A single step genetic evaluation including causal candidate SNPs for resistance to paratuberculosis in Holstein cattle

M P Sanchez, T Tribout, S Fritz, Raphaël Guatteo, C Fourichon, L Schibler, A Delafosse, D Boichard

A single step genetic evaluation including causal candidate SNPs for resistance to paratuberculosis in Holstein cattle

M.P. Sanchez[1]*, T. Tribout[1], S. Fritz[1,2], R. Guatteo[3], C. Fourichon[3], L. Schibler[2], A. Delafosse[4], D. Boichard[1]

[1]Université Paris Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France; [2]Allice, 75595 Paris, France; [3]INRAE, Oniris, BIOEPAR, 44300 Nantes, France; [4]GDS Orne, 61000 Alençon, France; *marie-pierre.sanchez@inrae.fr

Abstract

Bovine paratuberculosis, or Johne's disease (JD), is a contagious, incurable and very difficult to control disease due to *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Using a single step GBLUP, including candidate SNPs for resistance to JD, applied to a pedigree of 161,253 Holstein animals (56,766 with serological phenotypes and 12,431 with genotypes), we estimated a moderate heritability value (0.14) of this trait and detected genomic regions associated to resistance to JD on chromosomes 23, 21, 12, 3, 7, 27, 5, 10, and 20. We also estimated a slightly favorable genetic trend of resistance to JD over the last two decades and relatively reliable genomic predictions (CD = 0.55) in a validation population allowing the identification of cows at high risk of infection. Genomic predictions should therefore rapidly become an effective tool for controlling paratuberculosis in French Holstein cow farms.

Introduction

Bovine paratuberculosis, also referred as Johne's disease (JD), is a contagious and incurable disease, caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). After being contaminated *in utero* or more frequently via the intake of MAP-infected feces as a young calf, animals undergo a long latency phase that can last several years. Subclinical symptoms of JD are weight loss and reduced milk production together with an inconsistent humoral immune response and fecal shedding. Subclinical cases can therefore contaminate their environment. Clinical cases finally develop chronic diarrhea and severe emaciation, ending in death. JD, which has adverse effects on animal welfare and serious economic consequences, has no effective treatment and access to vaccines is restricted, it is therefore very difficult to control. A number of studies have shown the role of host genetics in resistance to MAP (Brito et al., 2018) but their results are not fully consistent mainly due to the various definitions of the MAP-resistant phenotype.

In genome-wide association studies (GWAS) conducted on imputed whole genome sequences of 1644 Holstein cows with an accurately defined status for MAP, we previously identified various candidate SNPs explaining around 20% of the phenotypic variance of resistance to MAP infection (Sanchez et al., 2020; 2021). Using these data together with additional data collected in infected herds and a ssGBLUP model including the best candidate SNPs, we investigate here the genetic determinism of resistance to JD and estimate genetic trends, reliability and risk factors associated to genomic predictions.

Materials & Methods

Phenotypes and genotypes. Data of 247,375 Holstein cows were collected in northwestern French herds enrolled in JD control plans. These data included MAP statuses of 4100 cows obtained from serum ELISA and PCR on faeces in the PARADIGM project (see details in

Sanchez et al., 2020) and MAP statuses of 243,275 cows deduced from serological data (Idexx or Idvet ELISA test, Montpellier, France) routinely recorded since 2015. These cows were non-infected (NI; n=228,337) or infected (I; n=19,038) by MAP. We applied different filters to these data. First, to exclude cows still potentially in the latency period, I and NI cows younger than 24 and 36 months old, respectively, were discarded. Then, to maximize the probability of exposure to MAP, we kept only cows in herds with at least one I cow and one NI cow born in the same year. These filters resulted in a drastically reduced dataset, which ultimately contained 56,766 Holstein cows (42,829 NI and 13,937 I) from 3114 herds. The pedigree was traced over four generations and contained 161,253 animals, including 12,431 genotyped individuals, among which 4031 had phenotypes (2787 NI and 1244 I). The 12,431 animals were genotyped with different SNP Beadchip versions, including the EuroGMD Beadchip that contains an add-on for selected SNPs predictive for resistance to JD. In total, 53,469 SNPs were imputed using FImpute (Sargolzaei et al., 2014).

Model and methods. To estimate ssEBVs, resistance to JD was modelled as follows:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Za} + \mathbf{e} \quad (1)$$

where $\mathbf{y}$ is the vector of phenotypes, *i.e.* 1 for NI cows and 0 for I cows, $\mathbf{a} \sim \mathbf{N(0, H\sigma_a^2)}$ is the vector of random additive genetic effects and $\mathbf{e} \sim \mathbf{N(0, I\sigma_e^2)}$ is the vector of random residual effects. The $\boldsymbol{\beta}$ vector included the fixed effects of herd x birth year, birth month, and ELISA test to account for variability of exposure to MAP in the environment between farms and in time, and for differences in test sensitivities and specificities. $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices, $\mathbf{H}$ is the relationship matrix among individuals and $\mathbf{I}$ the identity matrix, σ_a^2 and σ_e^2 are the additive genetic and residual variances, respectively. Model (1) was applied to the Hybrid Single Step method, proposed by Fernando et al. (2016) and implemented in the HSSGBLUP software (Tribout et al., 2020). This method directly provides SNP effect estimates and ssEBVs of all animals in the pedigree, genotyped or not. ssEBVs were then expressed in genetic standard deviation units by dividing predictions on the raw scale by 0.15.

Variance components were estimated with WOMBAT software (Meyer, 2007) using a similar model but only with pedigree information to build the relationship matrix $\mathbf{A}$. Heritability h^2 for resistance to JD was calculated as follows: $h^2=\sigma_a^2/(\sigma_a^2+\sigma_e^2)$.

Genetic trends of resistance to JD in females and males were assessed separately by the annual average ssEBVs of cows with phenotypes born between 2008 and 2018 and of bulls born between 2000 and 2014 with at least 10 daughters with MAP statuses.

Reliability of ssEBVs was estimated in a validation (VAL) population of 907 cows with phenotypes (715 NI and 192 I) and genotypes born between 2015 and 2018. For this study, model (1) was run on a truncated dataset constituted by phenotypes of the 41,774 cows (31,115 NI and 10,659 I) born before 2015, among which 3118 had genotypes (2069 NI and 1049 I). Reliability was estimated by $CD=r^2/h^2$, with r the correlation between phenotypes adjusted for non-genetic effects estimated with model (1) on the complete dataset and genomic predictions (GP) of VAL animals predicted using the truncated dataset. Risk factor was also assessed in the VAL set. GP were grouped in 0.5 point classes and the relative risk factor in class i was estimated by the ratio of the proportion of infected animals in class i to that in the central 0 class.

Finally, considering SNP effects ($\hat{\mathbf{s}}$) estimated in HSSGBLUP, the % of genetic variance explained by the i^{th} genomic region of 25 adjacent SNPs was calculated as $\text{var}(\hat{\mathbf{a}}_i)/\sigma_a^2$ where $\hat{\mathbf{a}}_i = \mathbf{M}_i\hat{\mathbf{s}}_i$ is the vector of genomic values for the i^{th} region, $\mathbf{M}_i$ the matrix of SNP content in region i, and $\hat{\mathbf{s}}_i$ the effects of the 25 SNP of the region.

Results

Estimates of genetic and residual variances were 0.019 and 0.113, respectively, corresponding to a moderate h^2 estimate of 0.143±0.01 for resistance to JD.

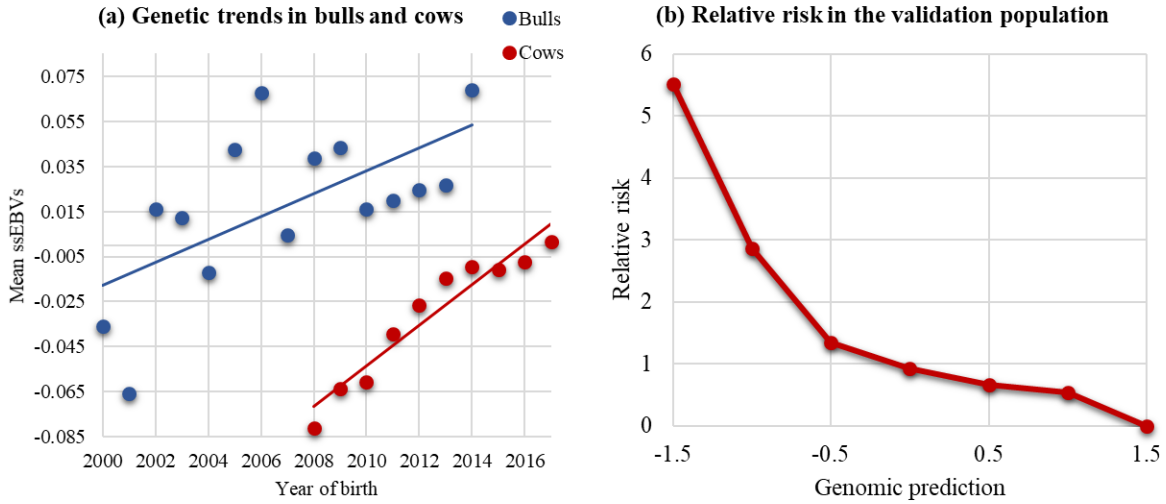

Figure 1. Evaluation results: (a) genetic trend of resistance to JD and (b) relative risk of being infected by MAP

Estimated genetic trends of resistance to JD in bull and cow populations (31 to 117 bulls and 2356 to 8674 cows depending on the year of birth) were favourable (+0.6 σ_a and +0.8 σ_a, respectively over the period considered), the younger animals being genetically more resistant than the older ones (Figure 1a). Correlation between (GP) and phenotypes adjusted for fixed effects in VAL was equal to 0.28 (SE=0.03); the corresponding reliability was therefore ~ 0.55. To illustrate the effect of genetics on the phenotype, we used the VAL population to estimate the relative risk for cows to be infected based on their GP (Figure 1b). A strong effect of genetics was observed in this validation dataset. Assuming the relative risk equal to 1 for GP=0, this risk was half as big and three times higher for cows with GP=1 and -1, respectively. Although numbers are limited, effects were even much stronger for cows with GP equal to +/-1.5.
The estimation of percentage of genetic variance explained by overlapping genomic regions of 25 SNPs revealed a high peak on chromosome 23 (max. 0.32%) and lower peaks explaining between 0.10% to 0.15% of the total genetic variance, on chromosomes 21, 12, 3, 7, 27, 5, 10, and 20 (Figure 2).

Discussion

This study, based on a single step SNP GBLUP including candidate SNPs for resistance to JD, provides new insights into the genetic determinism of this trait and gives very encouraging results for the implementation of a genomic evaluation in Holstein cows using MAP statuses deduced from serological data routinely recorded.
With respect to the genetic determinism of resistance to JD, we herein obtain results quite different from those found in our previous study conducted on imputed whole-genome sequences of 1644 Holstein cows with accurate extreme phenotypes (Sanchez et al., 2020; 2021). In the present study, we first estimated a moderate h^2 value (0.14) much lower than the one we reported before (0.57) but which agrees with h^2 values reported in various studies (0.03-0.27; Brito et al., 2018). The multi-SNP GWAS approach taking into account a total of 161,253 animals in the pedigree –56,766 with phenotypes and 12,431 with genotypes– confirms effects

of genomic regions located on chromosomes 23, 12, 3, and 5 and revealed new regions affecting resistance to JD on chromosomes 21, 7, 27, 10, and 20. However, the major effect detected on chromosome 13 in our previous single-SNP GWAS is not confirmed here.
Trends of genomic predictions, estimated over the last two decades, show a favourable genetic evolution of resistance to JD. In a validation population, we highlight relatively reliable genomic predictions for resistance to JD (CD = 0.55) estimated from a reduced reference population and we show that genomic predictions can already be of great help in identifying cows at high risk of infection. These breeding values can be used at two levels: (1) in infected herds, this information will allow orientating early culling of highly susceptible animals and planning matings with resistant bulls; (2) at the breeding scheme level, these breeding values can be integrated into the breeding goal and, more importantly, used to produce resistant artificial insemination bulls for matings in infected herds. We can anticipate that as the number of genotyped and phenotyped cows increases continuously, genomic predictions will rapidly become an effective tool for controlling paratuberculosis in French Holstein farms. The same strategy is currently being developed in the Normande breed and could be extended to other populations in the future.

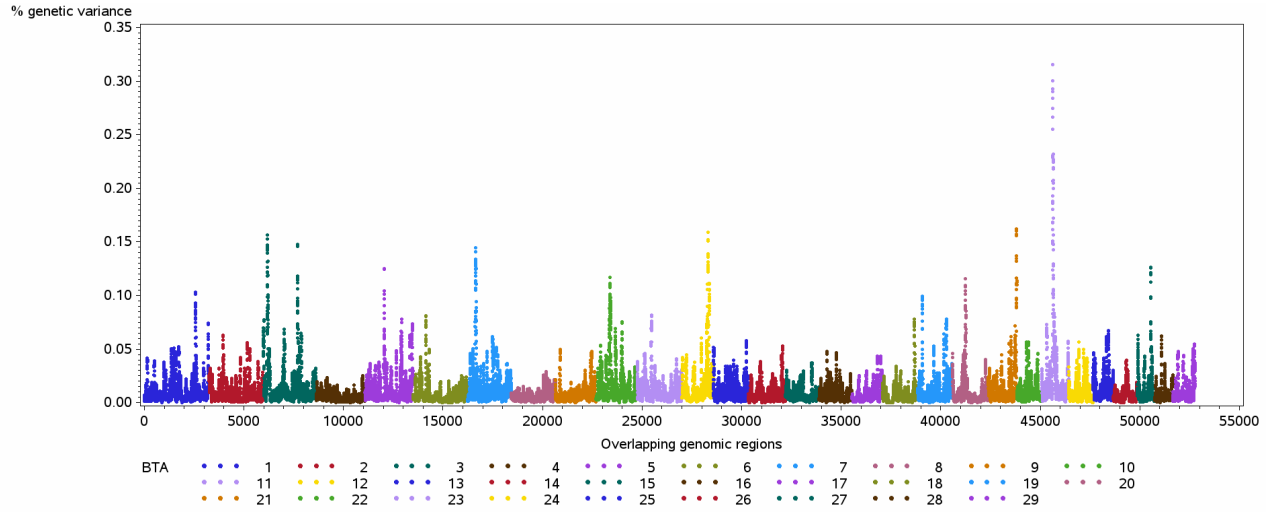

Figure 2. % of genetic variance explained by overlapping regions of 25 adjacent SNPs.

References

Brito L., Mallikarjunappa S., Sargolzaei M., Koeck A., Chesnais J., *et al.* (2018) J. Dairy Sci. 101(11):10062-10075. https://doi.org/10.3168/jds.2017-14250

Fernando R., Cheng H., Golden B., and Garrick D. (2016) Genet. Sel. Evol. 48:96. https://doi.org/10.1186/s12711-016-0273-2

Meyer K. (2007) J. Zhejiang Univ. Sci. B 8:815-821. https://doi.org/10.1631/jzus.2007.B0815

Sanchez M.P., Guatteo R., Davergne A., Saout J., Grohs C., *et al.* (2020) Genet. Sel. Evol. 52:14. https://doi.org/0.1186/s12711-020-00535-9

Sanchez M. P., Guatteo R., Delafosse A., Davergne A., Grohs C., *et al.* (2021) Proc. of the 72th EAAP, Davos, Switzerland.

Sargolzaei M., Chesnais J., and Schenkel F. (2014) BMC Genomics 15(1):478. https://doi.org/10.1186/1471-2164-15-478

Tribout T., Ducrocq V., and Boichard D. (2020) Proc. of the 6th ICQG, Brisbane, Australia.