doi: 10.17221/8732-CJAS

Genetic relationship between management units of Czech dam pig breeds based on various types of data and pedigree information

E. Krupa[1], E. Žáková[1], Z. Krupová[1], R. Kasarda[2], A. Svitáková[1]

[1]Institute of Animal Science, Prague-Uhříněves, Czech Republic
[2]Slovak University of Agriculture, Nitra, Slovak Republic

ABSTRACT: Knowledge of genetic relationship is an important control mechanism for animal performance-testing schemes. Genetic relationship between and within pig herds was calculated for two dam breeds, Czech Large White (CLW) and Czech Landrace (CLA). The impacts of different field data types (production and reproduction) and various numbers of generations within the pedigrees on genetic relationship were studied. The degree of genetic relationship between analyzed herds was generally low. It ranged from 1.01% (for CLW based on reproduction data and considering three generations of ancestors within the pedigree) to 2.57% (for CLA based on production data with seven generations of ancestors in the pedigree). In contrast, relationship within herds was high and ranged from 16.62% to 44.69% (when three and seven generations within the pedigree were taken into account, respectively), both for production data of the CLA breed. When considering the type of data, an impact on the observed genetic relationship between and within herds was found. Slightly higher genetic relationship between herds was determined in both breeds when using production data (1.64%) compared to reproduction data (1.40%). In contrast, a negligible influence between herds on genetic relationship was found from the number of ancestors' generations included into the calculations. That was especially so after five or six generations. Our results show that the relationship between herds is population specific and, consequently, must be analyzed on a case-by-case basis. Knowledge of genetic relationship between and within herds should be taken into account in regard to the complexity of genetic evaluation.

Keywords: production data; reproduction data; number of generations

INTRODUCTION

The best linear unbiased prediction (BLUP) animal model has been applied in the Czech Republic for genetical evaluating pig populations since 1998. The model enables taking into account differences among individual management units and thus comparing estimated breeding values across these units. The assumptions of the model are fulfilled only if all contrasts between management units are estimable, i.e. only if a connection (genetic relationship) exists between them (Kennedy and Trus 1993). Therefore, genetic relationship is examined first so as to quantify its impact on the actual genetic parameters of the investigated populations (Colleau 2002). Moreover, knowledge of genetic relationship is an important and useful control mechanism in animal performance-testing schemes. It enables making appropriate interventions into the performance-testing scheme and revising them. In other words, all interventions into performance testing (e.g. changes in data structure, inclusion of new traits and herds) should be evaluated to define their potential influence on the accuracy of estimated breeding values. Several methods have been developed to evaluate genetic relationship (e.g. expressed as degree of connectedness) between pairs of management units using statistical methods of various types (Foulley et al. 1992; Kennedy and Trus 1993; Lewis et al. 1999;

Supported by the Ministry of Agriculture of the Czech Republic (Project No. QJ1310109).

doi: 10.17221/8732-CJAS

Mathur et al. 2002; Fouilloux et al. 2008). All the aforementioned authors preferred methods involving the solution of systems of linear equations. Salaris et al. (2009) points to the fact that criteria for these methods are often difficult to achieve when large data sets with a great number of management units are applied. Zhang et al. (2004) compared five methods for calculating genetic connectedness. They found conflicting results by different authors, computing difficulty, and inconsistent correlations between some methods. They recommended using the average genetic covariance based on a numerator relationship matrix as the method for measuring genetic connectedness. To the best of our knowledge, assessment of genetic relationship between and within herds without the necessity of solving linear models has so far been investigated in only a few papers (Fernando et al. 1983; Banos and Cady 1988; Zhang et al. 2004; Salaris et al. 2009). With insufficient genetic connectedness of a data set, there can be a problem with biased values for genetic parameters and breeding values. The average animal without appropriate connectedness (range by BLUP animal model and expressed by breeding value) is average only within its herd group. There is no connection between its herd and others. This means we cannot make a decision regarding its range in the population as a whole. This animal's breeding value can be overestimated or underestimated based on the population average. The aim of the present study, therefore, was to calculate genetic relationship between and within herds while using production and reproduction data. The effect of the number of generations taken into account in calculating genetic relationship was analyzed, too.

MATERIAL AND METHODS

Pedigree information as well as production and reproduction field test data for the Czech Large White (CLW) and Czech Landrace (CLA) pig breeds included into the Czech national pig breeding programme were used in the analyses. Genetic relationship was calculated for two types of data. Production data were represented by the trait average daily gain in field test and reproduction data sets comprised the number of piglets born alive at first parity. All the data were collected within the performance test for pig breeds in the Czech Republic and were provided by the Czech Pig Breeders Association. Individual data sets were defined to comply fully with the traits included in genetic evaluation of dam breeds in the Czech Republic (Wolf et al. 2005; Wolf and Smital 2009; Krupa and Wolf 2013). Preliminary analyses had shown that more than 91% of herds consisted of only a single breed and that such herds represent more than 87% of all tested animals. Therefore, genetic relationship was analyzed separately for individual breeds. More detailed information about the evaluated pig breeds containing quality of pedigree, average inbreeding and co-ancestry coefficients, proportion of inbred animals, and generation interval can be found in the study by Krupa et al. (2015).

For the data to be used for analysis, the following conditions had to be fulfilled: only animals from the reference population (born during 2000–2013) were taken into account; for all animals with performance data the herd and birth day had to be known. Two base data sets were prepared for each breed, one consisting of production field data and the other of reproduction field data. Next, each animal within each data set thus formed was associated with its appropriate pedigree. The influence of the amount of pedigree data on genetic relationship between and within herds was analyzed according to the addition of various numbers of ancestor generations. The minimum number of generations traced back was set to three. This was increased by one generation at a time up to the maximum seven generations of ancestors tracing back to approximately 1990. Altogether, ten data sets were created for each breed according to the procedure described above:

- Production field data for animals in reference populations and for each breed with pedigree information from the third to the seventh generations (ten data sets), and
- Reproduction field data for animals in reference populations and for each breed with pedigree information from the third to the seventh generations (ten data sets).

Basic characteristics of the reference population data sets are shown in Table 1. The total number of herds analyzed was 81 and 33 for the CLW and CLA breeds, respectively.

Genetic relationship between individuals of interest and the average relationship between and within herds were computed using the indirect method described by Colleau (2002). The methodology can be described as follows: let vector p_1 denote the positions filled by the first group in sparse vector x_1 and let vector

doi: 10.17221/8732-CJAS

Table 1. Basic characteristics of the reference population[a] of Czech Large White (CLW) and Czech Landrace (CLA) breed

Characteristic	Production data		Reproduction data	
	CLW	CLA	CLW	CLA
Number of animals with data	58 746	21 098	7 448	2 146
Total number of animals[b]	61 858	22 296	10 798	3 462
Number of herds	78	33	81	28
Proportions of inbred animals (%)	77	80	50	48
Number of founders	1 488	652	1 201	507
Proportion of animals with both parents known (%)	97	97	89	85

[a]values applied when the maximum number of ancestor's generations was considered in calculating genetic relationship
[b]number of animals with data and pedigree

p_2 denote the positions filled by the second group in sparse vector x_2. Then, positions p_1 of vector $m_2^{-1} Ax_2$ correspond to the vector of the average relationship between members of group 1 versus the entire group 2. In the same run, positions p_2 correspond to the vector of the average relationship between members of group 2 versus the entire group 2. Finally, after a second run where x_1 and x_2 are permuted, complete statistics between and within groups are obtained (Colleau 2002). The average genetic relationship (AGR) for herd *i* was assumed as the AGR between herd *i* and all other herds and also as the AGR between herd *i* and all related herds providing the calculation for each data set. Genetic relationship within the herd was also calculated. The degree of genetic relationship between and within herds was reported as a relative ratio in percentage terms.

The initial analyses of the raw field and pedigree data, together with evaluation of the results, was carried out in SAS software (Statistical Analysis System, Version 9.2, 2008) using the SAS/STAT module and SAS macro language. CFC Version 1.0 (Sargolzaei et al. 2006) was used for computing genetic relationships between and within all data sets.

RESULTS AND DISCUSSION

The initial analysis revealed the existence of a decreasing year-to-year trend in average herd size for both breeds. Such trend was not observed in the average number of herds where only one boar had offspring (Figure 1). The average number of herds where one sow has been kept was close to 1 (not presented in the results). Conversely, a given boar was used in 1.0–3.9 herds for CLW and CLA,

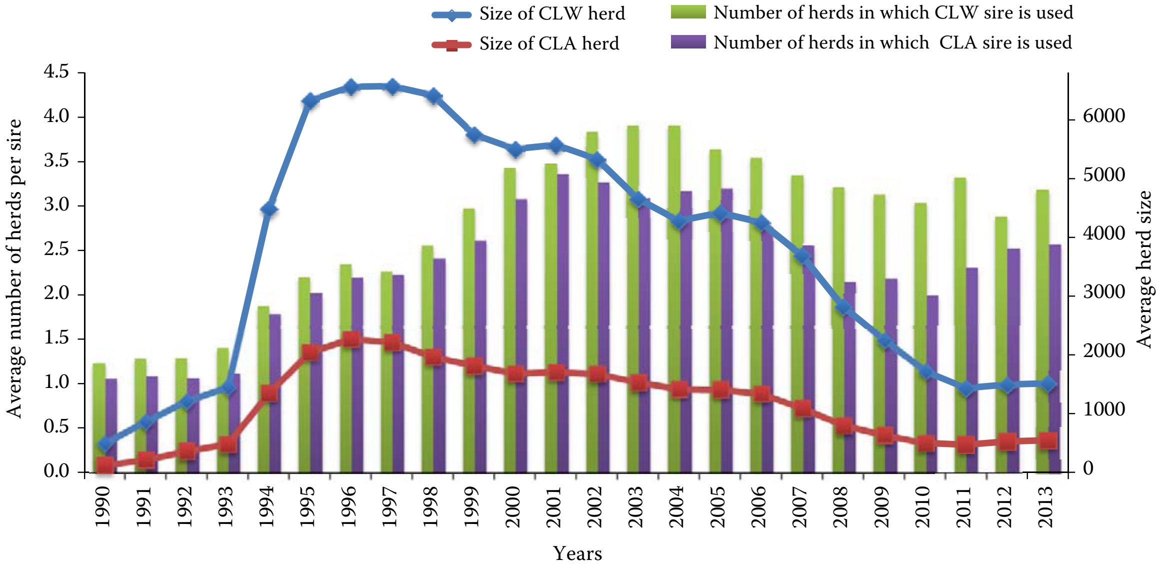

Figure 1. Average herd size (lines) and average number of herds where a given boar was used (bars) for Czech Large White (CLW) and Czech Landrace (CLA) breeds during 1990–2013

doi: 10.17221/8732-CJAS

Table 2. Characteristics of genetic relationship between herds for Czech Large White (CLW) breed

	Number of generations				
	3	4	5	6	7
Production data					
Total number of herds	75	77	78	78	78
Average number of herds with which a given herd is related	60.4	65.9	67.0	69.5	71.5
50th percentile (median) for number of herds related with other herds	67	72	73	74	75
Percentage of herds related with fewer than 50% of other herds	6.7	3.9	3.8	2.6	0.0
Percentage of herds related with 51–89% of other herds	46.7	31.2	32.1	28.2	25.6
Percentage of herds related with more than 90% of other herds	44.0	62.3	61.5	66.7	71.8
Reproduction data					
Total number of herds	69	74	77	79	81
Average number of herds with which a given herd is related	50.9	56.7	62.0	62.3	68.0
50th percentile (median) for number of herds related with other herds	58	65	69	71	75
Percentage of herds related with fewer than 50% of other herds	13.0	10.8	6.5	5.1	4.9
Percentage of herds related with 51–89% of other herds	78.3	51.4	44.2	45.5	38.3
Percentage of herds related with more than 90% of other herds	8.7	37.8	49.4	49.4	56.8

respectively. Summarizing all data from Figure 1, it can be concluded that even though the average herd size decreased rapidly over the past 15 years, the distribution of boars was relatively stable for both breeds and broader for CLW. In our opinion, variability of this type is probably specific for each livestock population.

The basic characteristics and distribution of the relationship among the data sets analyzed for CLW and CLA are summarized in Tables 2 and 3. On average for the CLW breed, a given herd was related with 60.4 or 50.9 other herds when the production and reproduction data sets with three generations of ancestors were applied. The relevant number of related herds increased from 50.9 to 68.0 when reproduction data was applied and from 60.4 to 71.5 herds for the production data set when larger numbers of generations were taken into account in calculating genetic relationship for CLW. For the CLA breed, the number of related herds increased with the number of ancestor generations when reproduction data were used (from 16.9 to 21.4)

Table 3. Characteristics of genetic relationship between herds for Czech Landrace (CLA) breed

	Number of generations				
	3	4	5	6	7
Production data					
Total number of herds	27	29	30	32	33
Average number of herds with which a given herd is related	24.0	23.7	22.2	23.4	23.5
50th percentile (median) for number of herds related with other herds	25	26	26	27	27
Percentage of herds related with fewer than 50% of other herds	0.0	3.4	12.5	9.4	12.1
Percentage of herds related with 51–89% of other herds	29.6	89.7	84.4	84.4	84.8
Percentage of herds related with more than 90% of other herds	66.7	3.4	0.0	3.1	0.0
Reproduction data					
Total number of herds	24	25	26	27	28
Average number of herds with which a given herd is related	16.9	19.6	20.9	20.8	21.4
50th percentile (median) for number of herds related with other herds	20	22	23	23	23
Percentage of herds related with fewer than 50% of other herds	16.7	8.0	3.8	7.4	10.7
Percentage of herds related with 51–89% of other herds	70.8	52.0	53.8	70.4	78.6
Percentage of herds related with more than 90% of other herds	12.3	40.0	42.3	22.2	10.7

doi: 10.17221/8732-CJAS

but decreased slightly (from 24.0 to 23.5 herds) when production data were applied.

Dodenhoff and Gotz (2010) analyzed the influence of different numbers of generations traced back on the relationship (connectedness) of production data between Saxon and Bavarian test stations for the Piétrain breed. Their results are not fully in agreement with ours and show a sharp increase in the relationship when two generations were used. In contrast, only a moderate strengthening in the relationship was found in their study when more than three generations were taken into account. Dissimilarities vis-à-vis our results were likely caused by a different structure of the data (station test vs field test) and a different sequence of generations traced back (range of 1 to 6 vs our 3 to 7).

Application of different numbers of ancestor generations had an impact not only on the number of related herds but also on the distribution of herds' genetic relationship for production and reproduction data. While the average number of herds which are related to a given herd was 68.0 (for CLW reproduction data when seven generations of ancestors were taken into account), the median was 75 herds. This means a higher number of larger related herds. There were 6.7% of herds related with fewer than 50% of other herds and 44.0% of herds were related with more than 90% of other herds when three generations of ancestors were taken into account, whereas these proportions changed to 0.0% and 71.8% when seven generations of ancestors were considered. Slightly lower numbers of related herds were obtained for reproduction data sets. Differences according to types of data could have been caused by there being different numbers of animals in the production and reproduction data sets.

Considering the number of related CLA herds, a difference was observed in the distributions of herd relationships according to type of data and number of ancestor generations when compared to the CLW breed. Most of the herds (ranging from 52.0 to 89.7% of all herds) were related to 51–90% of other herds. The only exception (66.7% of herds related with more than 90% of other herds) was found when production data and three generations of ancestors in the pedigree were used in the analysis of the CLA breed. The main differences in distributions of herd relationships between CLW and CLA dam breeds were probably caused by the differing dynamics of the breeds. CLW, as the most numerous dam breed, is a relatively closed population. Conversely CLA is a more open population and has been characterized by an intensive import of animals within the past few generations.

The average genetic relationship between herds for each breed and type of data set according to the number of ancestor generations is shown in Table 4. Although the proportions of related herds were relatively high in both breeds (as manifested in Tables 2 and 3), values of the average genetic relationship between herds were generally low for all investigated populations. Similar findings were reported by Banos and Cady (1988), who computed AGR between and within US and Canadian Holstein populations. In their study, Mathur et al. (2002) explained that whereas the average relationship gives an indication as to the accuracy of the estimated breeding value comparison between different

Table 4. Average percentage of genetic relationship between all herds and between only related herds for Czech Large White (CLW) and Czech Landrace (CLA) breeds

	Number of generations	CLW		CLA	
		production data	reproduction data	production data	reproduction data
All herds	3	0.94	0.75	1.58	1.14
	4	1.08	0.80	1.78	1.50
	5	1.17	0.91	1.77	1.66
	6	1.23	0.96	1.87	1.86
	7	1.25	0.97	1.88	1.88
Related herds	3	1.17	1.01	1.77	1.55
	4	1.26	1.04	2.18	1.83
	5	1.36	1.12	2.55	1.99
	6	1.38	1.15	2.55	2.32
	7	1.38	1.15	2.57	2.32

doi: 10.17221/8732-CJAS

Table 5. Average percentage of genetic relationship within herds for Czech Large White (CLW) and Czech Landrace (CLA) breeds

Number of generations	CLW		CLA	
	production data	reproduction data	production data	reproduction data
3	16.78	21.73	16.62	22.13
4	16.93	22.84	19.46	21.52
5	17.91	23.47	26.15	21.02
6	17.82	22.98	25.42	21.31
7	17.89	25.06	28.05	23.98

environments, a strong relationship with another herd enables effective selection from that herd.

When comparing the types of data in the two dam breeds, it can be stated that a slightly higher (by about one-fifth) genetic relationship between herds was achieved by using production data compared to reproduction data sets. This could have been caused by the fact that the total number of animals with production data was several times higher compared to those with reproduction data. It is also possible that the different pedigree completeness, which had been discovered in our previous study (Krupa et al. 2015), could also have played a role here. In contrast, a stronger impact from a different amount of pedigree information on inbreeding and on coancestry for the Arabian horse population in the Slovak Republic was found by Pjontek et al. (2009). Uimari and Tapio (2011) did not confirm this to be the case in their study of Finnish Landrace and Finnish Yorkshire pig breeds, however those authors investigated changes in effective population size while considering different numbers of pedigree generations (5 and 10) and different sizes of reference populations (3 and 5 years). Generally, their estimates were not sensitive to a reduced reference population and to decreased pedigree quality.

The AGR within herds for the analyzed breeds and the production and reproduction data sets is shown in Table 5. The AGR within analyzed herds (i.e. among animals of the given herd) ranged from 16.62 to 28.05%. This relationship was several times greater in comparison to the AGR between analyzed herds (Table 4). Banos and Cady (1988) also found strong differences when comparing AGR within and between cattle populations. In their study, AGR within groups (4.95×10^{-3}) was 100 times higher than was the relationship calculated between groups (4.6×10^{-5}).

Differences in genetic relationship within herds of the CLA dam breed were negligible (22.0 vs 23.1%). Kennedy and Trus (1993), cited by Roso et al. (2004), showed that an increase in the relationship between groups of interest has a positive impact on subsequent measurement of genetic connections between these groups. On the other hand, an increase in the relationship within groups resulted in a reduction of genetic relationship. A similar result was obtained also by Hanocq and Boichard (1999) for the French Holstein population. Zhang et al. (2004), meanwhile, reviewed five quantitative measures of genetic relationships (connectedness) from a national programme viewpoint and recommended the method based on AGR computation because of its easy implementation and lesser demand for computational capacity.

CONCLUSION

The proportions of related herds were relatively high in both breeds, but the degree of genetic relationship between the analyzed herds was generally low. The highest genetic relationship was found within the herds. The type of data used for the evaluation had an impact on the AGR between and within the analyzed herds, whereas the number of ancestor generations had a negligible influence on genetic relationship. Our results show that relationship between herds is population specific and consequently must be analyzed on a case-by-case basis. Our results also show that determining genetic connectedness is useful for successful and unbiased genetic evaluation. It is possible to evaluate the value of an animal based upon the average for the population as a whole. Knowledge of genetic relationship between and within herds should be taken into account in regard to the complexity of genetic evaluation. Measuring genetic relationship is very important for genetic improvement in a nucleus structure such as that of the Czech national genetic improvement program. In

doi: 10.17221/8732-CJAS

consequence of the high proportion of AGR within herds and the continuing trend toward reducing herd sizes, it will be necessary in future to start a mating programme focused on preventing increases in the relationship between animals within herds by, for example, better distributing artificial insemination doses of unrelated boars or monitoring the relationship to more than three generations back.

Acknowledgement. Thanks are due to the Czech Pig Breeders Association for making the performance data available and also to R. Prošková for technical assistance.

REFERENCES

Banos G., Cady R.A. (1988): Genetic relationship between the United States and Canadian Holstein bull populations. Journal of Dairy Science, 71, 1346–1354.

Colleau J.J. (2002): An indirect approach to the extensive calculation of relationship coefficients. Genetics Selection Evolution, 34, 409–421.

Dodenhoff J., Gotz K.U. (2010): Assessing connectedness among German swine herdbook populations. In: Proc. 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany (CD-ROM, Art. No. 813).

Fernando R.L., Gianola D., Grossman M. (1983): Identifying all connected subsets in two-way classification without interaction. Journal of Dairy Science, 66, 1399–1402.

Fouilloux M.N., Clement V., Laloe D. (2008): Measuring connectedness among herds in mixed linear models: from theory to practice in large-sized genetic evaluations. Genetics Selection Evolution, 40, 145–159.

Foulley J.L., Hanocq E., Boichard D. (1992): A criterion for measuring the degree of connectedness in linear models of genetic evaluation. Genetics Selection Evolution, 24, 315–330.

Hanocq E., Boichard D. (1999): Connectedness in the French Holstein cattle population. Genetics Selection Evolution, 31, 163–176.

Kennedy B.W., Trus D. (1993): Considerations on genetic connectedness between management units under an animal model. Journal of Animal Science, 71, 2341–2352.

Krupa E., Wolf J. (2013): Simultaneous estimation of genetic parameters for production and litter size traits in Czech Large White and Czech Landrace pigs. Czech Journal of Animal Science, 58, 429–436.

Krupa E., Zakova E., Krupova Z. (2015): Evaluation of inbreeding and genetic variability of five pig breeds in Czech Republic. Asian-Australasian Journal of Animal Sciences, 28, 25–36.

Lewis R.M., Crump R.E., Simm G., Thompson R. (1999): Assessing connectedness in across-flock genetic evaluations. In: Proc. British Society of Animal Science, Scarborough, UK, 121–122.

Mathur P.K., Sullivan B.P., Chesnais J.P. (2002): Measuring connectedness: concept and application to a large industry program. In: Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, 545–548.

Pjontek J., Kadlecik O., Kovacik E., Kasarda R. (2009): Estimation of inbreeding intensity from different quality of pedigree information. Acta Fytotechnica et Zootechnica, 4, 97–99.

Roso V.M., Schenkel F.S., Miller S.P. (2004): Degree of connectedness among groups of centrally tested beef bulls. Canadian Journal of Animal Science, 84, 37–47.

Salaris S., Fresi P., Casu S. (2009): Evaluation of connectedness in the registered population of the Sardinian dairy sheep breed. Italian Journal of Animal Science, 8, 141–143.

Sargolzaei M., Iwaisaki H., Colleau J.J. (2006): CFC: a tool for monitoring genetic diversity. In: Proc. 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, 27–28.

Uimari P., Tapio M. (2011): Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. Journal of Animal Science, 89, 609–614.

Wolf J., Smital J. (2009): Effects in genetic evaluation for semen traits in Czech Large White and Czech Landrace boars. Czech Journal of Animal Science, 58, 349–358.

Wolf J., Zakova E., Groeneveld E. (2005): Genetic parameters for a joint genetic evaluation of production and reproduction traits in pigs. Czech Journal of Animal Science, 50, 96–103.

Zhang H., Liu X.H., Wang C., Li J.Q., Chen Y.S. (2004): Study on the measures of connectedness between herds. Agricultural Sciences in China, 3, 143–148.

Received: 2015–06–26
Accepted after corrections: 2015–09–30

Corresponding Author

Ing. Emil Krupa, Ph.D., Institute of Animal Science, Přátelství 815, 104 00 Prague 10-Uhříněves, Czech Republic
Phone: +420 267 009 573, e-mail: krupa.emil@vuzv.cz