Acta fytotechn zootechn, 19, 2016(3): 103–105
http://www.acta.fapz.uniag.sk

Short Communication

Detection of SNP effects on feed conversion ratio in pigs based on entropy approach

Alicja Borowska[1]*, Henry Reyer[2], Klaus Wimmers[2], Patrick Varley[3], Tomasz Szwaczkowski[1]
[1]*Poznan University of Life Sciences, Department of Genetics and Animal Breeding, Division of Horse Breeding, Poznan, Poland*
[2]*Leibniz Institute of Farm Animal Biology, Institute of Genome Biology, Dummerstorf, Germany*
[3]*Hermitage Genetics, Kilkenny, Ireland*

Article Details: Received: 2016-04-27 | Accepted: 2016-05-05 | Available online: 2016-08-31

dx.doi.org/10.15414/afz.2016.19.03.103–105

The objectives of the study were to classify SNPs according to their contribution to the feed conversion ratio and to indicate interactions between the most informative SNPs using entropy analysis. The records of 1296 pigs were included. Two selection criteria for molecular data were applied: call rate 0.95 and minor allele frequency 0.05. After this, 50 951 SNPs were included into the entropy analysis. For each SNP entropy and conditional entropy were estimated. For interaction analyses the most informative SNPs were selected. For each pair of SNPs, the mutual information was assessed. A majority of the loci studied showed relatively small contributions. The most informative SNPs are mainly located on chromosomes: 1, 4, 7, 9 and 14. Whereas important interactions between SNP pairs were detected on chromosomes: 1, 14, 15 and 16. High mutual information was registered for SNPs located nearby.

Keywords: pigs, information theory, feed conversion, SNP

1 Introduction

The feed conversion ratio (FCR) is a measure of the amount of feed required to produce 1 kilogramme of pig live weight. The conversion of high quality feed into weight gain is importance in modern pig breeding as a broad indicator of efficiency and ecological-friendly pig production (Sahana et al. 2013, Do et al. 2014). More efficient use of feed has the potential to reduce greenhouse gas intensity and improve productivity (MacLeod M. et al. 2013). FCR is an index based on factors affecting the variables of growth rate and feed intake, like genetic background, feeding practice, environmental control or health status. Recently, several statistical approaches to detect the effects of SNPs on livestock performance traits have been described. The most common is the Genome-Wide Association Study method, so called genome scanning, applied to study the association between value of the trait and single nucleotide polymorphisms (SNPs) distributed throughout the genome. An alternative, new approach is entropy analysis addressed for categorical characters. Theoretical backgrounds of this methodology have been described among others by Shannon (1948). This methodology has been increasingly implemented in genetic studies especially in medicine (Moore et al. 2006) however the application of entropy analysis is still growing also in animal science (Borowska et al., 2014).

The objectives of the study were to classify SNPs according to their contribution to the feed conversion ratio, as well as to indicate interactions between the most informative SNPs using entropy analysis.

2 Material and methods

The records of 1296 pigs from the Hylean Maxgro™ line were included. The feed conversion ratio (FCR) was analysed. Entropy analysis requires discrete variables therefore the collected data wereclassified into four groups (according to the quartiles). In consequence, the traits can be treated s categorical. Molecular data included 61 565 SNPs. Two selection criteria of this data were applied: call rate 0.95 and minor allele frequency 0.05. After this, 50 951 SNPs were included into the entropy analysis.

The effects of studied loci were examined in two steps. Firstly, genotypes were classified according to their contribution to the traits. For each SNP, entropy and conditional entropy were estimated (see e.g. Borowska et al., 2014):

$$H(T) = -\sum_{i=1}^{4} p(t)\log p(t)$$

***Corresponding Author:** Alicja Borowska, Poznan University of Life Sciences, Department of Genetics and Animal Breeding, Division of Horse Breeding, Wolynska Str 33, 60-637 Poznan, Poland. E-mail: aborek@up.poznan.pl

Entropy [$H(T)$] being the measure of uncertainty, has the property of taking great values for very rare values, but for the most certain values, it is close to zero, where:

$p(t)$ – the probability that analysed trait (T) obtain a value of t, based on the available genotypes
n – the number of classes of trait (T)

Conditional entropy $H(T|S)$ quantifies the remaining uncertainty about trait (T) with the knowledge of SNP (S):

$$H(T|S) = -\sum_{i=1}^{4} p(t|s)\log p(t|s)$$

where:
$p(t|s)$ – the probability that analysed trait (T) obtain value of t under the condition that analysed SNP (S) obtain a genotype of s
n – the number of classes of trait T under analysed SNP S

In the second stage, interactions between SNPs were analysed. For this analysis the most informative SNPs were selected. In the first step, SNPs with the smallest values of conditional entropy (from first quartile) were chosen. Next, a quarter of the most informative SNPs for each chromosome were taken into interaction analysis. For each pair of selected SNPs, the mutual information (MI) were assessed:

$$MI(S_1|S_2) = H(S_1) + H(S_2) - H(S_1, S_1)$$

where:
$H(S_1)$, $H(S_2)$ – the entropies of the SNP S_1/S_2
$H(S_1, S_2)$ – the joint entropy between SNP S_1 and S_2

Analyses were performed by the use of the R 3.1 Statistical Software (2014).

3 Results and discussion

The participation of analysed loci in the feed conversion ratio varied. A majority of these loci showed relatively small contributions. Normed conditional entropy coefficients (ET – EC) for the most informative SNPs are visualized on a genome-wide scale in Figure 1.

Chromosome 1 harbours the largest number of SNPs with high conditional entropy values. The most informative SNPs are mainly located on chromosomes 1, 4, 7, 9 and 14. High mutual information was registered for SNPs located nearby. The most informative SNP pairs connected with feed conversion ratio are located on chromosomes 1, 14, 15 and 16. The dependencies among studied loci may be influenced by manifold genetic backgrounds, e.g. linkage disequilibrium, meiotic drive, etc. Previously, entropy method was also applied to classify SNPs according to their contribution to leanness of boars by Borowska et al. (2014). The authors revealed the largest effects for chromosomes 1 and 4. Prominent interactions between loci were also located on chromosome 1. Other studies indicated regions associated with FCR on chromosomes 4, 7, 9 and 14 (Onteru et al. 2013, Sahana et al. 2014). It is worth noting that two SNPs (ASGA0063199 and INRA0043941 on SSC14) among the fifty SNPs with highest values of conditional entropy are located in *TTC28* gene encoding the tetratricopeptide repeat domain 28. The orthologue gene in mouse is responsible for decreased lean body mass. Moreover, Do et al. (2014) identified *TTC29* as candidate gene for residual feed intake based on their genome-wide association analysis in pigs.

4 Conclusions

The most informative SNPs are mainly located on chromosomes: 1, 4, 7, 9 and 14. Whereas important interactions between SNP pairs are detected on chromosomes: 1, 14, 15 and 16. High mutual information was registered for SNPs located nearby. Moreover, entropy analysis is able to elucidate informative regions at the genome-wide level.

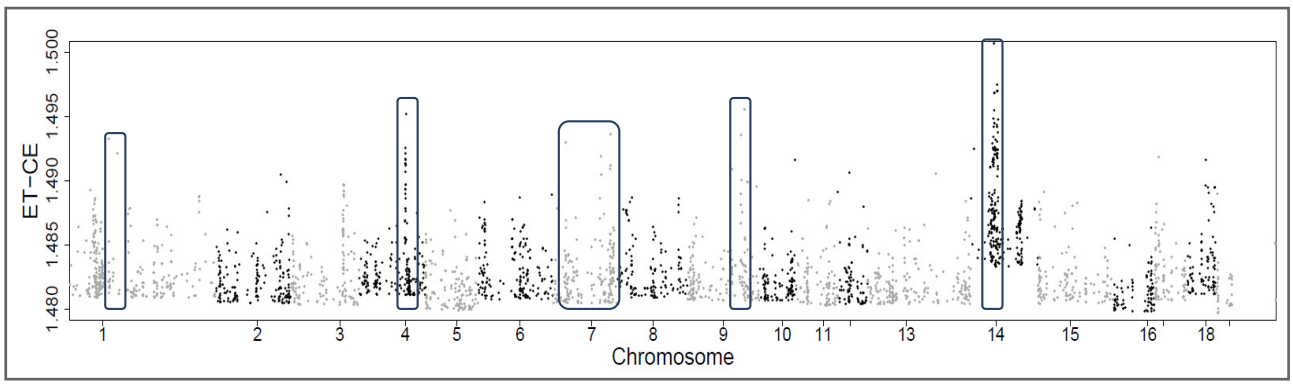

Figure 1 Localisation of the most informative SNPs for feed conversion ratio
ET – entropy of trait, CE – conditional entropy

Acknowledgments

The study was supported by the European Union Seventh Framework Programme as part of the ECO-FCE project under grant agreement No. 311794. Additionally, this scientific work is partly funded by the Ministry of Science and Higher Education from funds for science in the years 2015–2017 allocated to an international co-financed project (noW171.PR/2014).

References

BOROWSKA, A. et al. (2014) Associations of fifty single nucleotide polymorphisms within candidate genes with meatness in pigs. *Czech Journal of Animal Science,* vol. 59, pp. 227–237.

DO, D.N. et al. (2014) Genome-wide association and pathway analysis of feed efficiency in pigs reveal candidate genes and pathways for residual feed intake. *Frontiers in Genetics,* vol. 5, pp. 307. doi: http://dx.doi.org/10.3389/fgene.2014.00307

MaCLEOD, M. et al. (2013) *Greenhouse gas emissions from pig and chicken supply chains – A global life cycle assessment.* Rome: Food and Agriculture Organization of the United Nations (FAO).

MOORE, J.H. et al. (2006) A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *Journal of Theoretical Biology,* vol. 241, pp. 252–261. doi: http:/dx.doi.org/10.1016/j.jtbi.2005.11.036

ONTERU, S.K. et al. (2013) Whole genome association studies of residual feed intake and related traits in the pig. *PLoSONE*, vol. 8, e61756. http://dx.doi.org/10.1371/journal.pone.0061756

R CORE TEAM (2014) R: A *language and environment for statistical computing.* [Software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved October 12, 2015 from http://www.R-project.org/

SAHANA, G. et al. (2013) A genome-wide association scan in pig identifi es novel regions associated with feed efficiency trait. *Journal of Animal Science,* vol. 91, pp. 1041–105. doi: http://dx.doi.org/10.2527/jas2012-56430

SHANNON, C.E. (1948) A mathematical theory of communication. *Bell System Technical Journal,* vol. 27, pp. 379–423.

