Use of a reduced set of single nucleotide polymorphisms for genetic evaluation of resistance to *Salmonella* carrier state in laying hens

A. Legarra,*[1] F. Calenge,† P. Mariani,‡ P. Velge,§ and C. Beaumont†

*INRA, UR 631 SAGA, F-31326 Castanet Tolosan, France; †INRA, UMR 083 URA, F-37083 Nouzilly, France; ‡Parco Tecnologico Padano, CERSA, 26900 Lodi, Italy; and §INRA, UR 1282 IASP, F-37083 Nouzilly, France

ABSTRACT *Salmonella* propagation by apparently healthy chickens could be decreased by the selection and use of chicken lines that are more resistant to carrier state. Using a reduced set of markers, this study investigates, for the first time to the authors' knowledge, the feasibility of a genomic selection approach for resistance to carrier state in hen lines. In this study, commercial laying hen lines were divergently selected for resistance to *Salmonella* carrier state at 2 different ages: young chicks and adults at the peak of lay. A total of 600 birds were typed with 831 informative SNP markers and artificially infected with *Salmonella* Enteritidis. Phenotypes were collected 28 d (389 young animals) or 38 d (208 adults) after infection. Two types of variance component analyses, including SNP data or not, were performed and compared. The set of SNP used was efficient in capturing a large part of the genetic variation. Average accuracies from mixed model equations did not change between analyses, showing that using SNP data does not increase information in this data set. These results confirm that genomic selection for *Salmonella* carrier state resistance in laying hens is promising. Nevertheless, a denser SNP coverage of the genome on a greater number of animals is still needed to assess its feasibility and efficiency.

Key words: genomic selection, *Salmonella*, single nucleotide polymorphism, laying hen

2011 Poultry Science 90:731–736
doi:10.3382/ps.2010-01260

INTRODUCTION

Salmonella is the second leading cause of reported food poisoning cases in Europe, after *Campylobacter* (EFSA, 2010). Humans can be infected after the ingestion of contaminated poultry meat or eggs. The serotype *Salmonella enterica* Enteritidis in particular is responsible for most cases of human food poisoning outbreaks caused by *Salmonella*. This serotype can be carried by apparently healthy birds that cannot eliminate it and thus contribute to *Salmonella* propagation. In addition to prophylactic measures and vaccination, the selection and use of chicken lines more resistant to *Salmonella* carrier state could be a way to decrease the propagation of *Salmonella* in poultry stocks and hence a way to improve food safety.

The feasibility of selection for improved resistance to carrier state has been demonstrated by a divergent selection experiment from a laying hen commercial line, conducted in parallel on young chicks (7 d of age) and on adults hens at the peak of lay (Beaumont et al., 2009). Genetic heritabilities ranged from 0.14 to 0.23 according to the trait assessed and to the chicken's age. Selection efficiency was reduced because of the need to perform selection on siblings of the infected animals. Knowing the genes involved in the control of resistance to *Salmonella* carrier state could allow a marker-assisted selection performed directly on the infected animals, thus improving selection efficiency. In addition, artificial infections would not be needed at each generation because animals would be selected according to their genotype. Although many studies led to the identification of candidate genes or genomic regions controlling resistance to salmonellosis, few studies focused on genes controlling resistance to *Salmonella* carrier state (reviewed in Calenge et al., 2010). A QTL study identified QTL for resistance to carrier state on chromosomes 1, 2, 5, 11, and 16 in experimental chicken lines infected with *S.* Enteritidis at 7 d of age (Tilquin et al., 2005; Calenge et al., 2009). To assess the interest of these QTL for commercial selection, their effect on resistance to carrier state was tested on the above-mentioned divergent lines; the same study also investigated the effects of the candidate genes solute carrier family 11, member 1 (*SLC11A1*) and toll-like receptor 4 (*TLR4*; Calenge et al., 2009). The genomic regions carrying the candidate gene *SLC11A1* and 2 QTL on chromosomes

©2011 Poultry Science Association Inc.
Received November 26, 2010.
Accepted January 8, 2011.
[1]Corresponding author: andres.legarra@toulouse.inra.fr

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015

1 and 16 were found to be associated with resistance to carrier states in these lines (Calenge et al., 2009). Nevertheless, these loci account for only a small part of the total genetic variation observed in those lines, so a marker-assisted selection based on these genes alone would not be efficient. A more exhaustive study of the genes of markers controlling resistance to carrier state in these lines is needed to assess the feasibility of a marker-assisted selection.

In the present study, we used a low-density SNP chip covering the entire genome to perform a genetic evaluation of resistance to *Salmonella* carrier state in the divergent lines. The analyses performed were based on the first steps of the genomic selection methodology (Meuwissen et al., 2001). The objective of this work was to evaluate the interest of including SNP marker data in the genetic evaluation of fowls; it was based on the divergent lines and phenotypes measured after infection with *S.* Enteritidis. To reach this purpose, we compared, using estimates of genetic parameters and estimated reliabilities, the efficiency of a SNP-assisted genetic evaluation with the efficiency of a classical pedigree evaluation in these lines.

MATERIALS AND METHODS

An experiment of divergent selection for either an increased or a decreased resistance to *Salmonella* carrier state was achieved (for details see Beaumont et al., 2009). The base population was a commercial laying hen line. Selection was conducted in parallel on 7-d-old chicks and on adult hens at the peak of lay so that 4 lines were produced. Animals from the seventh breeding generation of those 4 divergent lines were challenged for the present study (389 chicks and 208 adult hens). Sires of the animals belonging to the chick line were also genotyped.

Salmonella Challenges

The *S.* Enteritidis nalidixic acid- and streptomycin-resistant PT 4 strain 1009 was used for all challenges. Chick resistance was assessed as described previously (Duchet-Suchaux et al., 1995). Chicks (7 d of age) were orally inoculated with 5×10^4 cfu of *S.* Enteritidis. Birds were slaughtered 35 d postinfection and bacteria were counted in the removed ceca. Cecal bacteria counts were expressed in log colony forming units per gram of ceca.

Resistance of adult hens was assessed as described previously (Protais et al., 1996). A total of 265 hens were orally inoculated at their peak of lay (140–168 d) with 10^9 cfu of *S.* Enteritidis. They were slaughtered 28 d postinfection and the liver, spleen, and ceca were removed for *Salmonella* counting. The presence or absence of *Salmonella* in the spleen, liver, and ceca and the animal contamination rate (0 for no organ contaminated, 1 for $\geq$1 organ contaminated) were considered for further analyses. A total of 5 traits, 1 describing chick carrier state and 4 describing adult hen carrier state, were thus used for further analyses.

Genotyping with SNP Markers

Genotypes for 1,536 SNP markers were obtained using Illumina Golden Gate (Ilumina, San Diego, CA) on a BeadExpress station (KOS Genetics, Milan, Italy). A total of 194 markers were chosen to specifically cover 3 previously identified QTL regions on chromosomes 1, 2, and 5 (Tilquin et al., 2005; Calenge et al., 2009), and the remaining 1,342 SNP markers were chosen to homogeneously cover the entire genome. Respectively, 389 and 208 individuals were genotyped in populations measured for chick and adult resistance. Most individuals were half-sibs, issued from 13 (12 of them genotyped) and 16 (none of them genotyped) sires for chick and adult resistance, respectively.

Parameters Estimation and Genetic Evaluation

For each of the 6 traits measured in the 4 divergent lines, 2 analyses were performed. The first did not take genotypes into account but rather used the usual pedigree-based relationship matrix **A**; this approach was referred to as control. The second analysis considered both pedigree and genotypes at the SNP sets and used 2 random effects, one with covariance matrix **A** (thus based on pedigree) and the other using a combined pedigree-genomic relationship matrix **H** (Legarra et al., 2009); this approach was referred to as combined (combined variance components). This matrix **H** is equivalent to predict genotypes for nongenotyped animals (Christensen and Lund, 2010). Two random effects were used for 2 reasons: 1) a priori the markers did not properly cover all genetic variability, and 2) this procedure made it possible to measure to what extent the phenotypic variation was better explained by pedigree and markers than by pedigree alone.

In a preliminary analysis, variance components were estimated by Bayesian methods and Gibbs sampling (Sorensen and Gianola, 2002) using flat priors. The mean of the posterior distribution was retained as the estimate of the parameters of interest. After variance component estimation and using the estimates of variance components, best linear unbiased prediction (**BLUP**; Henderson, 1973) predictions of genetic values were computed either using the pedigree-based relationship matrix **A** only (control) or also including the pedigree-genomic relationship matrix **H** (combined). To obtain only 1 (and not 2) estimated breeding value, a new relationship matrix was created weighting each matrix by its associated variance component, as ${}^{*}\sigma_g^2 = \mathbf{H}\sigma_h^2 + \mathbf{A}\sigma_u^2$, where $\sigma_g^2 = \sigma_h^2 + \sigma_u^2$ (Christensen and Lund, 2010), where σ_h^2 is variance associated with markers and σ_u^2 is variance associated with pedigree. In

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015

matrix **H**, a submatrix of genomic relationships was computed including all polymorphic SNP as $\mathbf{G} = \mathbf{ZZ}'/k$, where **Z** is an incidence matrix of SNP in animals coded for substitution effects of each SNP (i.e., $-2p$, $1–2p$, $2–2p$ for the AA, Aa, and aa genotypes, where p is the frequency of a; VanRaden, 2008). Coefficient k was obtained as the trace of $\mathbf{ZZ}'$ to obtain values of the diagonal of **G** similar to 1; thus, k scales **G** so that the variance of the genetic values implied by **G** is equivalent, on average, to the variance of genetic values implied by **A** (Gianola et al., 2009). Proper weighting of **G** is important when combining pedigree and genomic information (Van Raden et al., 2009; Aguilar et al., 2010). In both variance components and BLUP estimates, the BLUPF90 series of programs (Misztal et al., 2002) was used, with the modifications included to account for genomic relationship matrices as described by Aguilar et al. (2010).

The theoretical accuracy r was computed in both cases from the diagonal elements (PEV) of the inverse of the mixed model equations as $r = \sqrt{1 - \frac{PEV}{\sigma_g^2}}$, where σ_g^2 is the genetic variance.

RESULTS AND DISCUSSION

Salmonella Challenges

Presence of bacteria in liver and spleen was rare (frequencies of 1 and 3%, respectively), yet higher for ceca and overall (29 and 32%, respectively). Average chick load was 3.96 (±1.88) log cfu.

SNP Marker Genotypes Obtainment

From 1,536 original SNP, only 831 turned out to be polymorphic; the rest were discarded for further statistical analyses. A set of 141 SNP were not in Hardy-Weinberg equilibrium ($P < 10^{-6}$), but they were retained in the analysis because a selected population is not expected to behave in Hardy-Weinberg equilibrium. For example, the USDA procedures for genomic evaluation of dairy cattle do not filter SNP according to Hardy-Weinberg (Wiggans et al., 2010). Coverage of the genome is indeed scarce; for example, the squared correlation between consecutive markers is, on average, 0.05, whereas typical values in dairy cattle are 0.3 or higher (Sargolzaei et al., 2008).

Genetic Parameter Estimation and Evaluation

Table 1 shows estimates of genetic parameters. Results of the control analysis agree with estimates of Beaumont et al. (2009) for the same lines, except for adult liver, where our estimates were much lower. For most traits, standard errors of heritability were around 0.07, except for adult liver, where it was 0.01, probably because heritability was very low (0.01). On one hand, estimated residual variances do not change between control and combined analyses. This implies that none of the SNP captured a large effect on the trait, despite some being in QTL regions. On the other hand, it can be seen that in the second analysis (combined) most (but not all) heritability was captured by markers. Indeed, heritability explained by markers ranged between 40 and 90% of all heritability. This shows that markers were efficient in capturing genetic information, although it is difficult to distinguish whether it was because of strong associations with QTL or because they were surrogates of family information (Habier et al., 2007).

Table 2 and Figure 1 show theoretical accuracies computed from the inverse of the matrix of mixed model equations in each model. Average accuracies did not change, which confirms that the use of SNP did not really increase the information. However, for adult measures, their standard deviation increased, particularly for liver contamination. This can also be seen from the spread of accuracies in Figure 1: using SNP led to less uniform information from animal to animal.

Table 1. Estimates of genetic parameters of analyses with pedigree only (control) or pedigree and SNP markers (combined) for chicken cecal load, adult liver, spleen, or cecal contamination as well as animal contamination[1]

Method of genetic evaluation	Chicken cecal load	Adult liver	Adult spleen	Adult ceca	Animal contamination
Control					
Var(e)	1.716	0.011	0.022	0.170	0.187
h_u^2	0.165	0.039	0.171	0.181	0.212
Combined					
Var(e)	1.850	0.011	0.024	0.180	0.193
h_u^2	0.048	0.002	0.019	0.023	0.021
h_h^2	0.034	0.009	0.079	0.127	0.197

[1]Residual variance [Var(e)] and heritability explained by pedigree $\left(h_u^2\right)$ or markers $\left(h_h^2\right)$.

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015

Table 2. Mean ± SD of theoretical accuracies computed from mixed model equations using either classical selection based on phenotype and pedigree (control) or combined selection based on the former data and SNP markers (combined) for chicken cecal load, adult liver, spleen, or cecal contamination as well as animal contamination

Method of genetic evaluation	Chicken cecal load	Adult liver	Adult spleen	Adult ceca	Animal contamination
Control	0.41 ± 0.02	0.18 ± 0.03	0.45 ± 0.04	0.52 ± 0.04	0.59 ± 0.04
Combined	0.42 ± 0.05	0.20 ± 0.09	0.42 ± 0.08	0.51 ± 0.06	0.59 ± 0.05

Figure 1 also shows that for some animals, genetic evaluations were greatly dependent on the information that was used for genetic evaluation (i.e., control or combined). Using markers may allow for a more pertinent within-family evaluation, contributing to a more efficient selection.

Genomic selection relies on the exploitation of preexisting, historical linkage disequilibrium between SNP markers and the trait of interest. This whole-genome approach is more exhaustive than marker-assisted selection based on a small number of genes. Moreover, genetic evaluations can be performed directly in the populations under selection, discarding the need for confirmation of QTL effects in those populations. Genomic selection is thus a promising tool for those traits that are expensive or difficult to record in the candidates to selection (Goddard and Hayes, 2009). *Salmonella* carrier state resistance is undoubtedly one of these traits. *Salmonella* challenges are expensive and selection has to be conducted on sib animals, which considerably reduces the selection efficiency. In addition, the heritability of resistance to *Salmonella* carrier state is rather

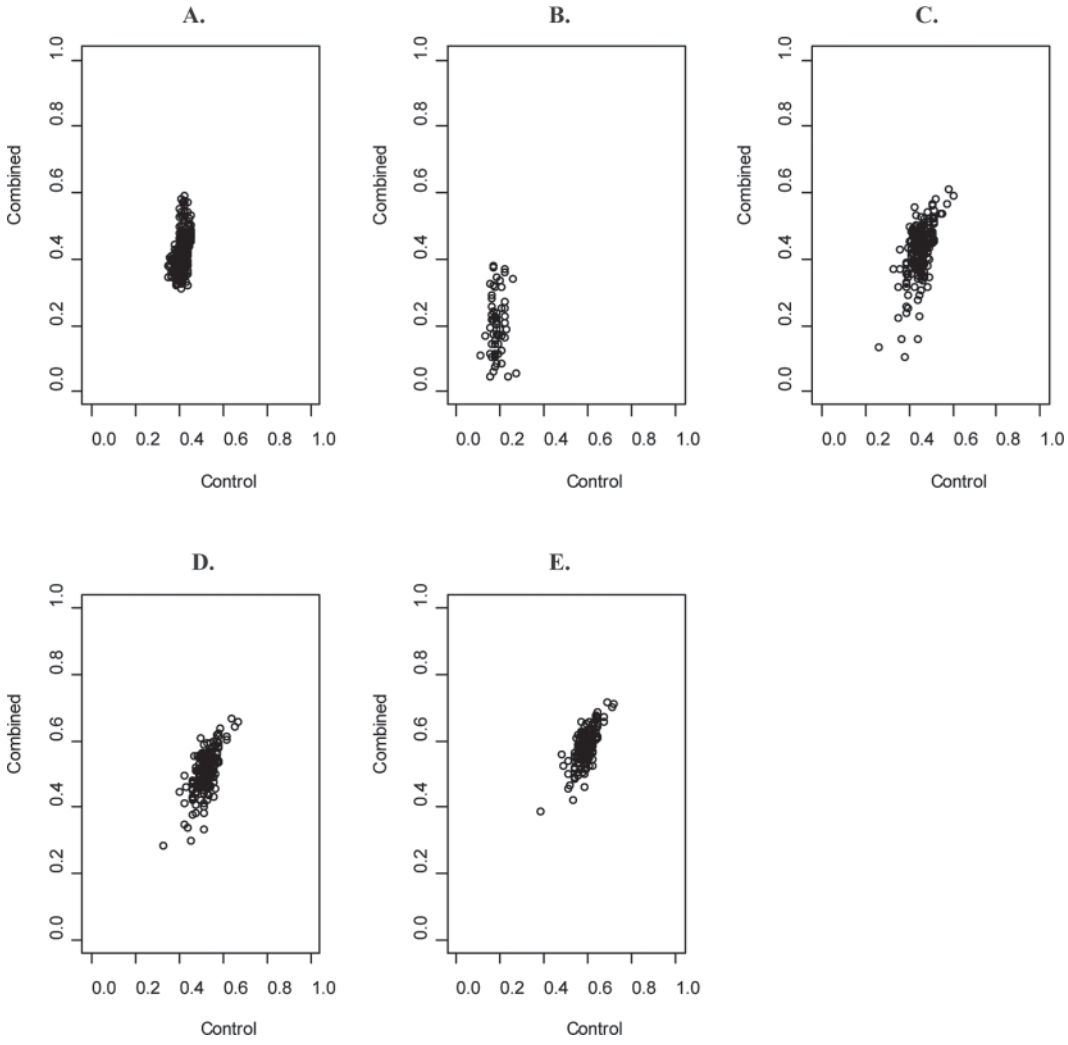

Figure 1. Theoretical accuracies from the inverse of the mixed model equations using either pedigree only (control) or pedigree and SNP markers (combined) for A) chicken cecal load, B) adult liver, C) spleen, or D) cecal contamination as well as E) animal contamination.

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015

weak (Berthelot et al., 1998; Girard-Santosuosso et al., 2002; Beaumont et al., 2009), so for this trait marker-assisted selection should be more efficient than classical pedigree selection.

Before applying genomic selection for resistance to carrier state in laying hens, this preliminary study intended to assess the interest of including marker data in the genetic evaluation of animals, using the first steps of the genomic selection methodology. We used the single step method (Legarra et al., 2009; Aguilar et al., 2010) to accommodate all available information (i.e., genotypes, pedigree, and the fact that some sires are genotyped whereas others are not). This study shows that such a procedure may be efficiently implemented.

The SNP markers were efficient in capturing genetic variation. This is probably partly attributable to the specific targeting of known QTL regions on chromosomes 1, 2, and 5 (Tilquin et al., 2005; Calenge et al., 2009). Nevertheless, none captured a large effect on the traits analyzed. This is most probably attributable to the small number of SNP used. Genomic selection relies on the use of SNP markers in close linkage disequilibrium with causal genes controlling the trait under interest. A few studies have been conducted to assess linkage disequilibrium in broiler or laying hen lines (Heifetz et al., 2005; Aerts et al., 2007; Andreescu et al., 2007). They all agree for a low extent of linkage disequilibrium, which implies that a very dense SNP coverage should be used to reliably cover the genome. For instance, Megens et al. (2009) recommend the use of 10,000 to 20,000 SNP markers in laying hens, provided the haplotype block structure is taken into account. Thus, possible improvements in accuracy will come from the use of many more SNP, each capturing a small part of genetic variation. Indeed, this is the trend in current studies (Avendaño et al., 2010; Chen et al., 2011)

Average accuracies do not change, which confirms that the use of SNP did not really increase the information. Those figures are theoretical accuracies. Cross-validation or measures of goodness of fit are difficult to compute because of the scarcity of available data. However, it has been shown (VanRaden et al., 2009) that increase of true accuracies (as assessed by cross-validation) by the inclusion of genomic information comes with an increase of this theoretical accuracy by providing additional genomic relationships. Thus, use of genomic data (as here) that do not increment r (the theoretical accuracy) in respect to a pedigree-based procedure is not expected to increase true accuracies.

In conclusion, this study shows that a SNP-assisted selection for resistance to carrier state may be implemented and even little SNP are efficient in capturing genetic variation. Overall, however, the use of SNP markers did not much change the picture of genetic evaluation. Using a denser SNP chip and more data would allow a more efficient evaluation of the possibilities of genomic selection for resistance to *Salmonella* carrier state.

ACKNOWLEDGMENTS

This study was funded by the Technology Transfer project of the EADGENE project (European Animal Disease Genomics Network of Excellence for Animal Health and Food Safety). Project partly supported by Toulouse Midi-Pyrénées bioinformatic platform (France).

REFERENCES

Aerts, J., H. Megens, T. Veenendaal, I. Ovcharenko, R. Crooijmans, L. Gordon, L. Stubbs, and M. Groenen. 2007. Extent of linkage disequilibrium in chicken. Cytogenet. Genome Res. 117:338–345.

Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. Lawlor. 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.

Andreescu, C., S. Avendano, S. Brown, A. Hassen, S. Lamont, and J. Dekkers. 2007. Linkage disequilibrium in related breeding lines of chicken. Genetics 177:2161–2169.

Avendaño, S., K. Watson, and A. Kranis. 2010. Genomics in poultry breeding—From utopias to deliverables. In 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany. CD-ROM comm 0049.

Beaumont, C., H. Chapuis, J. Protais, N. Sellier, P. Menanteau, P. Fravalo, and P. Velge. 2009. Resistance to *Salmonella* carrier state: Selection may be efficient but response depends on animal's age. Genet. Res. (Camb) 91:161–169.

Berthelot, F., C. Beaumont, F. Mompart, O. Girard-Santosuosso, P. Pardon, and M. Duchet-Suchaux. 1998. Estimated heritability of the resistance to cecal carrier state of *Salmonella enteritidis* in chickens. Poult. Sci. 77:797–801.

Calenge, F., P. Kaiser, A. Vignal, and C. Beaumont. 2010. Genetic control of resistance to salmonellosis and to *Salmonella* carrier-state in fowl: A review. Genet. Sel. Evol. 42:11.

Calenge, F., F. Lecerf, J. Demars, K. Feve, F. Vignoles, F. Pitel, A. Vignal, P. Velge, N. Sellier, and C. Beaumont. 2009. QTL for resistance to *Salmonella* carrier state confirmed in both experimental and commercial chicken lines. Anim. Genet. 40:590–597.

Chen, C. Y., I. Misztal, I. Aguilar, S. Tsuruta, T. H. E. Meuwissen, S. E. Aggrey, T. Wing, and W. M. Muir. 2011. Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens. J. Anim. Sci. 89:23–28.

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

Duchet-Suchaux, M., P. Léchopier, J. Marly, P. Bernardet, R. Delaunay, and P. Pardon. 1995. Quantification of experimental *Salmonella enteritidis* carrier state in B13 leghorn chicks. Avian Dis. 39:796–803.

EFSA. 2010. The community summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. EFSA J. 8:1496.

Gianola, D., G. de los Campos, W. Hill, E. Manfredi, and R. Fernando. 2009. Additive genetic variability and the Bayesian alphabet. Genetics 183:347–363.

Girard-Santosuosso, O., F. Lantier, I. Lantier, N. Bumstead, J.-M. Elsen, and C. Beaumont. 2002. Heritability of susceptibility to *Salmonella enteritidis* infection in fowls and test of the role of the chromosome carrying the NRAMP1 gene. Genet. Sel. Evol. 32:211–219.

Goddard, M. E., and B. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10:381–391.

Habier, D., R. L. Fernando, and J. C. M. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177:2389–2397.

Heifetz, E. M., J. Fulton, N. O'Sullivan, H. Zhao, J. Dekkers, and M. Soller. 2005. Extent and consistency across generations of linkage

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015

disequilibrium in commercial layer chicken breeding populations. Genetics 171:1173–1181.
Henderson, C. 1973. Sire evaluation and genetic trends. Pages 10–41 in Proc. Animal Breeding and Genetic Symposium in Honor of Dr Jay L. Lush. American Society of Animal Science, Champaign, IL.
Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92:4656–4663.
Megens, H. J., R. Crooijmans, J. Bastiaansen, H. Kerstens, A. Coster, R. Jalving, A. Vereijken, P. Silva, W. Muir, H. Cheng, O. Hanotte, and M. Groenen. 2009. Comparison of linkage disequilibrium and haplotype diversity on macro- and microchromosomes in chickens. BMC Genet. 10:86.
Meuwissen, T. H., B. Hayes, and M. Goddard. 2001. Prediction of total genetic value using genome wide dense marker maps. Genetics 157:1819–1829.
Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. H. Lee. 2002. BLUPF90 and related programs (BGF90). In Proc. 7th World Congress on Genetics Applied to Livestock Production. CD-ROM communication no. 28-07.
Protais, J., P. Colin, C. Beaumont, J. Guillot, F. Lantier, P. Pardon, and G. Bennejean. 1996. Line differences in resistance to *Salmonella enteritidis* PT4 infection. Br. Poult. Sci. 37:329–339.
Sargolzaei, M., F. S. Schenkel, G. B. Jansen, and L. R. Schaeffer. 2008. Extent of linkage disequilibrium in Holstein cattle in North America. J. Dairy Sci. 91:2106–2117.
Sorensen, D., and D. Gianola. 2002. Likelihood, Bayesian and MCMC Methods in Quantitative Genetics. Springer-Verlag, New York, NY.
Tilquin, P., P. Barrow, J. Marly, F. Pitel, F. Plisson-Petit, P. Velge, A. Vignal, P. Baret, N. Bumstead, and C. Beaumont. 2005. A genome scan for quantitative trait loci affecting the *Salmonella* carrier-state in the chicken. Genet. Sel. Evol. 37:539–561.
VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423.
VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.
Wiggans, G. R., P. VanRaden, L. Bacheller, M. Tooker, J. Hutchinson, T. Cooper, and T. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. J. Dairy Sci. 93:2287–2292.

Downloaded from http://ps.oxfordjournals.org/ at New York University on April 17, 2015