Assessment of heritability coefficient variation for various breeds of cattle

Jolanta Różańska-Zawieja*[1], Alicja Szabelska-Beręsewicz [2], Zbigniew Sobek[1], Joanna Zyprych-Walczak[2], Idzi Siatkowski[2]

[1]*Department of Genetics and Animal Breeding, Poznan University of Life Sciences, Wołyńska 33, 60-637 Poznań, Poland,*
[2]*Departament of Mathematical and Statistical Methods, Poznan University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, Poland*
 corresponding author: jolek@up.poznan.pl
Key words: dairy cattle, gestation length, heritability

1 RESUME
For over 2.5 million dairy cows (Jersey, Montbéliard, Polish Red, Polish Red-and-White Holstein-Friesian, Simmental, Polish Black-and-White, Polish Red-and-White and Polish Black-and-White HF) under assessment in Poland in 2005-2010, gestation length was analysed. This trait had a normal distribution. Linear mixed model with two random effects (herd and sire) was used for the analysis. Then the coefficient of heritability was calculated from components of variances. Obtained values of this coefficient were in the range of 0.047 to 0.243. Within the accessed values of heritability, coefficient Jersey breed could distinguished, which resulted with the highest value of 0.243. In the most numerous breed Polish Black-and-White Holstein-Friesian coefficient of heritability within the group after different breeds of sires was evaluated. Average gestation length was between 279.83 and 284.28. For the analyzed sire's breeds, the coefficient of heritability was in the range from 0.08 to 0.117. Calculations howed that the estimates of heritability coefficient characteristics of the gestation length, when considering the different breeds of cattle and different breeds of sires, have an impact on the size of the obtained assessment.

2 INTRODUCTION

Gestation length is a reproductive trait that significantly affects cattle breeding and production. Because of the linkage between length of the pregnancy and other properties, it is important to have a knowledge about the heritability of this property, however there frequently is the problem with selection of such model for analysis which could allow to obtain more accurate assessment. Systems for breeders allow collection of large data sets, which enable the use of different models. The basic problem is to find the best model and to indicate a sufficient size of the dataset to ensure the accuracy of assessments. Other authors often use typical models, even in situations when they have additional information that could be included in the model (Rönnegård and Lee 2013). Condition for the model selection must be completed for each population within a specific generation. The gestation length depends on many non-genetical factors, such as calving season, herd, lactation number, sex of the calf or the type of delivery (Norman *et al.* 2009, Petrović *et al.* 2010, Kumar *et al.* 2016). In the case of dairy cattle, it happens that heritability coefficient is calculated for the whole population without paying attention to the fact that there are different breeds. In such a situation most often to model, the breed effect is introduced in order to improve the overall assessment, and not in order to show the differences in the values heritability coefficient in the individual breeds

(subpopulations). According to Yang J. *et al.* (2014), linear mixed model association methods prevent false positive associations and increase power. In addition, Yang J. *et al.* (2014) stated that the choice of whether to apply linear mixed model or other methods should be a function of sample size and the severity of case-control ascertainment. Considered dataset is of size more than 1,000,000, which is appropriate for the proposed analysis. Purpose of the research was assessment of the heritability factor variability of the pregnancy length for various breeds of dairy cattle, which was submitted evaluation in Poland, and estimated by assorted models.

3   MATERIALS AND METHODS

Data for the research was acquired with PFHBiPM consent of Polish nationwide programme SYMLEK which has inventoried all cows rated of use value in Poland. Database contained information for more than 2.5 million Animals - dairy cattle belonging to eight different breeds. The gathered dataset covers the cows that gave birth during years 2005-2010. The analyzed feature was the gestation length. In the described dataset 8 breeds was included: Jersey (JE), Montbéliard (MO), Polish Red (RP), Polish Red-and-White Holstein- Friesian (RW), Simmental (SM), Polish Black-and-White (ZB), Polish Red-and-White (ZR), Polish Black-and-White Holstein-Friesian (PHF). The Polish Black-and-White Holstein-Friesian (PHF) breed is the biggest part of the dataset. For this breed we choose the records where bulls had at least 100 calves as well as where herds occur at least 20 times in the dataset. The criteria was fulfilled by 1 095 828 cows. Based on such dataset the linear mixed model was used:

$$Y_{ijklmnop} = s_i + L_j + G_k + M_l + P_m + o_n + B_o + R_p + e_{ijklmnop}, \qquad (1)$$

Where:

$Y_{ijklmnop}$ is the gestation length,

s_i is a random effect of i-th sires,

L_j is a fixed effect of j-th lactation number,

$j = 1, \dots, 11,$

G_k is a fixed effect of the sex of calf, $k = 1, 2,$

M_l is a fixed effect connected with the mass of l-th calf,

P_m is a fixed effect of m-th year,

$m = 2005, \dots, 2010,$

o_n is a random effect of n-th herd,

B_o is a fixed effect of o-th season of calf birth,

$o = 1, 2$, (summer, winter),

R_p is a fixed effect of p-th breed, $p = 1, \dots, 8,$

$e_{ijklmnop}$ is a random error.

The following phenotype variance (σ_Y^2) was derived from the estimated variance components:

$$\sigma_Y^2 = \sigma_s^2 + \sigma_o^2 + \sigma_e^2,$$

where σ_s^2 is a variance of a bull effect, σ_o^2 is a variance of the herd effect and σ_e^2 is a residual variance. Using these parameters the coefficient of heritability (h_s^2) can be calculated as follows:

$$h_s^2 = \frac{\sigma_s^2}{\sigma_Y^2}.$$

All the calculations were performed in R platform 3.1.3 (R Core Team, 2015) with usage of the packages lme4 (Bates et. al, 2014), data table (Dowle et. al, 2014) and dplyr (Wickham and Francois, 2015).

The calculations were performed for the whole dataset based on the model (1) – all breeds as

well as for the subsets within each breed separately with the reduced model without fixed

effect of the breed ($\mathbf{R_p}$). In mixed models, we took into account only those records for which

sires had at least 100 calves at the same time the herd was repeated in at least 20 observations.

4   RESULTS

Firstly, the analysis of the linear mixed model (1) was performed. The estimated values, standard errors as well as t values for the fixed effects are gathered in Table 1.

Table 1. The characteristics of the fixed effects for the full model (1).

Term	Estimate	Standard error	t value
PHF (Intercept)	281.112	0.065	4343.940
ZR	-0.661	0.106	-6.245
ZB	-1.308	0.132	-9.894
SM	2.871	0.091	31.492
RW	-1.001	0.059	-16.851
RP	1.404	0.193	7.275
MO	1.410	0.183	7.684
JE	-1.576	0.156	-10.074
P	0.062	0.007	8.957
B(W)	0.723	0.011	63.896
L	0.109	0.003	37.293
G(F)	-1.043	0.012	-89.871
M	0.147	0.001	103.860

Breeds are denoted as following: Jersey (JE), Montbéliard (MO), Polish Red (RP), Polish Red-and-White Holstein-Friesian (RW), Simmental (SM), Polish Black-and-White (ZB), Polish Red-and-White (ZR), Polish Black-and-White Holstein-Friesian (PHF). Besides the breeds, P denotes the year of birth from 0 to 6, where 0 is connected with year 2005 and 6 is connected with year 2010, B denotes the season of calf's birth (W is connected with winter), L denoted lactation number, G denotes the sex of the calf (F stands for female). M denotes the standardized mass of the calf.

The PHF (Intercept) is the fitted value of gestation length value for PHF breed, 2005 year, season – summer, male - calf, when lactation equals to 0 and average mass. The estimated value ZR is the difference between breeds ZR and PHF. The estimate equals to -0.661 means that the fitted estimated value in ZR breed is lower than 0.661 compared to PHF breed. The P term means that if we hold all the other factors constant the increase in 1 year leads to an increase 0.062 of a day of gestation length. If we increase lactation number (L) in 1, the gestation length will increase of 0.109 of a day. Similarly, if the mass (M) is increased by 1kg the gestation length will increase of 0.147 of a day. Finally, if we consider winter season, B (W), instead of summer the length of gestation will increase of 0.723 of a day and for a female instead of male calf, G (F), the length of gestation will decrease of 1.043 of a day. As the second step the model (1) was analyzed without the breed (R_p) factor. For each we calculated Akaike coefficient (AIC), (Akaike 1974). This criterion was used to compare the models that were evaluated. The lower value of AICs is a more appropriate model. The variance of the random variances was also calculated and the heritability coefficient was estimated Next, we calculated the same characteristics for the full model (1). All the results are gathered in Table 2.

Table 2: The characteristics for the full model (1) as well as models for specific breeds.

Breed	AIC	N	No of herds	σ_e^2	σ_s^2	σ_z^2	h_z^2
ZR	25911.11	4051	444	4.510	3.064	31.159	0.079
ZB	14671.23	2298	253	5.694	1.968	29.716	0.053
SM	86906.68	13607	906	5.668	5.346	31.651	0.125
RW	260341.87	40997	3488	3.877	4.754	30.893	0.120
RP	19407.81	3137	379	2.938	1.465	26.016	0.048
MO	10579.60	1661	72	7.902	1.917	30.947	0.047
JE	17319.50	2865	325	2.170	7.716	2 1.899	0.243
PHF	6336668.18	1026989	12474	2.342	3.486	27.188	0.106
All	6770779.85	1095605	18341	2.439	3.796	27.424	0.113

Jersey (JE), Montbéliard (MO), Polish Red (RP), Polish Red-and-White Holstein-Friesian (RW), Simmental (SM), Polish Black-and-White (ZB), Polish Red-and-White (ZR), Polish Black-and-White Holstein-Friesian (PHF).

The range of obtained values of the heritability coefficient is between 0.047 and 0.243. The maximum value (0.243) is evaluated for the breed Jersey. This value decisively differs from the values for the rest of breeds, which are in range between 0.047 to 0.125. For each breed in model (1) we estimated the mean value, standard error and bound of the 95% confidence intervals of gestation length for breed factor. The results are gathered in Table 3.

Table 3: The estimated values of mean, standard error and limits of the 95% confidence intervals of gestation length for each level of breed factor (the lower confidence limit is designated as LCL, whereas the upper confidence limit is labelled as UCL)

Breed	Mean	Standard error	LCL	UCL
ZR	280.74	0.12	280.52	280.97
ZB	280.10	0.15	279.81	280.38
SM	284.28	0.10	284.09	284.46
RW	280.41	0.06	280.30	280.51
RP	282.81	0.22	282.39	283.23
MO	282.82	0.21	282.41	283.22
JE	279.83	0.17	279.49	280.17
PHF	280.27	0.04	280.19	280.35

The evaluation of significance of the differences between each level of breed is presented in Figure 1.

Figure. 1: The significance of the mean differences in gestation length between each pair of levels of breed. The colours show significant (S) and not significant differences (NS).
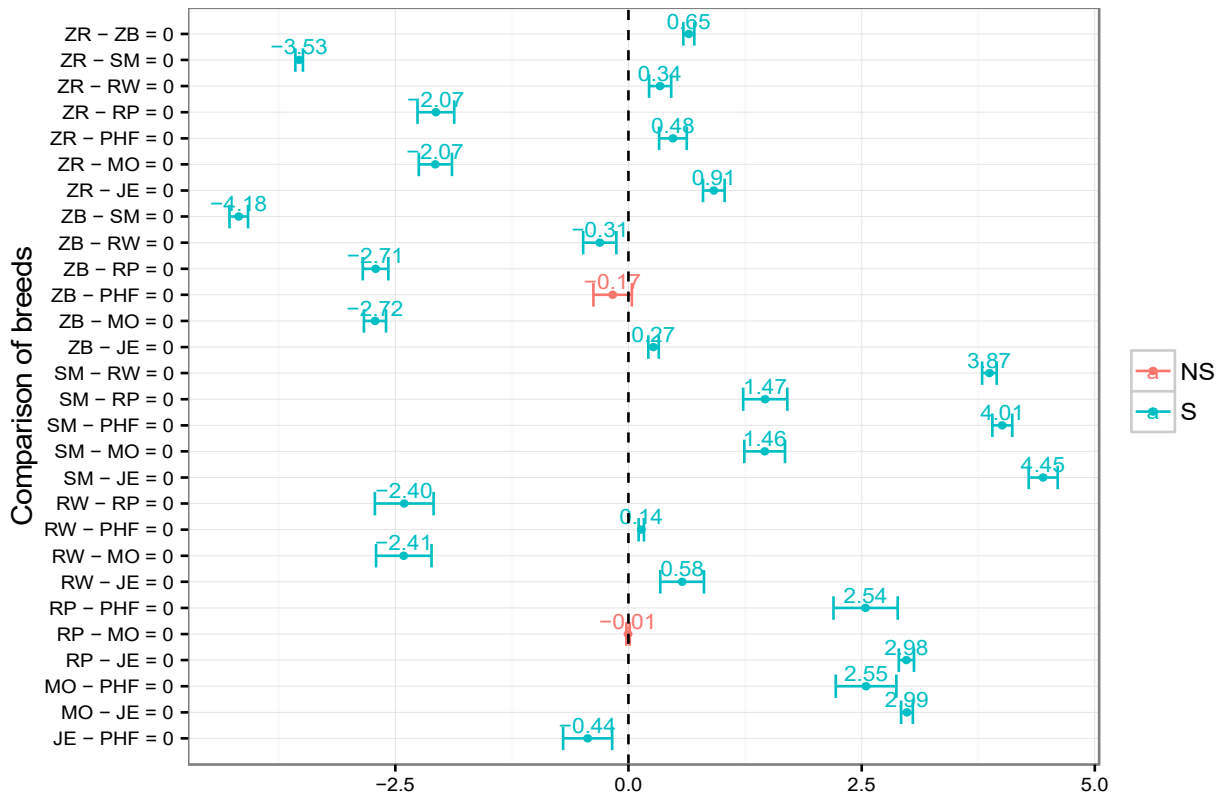
With the exception of pairs ZB-PHF and RP-MO, the comparison of the remaining pairs of breeds indicated that the differences for the mean value of the studied feature are statistically significant. As the PHF was the largest part of dataset, we wanted to have more insight about this breed. We divided the calves from this breed into the groups based on the breed of the bulls. We took into account only those breeds of bulls for which the number of calves were higher than 1000. We verified if the differences between mean values of gestation length for each pair of bull's breeds are significant with the usage of t test. The obtained p-values are presented in Table 4. GL of the calves born from PHF breed was analyzed with respect to breeds of sires. 1026989 cows of PHF breed gave birth to calves after 13 breeds of sires: LM (Limousine), MM (Mix beef breed), PHF (polish black-and-white Holstein-Friesian), BB (Belgian Blue), RW (polish red-and-white Holstein-Friesian), SM (Simmental), MO (Montbéliard), CH (Charolaise), PI (Piemontese), AN (Black Angus), AR (Read Angus), NR (Norwegian Red), JE (jersey) and SR (Swedish Red).

Table 4: The p-values of the t tests performed on the average gestation length for datasets obtained for each pair of bull breeds.

Breed	LM	MM	PHF	BB	RW	SM	MO	CH	PI	AN	AR	NR	JE
MM	0 .00												
PHF	0 .00	0 .02											
BB	0 .00	0 .00	0 .00										
RW	0 .00	0 .00	0 .00	0 .00									
SM	0 .00	0 .00	0 .00	0 .00	0 .00								
MO	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00							
CH	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00						
PI	0.15	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00					
AN	0 .00	0 .00	0 .00	0 .00	0 .02	0 .00	0 .00	0 .00	0 .00				
AR	0 .00	0 .00	0 .00	0 .00	0 .02	0 .00	0 .00	0 .00	0 .00	0 .86			
NR	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00		
JE	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	
SR	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .00	0 .43	0 .00

NS, $p<0.05$, $p<0.01$ - the highlighted p-values in yellow are significant on the 0.01 level, highlighted p-values in green are significant on the 0.05, highlighted p-values in purple are considered not significant.
Jersey (JE), Montbéliard (MO), Polish Red (RP), Polish Red-and-White Holstein-Friesian (RW), Simmental (SM), Polish Black-and-White (ZB), Polish Red-and-White (ZR), Polish Black-and-White Holstein-Friesian (PHF).

The analysis of the most numerous breed of cows (PHF), with respect to the influence of the sire's breed on the gestation length revealed that such influence exists for the vast majority of breeds. Only the pairs LM-PI, AR-AN and SR-NR did not show such differences. AS the next step we calculated the characteristics of gestation length in PHF breed for each breed of the bulls as well as we estimated the h_s^2 for each of these groups. The results are presented in Table 5.

Table 5: The characteristics of the PHF breed concern the calves and specific breeds of bulls.

Breed	N	Mean	Standard deviation	Min	Max	R	σ_o^2	σ_s^2	σ_e^2	h_s^2
LM	31949	285 .56	6 .05	254	297	43	2 .131	3 .893	30 .400	0 .107
MM	16853	279 .37	5 .54	254	297	43	2 .850	1 .956	25 .678	0 .064
PHF	902902	279 .27	5 .64	254	297	43	2 .333	2 .297	26 .942	0 .073
BB	3169	280 .82	6 .05	254	297	43	3 .113	1 .241	31 .654	0 .034
RW	16895	279 .99	5 .73	254	297	43	3 .261	3 .015	26 .119	0 .093
SM	21304	284 .01	5 .91	254	297	43	2 .917	2 .385	29 .060	0 .069
MO	7939	282 .58	5 .77	254	297	43	2 .576	2 .734	27 .587	0 .083
CH	6319	283 .32	6 .12	254	297	43	3 .287	4 .135	28 .830	0 .114
PI	5014	285 .69	5 .81	254	297	43	1 .647	2 .075	29 .900	0 .062
AN	1763	280 .33	5 .57	255	297	42	3 .435	0 .858	26 .626	0 .028
AR	2110	280 .29	5 .89	254	297	43	4 .034	3 .332	27 .000	0 .097
NR	1267	278 .33	5 .49	254	296	42	3 .818	3 .653	23 .706	0 .117
JE	4661	279 .69	5 .87	254	297	43	3 .298	2 .173	27 .990	0 .065
SR	2591	278 .48	5 .64	254	297	43	3 .085	1 .977	26 .441	0 .063

Jersey (JE), Montbéliard (MO), Polish Red (RP), Polish Red-and-White Holstein-Friesian (RW), Simmental (SM), Polish Black-and-White (ZB), Polish Red-and-White (ZR), Polish Black-and-White Holstein-Friesian (PHF).

DISCUSSION

Yang *et al.* (2014) stated that linear mixed model association methods could prevent false positive associations and increase power at reasonable computational cost. Following this suggestion, we applied linear mixed model for the analysis of gestation length, where two random effects were included: sire and herd. This choice was made since for both effects the observed number of levels is only subsets of all possible outcomes. At the same time, the number of levels for herd effect is considered large. Considering large amount of data we can as-

sume based on central limit theorem that the gestation length follows normal distribution. Hansen *et al.* (2004) concluded that the coefficient of heritability for GL is equal to 0.42, whereas Jamrozik *et al.* (2005) consider it is equal to 0.31. This suggests genetic possibility for the change in this feature. However, it is not suggested to conduct selection in the direction of change of the gestation length, since the extreme values of this feature have negative influence on cause dystocia or even stillbirths (Olson *et al.* 2009). In addition, Eaglen *et al.* (2013) demonstrate that the most preferable is GL similar to mean value, since such length is proven to result with low number of difficult births, as well as has positive effect on lactation. Numerous authors have shown that there is a relationship between GL and difficult birth or incidence of miscarriage. Eliminating these cows, which are characterized by unfavourable values of these characteristics, we actually perform indirect selection on the trait GL. Eaglen *et al.* (2013) clearly stated that often-difficult birth and more frequent miscarriages appear with atypical length of gestation. This means that the elimination of such cows. is a selection of defect stabilization leading to a reduction in spreading GL. Kumar *et al.* (2016) estimated the

coefficient of heritability for GL. For the JE breed on the level 0.24 ± 0.08, which is similar to the value obtained by our calculations in this paper. In turn, the coefficient of heritability for PHF breed described in Nogalski *et al.* (2012) was evaluated on the level 0.054–0.073, which again is close to our results presented in the paper. The results of cited authors above as well as our outcomes suggest that the two breeds JE and PHF should not be combined when h_s^2 is calculated.

The resulted assessment of heritability coefficient for dairy breeds of cows used in Poland is similar with the assessments in other countries (Hansen *et al.* 2004 Jamrozik *et al.* 2005, Olson *et al.* 2009). Eghbalsaied (2011) assessed the heritability for the population of HF breed in Iran for primiparas and multiparous for GL and obtained values 0.184 and 0.153, respectively. In contrast, Johanson *et al.* (2011) also for HF breed in Iowa, the United States the heritability coefficient has been rated at 0.07. The reasons for such differences in evaluations can be discerned either in that they are related to different populations, or that were performed using a variety of statistical methods.

CONCLUSION

The research on the Polish population of dairy cattle has shown that the evaluations of the heritability coefficients for GL trait for individual breeds are different. These results suggest that in the calculation, data should be limited to one breed and it should be avoided evaluate throughout the material using the effect of breed. If, within the population of one breed of cows, interbred hybrids were created with sires of other breeds, it had an effect on GL of hybrids. Also in this situation, assessment of heritability for total material is vague.

REFERENCES

Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. No. 19. pp. 716-723.

Bates D., Maechler M., Bolker B. and Walker S. 2014. _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-7. URL: http://CRAN.R-project.org/package=lme4.

Dowle M., Short T., Lianoglou S., Srinivasan A. with contributions from Saporta R. and Antonyan E. 2014. data.table: Extension of data.frame. R package version 1.9.4. URL: http://CRAN.R-project.org/package=data.table.

Eaglen S. A. E., Coffey M. P., Woolliams J. A. and Wall E. 2013. Direct and maternal genetic relationships between calving ease, gestation length, milk production,

fertility, type and lifespan of Holstein-Friesian primiparous cows. J. Dairy Sci. 96:4015–4025.

Eghbalsaied S. 2011. Estimation of genetic parameters for 13 female fertility indices in Holstein dairy cows. Trop. Anim, Health Prod 43: 811-816.

Hansen M., Lund M.S., Pedersen J., Christensen L.G. 2004. Gestation length in Danish Holsteins has weak genetic associations with stillbirth, calving, difficulty and calf size. Livestock Production Science 91: 23-33.

Jamrozik J., Fatehi J., Kistemaker G.J., Schaeffer L.R. 2005. Estimates of genetic parameters for Canadian Holstein female reproduction traits. J. Dairy Sci. 88: 2199-2208.

Johanson J.M., Berger P.J., Tsuruta S., Misztal I. 2011. A Bayesian threshold-linear model evaluation of perinatal mortality, dystocia, birth weight and gestation length in a Holstein herd. J. Dairy Sci. 94: 450-460.

Kumar A, Mandal A., Gupta A. K., Ratwan P. 2016. Genetic and environmental causes of variation in gestation length of Jersey crossbred cattle. Veterinary World, EISSN: 2231-0916, /Vol.9/April-2016/3.pdf, s. 351-355.

Nogalski, Z. and Piwczyński, D. 2012. Association of length of pregnancy with other reproductive traits in dairy cattle. Asian. Aust. J. Anim.*,* 25(1): 22-27.

Norman H. D., Wright M. T., Kuhn S.M., Hubbard J. B., Cole J. B., VanRaden P. M. 2009. Genetic and environmental factors that affect gestation length in dairy cattle. J. Dairy Sci. 92: 2259-2269.

Olson K.M., Cassel B.G., McAllister A.J., Washburn S.P. 2009. Dystocia, stillbirth, gestation length and birth weight in Holstein, Jersey and reciprocal crosses from a planned experiment. J. Dairy Sci. 92: 6167-6175.

Petrović M. D., Bogdanović V., Bogosavljević-Bošković S., Đokovoć R. 2010. Effect of Systematic Factors on Gestation Length in Simmental Cows. ActaAgriculturaeSerbica, XV(29): 31-37.

R Core Team 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/

Rönnegård, L.; Lee, Y. 2013. Exploring the potential of hierarchical generalized linear models in animal breeding and genetics. J. *Anim. Breed.* Genet., Vol. 130 Issue 6, 415-416.

Wickham H. and Francois R. 2015. dplyr: A Grammar of Data Manipulation. R pack-age version 0.4.1. URL: http://CRAN.R-project.org/package=dplyr

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., & Price, A. L. 2014. Advantages and pitfalls in the application of mixed-model association methods. *Nature genetics*, 46(2), 100-106.