Received: 29 March 2018 | Revised: 18 June 2018 | Accepted: 3 July 2018

DOI: 10.1111/jbg.12353

ORIGINAL ARTICLE

WILEY Journal of Animal Breeding and Genetics

Single step genomic evaluation for female fertility in Nordic Red dairy cattle

Kaarina Matilainen[1] | **Ismo Strandén**[1] | **Gert Pedersen Aamand**[2] | **Esa A. Mäntysaari**[1]

[1]Natural Resources Institute Finland (Luke), Jokioinen, Finland

[2]Nordic Cattle Genetic Evaluation, Århus N, Denmark

Correspondence
Kaarina Matilainen, Natural Resources Institute Finland (Luke), FI-31600 Jokioinen, Finland.
Email: kaarina.matilainen@luke.fi

Funding information
MAKERA foundation (Finnish Ministry of Agriculture and Forestry); Valio; Faba; Viking Genetics

Abstract

Joint Nordic (Denmark, Finland, Sweden) genetic evaluation of female fertility is currently based on the multiple trait multilactation animal model (BLUP). Here, single step genomic model (ssGBLUP) was applied for the Nordic Red dairy cattle fertility evaluation. The 11 traits comprised of nonreturn rate and days from first to last insemination in heifers and first three parities, and days from calving to first insemination in the first three parities. Traits had low heritabilities (0.015–0.04), but moderately high genetic correlations between the parities (0.60–0.88). Phenotypic data included 4,226,715 animals with records and pedigree 5,445,392 animals. Unknown parents were assigned into 332 phantom parent groups (PPG). In mixed model equations animals were associated with PPG effects through the pedigree or both the pedigree and genomic information. Genotype information of 46,914 SNPs was available for 33,969 animals in the pedigree. When PPG used pedigree information only, BLUP converged after 2,420 iterations whereas the ssGBLUP evaluation needed over ten thousand iterations. When the PPG effects were solved accounting both the pedigree and the genomic information, the ssGBLUP model converged after 2,406 iterations. Also, with the latter model breeding values by ssGBLUP and BLUP became more consistent and genetic trends followed each other well. Models were validated using forward prediction of the young bulls. Reliabilities and variance inflation of predicted genomic breeding values (values for parent averages in brackets) for the 11 traits ranged 0.22–0.31 (0.10–0.27) and 0.81–0.95 (0.83–1.06), respectively. The ssGBLUP model gave always higher validation reliabilities than BLUP, but largest increases were for the cow fertility traits.

KEYWORDS
dairy cattle, fertility, genetic groups, genomic evaluation

1 | INTRODUCTION

Genetics of fertility have been studied for decades, because fertile animals are basis for cost-effective production and breeding of dairy cattle (e.g. Mäntysaari & Van Vleck, 1989; Philipsson, 1981; Sewalem, Kistemaker, & Miglior, 2010; VanRaden et al., 2004). Fertility traits have considerable genetic variation but low heritability. Attributed to their unfavourable genetic correlation with milk yield, fertility needs to be selected in breeding. In the Nordic countries, the Total Merit Index has included female fertility for many decades (Mäntysaari & Van Vleck, 1989). The joint Nordic routine genetic evaluation for fertility in Denmark, Finland, and Sweden has been in use since 2005 (Fogh et al., 2003) and was updated in 2015 (Muuttoranta et al., 2015). In the new evaluations, heifer and cow fertility are

modelled by multitrait multilactation animal models in three separate trait groups. The first trait group has nonreturn rate, days from first to last insemination and days from calving to first insemination, and the second trait group has number of inseminations, heat strength, and interval from calving to first insemination. A new third trait group has conception rates (Tyrisevä et al., 2017).

Genotypes are a recent source of information in genetic evaluations. In order to increase accuracy of estimated breeding values, all phenotypic and genomic information should be included into the genetic evaluation. Single step genomic evaluation (ssGBLUP) is regarded as a simple and accurate approach (Aguilar et al., 2010; Christensen & Lund, 2010), which takes into account phenotypic, pedigree, and genomic data simultaneously. Solving of ssGBLUP is relatively easy by using the same procedures as for evaluation without genomic information (hereinafter traditional BLUP). However, calculation of the covariance structure of breeding values in ssGBLUP requires a lot of attention because it has pedigree and genomic relationship matrices that need to be compatible. For example, there are many alternative approaches to calculate allele frequencies that are needed in forming the genomic relationship matrix **G** Compatibility can be achieved by different kind of adjustments to the **G** matrix such as shifting or scaling (Chen, Misztal, Aguilar, Legarra, & Muir, 2011; Christensen, Madsen, Nielsen, Ostersen, & Su, 2012; Forni, Aguilar, & Misztal, 2011; Vitezica, Aguilar, Misztal, & Legarra, 2011) or weighting both the **A** and the **G** matrices (Tsuruta, Misztal, Aguilar, & Lawlor, 2011). Incompatibility has been recognized to be associated also with the genetic groups for the unknown parents (Misztal, Vitezica, Legarra, Aguilar, & Swan, 2013). Furthermore, Strandén, Matilainen, Aamand, and Mäntysaari (2017) observed that convergence of the used iterative solving method can be impaired if the inverse relationship matrices based on all animals ($\mathbf{A}^{-1}$) and the genotyped animals $(\mathbf{A}_{22})^{-1}$ in ssGBLUP are made inconsistently. An alternative approach is to adjust the **A** matrix to be compatible with the **G** matrix. This approach accounts for nonzero inbreeding and relationships among base animals using so called metafounders (Legarra, Christensen, Vitezica, Aguilar, & Misztal, 2015).

The ssGBLUP has been tested with promising results for milk production traits (e.g. Koivula, Strandén, Pösö, Aamand, & Mäntysaari, 2015; Lourenco et al., 2014). However, only one ssGBLUP fertility evaluation has been reported: Aguilar, Misztal, Tsuruta, Wiggans, and Lawlor (2011) estimated breeding values for conception rate in Holstein using multiple trait ssGBLUP. They found the genomic enhanced breeding values (GEBV) by ssGBLUP to be more accurate than corresponding breeding value estimates (EBV) by traditional BLUP. In this study, we will shortly describe the official Nordic evaluation model of Nordic Red dairy cattle for the first fertility trait group: nonreturn rate, days from first to last insemination and days from calving to first insemination. The official BLUP model is upgraded into a ssGBLUP evaluation. In the ssGBLUP, the **G** matrix was formed using either the base population or average genomic data allele frequencies. Furthermore, phantom parent groups (PPG) were included into the mixed model equations (MME) using QP-transformation (Quaas & Pollak, 1981) where the PPG equations used pedigree or both pedigree and genomic information. Feasibility and validity of large-scale ssGBLUP, as well as comparisons between GEBVs and EBVs are reported.

2 | MATERIALS AND METHODS

2.1 | Data

Data used in routine Nordic fertility evaluations for Nordic Red dairy cattle were obtained from the Nordic Cattle Genetic Evaluation (NAV, Aarhus, Denmark) in spring 2016. The Nordic fertility evaluations are calculated separately for three trait groups, but only the first trait group was analysed in this study. The multiple trait model had traits for heifers and cows (Table 1). The heifer traits were nonreturn rate within 56 days after first-service (NRR0) and days from first to last insemination (IFL0). The cow traits (considered to be different in first, second, and third parities) were nonreturn rate within 56 days after first-service (NRR1, NRR2, NRR3), days from first to last insemination (IFL1, IFL2, IFL3) and days from calving to first insemination (ICF1, ICF2, ICF3). Phenotypic data had over 4.2 million cows with records collected since year 1982 in Sweden, 1985 in Denmark, and 1992 in Finland until 2016. Total number of observations was approximately 29 million for the eleven traits. Pedigree consisted of over 5.4 million animals. There were 332 PPGs which accounted genetic level by breed, country of origin, and birth year. On average a group had almost 2,000 animals. Few groups had less than 10 animals, but the largest groups had about 40,000 animals. Genotype information was available for 33,969 animals in the pedigree, from which 6,072 were males. Bulls were genotyped using the Illumina BovineSNP50 and cows with BovineLD Bead Chips with the genotypes imputed to the 50K chip (Illumina Inc., San Diego, CA, USA). After applying editing criteria, of minor allele frequency of 0.01 and locus average GenCall score of 0.60, 46,914 markers were used in the analysis.

Validity of ssGBLUP evaluation was studied by the forward prediction approach (Mäntysaari, Liu, & VanRaden, 2010). To obtain enough information also for the third parity traits in validation bulls, the observations from the last

TABLE 1 Eleven traits in the female fertility evaluations of Nordic Red dairy cattle, their abbreviations (abb), number of observations (N), and heritabilities (h^2)

Parity no.	Trait	abb	N	h^2
Heifer	Nonreturn rate	NRR0	3,661,674	0.015
	Interval from first to last insemination	IFL0	3,500,021	0.015
1	Nonreturn rate	NRR1	3,435,069	0.015
	Interval from calving to first insemination	ICF1	3,464,583	0.040
	Interval from first to last insemination	IFL1	3,459,107	0.030
2	Nonreturn rate	NRR2	2,389,108	0.015
	Interval from calving to first insemination	ICF2	2,417,653	0.040
	Interval from first to last insemination	IFL2	2,415,235	0.030
3	Nonreturn rate	NRR3	1,395,081	0.015
	Interval from calving to first insemination	ICF3	1,415,438	0.040
	Interval from first to last insemination	IFL3	1,414,073	0.030

6 years in the full data were removed in the reduced data, i.e., the reduced data had records until 2010. This reduced the number of the heifer observations by 18%, and the first, second, and third lactation observations by 19%, 20%, and 21% respectively. The reduced data contained over 3.4 million cows with records and total of 23.4 million observations for the eleven traits. The same pedigree and genotype information were used for both the full data and the reduced data analyses.

2.2 | Model

The same multiple trait multilactation animal model and (co)variance parameters (Muuttoranta et al., 2015; Tyrisevä et al., 2016) were used for the GEBV and EBV calculations except that genomic information was included in the GEBV calculations. Model contained one general regression effect for total heterosis, three fixed effects, and random genetic animal effect for each trait. Fixed effects were age at first breeding and, depending on the trait, herd × birth year or herd × first calving year − interaction and month × insemination year or month × calving year − interaction. In solving the mixed model equations, the PPG solutions were regressed towards zero which is similar to regarding the groups as random effects. Heritabilities were 0.015 for the heifer traits and NRR traits, 0.03 for IFL traits of cows, and 0.04 for ICF traits (Table 1). Genetic correlations between the parties were moderately high (0.60–0.88).

The single step method (Aguilar et al., 2010; Christensen & Lund, 2010), was used. The method needs inverse of the single-step relationship matrix $\mathbf{H}$ which accounts for both the pedigree and the genomic information:

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

where $\mathbf{A}^{-1}$ and $\mathbf{A}_{22}^{-1}$ are inverses of pedigree based relationship matrices for all animals and for the genotyped animals only, respectively, and $\mathbf{G}_w = (1-w)\mathbf{G} + w\mathbf{A}_{22}$. The genomic relationship matrix $\mathbf{G}$ was based on method 1 in VanRaden (2008), where alternatively either the data or the base population allele frequencies (McPeek, Wu, & Ober, 2004) were used, and the w, representing the weight for the polygenic relationships, was 10%. To obtain values in the $\mathbf{G}$ and $\mathbf{A}_{22}$ matrices on the same scale, values in $\mathbf{G}$ were multiplied by the scaling factor trace($\mathbf{A}_{22}$)/trace($\mathbf{G}$). Inbreeding coefficients were used in forming the inverse of pedigree based relationship matrix $\mathbf{A}^{-1}$ to be compatible with $(\mathbf{A}_{22})^{-1}$ which included inbreeding.

It is important that genomic and pedigree relationship matrices are compatible (Chen et al., 2011; Christensen et al., 2012; Forni et al., 2011; Tsuruta et al., 2011; Vitezica et al., 2011), also when PPGs are used (Matilainen, Koivula, Strandén, Aamand, & Mäntysaari, 2016). Commonly the PPG equations are included into the pedigree based relationship matrix $\mathbf{A}^{-1}$ via QP-transformation (Quaas & Pollak, 1981). Thus, the genetic group equations are augmented into the inverse relationship matrix:

$$\mathbf{A}_A^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{Q} \\ -\mathbf{Q}'\mathbf{A}^{-1} & -\mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & -(\mathbf{A}^{11}\mathbf{Q}_1 + \mathbf{A}^{12}\mathbf{Q}_2) \\ \mathbf{A}^{21} & \mathbf{A}^{22} & -(\mathbf{A}^{21}\mathbf{Q}_1 + \mathbf{A}^{22}\mathbf{Q}_2) \\ -(\mathbf{Q}_1'\mathbf{A}^{11} + \mathbf{Q}_2'\mathbf{A}^{21}) & -(\mathbf{Q}_1'\mathbf{A}^{12} + \mathbf{Q}_2'\mathbf{A}^{22}) & \mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q} \end{bmatrix},$$

where matrix $\mathbf{Q}$ describes the proportion of contributions each animal receives from the genetic groups $\mathbf{g}$, so that

$$E[\mathbf{a}] = \mathbf{Q}\mathbf{g}$$

where $\mathbf{a}$ is vector of all breeding values, $\mathbf{Q}_2$ and $\mathbf{Q}_1$ are submatrices of $\mathbf{Q}$ that pertains to the genotyped and nongenotyped animals, respectively, and $\mathbf{A}^{xx}$ is a submatrix of $\mathbf{A}^{-1}$ with x referring to genotyped (2) or nongenotyped (1) animals. In most implementations of ssGBLUP, the PPG equations are included into the pedigree based

relationship matrix, as in $\mathbf{A}_A^{-1}$ above, but not into the genomic based relationship matrix $\mathbf{G}_w$. Thus, by defining $\mathbf{B} = \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1}$, the augmented inverse of the single-step relationship matrix $\mathbf{H}_A^{-1}$ becomes:

$$\mathbf{H}_A^{-1} = \mathbf{A}_A^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

This is incompatible, however, in a sense that genomic information is not accounted for in the PPG effects. The correct way would be to use (Misztal et al., 2013)

$$\mathbf{H}_{\mathrm{AB}}^{-1} = \mathbf{A}_A^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & -\mathbf{B}\mathbf{Q}_2 \\ \mathbf{0} & -\mathbf{Q}_2'\mathbf{B} & \mathbf{Q}_2'\mathbf{B}\mathbf{Q}_2 \end{bmatrix}$$

In our pedigree, the genotyped animals were associated with 262 of 332 genetic groups.

2.3 | Computations

2.3.1 | Comparison of full data solutions

Three different GEBV solutions were calculated for models with PPG: (a) $\mathrm{ssGBLUP_{Ad}}$: using $\mathbf{H}_A$ with $\mathbf{G}$ based on data allele frequencies, (b) $\mathrm{ssGBLUP_{Ab}}$: using $\mathbf{H}_A$ with $\mathbf{G}$ based on base population allele frequencies, and (c) $\mathrm{ssGBLUP_{AB}}$: using $\mathbf{H}_{\mathrm{AB}}$ with $\mathbf{G}$ based on base population allele frequencies. For comparison purposes, two additional genomic evaluations were made that did not include PPG: $\mathrm{ssGBLUP_d}$ and $\mathrm{ssGBLUP_b}$ used $\mathbf{H}$ with $\mathbf{G}$ based on either data or base population allele frequencies, respectively. Genetic trends were addressed using sire GEBVs and EBVs only, but the correlations between EBVs and GEBVs within birth years were calculated both for males and for females.

2.3.2 | Validation

Validations of the different genetic evaluations were based on regression as in Mäntysaari et al. (2010). Regression model for the validation was

$$\mathbf{DRP}_f = \mu + \delta\mathbf{x}_r + \mathbf{e},$$

where vector $\mathbf{DRP}_f$ contained deregressed genetic predictions (DRP) for bulls calculated by single-trait model from the EBVs based on the full data and $\mathbf{x}_r$ contained either GEBVs or EBVs from the reduced data. In the validation regression, the DRP for trait i of sire j was weighted by

$$r^2_{\mathrm{DRP}_{ij}} = \frac{\mathrm{ERC}_{ij}}{\mathrm{ERC}_{ij} + \lambda_i},$$

where ERC is the effective record contribution based on the full data and

$$\lambda_i = \frac{1 - h_i^2}{h_i^2}.$$

The validation regressions were fitted among 750 validation bulls which were selected to be young genotyped sires with $\mathrm{ERC}_{ij} = 0$ in the reduced data but $\mathrm{ERC}_{ij} > 10$ in the full data for every trait i. Thus, for the bulls in the validation cohort, the EBVs estimated from the reduced data were simple parent averages (PA).

Regression coefficient δ is expected to be one when the evaluation based on reduced data predicts the future differences between animals properly. Furthermore, validation reliability of the model was calculated as

$$R_i^2 = \frac{\tilde{R}_i^2}{\overline{r^2_{\mathrm{DRP}_i}}}$$

where $\tilde{R}_i^2$ is R-square value of the regression and $\overline{r^2_{\mathrm{DRP}_i}}$ is average of the weights for trait i. The higher the R-square value the better the model fits the data. Bootstrapping based on 10,000 replicates was used to estimate 95% confidence intervals for both the regression coefficients and the validation reliabilities.

2.3.3 | Software

Numerator relationship matrix $\mathbf{A}_{22}$ was calculated using RelaX2 (Strandén, 2014) and all GEBV and EBV analyses were done by MiX99 (Vuori, Strandén, Lidauer, & Mäntysaari, 2006). MiX99 uses iteration on data in preconditioned conjugate gradient (PCG) algorithm to solve the MME. For all analyses, PCG iteration was assumed to be converged when the square root of relative difference between consecutive solutions was smaller than 1.0^{-5}. DRP for the validation were calculated by MiX99 software using the bisection method (Strandén & Mäntysaari, 2010). These were calculated only for sires, using EBV and ERC based on the full data. ApaX99 program was used to calculate the Interbull ERCs (Strandén, Lidauer, Mäntysaari, & Pösö, 2000). Bootstrapping was made using package *boot* in R software (R Core Team, 2015). Ordinary nonparametric bootstrap resampling using boot-function was applied to generate replicates, after which the basic method of boot.ci-function calculated the confidence intervals.

3 | RESULTS

3.1 | Comparison of full data solutions

The traditional BLUP MME for solving EBVs converged after 2,420 PCG iterations, whereas solving GEBVs from the MMEs of $\mathrm{ssGBLUP_{Ad}}$ based on data allele frequencies needed 23,284 PCG iterations. Logarithm of the

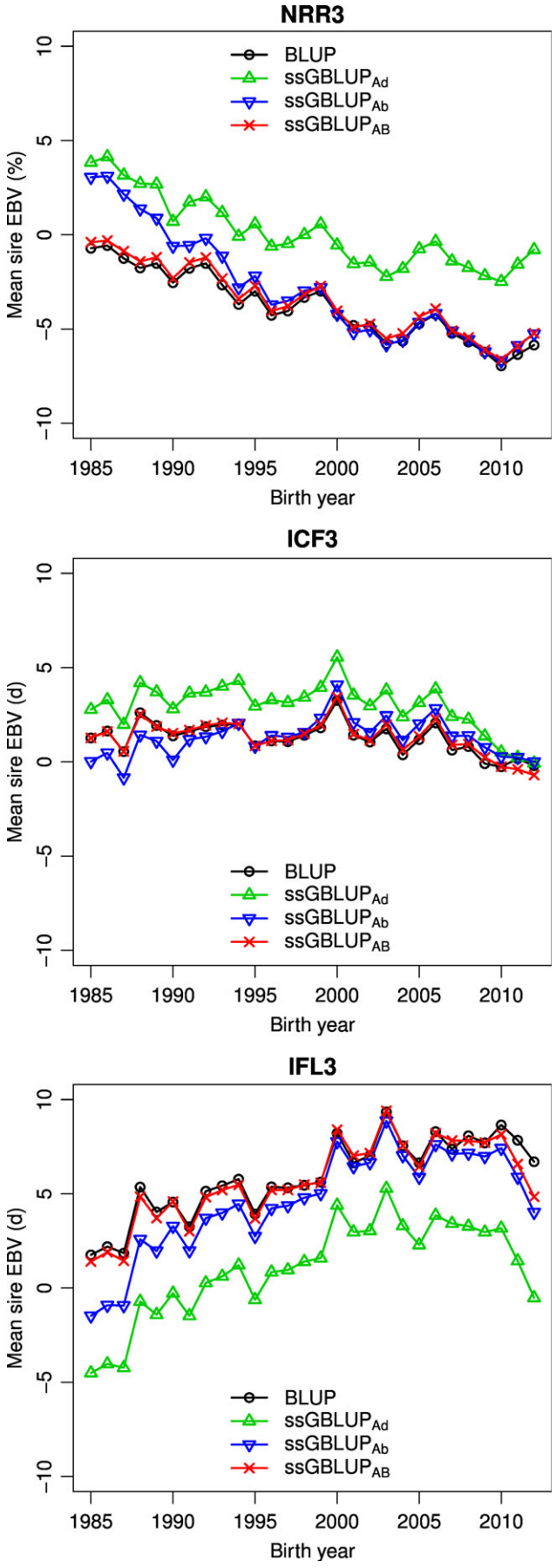

FIGURE 1 Genetic trends based on estimated breeding values from the animal model (BLUP) and genomic enhanced breeding values (GEBV) from the single-step genomic model (ssGBLUP) for RDC Nordic bulls having at least 50 daughters. Models for GEBVs used data (ssGBLUP$_{Ad}$) or base population (ssGBLUP$_{Ab}$ and ssGBLUP$_{AB}$) allele frequencies. All models had phantom parent groups (PPG) that used pedigree information, but model ssGBLUP$_{AB}$ used also genomic information in PPG estimation. Traits are nonreturn rate (NRR3), interval from calving to first insemination (ICF3) and interval from first to last service (IFL3) in the third parity [Colour figure can be viewed at wileyonlinelibrary.com]

convergence criteria had repetitive high peaks which may be an indication of incorrect model or MME. Nevertheless, genetic trends based on sire GEBVs and EBVs had the same pattern. Figure 1 illustrates the averages of nonstandardized sire EBVs and GEBVs for third parity traits by birth year. Although the genetic trends were similar, the individual EBV and GEBV estimates differed considerably. This can be seen in correlations between the estimates by trait and birth year. In general, IFL had correlations between 0.88 and 0.99, whereas NRR and ICF traits had correlations 0.65–0.96 and 0.55–0.99, respectively (Table 2).

Use of base population allele frequencies in the **G** matrix used in ssGBLUP had positive impact on the PCG convergence. The GEBV solutions for ssGBLUP$_{Ab}$ were obtained after 12,450 PCG iterations but the convergence criteria values did not decrease smoothly. The yearly sire means based on ssGBLUP$_{Ab}$ reached the means based on the traditional BLUP for younger year classes, although trend for older year classes was different (Figure 1). Again, correlations between EBVs and GEBVs differed considerably by trait and year (Table 2). Especially for NRR and ICF traits many birth year classes had correlations lower than 0.90. Correlations are shown by birth years for the third lactation traits in Figure 2. The genetic group solutions differed between genetic and genomic evaluations for some traits. For example, for the third lactation, clear differences between the Finnish Ayrshire genetic group solutions can be seen in NRR3 but some smaller in ICF3 and hardly any for IFL3 (Figure 3). The same can be seen even more dramatically for the reduced data set in the same figure.

Models without PPG converged faster than with PPG. EBV solutions were obtained after 1,562 PCG iterations. More dramatic reduction in the number of iterations was seen in ssGBLUP evaluation: GEBVs were obtained after 1,296 and 1,553 PCG iterations using either data (ssGBLUP$_d$) or base population (ssGBLUP$_b$) allele frequencies, respectively. In addition, the convergence criteria decreased smoothly in contrast with ssGBLUP with PPG. Especially, the genetic trends based on GEBV solutions by ssGBLUP$_b$ followed nicely the genetic trends based on EBV solutions (trends for the cow traits in the third

TABLE 2 Ranges of yearly correlations (among both the males and the females) between EBVs and GEBVs by traits. Models for GEBVs used data (ssGBLUP$_{Ad}$) or base population (ssGBLUP$_{Ab}$ and ssGBLUP$_{AB}$) allele frequencies. All models had phantom parent groups (PPG) that used pedigree information, but model ssGBLUP$_{AB}$ used also genomic information in PPG estimation

Trait	Parity	GEBV$_{Ad}$	GEBV$_{Ab}$	GEBV$_{AB}$
NRR	0	0.76–0.93	0.87–0.97	0.87–1.00
IFL	0	0.90–0.97	0.90–0.99	0.91–1.00
NRR	1	0.71–0.96	0.62–0.97	0.87–1.00
ICF	1	0.91–0.99	0.91–0.99	0.91–1.00
IFL	1	0.88–0.99	0.89–0.99	0.90–1.00
NRR	2	0.82–0.95	0.45–0.95	0.87–1.00
ICF	2	0.77–0.98	0.81–0.99	0.93–1.00
IFL	2	0.88–0.98	0.91–0.98	0.92–1.00
NRR	3	0.65–0.91	0.54–0.92	0.87–1.00
ICF	3	0.55–0.98	0.45–0.97	0.94–1.00
IFL	3	0.88–0.98	0.90–0.99	0.91–1.00

Note. ICF: interval from calving to first insemination; IFL: interval from first to last insemination; NRR: nonreturn rate.

lactation in Figure 4). Furthermore, EBV and GEBV correlations within birth year classes were between 0.83 and 1.00 (table or figure not shown, but follows ssGBLUP$_{AB}$ results in Figure 2).

Removing of PPG is unrealistic in practical genetic evaluations. It is important to take into account the correct genetic base of each animal. When the PPG coefficients in the MME of ssGBLUP included both the pedigree and the genomic based relationship matrices, the PCG method converged well. Solutions of GEBVs were obtained after 2,406 PCG iterations, which decreased wall clock computing time from about 11 days to 1.5 days. No peaks occurred in the logarithm of the convergence statistic during the iteration. Furthermore, there were no differences in the PPG solutions between the non-genomic and genomic evaluations, neither with full or reduced data (Figure 3). Genetic trends based on male GEBVs corresponded well with the trends based on EBVs (Figure 1) and correlations between EBVs and GEBVs by birth year were high (Table 2 and Figure 2). The correlations for NRR traits were somewhat lower than those for the other traits within

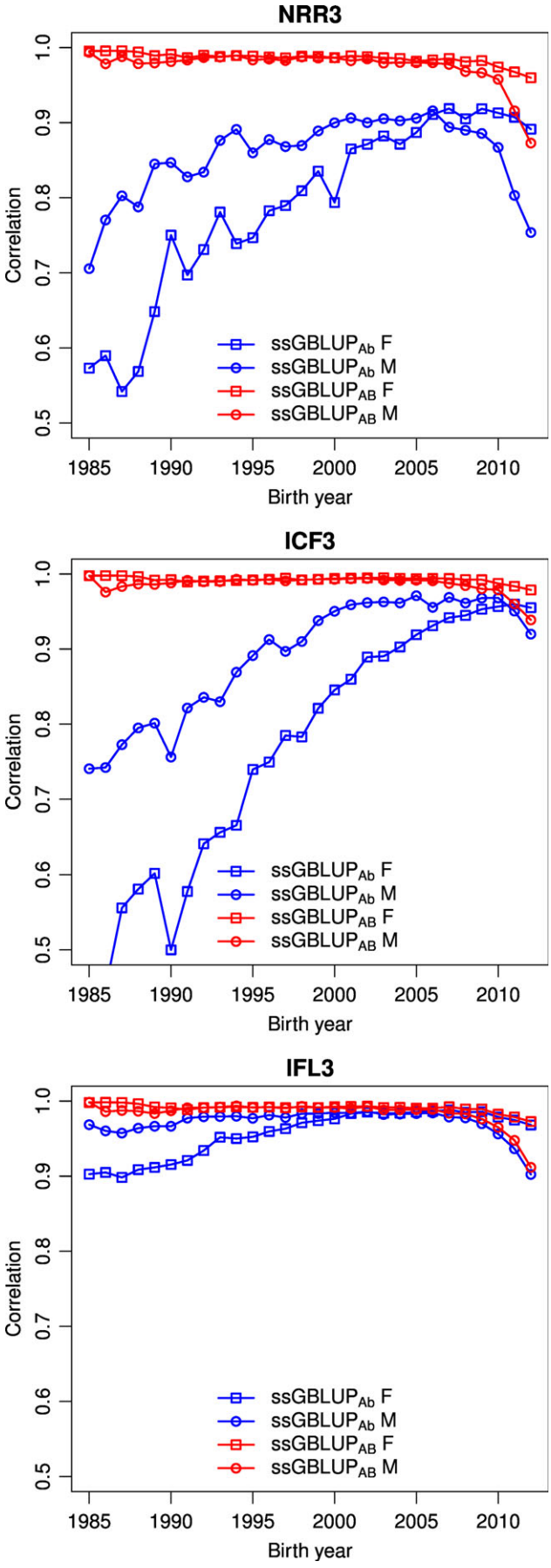

FIGURE 2 Yearly correlation between estimated breeding values by BLUP and genomic enhanced breeding values by single step (ssGBLUP) by sex (F for females and M for males) for third parity traits nonreturn rate (NRR3), days from first to last insemination (IFL3) and days from calving to first insemination (ICF3). Phantom parent group estimation used pedigree information in ssGBLUP$_{Ab}$ but genomic information as well in ssGBLUP$_{AB}$ [Colour figure can be viewed at wileyonlinelibrary.com]

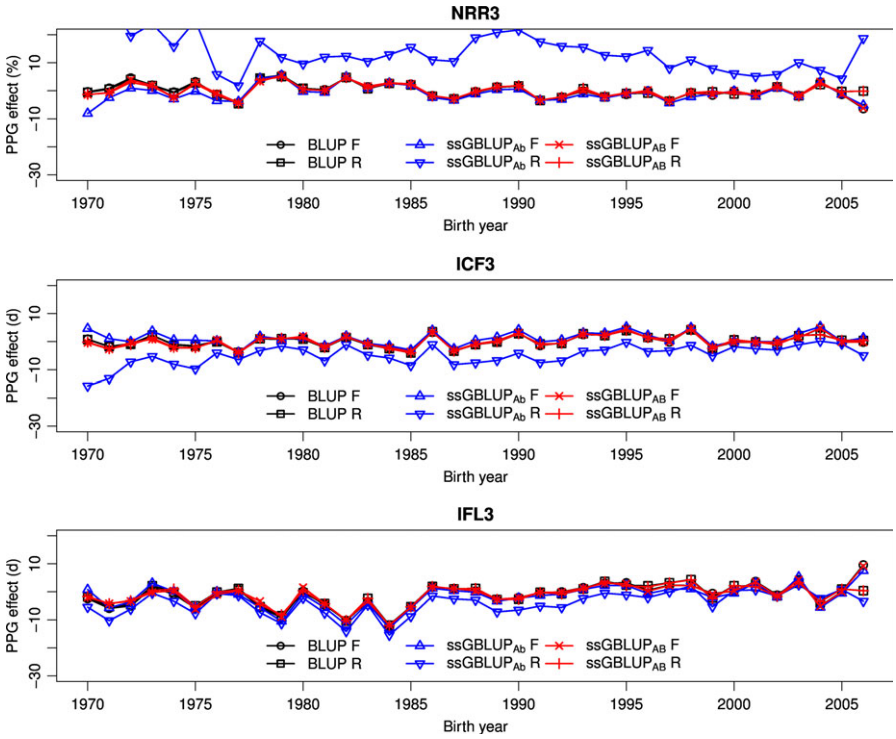

FIGURE 3 Estimated genetic group effects of group describing Finnish Ayrshire for the third parity traits nonreturn rate (NRR3), days from first to last insemination (IFL3) and days from calving to first insemination (ICF3) from the animal model (BLUP) and two genomic (ssGBLUP$_{Ab}$ and ssGBLUP$_{AB}$) evaluations based on both the full (F) and the reduced (R) data. Group equations were based on pedigree information in ssGBLUP$_{Ab}$ but genomic information as well in ssGBLUP$_{AB}$ [Colour figure can be viewed at wileyonlinelibrary.com]

the same parity. For example, NRR1 had correlations between 0.87 and 1.00 over all birth year classes both for males and females, whereas IFL1 and ICF1 had correlations ranging from 0.90 and 0.91 to 1.00. When the PPG coefficients were correctly taken into account in the MME (ssGBLUP$_{AB}$), use of data allele frequencies gave similar convergence and genetic trends as attained with base population allele frequencies, and correlations between GEBVs and EBVs by trait were almost one (results not shown).

Multiplication of $\mathbf{B}$ with the contribution matrix $\mathbf{Q}_2$ generated about 9 million new nonzero elements into $\mathbf{H}_{AB}^{-1}$ in comparison to $\mathbf{H}_A^{-1}$. This was only 1.5% of the original unique elements in $\mathbf{B}$. From the new nonzero coefficients, 99% of the values were between −0.1 and 0.1, but also higher contributions existed especially on the diagonals for the groups associated with the oldest year classes. Mean of the diagonal elements in $\mathbf{Q}_2'\mathbf{B}\mathbf{Q}_2$ was 1.6 and the highest coefficient was 88.6 for the PPG of Finnish RDC born in 1970. This is almost four times larger contribution than the value of 22.7 subtracted from the group equations due to $\mathbf{Q}_2'\mathbf{A}_{22}^{-1}\mathbf{Q}_2$. On average the diagonals and off-diagonals in $\mathbf{Q}_2'\mathbf{B}\mathbf{Q}_2$ were three and five times bigger, respectively, than in $\mathbf{Q}_2'\mathbf{A}_{22}^{-1}\mathbf{Q}_2$. The correlation of diagonals in $\mathbf{Q}_2'\mathbf{G}^{-1}\mathbf{Q}_2$ and $\mathbf{Q}_2'\mathbf{A}_{22}^{-1}\mathbf{Q}_2$ was 0.99 versus 0.97 for base population and data allele frequencies, respectively.

Descriptive statistics of diagonals and off-diagonals in $\mathbf{G}$ and $\mathbf{A}_{22}$ (in Table 3) do not change whether PPGs are included or not in the MME. However, correlations between groups of matrix elements of $\mathbf{G}$ and $\mathbf{A}_{22}$ were higher when $\mathbf{G}$ used base population than data allele frequencies (0.49 versus 0.04 for diagonal values and 0.88 versus 0.80 for off-diagonal values). In the nonscaled $\mathbf{G}$ matrix using base allele frequencies (data allele frequencies in parenthesis), averages of diagonal and off-diagonal values were 16,271.56 and 739.11 (15,532.00 and −0.46), respectively. After scaling of $\mathbf{G}$, the average of diagonal values was 1.02 and the average of off-diagonal values was near zero both in $\mathbf{G}$ and $\mathbf{A}_{22}$. The range of values in $\mathbf{A}_{22}$ was from zero to 1.29, but the range of values in $\mathbf{G}$ was from −0.10 to 1.28 with base allele frequencies and from −0.14 to 1.43 with data allele frequencies.

3.2 | Validation

Validation used DRPs calculated from the EBVs based on the full data. Mean reliabilities of DRPs as a predictor of breeding value for validation bulls by trait are in Table 4. The average reliability decreased as parity increased. Mean reliability for NRR0 was almost as high as reliability for NRR1. However, IFL0 had lower reliability than IFL1-3. This can be explained with lower heritability for IFL0 compared to IFL1-3.

Validation reliabilities and regression coefficients were calculated for PAs from BLUP and for GEBVs from ssGBLUP$_{AB}$ which included QP-transformation for full MME (Table 4). The validation reliabilities for PAs were low or moderate (0.10–0.27). In the case of heifer traits, validation reliabilities for GEBVs were only somewhat higher than for PAs. NRR0 and IFL0 had validation reliabilities 0.19 and 0.27 for PAs, respectively, but 0.23 and 0.29 for GEBVs. Furthermore, the validation reliabilities for both PAs and GEBVs had almost as wide 95% confidence

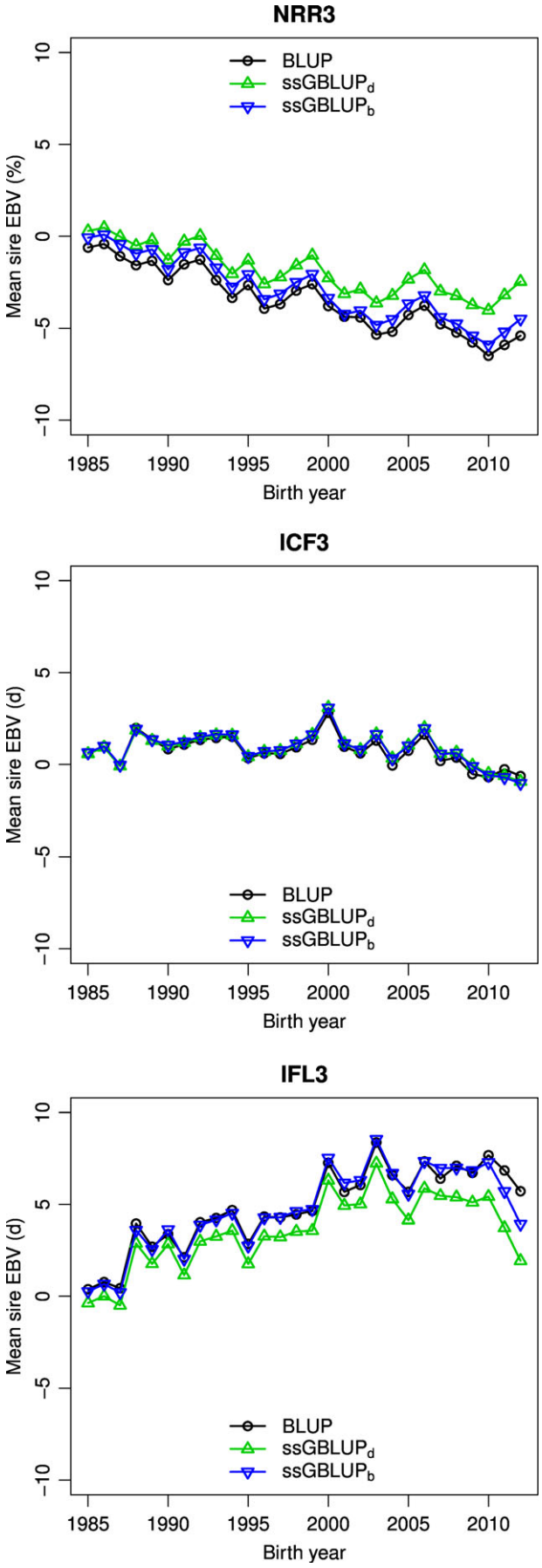

FIGURE 4 Genetic trend based on estimated breeding values from the animal model (BLUP) and genomic enhanced breeding values (GEBV) from the single-step genomic model (ssGBLUP) for RDC Nordic bulls having at least 50 daughters. Base population allele frequencies (ssGBLUP$_b$) or data allele frequencies (ssGBLUP$_d$) were used. None of the models had phantom parent groups. Traits are nonreturn rate (NRR3), interval from calving to first insemination (ICF3) and interval from first to last service (IFL3) in the third parity [Colour figure can be viewed at wileyonlinelibrary.com]

intervals for these traits. In contrast, advantage of the ssGBLUP model was clear for all cow traits. Validation reliabilities increased on average from 0.13 to 0.24 for NRR cow traits, on average from 0.18 to 0.29 for ICF cow traits, and on average from 0.18 to 0.30 for IFL cow traits. The lower limit of the 95% confidence interval for the GEBV validation reliabilities were closer to the corresponding higher limit than the lower limit for PA validation reliabilities.

Regression coefficients for the EBVs varied more than the regression coefficients for the GEBVs (Table 4). Largest differences were in the regression coefficients of the heifer traits, which were clearly higher for the EBVs than for the GEBVs (on average 1.03 versus 0.84). In fact, the regression coefficients for the GEBVs were close to the lower limit of 95% confidence interval for the EBV solutions. For the cow traits, regression coefficients were more similar for the two models. Regression coefficients were on average 0.91 for the EBV solutions and on average 0.90 for the GEBV solutions.

4 | DISCUSSION

Simplifications in MME are common in dairy cattle evaluations for computational reasons. For example, traditional BLUP evaluations often neglect inbreeding in $\mathbf{A}^{-1}$. A major challenge is making computationally feasible

TABLE 3 Mean, standard deviation (*SD*), minimum (Min) and maximum (Max) for diagonal and off-diagonal values of three different relationship matrices: Pedigree based relationship matrix related to genotyped animals ($\mathbf{A}_{22}$), and genomic based relationship matrices both with base and data allele frequencies ($\mathbf{G}_b$ and $\mathbf{G}_d$, respectively)

	Mean	*SD*	Min	Max
Diagonal($\mathbf{A}_{22}$)	1.02	0.01	1.00	1.29
Diagonal($\mathbf{G}_b$)	1.02	0.03	0.74	1.27
Diagonal($\mathbf{G}_d$)	1.02	0.05	0.81	1.43
Off-diagonal($\mathbf{A}_{22}$)	0.05	0.04	0.00	0.81
Off-diagonal($\mathbf{G}_b$)	0.05	0.05	−0.10	1.01
Off-diagonal($\mathbf{G}_d$)	0.00	0.05	−0.14	1.00

TABLE 4 Mean reliability of DRPs for validation bulls (r^2_{DRP}), and validation reliability (R^2) and regression coefficient (δ) of deregressed proofs on the parent averages (PA) and on the genomic enhanced breeding values ($GEBV_{AB}$). 95% bootstrap confidence intervals are in the parentheses

			PA		$GEBV_{AB}$	
Trait	**Parity**	$\overline{r^2_{DRP}}$	R^2	δ	R^2	δ
NRR	0	0.42	0.19 (0.09–0.28)	1.00 (0.75–1.26)	0.23 (0.14–0.32)	0.81 (0.63–0.98)
IFL	0	0.41	0.27 (0.16–0.37)	1.06 (0.83–1.28)	0.29 (0.19–0.39)	0.87 (0.70–1.04)
NRR	1	0.48	0.16 (0.09–0.23)	0.96 (0.74–1.18)	0.27 (0.17–0.36)	0.86 (0.70–1.02)
ICF	1	0.62	0.16 (0.09–0.22)	0.99 (0.77–1.20)	0.28 (0.20–0.36)	0.90 (0.76–1.04)
IFL	1	0.58	0.17 (0.10–0.24)	0.92 (0.73–1.10)	0.31 (0.22–0.39)	0.89 (0.77–1.02)
NRR	2	0.41	0.12 (0.05–0.18)	0.98 (0.71–1.25)	0.24 (0.14–0.34)	0.95 (0.74–1.14)
ICF	2	0.61	0.17 (0.10–0.24)	0.88 (0.69–1.07)	0.29 (0.20–0.37)	0.86 (0.72–0.99)
IFL	2	0.55	0.16 (0.09–0.22)	0.85 (0.67–1.03)	0.29 (0.21–0.38)	0.89 (0.75–1.02)
NRR	3	0.35	0.10 (0.03–0.17)	0.83 (0.53–1.12)	0.22 (0.11–0.32)	0.92 (0.69–1.13)
ICF	3	0.56	0.20 (0.12–0.27)	0.92 (0.73–1.11)	0.31 (0.22–0.39)	0.90 (0.77–1.04)
IFL	3	0.50	0.20 (0.12–0.27)	0.88 (0.71–1.05)	0.31 (0.22–0.41)	0.91 (0.77–1.04)

Note. ICF: interval from calving to first insemination; IFL: interval from first to last insemination; NRR: nonreturn rate.

ssGBLUP evaluations where contributions in MME due to genomic information are compatible with those from pedigree information. Without PPG in the model, compatibility was obtained easily resulting in similar genetic trends and high yearly correlations between EBV and GEBV. With PPG in the model, incompatible matrices that did not account properly coefficients pertaining to PPG in the MME caused the poor PCG method convergence of the ssGBLUP evaluations. The poor convergences lead to GEBVs that were very different from EBVs, although by chance the genetic trends may be similar.

Although the problem was mainly due to accounting only pedigree information in the coefficients of the genetic groups, also inclusion of the inbreeding coefficients in the pedigree based relationship matrix **A** improved the evaluations (Matilainen et al., 2016). Inclusion of genomic information on PPG equations guarantees correct expectation of breeding values for the genotyped animals, and not only the genotypic information, but also the pedigree based information have to be taken correctly into account. In other words, if $\mathbf{A}_A^{-1}$ includes the genetic groups, but $\mathbf{A}_{22}^{-1}$ is constructed assuming only one base population, the breeding value expectations of genotyped animals become incorrect.

When the coefficients for the PPG in the MME were properly accounted and inbreeding coefficients were included consistently in the inverses of the pedigree based relationship matrices, the convergence statistics decreased smoothly. Above all, EBV and GEBV solutions were concordant after the PPG equations of the MME accounted both the pedigree and the genomic relationship matrix information. Genetic trends followed well each other and yearly correlations were high for the old year classes. Correlations decreased within the younger year classes, but this is expected due to the significant additional genomic information for the young animals. Use of either data or base population allele frequencies had little effect on convergence and the genetic trends when pedigree and genomic based relationships were compatible, either with or without PPG.

The inclusion of contributions into PPG equations due to genomic information was found to be relatively easy. The group proportions for the genotyped animals in the contribution matrix $\mathbf{Q}_2$ were attained using readily available pedigree analysis software RelaX2 (Strandén, 2014). Thereafter, the only requirement was multiplication of **B** with the contribution matrix $\mathbf{Q}_2$. The values from this multiplication were included in the same file that provides the **B**-matrix for the MME of the genomic evaluation. Thus, no changes were needed in the MME solver software. Another way to correct inconsistency of the two relationship matrices could be the use of the metafounder approach (Legarra et al., 2015). This would, however, require a large number of group relationship parameters to be estimated for the 332 genetic groups.

Using only pedigree based information in PPG coefficients of MME had no remarkable influence in the ssGBLUP evaluations of dairy production traits (Koivula et al., 2015), but we observed poor convergence in the fertility trait evaluations. The effect may have been especially strong because some PPG were only two generations away from the genotyped animals in our data. Low heritabilities for the fertility traits may also have increased size of the effect which is due to inconsistent inclusion of the pedigree

and genomic relationship matrices. It is unclear if PPG should include contributions due to genomic relationships also in the evaluations when there are no convergence problems. Large coefficient values resulting in multiplication of **B** with contribution matrix $\mathbf{Q}_2$ could be a useful indicator on the need to include genomic information in the genetic group coefficients. Inadequate inclusion of the contributions gave PPG effect estimates that differed between the genomic evaluations of the full and the reduced data across the year classes. Hence, differences in the PPG solutions from full and reduced data analyses may be the first indicator for an unstable MME system and incorrectly accounted contributions into PPG. Also differences in the PPG solutions from the traditional genetic and the genomic evaluations within both the full and the reduced data may indicate the incorrectly accounted contributions. Exact limits for the differences are difficult to say, however, because these are likely to depend on data and population.

The validation reliabilities for PA of NRR traits in this study were higher than those for the conception rate traits (on average 0.06) reported in Aguilar et al. (2011). Main reason for this may be that in our study there were additional correlated traits in the multiple trait model. However, PA validation reliabilities for the fertility traits (on average 0.17) were lower compared to the milk production traits (on average 0.30) in Koivula et al. (2015). In absence of selection, the validation reliability is an estimate of the prediction reliability. With low heritability traits, the candidate PA reliability is roughly 5/16 of a reliability of progeny tested sires. This seems generally as a right assumption which can be verified using the average reliabilities of DRPs for validation bulls in Table 4.

For most of the traits the validation reliabilities for GEBVs were substantially higher than those for PAs. Also, the confidence intervals for the GEBV reliabilities were slightly larger than those for the PA reliabilities, although the increase in the confidence intervals was less than the increase in the reliabilities. The length of confidence intervals for the validation reliabilities were between 70 and 140 percent of the estimates for PA reliabilities, and between 55 and 95 percent of the validation reliabilities for GEBVs. For example, length of confidence interval of the validation reliability for NRR3 increased from 0.14 to 0.21, but the relative length decreased from 140 to 95 percent.

The improvement due to the genomic information in the validation reliability for the cow traits corresponded to approximately 51 additional ERCs. The biggest relative improvement due to genomic information was for the NRR cow traits, for which on average 101 additional ERCs would have been needed to attain similar increase in the validation reliability. The corresponding addition for the other cow traits would have been on average 27 ERCs. Improvement due to genomic information was less clear for the heifer traits. The estimates of R^2 for NRR0 and IFL0 GEBVs were within confidence interval of the R^2 for corresponding PA. Especially for IFL0, R^2 for GEBVs remained low while for the cow traits R^2 for GEBVs was almost double the R^2 for PA. Furthermore, R^2 for PAs were higher for the heifer traits compared to corresponding cow traits, whereas R^2 for GEBVs were lower for the heifer traits compared to corresponding cow traits. Overall, validation reliabilities were nearly equal across parities for both EBVs and GEBVs, although the model validation reliability has usually decreased as parity has increased (e.g. Aguilar et al., 2011; Lourenco et al., 2014).

Regression coefficient δ close to one indicates that the evaluation based on the reduced data predicts the future differences between animals properly. Based on this, GEBV estimates were more inflated than EBVs. This has been observed also for milk and protein yields in Koivula et al. (2015) for bull validation results when using the same polygenic proportion (10%) in the genomic relationship matrix as in our study. However, in all the validation tests, δ for the GEBVs was within the confidence interval of δ for the PAs. Moreover, confidence interval of δ for the GEBVs included 1.0 with every trait except NRR0 (upper limit 0.98) and ICF2 (upper limit 0.99), and, thus, these GEBVs can be considered unbiased according to δ. For the final score and type traits in US Holstein, Tsuruta et al. (2011) reported increase in the regression coefficients when using different weights for the genomic and pedigree relationship matrices. Choice of optimal weights increased regression coefficients also for milk production traits for Nordic Red dairy cattle in Koivula et al. (2015). However, we see only little gain to be achieved through such weights, as in nine traits of 11 the GEBVs were not statistically biased.

In this study, ssGBLUP for Red dairy cattle fertility traits in the joint Nordic evaluations was explored. Modelling was based on the multiple trait multilactation animal model and variance components used in the routine Nordic fertility evaluation. Large-scale ssGBLUP evaluation was found to be feasible when the coefficients of the genetic groups in the MME accounted both the pedigree and genomic information correctly. In this case faster convergence by the iterative solver and more reliable GEBVs were obtained based on the comparisons with the EBVs. Model validation showed that, especially for the cow traits, ssGBLUP improved the fertility evaluations compared to the traditional BLUP without genomic information. Based on the genetic trends, correlations between EBV and GEBV estimates, and validation results, genomic information improved the breeding value prediction in the youngest birth year classes.

ACKNOWLEDGEMENTS

Nordic Cattle Genetic Evaluation Ltd (NAV, Aarhus, Denmark), Viking Genetics (Randers, Denmark), Faba (Hollola, Finland), Valio (Helsinki, Finland), and MAKERA foundation (Finnish Ministry of Agriculture and Forestry) are acknowledged for financial support.

CONFLICT OF INTEREST

The authors have no conflict of interest.

ORCID

Kaarina Matilainen http://orcid.org/0000-0003-3539-6503
Ismo Strandén http://orcid.org/0000-0003-0161-2618

REFERENCES

Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, *93*, 743–752. https://doi.org/10.3168/jds.2009-2730

Aguilar, I., Misztal, I., Tsuruta, S., Wiggans, G. R., & Lawlor, T. J. (2011). Multiple trait genomic evaluation of conception rate in Holsteins. *Journal of Dairy Science*, *94*, 2621–2624. https://doi.org/10.3168/jds.2010-3893

Chen, C. Y., Misztal, I., Aguilar, I., Legarra, A., & Muir, W. M. (2011). Effect of different genomic relationship matrices on accuracy and scale. *Journal of Animal Science*, *89*, 2673–2679. https://doi.org/10.2527/jas.2010-3555

Christensen, O. F., & Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, *42*, 2. https://doi.org/10.1186/1297-9686-42-2

Christensen, O. F., Madsen, P., Nielsen, B., Ostersen, T., & Su, G. (2012). Single-step methods for genomic evaluation in pigs. *Animal*, *6*, 1565–1571. https://doi.org/10.1017/S1751731112000742

Fogh, A., Roth, A., Pedersen, O. M., Eriksson, J.-Å., Juga, J., Toivonen, M., … Aamand, G. P. (2003). A joint Nordic model for fertility traits. *Interbull Bulletin*, *31*, 52–55.

Forni, S., Aguilar, I., & Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution*, *43*, 1. https://doi.org/10.1186/1297-9686-43-1

Koivula, M., Strandén, I., Pösö, J., Aamand, G. P., & Mäntysaari, E. A. (2015). Single-step genomic evaluation using multitrait random regression model and test-day data. *Journal of Dairy Science*, *98*, 2775–2784. https://doi.org/10.3168/jds.2014-8975

Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics*, *200*, 455–468. https://doi.org/10.1534/genetics.115.177014

Lourenco, D. A. L., Misztal, I., Tsuruta, S., Aguilar, I., Ezra, E., Ron, M., … Weller, J. I. (2014). Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. *Journal of Dairy Science*, *97*, 1742–1752. https://doi.org/10.3168/jds.2013-6916

Mäntysaari, E. A., Liu, Z., & VanRaden, P. (2010). Interbull validation test for Genomic evaluations. *Interbull Bulletin*, *41*, 17–22.

Mäntysaari, E. A., & Van Vleck, L. D. (1989). Estimation of genetic parameters for production and reproduction in Finnish Ayrshire cattle. *Journal of Dairy Science*, *72*, 2375–2386. https://doi.org/10.3168/jds.S0022-0302(89)79370-X

Matilainen, K., Koivula, M., Strandén, I., Aamand, G. P., & Mäntysaari, E. A. (2016). Managing genetic groups in single-step genomic evaluations applied on female fertility traits in Nordic Red Dairy cattle. *Interbull Bulletin*, *50*, 71–75.

McPeek, M. S., Wu, X., & Ober, C. (2004). Best linear unbiased allele-frequency estimation in complex pedigrees. *Biometrics*, *60*, 359–367. https://doi.org/10.1111/j.0006-341X.2004.00180.x

Misztal, I., Vitezica, Z. G., Legarra, A., Aguilar, I., & Swan, A. A. (2013). Unknown-parent groups in single-step genomic evaluation. *Journal of Animal Breeding and Genetics*, *130*, 252–258. https://doi.org/10.1111/jbg.12025

Muuttoranta, K., Tyrisevä, A.-M., Mäntysaari, E. A., Pösö, J., Aamand, G. P., Eriksson, J.-Å., … Lidauer, M. H. (2015). Genetic Parameters for Female Fertility in Nordic Dairy Cattle. *Interbull Bulletin*, *49*, 32–35.

Philipsson, J. (1981). Genetic aspects of female fertility in dairy cattle. *Livestock Production Science*, *8*, 307–319. https://doi.org/10.1016/0301-6226(81)90049-X

Quaas, R. L., & Pollak, E. J. (1981). Modified equations for sire models with groups. *Journal of Dairy Science*, *64*, 1868–1872. https://doi.org/10.3168/jds.S0022-0302(81)82778-6

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sewalem, A., Kistemaker, G. J., & Miglior, F. (2010). Relationship between female fertility and production traits in Canadian Holsteins. *Journal of Dairy Science*, *93*, 4427–4434. https://doi.org/10.3168/jds.2009-2915

Strandén, I. (2014). *RelaX2: program for pedigree analysis*. Version 1.65. Jokioinen, Finland: Natural Resources Institute Finland (Luke).

Strandén, I., Lidauer, M., Mäntysaari, E. A., & Pösö, J. (2000). Calculation of Interbull weighting factors for the Finnish test day model. *Interbull Bulletin*, *26*, 78–81.

Strandén, I., & Mäntysaari, E. A. (2010). A recipe for multiple trait deregression. *Interbull Bulletin*, *42*, 21–24.

Strandén, I., Matilainen, K., Aamand, G. P., & Mäntysaari, E. (2017). Solving efficiently large single-step genomic best linear unbiased prediction models. *Journal of Animal Breeding and Genetics*, *134*, 264–274. https://doi.org/10.1111/jbg.12257

Tsuruta, S., Misztal, I., Aguilar, I., & Lawlor, T. J. (2011). Multiple-trait genomic evaluation of linear type traits using genomic and phenotypic data in US Holsteins. *Journal of Dairy Science*, *94*, 4198–4204. https://doi.org/10.3168/jds.2011-4256

Tyrisevä, A.-M., Muuttoranta, K., Pösö, J., Nielsen, U. S., Eriksson, J.-Å., Aamand, G. P., … Lidauer, M. (2017). Evaluation of conception rate in Nordic dairy cattle. *Interbull Bulletin*, *51*, 57–62.

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M., Sanders, A. H., Tooker, M. E., Miller, R. H., Norman, H. D., Kuhn, M. T., & Wiggans, G. R. (2004). Development

of a national genetic evaluation for cow fertility. *Journal of Dairy Science*, *87*, 2285–2292. https://doi.org/10.3168/jds.S0022-0302(04)70049-1

Vitezica, Z. G., Aguilar, I., Misztal, I., & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, *93*, 357–366. https://doi.org/10.1017/S001667231100022X

Vuori, K., Strandén, I., Lidauer, M., & Mäntysaari, E. A. (2006). MiX99 - effective solver for large and complex linear mixed models. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006, pp.27–33

How to cite this article: Matilainen K, Strandén I, Aamand GP, Mäntysaari EA. Single step genomic evaluation for female fertility in Nordic Red dairy cattle. *J Anim Breed Genet*. 2018;135:337–348. https://doi.org/10.1111/jbg.12353