Chandrababu-Shailaja C., Argente, M., García, M., Hernández, P., Martínez-Álvaro, M. (2026) Metabolic signatures after genetic selection of litter size homogeneity and resilience in a maternal rabbit line, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23719: https://doi.org/10.31274/wcgalp.23719 ©2026 Chandrababu-Shailaja, et al. CC BY-NC-ND

Metabolic signatures after genetic selection of litter size homogeneity and resilience in a maternal rabbit line

C. Chandrababu-Shailaja[1*], M.J. Argente[2], M.L. García[2], P. Hernández[1], and M. Martínez-Álvaro[1]

[1]*Institute of Animal Science and Technology, Universitat Politècnica de Valéncia, Camí de Vera, s/n, Algirós 46022 Valencia, Spain;* [2]*Institute for Agri-Food and Agro-Environmental Research and Innovation (CIAGRO-UMH), Miguel Hernández University of Elche, Orihuela, Spain.* [*]*Corresponding author: cchasha@upv.edu.es*

Introduction

Animal resilience is defined as an animal's ability to cope with short-term environmental fluctuations and rapidly return to its pre-affected state (Colditz and Hine, 2016). In the face of global climate change, selection for resilience has become an attractive breeding goal in livestock farming to ensure productivity under unpredicted environmental challenges and long-term welfare. Characterized by a dynamic and multifactorial nature, the trait remains inaccessible to direct measurement at the individual level. In practice, the trait is assessed indirectly through indicator traits that capture deviations from normal performance trajectories or immune responsiveness to internal or external stimuli. It is critical to understand how genetic selection reshapes resilience in selected populations, where intense selection on specific traits may indirectly impact how the animals respond to diverse stressors (Colditz and Hine, 2016). In a divergent selection experiment in rabbits, resilience was quantified as the residual variance of litter size (*Ve*) over 17 generations (Blasco *et al.*, 2017), establishing a homogenous line (low *Ve*) and a heterogenous line (high *Ve*). By generation 8, the homogenous line showed improved on-farm resilience indicators (lower involuntary culling rate, and higher survival at birth), as well as reduced levels of biochemical stress biomarkers (cortisol, and C-reactive protein) under both basal conditions and following a vaccination challenge against viral haemorrhagic disease or myxomatosis (Argente *et al.*, 2019). Corroborating this, Beloumi *et al*. (2020) observed that the homogenous line at generation 13 exhibited lower inflammatory responses to post-partum stress than the heterogeneous line, suggesting their lower disease susceptibility and enhanced immunocompetence in the homogenous line (for clarity, the homogenous and heterogenous lines are hereafter referred to as the high- and low-resilient lines, respectively, Argente *et al.*, 2019). Although companion studies have demonstrated relevant differences between lines in specific plasma biomarkers indicative of diverse stress responses, the cross-generational consistency of these markers remains unverified. Moreover, the effect of genetic differences in resilience on an untargeted, comprehensive screening of plasma metabolites has not yet been explored. This study aims to compare untargeted plasma metabolomic profiles between genetically resilient and non-resilient lines and to identify biomarkers of resilience that exhibit directional consistency across generations.

Materials & Methods

Animals and metabolomic analysis. A divergent selection experiment based on the residual variance of litter size (*Ve*), after correcting litter size by year-season and lactation, was conducted in rabbits at Miguel Hernández University (Orihuela, Spain) for 17 generations. In total, the dataset comprised of 4,421 records of *Ve* and a pedigree of 5,763 animals and genomic data (134,132 SNPs) from 455 animals belonging to the base, 11th, 12th and 13th generations. Details about the selection methods can be found in (Blasco *et al.*, 2017). Response to selection

for *Ve* was calculated by averaging the breeding values in each generation. Breeding values were estimated by a univariate model with single step genomic BLUP, including a fixed generation effect (18 levels), and random genetic and residual effects. The analysis was performed using Bayesian Inferences (software GIBBSF90+, Misztal *et al*., 2014).
For metabolomic analysis, blood samples were collected from 80 does balanced by line and generation (20 per line, 40 per generation), 24 h after first kindling to ensure a uniform physiological status. Plasma samples were subjected to untargeted semipolar LC–MS metabolomic analysis performed at MS-Omics laboratory, Biogenity (Denmark), and metabolites were quantified as relative abundances with respect to total sum of ion counts per samples. Metabolites detected in at least 80% of samples across all animals or 75% within lines were retained. The final data contained relative abundances, of 1,397 metabolites in generation 13 (1.6% of missing data) and 1,170 in generation 17 (no missing data).

Identification of resilience metabolic biomarkers. The identification of resilience biomarkers was conducted using metabolomic data from generation 17, as *Ve* data was not available for does with metabolomic data in generation 13. Before applying more sophisticated algorithms, metabolites were pre-filtered according to their differences between lines in relative abundances means and variances. The metabolites with relevant mean differences between lines were detected using a linear model with each metabolite as independent variable and line as a fixed effect using Bayesian inference with R package *RabbitR*. Metabolites were selected if their ratio between lines showed a posterior probability of relevance (P_r) ≥90% (Blasco, 2017), assuming 10% (1.1 or 1/1.1) as a relevant value. Differences between lines in metabolite variability were assessed using the univariate Fisher's F test ($p < 0.005$). The final pre-filtered set of metabolites comprised unique metabolites resulting from these two filtering steps. The pre-filtered metabolites abundances were subjected to alternative transformations: chi-squared normalization, and several log-ratio transformations (Pair-wise log-ratio (LR), centered log-ratio (CLR), and additive log-ratio (ALR)), in which missing values were imputed using a random forest algorithm (*missForest* R package). All log-ratio transformations were applied using the *easyCODA* R package. For ALR, *Xanthine* ($C_6H_{13}NO_3$) was selected as the reference metabolite, common between generation 13 and 17, due to its high Procrustes correlation (0.99, 0.95) and comparatively low variances (coefficient of variation is 40% in generation 13, and 47% in generation 17) in both generations (Greenacre *et al.*, 2021). An additional pre-filtering step was performed for pairwise LR-transformed metabolites to improve computational efficiency in the subsequent algorithms, by fitting each LR as independent variable and line as a fixed effect in a linear model and retaining those whose mean ratio between lines had a $P_r \geq$ 90%.
Three classification algorithms namely logistic regression (LGR), partial least squares discriminant analysis (PLS-DA), and random forest (RF), were used to discriminate lines based on pre-filtered metabolic variables, each fitted separately for each data transformation. The dataset was split into 70:30 training and testing sets, repeated 100 times. Hyperparameters were optimized within each training set using cross-validation: *log-loss* for LGR, number of components (*ncomp*) for PLS-DA, and *ntree*, *mtry*, *node size*, and *maxnodes* for RF. Metabolic variables selection was performed using stepwise AIC minimization (LGR), VIP scores and jack-knifing with 95% confidence interval (PLS-DA), and the *Boruta* algorithm (RF). Metabolic variables selected in at least 75% of the 100 models were retained. Final models were refitted using only selected metabolic variables, and performance was evaluated using classification accuracy and area under the curve (AUC) of test sets (expressed as median and interquartile range (IQR)). Subsequent predictive analyses focused on the specific algorithm that yielded the highest classification accuracy. We assessed the ability of metabolic variables to predict *Ve* as a continuous trait by fitting the regression version of the best-performing

classification algorithm to the combined data from both lines. A 75:25 train–test split with 100-repetitions cross-validation was applied, and predictive ability was assessed as the correlation between observed and predicted values (r) and root mean squared error (RMSE) in the testing sets (expressed as median and IQRs). The final subset of metabolic variables relevant to resilience were those with ability to discriminate between lines and predict *Ve*.
From the final subset of metabolomic variables from generation 17, we identified those that overlapped with generation 13, calculated their differences between lines, and screened the variables with consistent directional changes across generations.

Results & Discussion

To our knowledge, this study represents the first cross-generational metabolomic analysis of genetic resilience. By generation 13, the two lines differed by 1.32±0.19 kits^2 (≈0.41 standard deviation units (SD)) and by generation 17, the two lines differed by 1.58±0.22 kits^2 (≈0.5 SD, Figure 1A). In generation 17, PLS-DA was the best-performing algorithm for classifying lines based on metabolomic data. In the testing set, it achieved identical values for classification accuracy and AUC values of 0.92[IQR: 0.83–0.92], with similar performance across all log-transformations when models were fitted using 515 LRs, 15 CLRs, and 12 ALRs. The performance of this algorithm was much lower when the metabolic information was chi-square normalized (0.58[0.50-0.63] with 12 chi-transformed metabolites). The RF and LGR algorithms achieved maximum accuracies with median of 0.83[0.81, 0.92] when using 613 LRs in RF, and 0.75[0.33-0.75] when using CLR in a stepwise LGR. Subsequently, a PLS model was fitted to explain *Ve* based on LR, CLR, or ALR variables, yielding accuracies (mean and IQR) of r = 0.57[0.44-0.69] (RMSE = 2.02[1.70-2.51]) using 390 LRs; r = 0.33[0.13-0.53] (RMSE = 2.34[1.74-2.84]) using 22 CLRs; and r = 0.08[0.02-0.28] (RMSE = 2.20[1.78-2.72]) using 21 ALRs. The combined approach reduced the set of selected of metabolic variables to 167 LRs, 7 CLRs, and 5 ALRs. Of these, only 17 LRs, 1 ALR and 1 CLR were found in generation 13, and only 2 LRs (*creatine/*($C_{10}H_8N_8O_2$; molecular mass 67.9875 Da) and *3-methylglutarylcarnitine/N-(3-acetamidopropyl)pyrrolidin-2-one* [3-MGC/NAAP]), exhibited consistent differences between lines. Interestingly, the magnitude of the differences between lines increased across generations in both lines (Figure 1B).

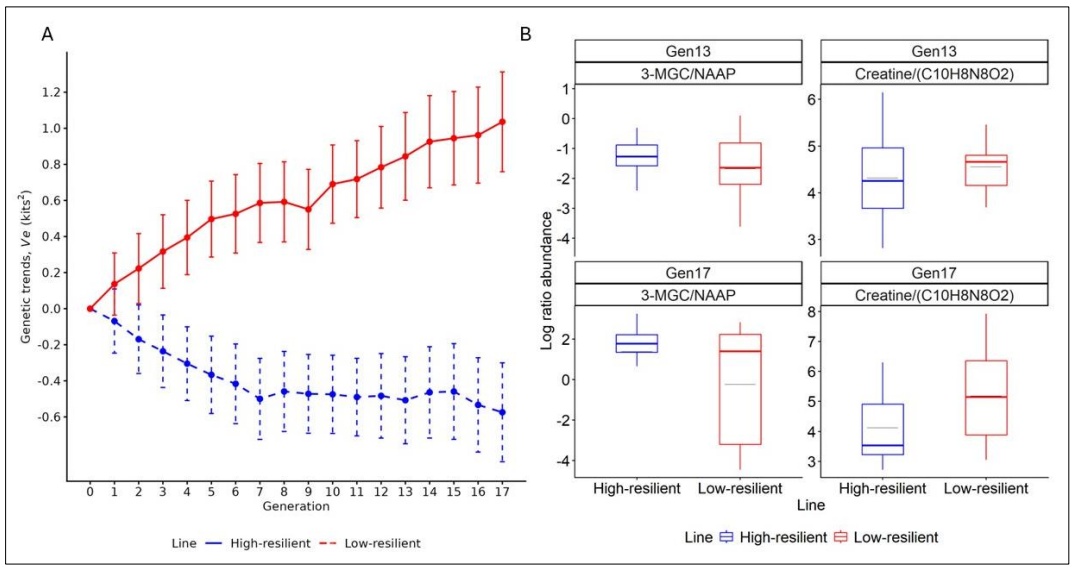

Figure 1. Response to selection for residual variance of litter size (*Ve*) across 17 generations, showing genetic means and standard deviations of posterior distribution of genetic means (A), and changes in consistent metabolic biomarkers across generations in high-resilient and low-resilient lines (B). Mean values are indicated by grey crossbars in each boxplot.

In low-resilient line, the mean relative abundance of *creatine/*($C_{10}H_8N_8O_2$) was higher (mean difference between low-resilient and high-resilient line, D_{L-H}=0.23 in generation 13; D_{L-H}=1.1 in generation 17), whereas *3-methylglutarylcarnitine/N-(3-acetamidopropyl)pyrrolidin-2-one* was lower (D_{L-H}= -0.42 in generation 13; D_{L-H}= -1.65 in generation 17), in comparison to the high-resilient line. While biological interpretation of individual log-ratios is complex, they provide a more robust representation of the systemic metabolic changes for a resilient response compared to absolute abundances. Each metabolite composing the log-ratios play roles related to biological pathways of energy metabolism and immune defence, components enabling the animals to respond quickly to the environmental perturbations. *Creatine* is important for cellular energy homeostasis through rapid ATP regeneration through phosphocreatine enabling muscles to cope up with sudden energy crisis and contributes to antioxidant defence (Nedeljkovic *et al.*, 2024). Higher *creatine*-related protein levels in low-resilient line may therefore reflect the increased energetic costs, consistent with increased corticotrope and inflammatory responses in the line (Argente *et al.*, 2019). *3-methylglutarylcarnitine*, belonging to carnitines, is involved in fatty acid transport and mitochondrial β-oxidation, thereby supporting energy production (Hoppel, 2003; Houten *et al.*, 2020). The other metabolite, *N-(3-acetamidopropyl)pyrrolidin-2-one* (a polyamine catabolite) is involved in cellular proliferation and immune regulation (Puleston *et al.*, 2021).

This study suggests that certain metabolomic variables as log-ratios are valuable sources of biological information for resilience to postpartum stress. However, the small sample size and lack of phenotypic data for generation 13 limited full cross-generational validation and the assessment of incorporating metabolomic data into genomic predictions.

Conclusions

To our knowledge, this was the first cross-generational metabolomic study addressing genetic resilience. The key biomarkers, *creatine*/$C_{10}H_8N_8O_2$ and *3-methylglutarylcarnitine*/*N-(3-acetamidopropyl)pyrrolidin-2-one,* showed consistent changes in magnitude and direction across generations and were implicated in energy metabolism, and immune regulation. Selection for residual variance of litter size produced correlated responses in the metabolic physiology associated with stress resilience.

References

Argente M.J., García M.L., Zbyňovská K., et al. (2019) Anim. 13:2348–2355. doi.org/10.1017/S1751731119000302

Beloumi D., Blasco A., Muelas R., et al. (2020) Animals 10:1540. doi.org/10.3390/ani10091540

Blasco A. (2017). Bayesian data analysis for animal scientists. Springer International Publishing, New York, USA.

Blasco A., Martínez-Álvaro M., García M.L., et al. (2017) Genet. Sel. Evol. 49:48. doi.org/10.1186/s12711-017-0323-4

Colditz I.G., and Hine B.C. (2016). Anim. Prod. Sci. 56:1961-1983. doi.org/10.1071/AN15297

Greenacre M., Martínez-Álvaro M., and Blasco A. (2021) Front. Microbiol. 12:727398. doi.org/10.3389/fmicb.2021.727398

Hoppel C. (2003) Am. J. Kidney Dis. 41:S4–S12. doi.org/10.1016/S0272-6386(03)00112-4

Houten S.M., Wanders R.J.A., and Ranea-Robles P. (2020) Biochem. Biophys. Acta Mol. Basis Dis. 1866:165720. doi.org/10.1016/j.bbadis.2020.165720

Misztal I., Tsuruta S., Lourenco D.A.L., et al. (2014). Manual for BLUPF90 family programs. University of Georgia. Available at: http://nce.ads.uga.edu/wiki/doku.php?id=documentation, visited on: 23-04-2025.

Nedeljkovic D.D., and Ostojic S.M. (2024) Clin. Bioenerg. 1:2. doi.org/10.3390/clinbioenerg1010002

Puleston D.J., Baixauli F., Sanin D.E., et al. (2021) Cell 184(16):4186–4202. doi.org/10.1016/j.cell.2021.06.007

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/, visited on: 23-04-2025.