Effect of herd prevalence on heritability estimates of antibody response to Mycobacterium avium subspecies paratuberculosis

K. J. E. van Hulzen,*1 M. Nielen,* A. P. Koets,† G. de Jong,‡ J. A. M. van Arendonk,§ and H. C. M. Heuven#
*Department of Farm Animal Health, Utrecht University, PO Box 80.151, 3508 TD Utrecht, the Netherlands
†Department of Infectious Diseases and Immunology, Utrecht University, PO Box 80.165, 3508 TD Utrecht, the Netherlands
‡CRV Animal Evaluation Unit, PO Box 454, 6800 AL Arnhem, the Netherlands
§Animal Breeding and Genomics Centre, Wageningen University, PO Box 338, 6700 AH Wageningen, the Netherlands
#Clinical Sciences of Companion Animals, Utrecht University, PO Box 80.163, 3508 TD Utrecht, the Netherlands

ABSTRACT

Worldwide, classical control strategies based on hygiene and culling of infected animals have been implemented to eradicate Johne’s disease. Breeding for disease resistance may be a useful additional tool to control the disease. The aim of this study was to estimate genetic parameters for the presence of a Mycobacterium avium ssp. paratuberculosis specific antibody response in milk of Dutch Holstein-Friesian cows using subsets of data based on within-herd test prevalence. The analyzed data set consisted of milk samples of 684,364 animals from 12,077 herds collected during the routine milk production scheme. Milk samples were tested for antibodies specific for Johne’s disease by an ELISA test. Heritability estimates were calculated for 4 different subsets of data to determine the sensitivity of heritability for within-herd test prevalence. Results expressed as percentage of the sample to positive ratio were analyzed with a sire-maternal grandsire model with fixed effects for parity, year of birth, lactation stage, and herd; a covariate for milk yield at test day; and random effects for sire, maternal grandsire, and error. The estimated heritability ranged from 0.031 for the complete data set to 0.097 for herds with a test prevalence of at least 10%. Cross-validation was applied to determine which of the subsets of data produced the most accurate estimated breeding values. Results showed that for genetic selection to contribute to disease control, breeding values were estimated most accurately from herds with at least 2 animals that tested positive. In this subset the heritability was 0.041.

Key words: Johne’s disease, heritability, ELISA, dairy cow

INTRODUCTION

Johne’s disease, also known as paratuberculosis, is characterized by granulomatous lesions in the distal part of the ileum. Ileal lesions elicit a deterioration of nutrient uptake; therefore, animals suffering from Johne’s disease show weight loss, diarrhea, and reduced milk production, and eventually they die. The causative agent of Johne’s disease is Mycobacterium avium ssp. paratuberculosis (MAP). Most animals infected with MAP are not able to clear the infection and as a result, Johne’s disease has a substantial economical impact on farms where the disease is endemic (Kreeger, 1991). Classical control strategies on the farm are based on management measures to improve hygiene and culling of infected animals, but eradication of MAP has been shown to be difficult (Beyerbach et al., 2001). Additional approaches to address Johne’s disease are needed. One approach is genetic selection for animals resistant to Johne’s disease.

Earlier research showed heritability estimates of susceptibility to Johne’s disease of 0.060 to 0.159 (Koets et al., 2000; Mortensen et al., 2004; Gonda et al., 2006; Hinger et al., 2008; Attalla et al., 2010). Large variation in estimated heritability in these studies could be due to (1) differences in incidence of Johne’s disease in the research populations: 0.077 (Gonda et al., 2006) versus 0.310 (Koets et al., 2000) before the start of a vaccination trial; (2) differences in sample sizes: 4,524 animals (Hinger et al., 2008) versus 21,514 animals (Attalla et al., 2010); (3) differences in diagnostic methods: determining infection status based on clinical inspection after slaughter (Koets et al., 2000) versus determining infection status based on ELISA to detect MAP-specific antibodies in milk or blood (Mortensen et al., 2004; Gonda et al., 2006; Hinger et al., 2008). Information from population-wide screening is lacking, and knowledge of the effect of within-herd exposure on estimates of heritability is limited.
The objective of this study was to estimate genetic parameters for the presence of a MAP-specific antibody response in milk based on population-wide screening of the Holstein-Friesian cows in the Netherlands. Further, the effect of within-herd exposure on estimate of heritability will be quantified using subsets of data based on within-herd test prevalence. Finally, opportunities to reduce test prevalence of Johne’s disease by selective breeding will be quantified.

MATERIALS AND METHODS

Data

Milk samples were collected from lactating cows during the routine milk production scheme. Collection, transportation, and storage of milk samples were accomplished as described by van Hulzen et al. (2009). From January until December 2008, milk samples could be sent to the Dutch Animal Health Service (Deventer, the Netherlands), upon the decision of the farmer, to test for antibodies specific for Johne’s disease using an ELISA. Testing individual cows with the ELISA was financially supported to encourage Dutch dairy farmers to participate in the program. Pedigrees and milk production records of the animals were provided by the Dutch Cattle Improvement Organization (CRV, Arnhem, the Netherlands). The data set was edited to ensure that the requirements of at least 20 animals per herd and a minimum of 5 daughters plus granddaughters per sire were met. Only cows with at least 75% Holstein-Friesian genes were included. In 2008, 20,750 dairy farms were present in the Netherlands (CRV, 2009), of which 12,077 herds with 684,364 animals were included in the data (58%).

Antibody Detection

All samples were tested using a commercially available ELISA kit (ELISA Paratuberculosis Antibody screening, Institut Pourquier, Montpellier, France) according to the instructions of the manufacturer. Results were expressed as percentage of the sample to positive ratio \( S/P \), calculated as 100 \( \times \) \{[the optical density (OD) value of the sample – the OD value of the negative control]/[the OD value of the positive control – the OD value of the negative control]\} (van Weering et al., 2009).

Classification of Within-Herd Test Prevalence

To quantify the effect of within-herd exposure on genetic parameter estimation, estimates of variance components were calculated for 4 subsets of data based on levels of within-herd test prevalence: (1) the complete data set, (2) herds with at least 2 test-positive animals, (3) herds with a within-herd test prevalence of at least 5%, and (4) herds with a within-herd test prevalence of at least 10%.

To define the various subsets of within-herd test prevalence, each individual animal needed to be classified as positive or negative. Cut-off value used for a positive test result was 25\% \( S/P \), which was lower than the cut-off value recommended by the manufacturer (40\% \( S/P \)). Because of our lowered cut-off value for the individual animal, we increased the sensitivity but also increased the likelihood of a false-positive result. Therefore, only herds with at least 2 test-positive animals were considered to have MAP present on the farm (data set 2).

Statistical Analysis

Variance components for the genetic effect were estimated using a sire-maternal grandsire model in ASReml (Gilmour et al., 2006):

\[
\ln(Y_{ijklmno} + 50) = \mu + P_i + YB_j + LS_k + b_l MY_l + HERD_m + sire_n + mgs_o + e_{ijklmno}
\]

where \( Y_{ijklmno} \) is the ELISA test result of the individual animal. The ELISA test result used in the model is the log-transformed (percentage \( S/P + 50 \)) to approach the normal distribution; \( \mu \) is the general mean; \( P_i \) is the fixed effect of the \( i \)th parity (\( i = 1, 2, 3, 4, 5 \) or \( >5 \)); \( YB_j \) is the fixed effect of the \( j \)th year of birth (\( j = <1993, 1993 \) to 2007 as separate classes); \( LS_k \) is the fixed effect of the \( k \)th stage of lactation with 6 classes (wk 1 and 2, wk 3–12, wk 13–28, wk 29–44, wk 45–60, and \( >60 \) wk); \( MY_l \) is the \( l \)th milk yield on test day; \( b_l \) is the regression coefficient for \( MY_l \); \( HERD_m \) is the fixed effect of the \( m \)th herd; \( sire_n \) is the random effect of the \( n \)th sire; \( mgs_o \) is the random effect of the \( o \)th maternal grandsire; and \( e_{ijklmno} \) is the residual component. The following distributional assumptions were made for the random effects:

\[
s \sim N\left(0, \frac{1}{2} \sigma^2_s\right),
\]

\[
mgs \sim N\left(0, \frac{1}{4} \sigma^2_s\right),
\]

\[
e \sim N\left(0, \sigma^2_e\right),
\]

where \( s \) and \( mgs \) are the vectors included in the sire additive genetic effects, \( e \) is the vector of residual ef-

fects, \( A \) is the sire relationship matrix, \( \sigma_s^2 \) represents the sire variance, \( I \) is the identity matrix, and \( \sigma_e^2 \) represents the residual variance. Variance components were used to estimate the heritability \( (h^2) \), which was defined as follows:

\[
h^2 = 4\sigma_s^2 / \sigma_p^2,
\]

where \( \sigma_p^2 \) is the phenotypic variance:

\[
\sigma_p^2 = \sigma_s^2 + \frac{1}{4} \sigma_s^2 + \sigma_e^2.
\]

**Cross-Validation**

Heritability, as well as the number of progeny per sire, varied over the different subsets of data, and cross-validation was applied to determine which of the subsets of data produced the most accurate EBV. Accuracy was estimated by determining the correlation between predicted breeding values and observed phenotypes using the following steps: (a) for each data set, 20% of the phenotypes of each herd were set as missing, which resulted in a subset containing 80% of the data of each herd. This process was nonrandomly repeated 5 times. In this way, each observation was missing once. (b) Sire and maternal grandsire breeding values were predicted by the remaining 80% of data in the corresponding data set using the linear mixed model. (c) Breeding values for 20% of the cows that were set as missing were estimated as follows:

\[
EBV_{cow} = \frac{1}{2} EBV_{sire} + \frac{1}{4} EBV_{mgs}.
\]

This procedure was repeated 5 times to obtain an EBV for each animal in the data set. The final step (d) was to calculate the correlation coefficient between the predicted breeding values and the observed phenotypes of cows. Correlation coefficients were also expressed as percentage of maximum correlation, which was calculated as

\[
\sqrt{\frac{\sigma_s^2 + \frac{1}{4} \sigma_s^2}{\sigma_p^2}}.
\]

Additionally, correlation coefficients were calculated between predicted breeding values from 1 subset of data based on levels of within-herd test prevalence and observed phenotypes from the remaining 3 subsets of data based on levels of within-herd test prevalence to investigate the ability of each subset of data to predict observed phenotypes in the other subsets.

**Potential Genetic Gain**

To determine how much potential genetic progress could be made by single trait selection, the number of test-positive animals and the number of herds without test-positive animals in the complete data set were calculated. Subsequently, the estimated sire breeding values were used to eliminate the 10% of sires most susceptible to Johne’s disease. Progeny of those sires were removed from the data and the number of test-positive animals and the number of herds without test-positive animals were calculated in the reduced data set. The same procedure was repeated for the elimination of the 20% of sires most susceptible to Johne’s disease based on their EBV.

**RESULTS**

**Within-Herd Test Prevalence**

Prevalence of infection as measured by a positive ELISA test in milk according to our test interpretation demonstrated that of the 12,077 herds participating in the program, 6,438 herds had no test-positive animals and 1,712 herds contained 1 test-positive animal. In 2,153 herds, at least 2 test-positive animals were detected but the within-herd test prevalence remained <5%. In 1,232 herds, a within-herd test prevalence between 5 and 10% was found, whereas 542 herds had a within-herd test prevalence of at least 10%.

**Heritability**

Genetic parameters for susceptibility to Johne’s disease were estimated in 4 sets of data (Table 1). The estimated heritability ranged from 0.031 for the complete data set to 0.097 for herds with a minimum test prevalence of 10% (data set 4). Table 1 shows an increase in phenotypic variance concurrently with an increase in within-herd test prevalence. This can be explained by a relative increase in the number of high ELISA test results. Additionally, the heritability increased, which indicated a relatively larger increase in genetic variance compared with phenotypic variance.

**Cross-Validation**

The ability of the different subsets of data to predict observed phenotypes in the specific data set itself increased with increasing within-herd test prevalence (Table 2; diagonal values). Higher heritability seemed
more important than the number of progeny per sire to obtain accurate breeding values.

The ability of the different subsets of data based on levels of within-herd test prevalence to predict observed phenotypes in the other subsets of data are depicted in Table 2 (off-diagonal values). Breeding values predicted by data set 1 provided high correlation coefficients with observed phenotypes from data sets 1 and 2. Breeding values predicted by data set 2 provided high correlation coefficients with observed phenotypes from data sets 2, 3, and 4. Breeding values predicted by data set 3 provided high correlations with observed phenotypes from data sets 3 and 4. Breeding values predicted by data set 4 only provided a high correlation with observed phenotypes from data set 4.

**Potential Genetic Gain**

In Table 3, a decrease in the number of test-positive animals due to sire elimination can be observed. The complete data set contained 16,627 test-positive animals, whereas in the data set with 10% sire elimination, 11,438 test positive animals were included. After implementation of 20% sire elimination, 10,112 test-positive animals were left in the data set. In other words, this theoretical approach to determine how much genetic progress could be made by single trait selection showed, by eliminating 10% of the sires that were most susceptible to Johne’s disease based on their EBV were eliminated from the data, 7,248 herds out of 12,074 had a within-herd test prevalence of zero. Elimination of progeny of 20% of the most susceptible sires from the data resulted in 7,576 herds out of 12,073 having a within-herd test prevalence of zero.

**DISCUSSION**

This study estimated genetic parameters for the presence of a MAP-specific antibody response in milk of Dutch Holstein-Friesian cows. Besides the genetic capacity of the animal to resist infection to Johne’s disease, an important environmental factor for infection is the degree of exposure to MAP (Gonda et al., 2006). A high level of exposure increases the likelihood of becoming infected; therefore, a high level of exposure is expected to increase the within-herd test prevalence. In this study, subsets of data based on within-herd test prevalence were used to investigate the effect of exposure to MAP on genetic parameter estimation. The estimated heritability increased from 0.031 for the complete data set to 0.097 for herds with a minimum within-herd test prevalence of 10%. Selection on within-herd test prevalence decreased the number of test-positive animals of 8%, whereas the number of animals in the data set decreased by 6%.

Table 3 shows that 6,438 out of 12,077 herds had a within-herd test prevalence of zero. When progeny of 10% of the sires that were most susceptible to Johne’s disease were eliminated from the data, 7,248 herds out of 12,074 had a within-herd test prevalence of zero. Elimination of progeny of 20% of the most susceptible sires from the data resulted in 7,576 herds out of 12,073 having a within-herd test prevalence of zero.

**Table 1.** Sire variance \( (\sigma_s^2) \), residual variance \( (\sigma_e^2) \), and within-herd heritability \( (h^2) \) for the presence of a *Mycobacterium avium* ssp. *paratuberculosis*-specific antibody response in complete data set and subsets based on test prevalence

<table>
<thead>
<tr>
<th>Data set</th>
<th>Animals, n</th>
<th>Herds, n</th>
<th>Sires, n</th>
<th>( \sigma_s^2 )</th>
<th>( \sigma_e^2 )</th>
<th>( h^2 ) (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Complete</td>
<td>684,364</td>
<td>12,077</td>
<td>9,870</td>
<td>0.215 (^{-3})</td>
<td>0.273 (^{-1})</td>
<td>0.031 (0.002)</td>
</tr>
<tr>
<td>2: At least 2 test-positive animals in herd</td>
<td>265,290</td>
<td>3,927</td>
<td>7,021</td>
<td>0.606 (^{-1})</td>
<td>0.578 (^{-1})</td>
<td>0.041 (0.004)</td>
</tr>
<tr>
<td>3: Prevalence ≥5%</td>
<td>104,382</td>
<td>1,774</td>
<td>4,570</td>
<td>0.150 (^{-2})</td>
<td>0.984 (^{-1})</td>
<td>0.060 (0.006)</td>
</tr>
<tr>
<td>4: Prevalence ≥10%</td>
<td>28,916</td>
<td>542</td>
<td>1,851</td>
<td>0.383 (^{-2})</td>
<td>0.153</td>
<td>0.097 (0.014)</td>
</tr>
</tbody>
</table>

\[ 1 \sigma_p^2 \text{(phenotypic variance)} = \left[ \sigma_s^2 + \frac{1}{4} \sigma_e^2 \right] + \sigma_e^2; h^2 = 4\sigma_s^2 / \sigma_p^2. \]

**Table 2.** Correlations between breeding values predicted by complete data set and subsets based on within-herd test prevalence and observed phenotypes in the data sets (correlations expressed as percentage of maximum correlation in parentheses)

<table>
<thead>
<tr>
<th>Data set</th>
<th>Animals, n</th>
<th>( h^2 )</th>
<th>Observed phenotypes from data set</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Complete</td>
<td>684,364</td>
<td>0.031</td>
<td>0.072 (73)</td>
</tr>
<tr>
<td>2: At least 2 test-positive animals in herd</td>
<td>265,290</td>
<td>0.041</td>
<td>0.059 (60)</td>
</tr>
<tr>
<td>3: Prevalence ≥5%</td>
<td>104,382</td>
<td>0.060</td>
<td>0.048 (49)</td>
</tr>
<tr>
<td>4: Prevalence ≥10%</td>
<td>28,916</td>
<td>0.097</td>
<td>0.037 (37)</td>
</tr>
<tr>
<td>Maximum correlation</td>
<td>—</td>
<td>—</td>
<td>0.099</td>
</tr>
</tbody>
</table>

in-herd test prevalence may also induce sire selection. Specific sires may produce susceptible progeny and this will increase the herd prevalence, resulting in an under-estimation of the heritability. However, this was not the case in this study. In the Netherlands, unproven sires are randomly used for breeding. That this is the case becomes apparent in Table 3: when progeny of the 10 and 20% most susceptible sires were removed from the data, the number of herds present in the data declined only negligibly (from 12,077 to 12,074 to 12,073 herds, respectively).

To obtain reliable genetic parameter estimates, the use of an animal model is preferred over a sire model. A sire model does not account for the genetic merit of the dam. Within a sire model, it is assumed that all dams are of similar genetic merit and this can result in bias in the predicted breeding values if there is preferential mating (Mrode, 2005). Everett et al. (1979) and Schaeffer (1983) showed that including the maternal grandsire of the animal could reduce bias compared with the sire model. In this study, a sire-maternal grandsire model was implemented because of limits in computing resources. Besides, infection with Johne’s disease may take place early in life as well as intrauterine (Whittington and Windsor, 2009). Using an animal model in the case of intrauterine infection will probably give an overestimation of the heritability because intrauterine infection is included in the genetic effect, whereas the sire-maternal grandsire model accounts for the genetic merit of the dam but excludes the intrauterine infection effect.

In the case of Johne’s disease, different diagnostic methods can be applied to determine infection status of the animal. Sensitivity of an ELISA test is low, indicating that in contrast to determination of infection status at slaughter, an ELISA test to detect MAP-specific antibodies in milk or blood is not necessarily positive when the animal is infected. An ELISA test measures the humoral immune response, which is often induced when a late-stage disease type 2-like response, characterized by production of immunoglobulin G1 antibodies, predominates. This shift in predominant immune response is often associated with progression to clinical disease (Stabel, 2000; Coussens, 2004). Using an ELISA test in milk facilitates quick testing in many animals and herds but the sensitivity of the test makes it hard to detect all infected animals. Determination of infection status at slaughter or repeated fecal cultures is more sensitive but has prohibitive costs for sampling and testing. In this study, ELISA in milk enabled the analysis of a large sample size and facilitated the estimation of reliable genetic parameters for the presence of a MAP-specific antibody response.

The accuracy of breeding value estimation from the different subsets of data was assessed using cross-validation. Accuracy was estimated by determining the correlation between predicted breeding values and observed phenotypes. On 2 occasions the percentage of the maximum correlation exceeded 100%. This may be the result of rounding errors and underestimation or overestimation of sire breeding values. Because breeding values predicted based on data set 2 gave high correlations with observed phenotypes in the different subsets of data independent of within-herd test prevalence, to reduce the incidence of paratuberculosis by selection it is optimal to estimate breeding values based on herds with at least 2 test-positive animals, resulting in a heritability of 0.041. Including cows from herds in which exposure to MAP is absent will not add information to determine the genetic ability of sires. Including all herds where exposure to MAP is present (in this study, defined as a herd with at least 2 test-positive cows) allows the maximum number of daughters to be included in a sire’s proof. Using more daughters is preferred over using only herds with at least 5 or 10% test-positive cows, although the heritability in these subsets is higher.

This study showed that elimination of progeny of the 10% of sires most susceptible to Johne’s disease based on their EBV has a strong decreasing effect on the number of test-positive cows in the population and showed an increase of the number of herds without test-positive animals. Additionally, genetic-epidemiological models demonstrated that altering the host genotype

### Table 3. Number of test-positive animals and number of herds without test-positive animals 1) without the implementation of sire elimination, 2) with elimination of progeny of the 10% of sires most susceptible to Johne’s disease based on their EBV, and 3) with elimination of progeny of the 20% of sires most susceptible to Johne’s disease based on their EBV

<table>
<thead>
<tr>
<th>Variable</th>
<th>No sire elimination</th>
<th>10% sire elimination</th>
<th>20% sire elimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sires, n</td>
<td>9,870</td>
<td>8,878</td>
<td>7,885</td>
</tr>
<tr>
<td>Animals, n</td>
<td>684,364</td>
<td>539,896</td>
<td>498,666</td>
</tr>
<tr>
<td>Test-positive animals, n</td>
<td>16,627</td>
<td>11,438</td>
<td>10,112</td>
</tr>
<tr>
<td>Herds, n</td>
<td>12,077</td>
<td>12,074</td>
<td>12,073</td>
</tr>
<tr>
<td>Herds without test-positive animals, n</td>
<td>6,438</td>
<td>7,248</td>
<td>7,576</td>
</tr>
</tbody>
</table>
for disease resistance will alter the transmission of disease through the population, and hence the challenge faced by each animal (MacKenzie and Bishop, 1999). Therefore, effects of selection might be larger than predicted by the quantitative genetic theory (Bishop and Stear, 1997). Results show potential for genetic improvement but additional cost-benefit analysis should precede selection, and should also consider the consequences of selection for resistance on other traits. Mortensen et al. (2004) showed a nonsignificant negative genetic correlation between daily milk yield and ELISA OD value for antibodies specific against Johne’s disease in milk. Attalla et al. (2010) showed statistically significant negative correlations of productive life and net merit with sire breeding values for OD value, which suggests that selection for productive life and net merit will result in cows genetically more resistant to disease. Implementation of EBV of sires for susceptibility to Johne’s disease in the breeding program can be done in several ways: (1) by breeding organizations for their approval decisions; (2) by farmers for their mating decisions; (3) in combination with other disease breeding values to create a general resistance parameter; and (4) in combination with other sustainability parameters to create a general sustainability parameter.

CONCLUSIONS

Differences were observed between cows in their genetic ability to produce antibodies against Johne’s disease. This study shows that it is optimal to estimate breeding values based on herds with at least 2 test-positive animals, which resulted in a heritability of 0.041. Although heritability is low, breeding for disease resistance to contribute to a more effective control of Johne’s disease seems feasible.

ACKNOWLEDGMENTS

This project was supported by the “Steering Committee Paratuberculosis and Salmonellosis,” which includes members of the Dutch Dairy Board (PZ, Zoetermeer, the Netherlands), Dutch Dairy Association (NZO, Zoetermeer, the Netherlands), and the Dutch organization for Agriculture and Horticulture (LTO Nederland, Den Haag, the Netherlands). The authors thank the farmers for participating in the program, the “Steering Committee Paratuberculosis and Salmonellosis” for putting the data at our disposal, CRV (Arnhem, the Netherlands) for supplying pedigrees, milk production records, and an informative work environment, and P. Bijma (ABGC, Wageningen, the Netherlands) for his valuable contribution to the cross-validation analysis.

REFERENCES