaisempi, ja tästä syystä näillä sonneilla EBV:n hajonta on pienempi kuin GEBV:n hajonta (Kuva 2).

Myös kandidaattisonneilla, EBV:den korrelaatio eri sukupuutiedoilla oli yksi, mutta eri perustein laskettujen GEBV:den korrelaatio eri genomitiedolla putosi 0.98:aan (Taulukko 2). Toisin sanoen lisättäessä karsittujen sonnien genomitiedot genomiseen eläinmalliin, nuorten kandidaattisonnien GEBV:t muuttuivat jonkin verran. Lisäksi näyttää siltä, että ottamalla genomiseen eläinmalliarvosteluun mukaan karsittujen sonnien genotyyppit nousevat nuorimpien kandidaattisonnien ikäluokkakieskiarvot (Kuvio2).

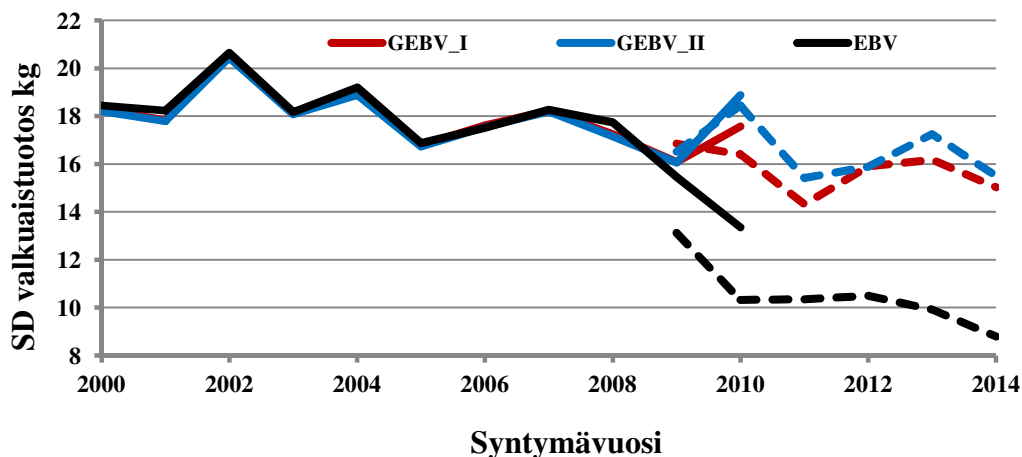
Taulukko 2. Valkuaistuotoksen EBV:den ja GEBV:den välinen korrelaatio eri koelypsymalleissa. EBV_I ja GEBV_I ovat laskettu aineistoista joissa mallissa ovat vain keinosiemennykseen hankitut sonnit, EBV_II ja GEBV_II ovat laskettu aineistosta jossa ovat mukana myös karsitut sonnit.

		EBV_II	GEBV_I	GEBV_II
Referenssisonnit	EBV_I	1.000	0.997	0.996
	EBV_II		0.997	0.996
	GEBV_I			0.999
Kandidaattisonnit	EBV_I	1.000	0.560	0.534
	EBV_II		0.560	0.534
	GEBV_I			0.984

Kuva 1. Perinnöllinen edistyminen valkuaistuotoksessa laskettuna tavallisista jalostusarvoista (EBV), sekä genomisista jalostusarvoista (GEBV_I ja GEBV_II) käyttäen eri genomitietoja. GEBV_II:ssa mukana ovat karsittujen sonnien genotyypit ja GEBV_I:ssä ei. Yhtenäinen viiva on referenssisonnien trendi ja katkoviiva on kandidaattisonnien trendi.



Kuva 2. Valkuaistuotoksen jalostusarvojen keskihajonta (SD) syntymävuosittain laskettuna tavallisista jalostusarvoista (EBV), sekä genomisista jalostusarvoista (GEBV_I ja GEBV_II) käyttäen eri genomitietoja. GEBV_II:ssa mukana ovat karsittujen sonnien genotyypit ja GEBV_I:ssä ei. Yhtenäinen viiva referenssisonnien trendi ja katkoviiva kandidaattisonnien trendi.



Johtopäätökset

Karsittujen sonnien genomitiedot voidaan jättää pois genomisesta eläinmallissa ilman negatiivista vaikutusta referenssisonnien arvosteluihin. Referenssisonnien GEBV:t pysyvät samoina, eivätkä myöskään geneettiset trendit muuttuneet. Sen sijaan karsittujen sonnien genomitiedot vaikuttavat nuorien kandidaattisonnien arvosteluun. Mikäli karsittujen sonnien genotyypit poistetaan arvosteluista muuttuva hankittujen kandidaattisonnien valkuaistuotosgenomi-arvostelut (korrelaatio 0.98), ja niiden taso laskee noin 2-3 indeksipistettä. Validaatiotutkimuksissa on huomattu, että kandidaattisonnien GEBV trendi on yleensä yliestimoitu verrattaessa EBV trendiin (Kuva 1). Tämä korostuu entisestään, jos karsittujen sonnien genomitiedot ovat mukana ssGBLUP:ssa. Ilmeisesti nuoret sonninit ja varsinkin sonninemät saavat liian suuret GEBV:t, jos niiltä on arvostelussa mukana karsittujen poikien genotyypit. Siksi siirryttäessä genomisiin eläin-arvosteluihin voisi olla jopa järkevää jättää karsittujen sonnien genomitiedot pois. Paitsi että se vähentää genomisen eläinmallin laskentapainetta, se myös parantaa arvosteluiden luotettavuutta.

Kirjallisuus

Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A. & Tsuruta, S. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743-752.

Christensen, O.F. & Lund, M.S. 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol.* 42:2.

Lidauer, M., Pösö, J., Pedersen, J., Lassen, J., Madsen, P., Mäntysaari, E.A., Nielsen, U.S., Eriksson, J-Å., Johansson, K., Pitkänen, T. & Strandén, I. 2015. Across-country test-day model evaluations for Holstein, Nordic Red Cattle, and Jersey. *J. Dairy Sci.* 98:1296-1309.

Strandén, I. & Lidauer, M. 1999. Solving large mixed models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82: 2779-2787.

Vuori, K., Strandén, I., Lidauer, M., & Mäntysaari, E.A. 2006. MiX99 – Effective solver for large and complex linear mixed models. *Proc. 8th World Congr. Genet. Appl. Livest. Prod. Belo Horizonte, MiG, Brazil*, ss. 27-33.