Bovo, S., Bolner, M., Holl. J., Valente, B., Lewis, C., Schiavo, G., Bertolini, F., Fontanesi, L. (2026) Integrative metabolomic-genomic analyses uncover key regulators of metabolism in growing Duroc male pigs, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23755: https://doi.org/10.31274/wcgalp.23755

Integrative metabolomic-genomic analyses uncover key regulators of metabolism in growing Duroc male pigs

S. Bovo[1*], M. Bolner[1], J. Holl[2], B. Valente[2], C. Lewis[3], G. Schiavo[1], F. Bertolini[1], and L. Fontanesi[1]

[1]University of Bologna, Bologna, Italy; [2]PIC North America - Genus, Hendersonville, United States of America; [3]PIC Europe - Genus, Barcelona, Spain; [*]Corresponding author: samuele.bovo@unibo.it

Introduction

Pig breeding plans are increasingly looking for novel phenotypes that go beyond traditional production traits to enhance selection strategies. In this context, phenomics approaches have emerged to allow for large-scale phenotypic characterization through digital technologies (e.g., sensors and video/imaging systems) as well as high-throughput molecular profiling (Koltes *et al.,* 2019). Among these, molecular phenotyping is particularly valuable because it not only provides new informative traits but also helps in understanding the biological mechanisms behind complex traits. Metabolomics is a key molecular approach as it enables the comprehensive profiling of small molecules that define metabolic pathways and allows for the definition of metabotypes (Fontanesi, 2016). Furthermore, metabolomic traits are generally more closely related to the genomic space than complex phenotypes, and exhibit moderate to high heritability (Bovo *et al.,* 2025a). While metabolite genome-wide association studies (mGWAS) have been widely used in human populations, their application in livestock is still limited. In pigs, only a few studies have explored the genetic architecture of the metabolome, focusing on different breeds or physiological conditions (e.g. Ponsuksili *et al.,* 2025; Bovo *et al.,* 2025a). In this study, we combined high-throughput plasma metabolomic profiles and genotyping data from growing Duroc male pigs using an mGWAS approach to identify genetic regulators of pig metabolism.

Materials & Methods

Animals, blood collection and plasma preparation. The study included 274 male Duroc pigs sampled in two batches (October and November 2023). Pigs were selected from the extremes of feed intake in a larger population. Feed intake was recorded from 62-71 to 91-133 days of age; blood was collected at 91-133 days of age. All pigs were raised under identical conditions. Blood was collected into EDTA tubes, processed for DNA extraction, and plasma metabolomics. Pigs were maintained in compliance with European legislation for pig production. The procedures described adhered to the regulations set forth by US law and the European Union regarding animal welfare. The pigs were not specifically bred or treated for the purposes of this research, thus no further ethical statement is necessary.

Genotyping. Genomic DNA was extracted from whole blood using the Wizard Genomic DNA Purification Kit and genotyping was carried out with a medium-density SNP panel (40K). Quality control was conducted using PLINK v1.09 (Chang *et al.,* 2015), which involved excluding SNPs with a call rate < 0.90, minor allele frequency < 0.05, Hardy-Weinberg disequilibrium ($P < 0.0001$) and non-autosomal markers. After filtering, 33,285 markers were retained.

Metabolomic profiling and data processing. Untargeted plasma metabolomics was performed using the Metabolon HD4 platform. A total of 776 metabolites across different metabolic super-pathways were identified, with xenobiotics (n = 70) being excluded. Data quality control followed established procedures (Bovo *et al.,* 2025b), including outlier detection, filtering and imputation. The data were Box-Cox normalized and adjusted for confounders via linear models (age, sampling date, room) (Bovo *et al.,* 2025a). The final dataset comprised 274 animals × 592 metabolites. Analyses were run in R v.4.2.3 (R Core Team, 2022).

mGWAS, peak annotation and heritability. Genome-wide association analyses of adjusted metabolomic traits were conducted using linear mixed models as implemented in GEMMA v.0.94.1 (Zhou and Stephens, 2014). Significance was assessed using Wald tests, with multiple-testing correction applied using a Bonferroni threshold. Significant loci were annotated using the Sscrofa11.1 genome and the functional relevance of genes was evaluated through an analysis of the scientific literature, the Gene Cards database (https://www.genecards.org/) and known metabolite-gene associations retrieved from the GWAS Catalog (https://www.ebi.ac.uk/gwas/), the KEGG (https://www.genome.jp/kegg/) and the HMDB (https://www.hmdb.ca/) databases. Genomic heritability (h^2_G) was estimated from the mixed models, and the percentage of variance in phenotypes explained by a given marker was computed as described by Shim *et al.* (2015).

Results

Description of the plasma metabolome. The plasma metabolome consisted of 776 metabolites: 627 (81%) were of endogenous origin, 70 (9%) were of xenobiotic origin, and 79 (10%) were unnamed metabolites. Endogenous metabolites covered eight super-pathways, with lipids (49%) and amino acids (30%), being the most dominant. Following quality control, 35 metabolites were excluded, leaving a total of 592 metabolites for genomic analysis (Table 1).

Table 1. Summary statistics of the analysed metabolites.

Super-pathway	Heritability		mGWAS	
	Metabolites (no.)[1]	$h^2_G \pm$ std[2]	Metabolites (no.)[3]	mQTLs (no.)[4]
Peptide	22	0.35 ± 0.22	13	22
Cofactors and Vitamins	18	0.17 ± 0.18	3	4
Amino Acid	170	0.17 ± 0.14	31	33
Unnamed	54	0.15 ± 0.14	8	9
Lipid	254	0.15 ± 0.13	36	38
Partially Characterized Molecules	4	0.13 ± 0.13	1	1
Nucleotide	40	0.12 ± 0.14	4	4
Carbohydrate	23	0.10 ± 0.11	5	5
Energy	7	0.07± 0.14	0	0

[1] Number of metabolites retained after the quality control and entering the analyses
[2] Average and standard deviation of genomic heritability estimates.
[3] Number of significantly associated metabolites.
[4] Number of top SNP-metabolite associations.

Heritability estimates of plasma metabolites. Genomic heritability estimates have revealed a significant genetic contribution to the variability of plasma metabolites (Table 1). The

estimates ranged from 0.0 to 0.83 ± 0.08, with an average of 0.16 (SD = 0.14). Overall, approximately 60% of the metabolome exhibited $h^2_G \geq 0.10$ and 32% $h^2_G \geq 0.20$. Peptides displayed the highest genomic heritability, followed by cofactors and vitamins, amino acids, and lipids. At the sub-pathway level, fatty acid-related amino metabolites, hexosylceramides, and gamma-glutamyl amino acids showed the most significant genetic control.

mGWAS results. Association analyses identified significant SNP-metabolite associations across all autosomes except *Sus scrofa* chromosome (SSC) 11. The association signals were sharp and well defined, as shown in Figure 1. In total, 116 peaks of associations were detected for 101 metabolites (15.0% of those analyzed; Table 1), with most metabolomic traits showing a single metabolite quantitative trait locus (mQTL). Two major regions on SSC2 and SSC14 were associated with multiple metabolites across different pathways, suggesting pleiotropic regulation. The top associated SNPs explained, on average, 15.34 ± 8.38% of the phenotypic variance, with values ranging from 8.1% to 43.5%.

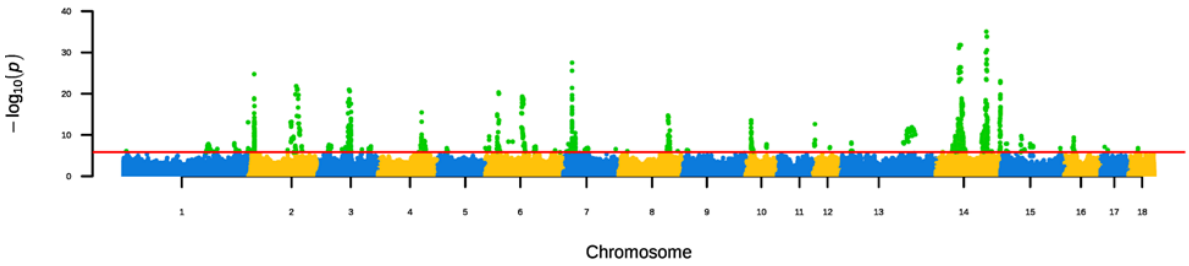

Figure 1. Joint Manhattan plot showing association results for metabolites. Each dot represents a SNP. The red horizontal line indicates the Bonferroni significance threshold ($P = 1.50 \times 10^{-06}$), and significantly associated SNPs are highlighted in green.

Functional interpretation of mQTLs. Functional annotation of the 116 significant SNP-metabolite associations revealed several overlaps with human metabolomic-genomic results. For about 60% of these associations, mQTLs annotated candidate genes coding for enzymes, metabolite transporters, or regulatory proteins directly involved in the metabolism of the associated compounds. Several mQTLs influenced multiple metabolites within and across metabolic pathways, enabling the transfer of functional annotations among related molecules. For the remaining 40% of the associations, no biologically plausible candidate genes were identified. The lipid super-pathway accounted for the largest number of mQTLs, often involving loci affecting multiple metabolites. For example, lipid-related mQTLs included the Sphingomyelin Phosphodiesterase 3 (*SMPD3*) and the Lysophosphatidylcholine Acyltransferase 2 (*LPCAT2*) genes. Several other associations, identified in humans, were confirmed also in pigs, including the Serpin Family A Member 6 (*SERPINA6*) gene that was associated with the level of cortisol. Within the tryptophan metabolism pathway, we identified associations with kynurenine-related metabolites, including the Kynurenine 3-Monooxygenase (*KMO*) gene. For carbohydrates, cofactors, nucleotides, and peptides, several associations replicated known human mQTLs, such as the Gamma-Glutamyltransferase 1 (*GGT1*) gene with several gamma-glutamyl amino acids.

Discussion

This study represents one of the largest metabolomic-genomic investigations conducted to date in pigs, in terms of the number of known metabolites analyzed per animal. When

considering the level of heritability, results confirmed that most plasma metabolites are genetically regulated intermediate phenotypes, further supporting the concept that metabolites are closer to the genomic space than complex production traits (Bovo *et al.,* 2025). The mGWAS results also revealed a relatively simple genetic architecture for many metabolites, showing that a limited number of key regulatory loci play a major role in shaping the anabolic and catabolic processes related to specific metabolites. In addition, the identification of pleiotropic regions affecting multiple metabolites across different pathways supports the existence of coordinated genetic regulation of interconnected metabolisms. Among the most relevant findings, mQTLs of the tryptophan-kynurenine pathway deserved particular attention. Our previous study in a Duroc population identified *KMO* as a major gene affecting kynurenine levels (Bovo *et al.,* 2025). In the Duroc line investigated here, kynurenine was associated not only with the *KMO* gene but also with other loci, suggesting a population-specific genetic control of this pathway. This result is particularly relevant given the central role of kynurenine in modulating immune responses and its involvement in the regulation of reproductive and central nervous system functions (Savits *et al.,* 2020).

Conclusions

This integrative analysis identified key genetic regulators of plasma metabolites in growing Duroc pigs. From an applied perspective, these identified metabolites are highly relevant for breeding strategies, as they may serve as biomarkers of metabolic efficiency, robustness, and responsiveness to nutritional interventions.

Acknowledgements

This study received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101059609 (Re-Livestock project).

References

Bovo S., Ribani A., Fanelli F., *et al.* (2025a) Genet Sel Evol 57:11. doi.org/10.1186/s12711-025-00960-8
Bovo S., Bolner M., Schiavo G., *et al.* (2025b) Animal 19:101393. doi.org/10.1016/j.animal.2024.101393
Chang C.C., Chow C.C., Tellier L.C., *et al.* (2015) Gigascience 4:7. doi.org/10.1186/s13742-015-0047-8
Fontanesi L. (2016) Anim Front 6:73-79. doi.org/10.2527/af.2016-0011
Koltes J.E., Cole J.B., Clemmens R., *et al.* (2019) Front Genet 10:1197. doi.org/10.3389/fgene.2019.01197
Ponsuksili S., Murani E., Fuchs B., *et al.* (2025) Open Biol 15:240285. doi.org/10.1098/rsob.240285
R Core Team (2022) R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Savitz J. (2020) Mol Psychiatry 25:131-147. doi.org/10.1038/s41380-019-0414-4
Shim H., Chasman D.I., Smith J.D., *et al.* (2015) PLoS One 10:e0120758. doi.org/10.1371/journal.pone.0120758
Zhou X., and Stephens M. (2014) Nat Methods 11(4):407-409. doi.org/10.1038/nmeth.2848