IMMUNOLOGICAL GENES SELECTED FOR ENVIRONMENTAL VARIANCE COULD CONTROL THE ANIMAL RESILIENCE

Casto-Rebollo[1] C., Argente[2] M.J., Garcia[2] M.L., Blasco[1] A., Ibáñez-Escriche[1*] N.

[1]Institute for animal Science and Technology, Universitat Politècnica de València, València, Spain.
[2]Departamento de Tecnología Agroalimentaria, Universidad Miguel Hernández de Elche, Orihuela, Spain.
Correspondence to: noeibes@dca.upv.es

ABSTRACT

Environmental variance of traits (V_E) has recently been related with resilience. Thus, a greater knowledge of the genetic background of V_E could help to understand better the animal resilience. A successful selection experiment in rabbits for a high and a low V_E of litter size (LS) allowed to identify differences in resilience between animals. The line with a low V_E of LS seemed to cope better with the environmental disturbances than the line with a high V_E of LS. The aim of this study was to identify genomic regions modified by selection of V_E and that could affect the animal resilience. For that, genotypes from 91 does of base population, and 142 of the line with a high V_E of LS and 134 of the line with a low V_E of LS at generation 11 were used to identify signatures of selection. 93 genotypes at generation 13 were used to validate the results. The signatures of selection were identified using three complementary analysis: runs of homozygosity (ROH), variations of linkage disequilibrium (VarLD) and fixation index (F_{ST}). A whole-genome sequencing (WGS) analysis was performed on 54 animals at generation 10 to highlight the genes with functional mutations. We identified 311 candidate genes with relevant functional mutation in their transcription unit. 107 of them had functions related to the stress response, reproduction and embryo development, carbohydrate and lipid metabolism, and/or immune system. Functional mutations fixed in one of the rabbit lines and absent in the other were identified in the genes *C3orf20, GRN, EPCAM, ENSOCUG00000017494, ENSOCUG00000024926, ENSOCUG00000026560, MYLK, HECA and NMNAT3*. The biological pathways of candidate genes explain the differences found between the rabbit lines in immune response biomarkers (plasma cortisol, leukocytes, and acute-phase protein levels), in plasma concentrations of cholesterol and triglycerides, mortality and resilience. Also, these results could explain the correlated response of the V_E of LS with embryo implantation, embryo survival and LS. However, the real implications of these genes for V_E and animal resilience must still be unravelled through their functional analysis.
Keywords: Resilience, environmental variance, selection signature, functional mutation, rabbit

INTRODUCTION

Environmental variance (V_E) of traits has been recently related with animal resilience (Berghof et al., 2019). Resilience is the animal's ability to maintain or rapid recovery of its productive performance after an environmental disturbance. Study resilience could help to improve the animal welfare in the farms (Colditz and Hine, 2016). A recent study showed that lines successful selected for a high and a low V_E of LS in rabbits during 13th generations (Blasco et al., 2017) showed differences in biomarkers of immune and stress response as well as in mortality (Argente et al., 2019; Beloumi et al., 2020). This suggested that the line with a low V_E of LS cope better with the environmental stressors than the line with a high V_E of LS. The aim of this work was to identify what genomic regions were modified due to the pressure of selection applied during 13th generations in these divergent rabbit lines. For that, we studied the signatures of selection identifying contiguous homozygous segments (ROH), variations in the extend of LD patterns (VarLD) and differences in allele frequencies (F_{ST}) between the rabbit populations. According to Cadzow et al. (2014), selection can generate different patterns of genetic variation. So, multiple methods with different assumptions are needed to detect a wide range of the genetic changes considered signatures of selection. Thus, the identification of signatures of selection in these rabbit lines could help us to understand the selection process of the V_E and how it could affect the resilience of the animals.

MATERIALS AND METHODS

Data

We used rabbits from generation 11, 13 and base population of a divergent selection experiment for a high and a low V_E of LS carried out by the University Miguel Hernandez in Elche, Spain (see Blasco et al., 2017). Genotyping was performed with the 200K Affymetrix Axiom OrcunSNP Array (ThermoFisher Scientific) from blood samples of 473 does: 96 from base population, 282 from generation 11 (147 from the line with a high V_E of LS and 135 from the line with a low V_E of LS) and 95 from generation 13 (46 from the line with a high V_E of LS and 49 from the line with a low VE of LS). Quality control removed animals with a call rate <97% and SNPs with minor allele frequency <0.05, missing genotype >0.05 and unknown positions on rabbit genome (OryCun v2.0.9). At the end, 452 animals and 97,155 SNPs remained in the data set.

Signature of selection

Selection signatures were searched using 276 genotypes from generation 11 (the inbreeding coefficient of the lines with a high and a low VE of LS was IC=0.084 and IC=0.087, respectively) and 91 genotypes from the base population (IC=0.077). The 93 genotypes from generation 13 remained for the validation analysis.

Detection of runs of homozygosity

The ROH was performed using the PLINK v1.90 software (Chang et al., 2015). The parameters used to calculate a ROH were set based on Ceballos *et al* (2018). The algorithm searched for homozygous segments in each chromosome using sliding windows of 500 kb with almost 50 SNPs. Heterozygous SNPs and missing calls were not allowed. A homozygous segment was considered a ROH if the number of consecutives SNPs was >50 and the density was above 1 SNP in 30 Kb. A ROH was considered a signature of selection if was a consensus genomic region in almost 50% of the animals in the line with a low V_E of LS, in 50% of the animals in the line with a high V_E of LS, and was in the base population.

Quantification of VarLD scores

Linkage disequilibrium patterns between populations were performed using the program VarLD (see more details in Teo et al., 2009). VarLD scores were calculated for pairwise comparisons between the three populations (B/H, B/L and H/L) and quantified for each window of 50 SNPs sliding by one SNP. All scores were standardized within pairwise comparison to reduce the effect of the window size and the population LD background. Candidate selection signatures were those windows with a VarLD score equal or higher to the VarLD score at percentile 99.9% and presenting in the B/H and B/L comparison. Windows shared between B/H or B/L and H/L comparisons were considered effect of gene drift.

Estimating of fixation index

The fixation index was calculated using the Weir and Cockerham's pairwise estimator method (Weir and Cockerham's, 1984) implemented in VCFtools v.1.16 software (Danecek et al., 2011). This method was applied to identify selection signatures between the lines with a high and a low V_E of LS. The FST values were calculated using overlapping windows of 500 kb sliding by steps of 250 kb. Windows with less than ten SNPs were excluded. We considered a candidate selection signature if the window overcame the weighted FST value at percentile 99.9%, and showed divergent changes in its MAF between the rabbits' lines, regarding the base population.

Validation

The candidate signatures of selections were validated using the base population and the animals from generation 13, applying the methods described above. Only the candidates replicated in both analyses (at generation 11 and 13) were proposed as true selection signatures.

Gene identification

Candidate genes were identified in the genomic regions proposed as true selection signatures using whole-genome sequencing (WGS) data to search for functional mutations. WGS data belong to two pools of DNA from all the sires of animals from generation 11 (27 animals per line). Data pre-processing and variant calling were performed following Elston *et al.* (2017) and the GATK Best Practices pipeline (McKenna et al., 2010), respectively. Variants affecting the transcript unit of a gene were considered a functional mutation (for further information see Casto-Rebollo *et al*., 2020). The gene ontologies of each candidate gene were extracted with the R/Bioconductor package biomaRt (Durinck et al., 2009).

RESULTS AND DISCUSSION

A total of 726-candidate selection signatures were identified in animals at generation 11, and 134 of them were validated at generation 13; 129 ROH, two VarLD regions on OCU13 at 89.31-90.54 Mb and OCU14 at 0.014-2.27 Mb, and the FST regions on OCU2 at 104.5-105 Mb (0.56), OCU12 at 8.75-9.5 and OCUX at 81-81.75. Non-overlapping selection signatures between methods were identified. The methods achieved a low correlation between their results (González-Rodríguez et al., 2016; Sosa-Madrid et al., 2020), making difficult overlapping between selection signatures of each method. We identified 815 genes in the 134 genomic regions with true selection signatures, but only 311 of them presented functional mutations affecting their transcript unit. These genes are involved in a wide range of biological processes, hindering the identification of the direct molecular mechanism involved in the V_E of LS. However, we highlighted 65 genes related to the immune response, five to the stress response, and 50 to energy, carbohydrate and lipid metabolism. These genes could explain the differences found on the rabbit lines for immune response biomarker, plasma concentrations of cholesterol and triglycerides, mortality, and resilience (Argente et al., 2019; Beloumi et al., 2020). Moreover, 29 genes involved in reproduction and embryo development could justify the correlated response of V_E of LS with embryo survivor and implantation, and litter size (Argente et al., 2017; Calle et al., 2017). Eight promising genes were identified (*C3orf20*, *GRN*, *EPCAM*, ENSOCUG00000017494, *ENSOCUG00000024926*, *ENSOCUG00000026560*, *MYLK*, *HECA,* and *NMNAT3)* with INDELs and/or SNVs with the alternative allele fixed in one rabbit line and absent in the other. The genes *GRN*, *MYLK*, and *NMNAT3* are also functions involved in the immune response. Our results agree with genome-wide associations findings which would support a relation between the inflammatory response and the V_E (Iung et al., 2019; Casto-Rebollo et al., 2020). The immune system was also related with the animal resilience (Colditz et al., 2016), so its control and modulation could explain the relation between the V_E and the animal resilience.

CONCLUSIONS

In this study we identified several genes that could explain the differences found between the lines in immune response biomarker, plasma concentrations of cholesterol and triglycerides, mortality, and resilience. We also identified promising genes involved in the immune response. That genes had functional mutations fixed in one of the rabbit lines and absent in the other. The molecular mechanism related with the immune system could be the link between the animal resilience and the V_E. However, further studies are necessaries to know the real implication of these genes in the V_E of LS.

REFERENCES

Argente, MJ, Calle, EW, García, ML, Blasco, A. Correlated response in litter size components in rabbits selected for litter size variability. J Anim Breed Genet. 2017; 134: 505– 511.

Argente M, García M, Zbyňovská K, Petruška P, Capcarová M, & Blasco A. Correlated response to selection for litter size environmental variability in rabbits' resilience. Animal 2019;13(10):2348-55.

Beloumi D, Blasco A, Muelas R., Santacreu MA, García ML, Argente MJ, Inflammatory Correlated Response in Two Lines of Rabbit Selected Divergently for Litter Size Environmental Variability. Animals, 2020;10:1540.

Berghof TVL, Poppe M, Mulder HA. Opportunities to Improve Resilience in Animal Breeding Programs. Front. Genet. 2019;9:1664-8021.

Blasco A, Martínez-Álvaro M, García ML, Ibáñez-Escriche N, Argente MJ. Selection for environmental variance of litter size in rabbit. Genet. Sel. Evol. 2017;49:48.

Cadzow M, Boocock J, Nguyen HT, Wilcox P, Merriman TR, Black MA. A bioinformatics workflow for detecting signatures of selection in genomic data. Front Genet. 2014;5: 293 10.3389/fgene.2014.00293.

Calle EW, García ML, Blasco A, Argente MJ. Correlated response in early embryonic development in rabbits selected for litter size variability. World Rabbit Science, 2017;25(4):323-327.

Casto-Rebollo C, Argente MJ, GarcíaML, Pena R, Ibáñez-Escriche N. Identification of functional mutations associated with environmental variance of litter size in rabbits. Genet Sel Evol. 2020:52, 22.

Ceballos FC, Hazelhurst S and Ramsay M. Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. BMC Genomics 2018;19:106. doi:10.1186/s12864-018-4489-0.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience. 2015;4:1–16.

Colditz IG, and Hine BC. Resilience in farm animals: biology, management, breeding and implications for animal welfare. Anim. Prod. Sci. 2016;56:1961–83.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al.. The variant call format and VCFtools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2011, 27:2156-8. doi: 10.1093/bioinformatics/btr330.

Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols. 2009;4:1184-1191.

Elston RC. Preprocessing and quality control for whole-genome sequences from the Illumina HiSeq X platform. In: Wright MN, Gola D, Ziegler A, editors. Statistical human genetics. Methods in Molecular Biology, vol 1666. New York: Humana Press; 2017. p .629-47.

González-Rodríguez A, Munilla S, Mouresan EF, et al. On the performance of tests for the detection of signatures of selection: a case study with the Spanish autochthonous beef cattle populations. Genet Sel Evol. 2016;48(1):81.

Iung LHDS, Carvalheiro R, Neves HHDR, Mulder HA. Genetics and genomics of uniformity and resilience in livestock and aquaculture species: A review. J Anim Breed Genet. 2020; 137: 263– 280.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297-303.

Morgante F, Sørensen P, Sorensen DA, Maltecca C, Mackay TFC. Genetic architecture of micro-environmental plasticity in Drosophila melanogaster. Sci. Rep. 2015;5:9785.

Sosa-Madrid BS, Varona L, Blasco A, Hernández P, Casto-Rebollo C, Ibáñez-Escriche N. The effect of divergent selection for intramuscular fat on the domestic rabbit genome. Animal. 2020;14(11):2225-2235.

Teo YY, Fry AE, Bhattacharya K, Small KS, Kwiatkowski DP and Clark TG. Genome-wide comparisons of variation in linkage disequilibrium. Genome Res. 2009;19:1849-60. doi: 10.1101/gr.092189.109.

Weir BD, Cockerham CC. 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution, 38(6):1358-70. doi: 10.2307/2408641.

Wijga S, Bastiaansen JWM, Wall E, Strandberg E, de Haas Y, Giblin L, Bovenhuis H. Genomic associations with somatic cell score in first-lactation Holstein cows. J. Dairy Sci. 2012;95:899–908.