Computational Molecular Biology 2025, Vol.15, No.5, 235-244
http://bioscipublisher.com/index.php/cmb

Research Insight **Open Access**

Integrating Multi-Omics Data to Explore the Genetic Basis of Milk Production in Dairy Cattle

Jingya Li, Jun Li
Animal Science Research Center, Cuixi Academy of Biotechnology, Zhuji, 311800, Zhejiang, China
Corresponding author: jun.li@cuixi.org

Computational Molecular Biology, 2025, Vol.15, No.5 doi: 10.5376/cmb.2025.15.0023
Received: 19 Jul., 2025
Accepted: 31 Aug., 2025
Published: 23 Sep., 2025

Preferred citation for this article:
Li J.Y., and Li J., 2025, Integrating multi-omics data to explore the genetic basis of milk production in dairy cattle, Computational Molecular Biology, 15(5): 235-244 (doi: 10.5376/cmb.2025.15.0023)

Abstract The milk production trait of dairy cows is a complex quantitative trait with high economic value. With the development of high-throughput technology, integrating multi-omics data such as genomics, transcriptomics, epigenomics, proteomics and metabolomics has become an important means to reveal the genetic basis of milk production traits. This study, through a multi-omics integration strategy, explored the core genes and regulatory networks that affect milk production and its related traits, reviewed the main milk production traits and their related quantitative trait loci (QTLS), systematically sorted out the characteristics of various omics data and their integrated analysis methods, such as weighted gene co-expression network analysis (WGCNA) and multi-omics factor analysis (MOFA), the focus is on key genes (such as *STAT5A*, *CSN2*) in the breast development and lactation regulatory pathways, the interaction network between miRNA and mRNA, as well as functional enrichment pathways. This study demonstrates the potential of multi-omics integration in analyzing complex traits, providing a scientific basis for promoting the intelligence and efficiency of molecular breeding in dairy cows, and also offering ideas for future functional gene verification and optimization of precise breeding strategies.

Keywords Dairy cows; Milk production characteristics; Multi-omics integration; Gene regulatory network; Candidate gene identification

1 Introduction

The fact that cows produce milk sounds simple, but the things involved behind it are actually quite complicated. It is not only related to how much money a ranch can earn in a year, but also to the stable development of the global dairy industry and even how we understand the complex traits of mammals (Brito et al., 2021; Ristanic et al., 2024). Making milk more abundant and of better quality has long been the goal of almost all breeding projects. But things are not that easy. Milk production is essentially a polygenic trait, not only controlled by hundreds or even thousands of genes, but also influenced by environmental fluctuations. In the past, most breeding relied on "visible and tangible" phenotypes and family trees to predict the quality of genes. Indeed, this method has greatly increased the overall yield of dairy cows, but there are also many problems, such as low heritability, slow intergenerational turnover, and the most troublesome ones are those genes with small effects and vulnerable to external influences, which are often overlooked.

The wind direction has started to change in recent years. The popularization of high-throughput omics - from the genome to the transcriptome, proteome, and then to epigenetics and metabolomics - has brought about new breakthroughs in the study of milk production. Rather than focusing on a single data layer, it is better to conduct multiple omics simultaneously. Only in this way can those variant loci, regulatory switches and pathway networks that truly affect milk production be identified (Nasir et al., 2024; Cai et al., 2025). Some analyses have shown that not only structural genes, but also non-coding Rnas specifically expressed in the mammary gland, epigenetic modifications, and even larger regulatory networks all have a say in milk production and component control. If all these omics layer data are integrated, not only can milk production be predicted more accurately, but the improvement direction will also be clearer (Zheng et al., 2025).

This study aims to establish a robust framework for integrating multi-omics data to analyze the genetic and molecular mechanisms of milk production in dairy cows, identify key genetic variations and regulatory elements related to milk yield and milk composition, clarify the biological pathways and cell types involved in milk synthesis, and evaluate the potential of multi-omics integration in enhancing the reliability of genomic selection. By bridging the gap between genotypes and phenotypes, this study aims to significantly accelerate genetic improvement, optimize breeding strategies, and ensure the long-term sustainability of dairy cow production systems.

2 Overview of the Genetic Basis of Milk Production Traits in Dairy Cattle

2.1 Major milk production traits and their genetic parameters

In terms of milk production by dairy cows, indicators such as milk yield (MY), milk fat production (FY), and milk protein production (PY) have long been studied as the main traits. Milk fat percentage and milk protein percentage are also often included in the improvement targets. Although these traits are regulated by multiple genes, their heritability is not uniform. For example, the estimated value of MY is usually between 0.12 and 0.35, and the same is true for FY and PY. However, the heritability of milk fat percentage and milk protein percentage is generally lower, usually within the range of 0.06 to 0.16 (Lu et al., 2022). Interestingly, these yield-related traits are often positively correlated with each other, but when yield and percentage traits are considered together, a negative correlation sometimes occurs. Incidentally, MY repeatability reaches 0.47, that is to say, although genes play a role, the influence of the environment is also considerable (Fotso-Kenmogne et al., 2025).

2.2 Key genes and QTL regions influencing milk yield

Many people think that "high milk production" depends on the "average strength" of the entire genome. But in fact, key genes like *DGAT1*, which are frequently mentioned, show stable effects on milk production (MY), milk fat (FY), and milk protein (PY), especially on chromosome 14 (Kim et al., 2021; Tahir et al., 2025). However, do not overlook other genes, such as *MGST1*, *GHR*, *ABCG2*, *CSN1S1*, as well as the newly emerged candidate genes in recent years, like *PDE4B*, *ANO2* and *NCOA6*. Nowadays, many QTLS have almost been labeled on all autosomes, especially the bunch on chromosome 14 that is closely related to milk fat production and fat percentage (Bekele et al., 2023; Cortes-Hernandez et al., 2025). Meanwhile, multi-omics data began to come into play. For instance, when mammary mammary specific expression, non-coding RNA, regulatory factors and epigenetic modifications were all taken into account, it was found that more key variations could be identified and the prediction accuracy improved significantly (Cai et al., 2025; Križanac et al., 2025).

2.3 Interaction between genetic background and environmental factors on milk performance

Ultimately, genes are not the only variable. The performance of cattle often depends on the weather, feeding conditions and the nutrition of the feed. The interaction between genotype and environment (GxE) may lead to different manifestations of the same genotype under different conditions, and even the estimated heritability may vary as a result (Mancin et al., 2023). For instance, in hot or malnourished environments, the heritability of MY may be lower than expected, while certain genotypes may perform better under these conditions (Martins et al., 2025). In addition, the traits of reproductive capacity and calving interval themselves have low heritability and are highly sensitive to management impacts. While "survival traits" such as stress resistance and heat resistance are not necessarily directly related to milk production, they are now also included in the selection index to improve adaptability (Oloo et al., 2025). Overall, if the GxE effect is not taken into account, the selected cattle may have "good paper performance", but they will "fail" once the environment changes.

3 Types and Characteristics of Multi-Omics Data

3.1 Genomic data (SNP, QTL, GWAS)

Genetic variations related to milk production were first explored through genomic data. In fact, not only SNPS in the coding region, but also in many non-coding and regulatory regions can be detected through high-throughput typing and whole-genome sequencing (Buaban et al., 2021; Ristanic et al., 2024; Venkatesan et al., 2025). It is not surprising that genes like *DGAT1*, *GHR* and *CTNNA3*, which are "familiar faces", have been repeatedly captured in GWAS or QTL analyses of different varieties. However, to more precisely identify these sites, a sufficient number of samples are needed and the annotations must be detailed. Especially for those variations located in regulatory regions with unknown functions, it is even more necessary to integrate functional information such as mammary specific genes, expression QTL, and splicing QTL in order to possibly improve the prediction accuracy (Križanac et al., 2025).

3.2 Transcriptomic data (mRNA, lncRNA, miRNA)

When it comes to the gene regulation related to lactation, relying solely on genomic information is clearly insufficient. At the transcriptome level, mRNA is undoubtedly the main character, but non-coding molecules such as lncRNA and miRNA should not be ignored either. Between high-yield dairy cows and low-yield dairy cows, differentially expressed genes (DEGs) are often enriched in pathways such as immunity, metabolism, and mammary gland function (Nguyen, 2025). Some specific lncrnas and mirnas are also considered to possibly be key nodes regulating milk fat and milk protein synthesis (Shin et al., 2025). Interestingly, the expression patterns of these non-coding Rnas often coincide with GWAS signals and QTL regions, which also indicates that they play a more important role in the regulation of milk production traits than imagined.

3.3 Functional information from epigenomic, metabolomic, and proteomic data

Not all variations related to milk production are written on the DNA sequence; some "hidden" regulations come from the epigenetic level. For instance, epigenetic markers such as DNA methylation or histone modifications (like H3K27ac, H3K4Me1) can be associated with milk production traits in breast tissue (Figure 1) (Dong et al., 2021; Cai et al., 2025). When it comes to the metabolome, milk production is essentially an energy-intensive process. The changes in metabolites along the energy metabolism or amino acid synthesis pathways often vary significantly between high-yield and low-yield dairy cows. Proteomics is similar. It can reveal the key proteins involved in lactation and even the signals of certain post-translational modifications. These seemingly scattered data layers may have limited significance when viewed separately, but once they are analyzed in series with the genome and transcriptome, they often lead to a more comprehensive understanding and also provide support for precision breeding (Zhang and Lin, 2025).

4 Integration Strategies and Analytical Methods for Multi-Omics Data

4.1 Data normalization and quality control procedures

High-throughput platforms have a large amount of data, diverse types, and inconsistent data types. If you want to analyze these pieces of information from different "omics" together, the initial quality control and standardization steps are basically unavoidable. After receiving the sequencing read segments, the first steps are cleaning, comparison, and feature screening - these processes may sound highly technical, but in fact, they are all aimed at ensuring comparability among different samples. In data such as metagenomic and metabolomics, commonly used methods include accumulation and scaling, quantile normalization, etc., which can calibrate the deviations brought by batches (Ravelo et al., 2024). However, standardization is not as easy as simply aligning formats. Problems such as false positives, duplicate signals, and file structure differences are sometimes not sufficient to be automatically handled by software alone. When integrating across omics, if this step is not handled cleanly, no matter how fancy the subsequent analysis results are, they will not be reliable.

4.2 Multi-omics correlation and network modeling approaches

When it comes to how to view data from different omics groups together, the set of network modeling comes in handy. Methods like WGCNA are not a recent trend; they have been widely used in multi-omics research for several years. It can cluster those gene, metabolite or microbial data into modules and then link them to specific phenotypes (such as milk production) (Zhang et al., 2023). The advantage of this kind of scale-free network analysis lies in its ability to clearly see the overall structure, not just focusing on a single gene. Integration frameworks like MOFA are also following suit. They not only look at common patterns but also pay attention to the sources of differences specific to each discipline (Wang et al., 2024). These models are more like "interpreting" the relationships between omics rather than simply piecing together a jigsaw puzzle.

4.3 Phenotype association analysis using machine learning and AI techniques

The integration of AI into omics is actually no longer a novelty. Whether it is milk production, milk fat ratio or protein level, it is always somewhat difficult to predict these traits with traditional statistical methods. So, machine learning began to take the lead. Regression methods like ridge regression and LASSO, or ensemble models such as random forests and lift algorithms, often outperform old-school models in predicting milk production phenotypes. Although neural networks and deep learning are complex, they perform well in high-dimensional data analysis, especially in feature selection, and can screen out which SNPS, genes or metabolites have the greatest impact on traits (Frizzarin et al., 2021). Of course, it's not the case that the more complex the model is, the better. The key is to ensure that the results are stable. Therefore, if multi-omics data can be reasonably combined with AI methods, it is not difficult to make stable phenotypic predictions and precise breeding decisions.

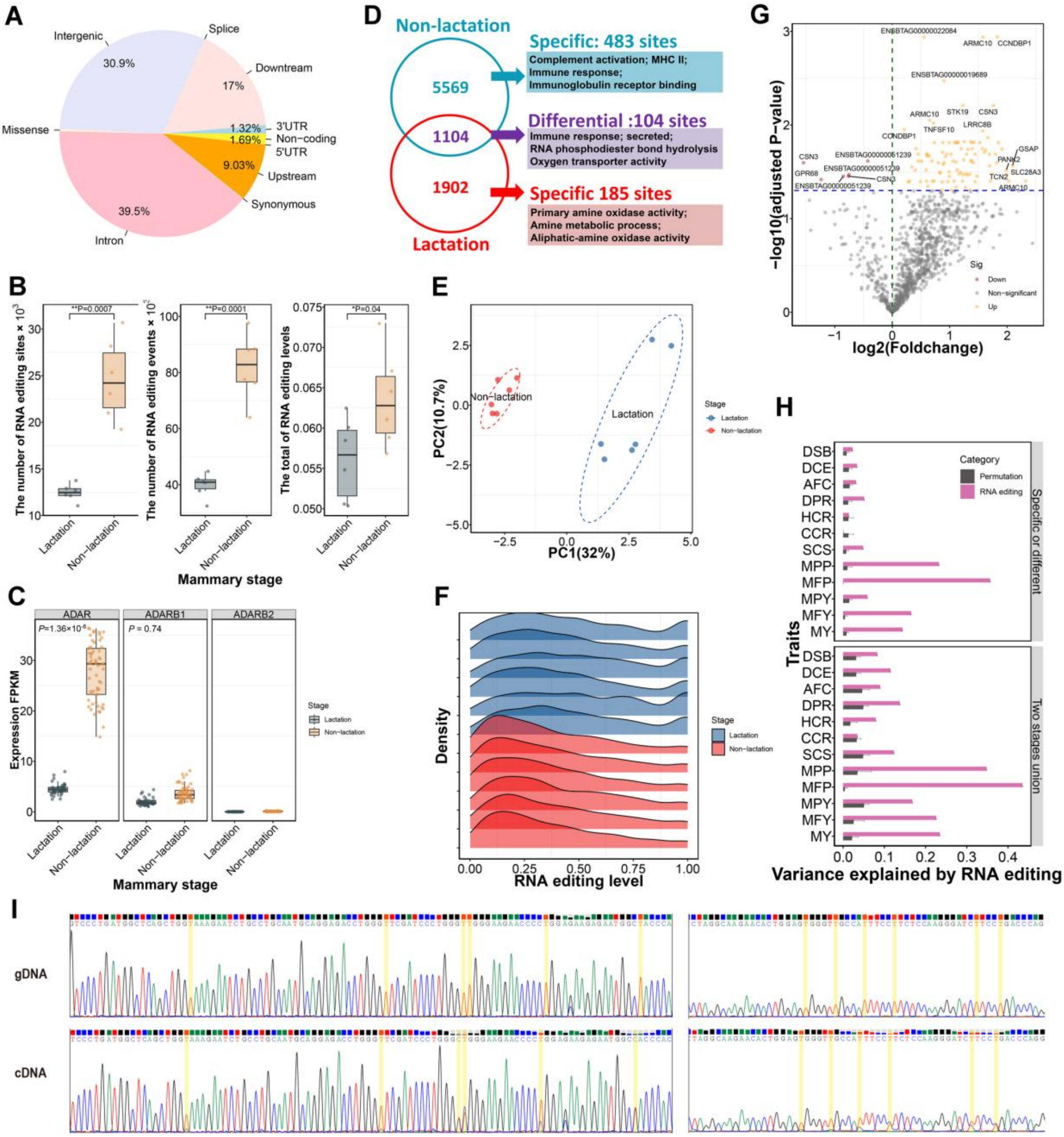

Figure 1 The proportion of variance explained by lactation associated RNA editing sites (Adopted from Cai et al., 2025)

5 Identification of Key Pathways and Regulatory Networks

5.1 Signaling pathways in mammary gland development and lactation regulation

During the process of breast development and lactation, more than one signaling pathway is at play. Pathways such as PI3K-Akt, AMPK, insulin and PPAR are often simultaneously involved in cell proliferation, fat synthesis and the regulation of milk secretion. For instance, COL6A1 and NFKB2 have been found to be closely related to milk fat synthesis, and they exert regulatory functions on mammary epithelial cells through these pathways (Han et al., 2025). However, the role of the JAK-STAT pathway cannot be ignored. It is usually activated under the action of prolactin or certain cytokines and plays a key role in the immune regulation and lactation function of the breast. As for pathways such as MAPK, mTOR, Rap1 and TGF-β, they are more involved in milk protein synthesis, cell cycle control, and even reflected in breast structure adjustment (Zheng et al., 2025). Not every pathway plays the same role at all stages. Such regulation is mostly phased and cross-cutting.

5.2 Identification of core regulatory factors in gene networks

In the complex gene expression network, those that truly play a "key role" may only be in the minority. Genes such as *GAPDH*, *KDR*, *CSF1*, *PYGM*, *RET*, *PPP2CA*, *GUSB* and *PRKCA* are frequently named in the lactation-related pathways (Roudbari et al., 2023). Some transcription factors - such as HHEX, HOXA7, NOTCH3/4 and members of the SMAD family - usually appear at the core position in the miRNA and mRNA co-expression network. They affect not only milk production but also indicators such as lactose content and somatic cell count. These regulatory factors do not act alone and are also strongly associated with the expression of mammary mammary specific genes (such as *DGAT1*, *GLYCAM1*) and certain specific lncrnas (Cai et al., 2025). Network analysis merely offers a starting point; more verification still relies on functional experiments.

5.3 Interaction patterns between miRNAs, mRNAs, and metabolites

Sometimes, it is difficult to explain the mechanism of lactation regulation by looking at a single gene or pathway, especially when it involves the interaction of multiple molecular layers. Mirnas such as miR-148a, miR-186, miR-200a and miR-152 often affect fat synthesis, milk production and breast development through signaling pathways such as PTEN, PI3K-Akt or PPAR (Xia et al., 2021; Li et al., 2024). WGCNA et al. Co-expression analysis methods have identified miRNA modules closely related to milk volume, protein, milk fat and even lactose, as well as their target mrnas. However, it is not only miRNA but also lncRNA and circRNA that are frequently involved in regulation. ceRNA networks such as lincRNA TCONS_00082721 combined with FABP4 or circ11103/miR-128/PPARGC1A are structures that cannot be avoided when studying milk fat metabolism. These interaction patterns are not necessarily all direct; sometimes they are chain-like or feedback-like, and as a whole, they demonstrate the high complexity of the lactation regulation mechanism.

6 Case Study: Multi-Omics Integration Reveals Key Genes for Milk Production

6.1 Multi-omics sample collection and analysis during lactation in Holstein cattle

Omics analysis does not start from a single perspective but requires repeated sampling from multiple tissues and at multiple time points. Tissues such as the mammary gland, liver and gastrointestinal tract are often collected both during lactation and non-lactation periods, and the collection involves more than just the transcriptome and genome. Epigenetic information such as DNA methylation, histone modification, and non-coding RNA is also crucial (Gao et al., 2024; Zheng et al., 2025). Especially in the large-scale sample study of Holstein dairy cows, researchers have dealt with millions of variant sites (Figure 2). Although these data are complex in hierarchy, the ultimate goal is the same - to identify which genes and regulatory elements play a leading role in the lactation stage, especially those candidate factors specifically expressed in the breast.

6.2 Omics comparison between high-and low-yielding dairy cattle groups

Sometimes, the difference between high and low milk production is not only reflected in the quantity, but also at the molecular level behind it. It can be seen from the comparison of transcriptome and metabolome that the pathways related to energy metabolism and amino acid synthesis are often upregulated in the mammary glands and livers of high-yield dairy cows. Conversely, those individuals with lower milk production were more active in immune pathways or certain splicing forms (Nguyen, 2025). Furthermore, some key metabolites, such as citrulline and N-acetylornithine, were significantly elevated in the high-yield group, which might imply stronger gluconeogenesis and energy utilization capabilities. However, not all differences can be directly attributed to the expression of a certain gene. It is necessary to consider the regulatory network together.

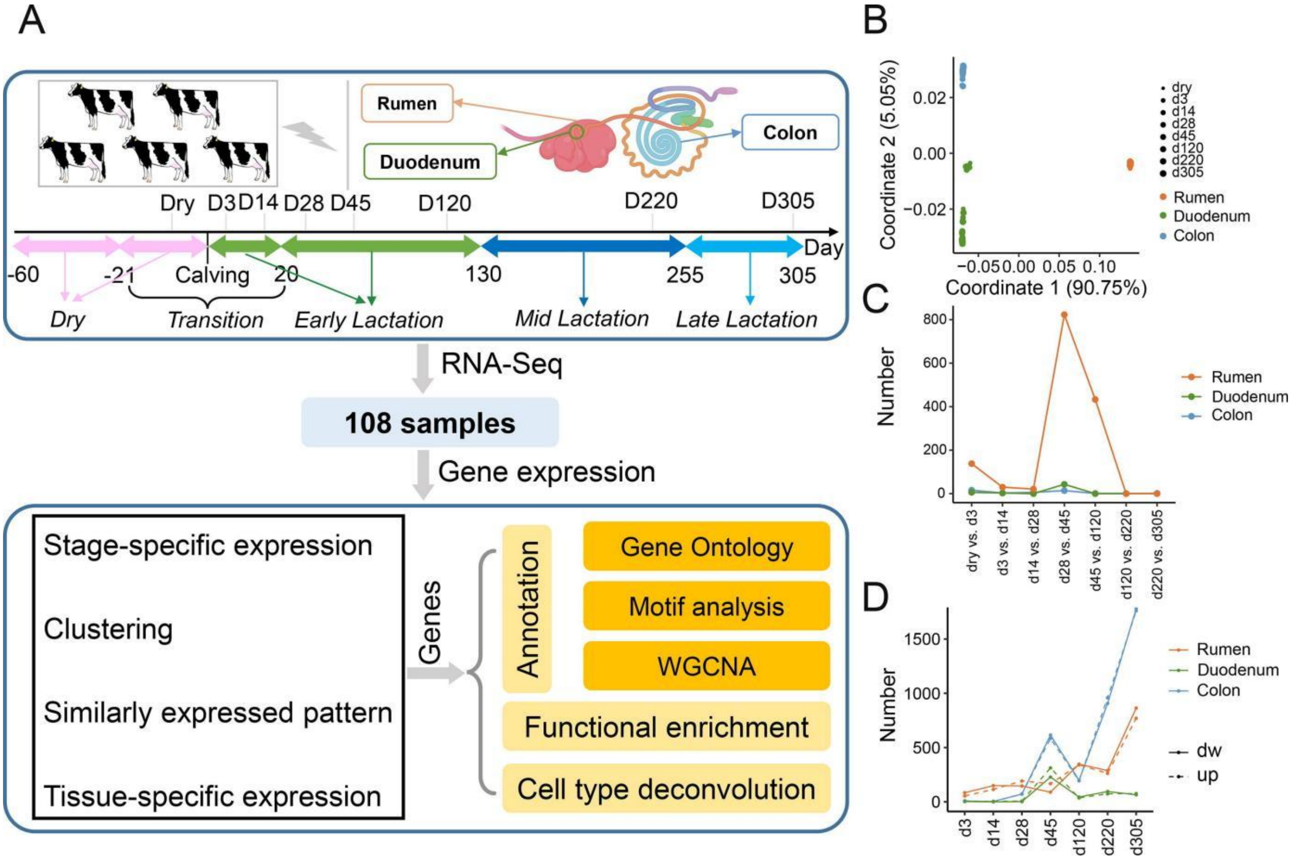

Figure 2 (A) Tissue collection and global analysis of the cattle transcriptomes. We explored expression specificity patterns by multiple analyses (differentially expressed gene analysis, stage-/tissue-specific expression, co-expression network analysis, functional enrichment, deconvolution analysis, etc.) using 108 samples from three gastrointestinal tissues (rumen, duodenum and colon) from eight lactation periods of dairy cattle. (B) The Principal Component Analysis (PCA) of three tissues across eight stages. Tissues are represented by different colors, whereas the size of the dots indicates different lactation stages. (C) The number of genes differentially expressed between adjacent stages for three tissues. (D) The number of genes differentially expressed between dry and other stages for three tissues (Adopted from Gao et al., 2024)

6.3 Identification and validation of candidate genes

To identify the core factors that truly influence milk production, relying solely on the omics level is far from sufficient. Genes related to milk protein synthesis, such as *CSN2* and *STAT5A*, have been involved in past studies (Zheng et al., 2025). But more confirmation was actually achieved through the superimposed analysis of GWAS, QTL mapping and RNA sequencing. For instance, classic genes such as *DGAT1*, *GHR*, and *ABCG2* have recurred in almost all studies related to milk production. Recently noticed ones such as *EFNA1*, *ERBB3* and *CIDEA* have only gradually demonstrated their potential in regulating milk components through the integration of multiple omics. Furthermore, variations such as those in the 5’utr region, splicing sites, and even mammary mammary specific enhancers are also considered to account for a considerable proportion of trait variations (Križanac et al., 2025). To verify the functions of these genes, it is generally necessary to examine whether they overlap with QTL, whether there is functional annotation support, and the changes in expression activity under different milk production backgrounds.

7 Application Prospects of Multi-Omics Information in Molecular Breeding

7.1 Value of candidate genes and biomarkers in genomic selection

Although traditional methods can capture the major effect sites, when dealing with complex traits, they often struggle to take into account changes in non-coding regions or regulatory levels. The emergence of multi-omics has precisely filled this gap. Once functional information such as genes specifically expressed in the breast, mirnas, differential expression patterns and even epigenetic markers is incorporated into genomic selection models, the prediction results will be more stable (Silpa et al., 2021). Not only that, these multi-layered validated biomarkers also provide a clear direction for marker-assisted selection, enabling breeding to no longer rely solely on phenotypes but to preferentially target those gene loci with a clear functional background (Wadood et al., 2025). Of course, some regulatory variations can easily be overlooked if only relying on traditional SNP models.

7.2 Construction of precision breeding models based on multi-omics support

Having data alone is not enough; the key lies in using it skillfully. Integrating the transcriptome, proteome and metabolome into breeding models is no longer a novelty, but the effectiveness still depends on the models themselves. Methods like MultiBLUP, BayesRC and kernel function have demonstrated higher predictive power than traditional genomic selection in multiple studies (Liang et al., 2022). The advantage of these models lies in their ability to handle network hierarchies, functional annotations, and even complex interactions between genes and the environment (Cai et al., 2025). In some groups, personalized breeding strategies have also gradually been put on the agenda, such as customized breeding of different varieties, or improving efficiency in combination with genome editing and assisted reproduction technologies (Xiong et al., 2025).

7.3 Data sharing across populations and development of breeding platforms

Scattered data and inconsistent standards are a major bottleneck restricting the wide application of multi-omics. Especially in small groups or non-mainstream varieties, information sparsity often directly affects the performance of the prediction model. For this reason, promoting cross-group data integration becomes particularly important. Multi-variety reference panels and open shared databases have been proven to significantly improve generalization ability (Jin et al., 2024). But to truly bring these data to life, it still relies on cloud platforms, open-source tools and unified formats. These technologies enable real-time connectivity of research data and make the transformation process from research to actual breeding smoother (Gong et al., 2022; Amin et al., 2025). As the platform architecture gradually matures, the threshold for multi-omics integration and meta-analysis is also decreasing.

8 Conclusion and Outlook

To understand how milk is produced and why its components are different, merely looking at genes is far from enough. Nowadays, it is generally agreed that multi-omics integration is a more reliable approach. From the genomic and transcriptomic levels all the way to epigenetics, proteins, and metabolism, it can dig deeper into the key points related to the lactation performance of dairy cows (for example, genes like *EFNA1*, *DGAT1*, and *GLYCAM1* are often not the "main characters" that can be seen at first glance). When the expression genes specific to the breast, the regulatory factors activated only during lactation, and even the epigenetic markers are all incorporated into the model, the estimation of breeding values is no longer a "guess". In other words, with precision breeding truly having a handle, there will be a solid basis for improving the quality and yield of milk.

However, one should not speak too confidently. Technology has indeed advanced, but when it comes to actually implementing it, multi-omics is not a "panacea". First of all, the data is mixed and the format is disordered. If not careful during integration, false positive results will pop up. Statistical analysis must keep up. As for the sample size, small-sample studies still account for a considerable proportion, and the truly differentially expressed genes that can be discovered are limited, which naturally reduces the generalisability. The lack of uniformity in sequencing platforms and the differences in experimental batches can be quite annoying due to batch effects, especially when comparing across teams and projects. Integration often fails to work together. Another practical issue: burning money. High-throughput omics technology is costly and requires supporting equipment and professional talents, which small breeding farms cannot afford. Not to mention that some omics levels have not been given due attention at all, such as metabolomics, proteomics and microbiome. Although they also have an impact on the composition of milk, the results are often overlooked.

How should the future be headed? Many people actually have a clear idea in their hearts, but to truly implement it, it still needs to be gradually advanced from several aspects. Functional verification of candidate genes and regulatory elements cannot be omitted, and neither wet laboratories nor animal models can be missing. Single-cell omics has also emerged in recent years. It enables us to observe the "division of labor" of different cells in the mammary gland under different conditions, which was unimaginable in the past. Another issue is the tools. The platform should not be too academic. It's better to be accessible to everyone for promotion, especially when collaborating across breeding units. Furthermore, the sample size still needs to be expanded, and multiple varieties should be studied simultaneously; otherwise, the conclusion will always be one-sided. Finally, artificial intelligence and the Internet of Things have entered the field. After the real-time cloud upload of breeding data, the efficiency of feedback and modeling has also improved to a new level. In the future, making dynamic adjustment breeding decisions will no longer be a dream. Multi-omics is not the end; it is merely a means. By integrating these "visible" technologies with "practical" breeding practices, future dairy cow breeding will not only focus on science but also on efficiency and sustainability.

Acknowledgments

We would like to express our gratitude to the reviewers for their valuable feedback, which helped improve the manuscript.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Amin A., Zaman W., and Park S., 2025, Harnessing multi-omics and predictive modeling for climate-resilient crop breeding: from genomes to fields, Genes, 16(7): 809.
https://doi.org/10.3390/genes16070809

Bekele R., Taye M., Abebe G., and Meseret S., 2023, Genomic regions and candidate genes associated with milk production traits in holstein and its crossbred cattle: a review, International Journal of Genomics, 2023: 8497453.
https://doi.org/10.1155/2023/8497453

Brito L., Bédère N., Douhard F., Oliveira H., Arnal M., Peñagaricano F., Schinckel A., Baes C., and Miglior F., 2021, Review: Genetic selection of high-yielding dairy cattle toward sustainable farming systems in a rapidly changing world, Animal, 15(Suppl 1): 100292.
https://doi.org/10.1016/j.animal.2021.100292

Buaban S., Lengnudum K., Boonkum W., and Phakdeedindan P., 2021, Genome-wide association study on milk production and somatic cell score for Thai dairy cattle using weighted single-step approach with random regression test-day model, Journal of Dairy Science, 105(1): 468-494.
https://doi.org/10.3168/jds.2020-19826

Cai W., Cole J., Goddard M., Li J., Zhang S., and Song J., 2025, Mammary gland multi-omics data reveals new genetic insights into milk production traits in dairy cattle, PLOS Genetics, 21(4): e1011675.
https://doi.org/10.1371/journal.pgen.1011675

Cortes-Hernández J., García-Ruiz A., Peñagaricano F., Montaldo H., and Ruiz-Lopez F., 2025, Uncovering the genetic basis of milk production traits in Mexican Holstein cattle based on individual markers and genomic windows, PLOS ONE, 20(2): e0314888.
https://doi.org/10.1371/journal.pone.0314888

Dong W., Yang J., Zhang Y., Liu S., Ning C., Ding X., Wang W., Zhang Y., Zhang Q., and Jiang L., 2021, Integrative analysis of genome-wide DNA methylation and gene expression profiles reveals important epigenetic genes related to milk production traits in dairy cattle, Journal of Animal Breeding and Genetics, 138(5): 562-573.
https://doi.org/10.1111/jbg.12530

Fotso-Kenmogne P., Carneiro P., Silva D., Cobuci J., Aponte P., Oliveira H., and Brito L., 2025, Genomic-based genetic parameters for daily milk yield and various lactation persistency traits in American Holstein cattle, Journal of Dairy Science, 108(7): 7329-7344.
https://doi.org/10.3168/jds.2024-25836

Frizzarin M., Gormley I., Berry D., Murphy T., Casa A., Lynch A., and McParland S., 2021, Predicting cow milk quality traits from routinely available milk spectra using statistical machine learning methods, Journal of Dairy Science, 104(7): 7438-7447.
https://doi.org/10.3168/jds.2020-19576

Gao Y., Liu G., Ma L., Fang L., Li C., and Baldwin R., 2024, Transcriptomic profiling of gastrointestinal tracts in dairy cattle during lactation reveals molecular adaptations for milk synthesis, Journal of Advanced Research, 71: 67-80.
https://doi.org/10.1016/j.jare.2024.06.020

Gong L., Lou Q., Yu C., Chen Y., Hong J., Wu W., Fan S., Chen L., and Liu C., 2022, GpemDB: a scalable database architecture with the multi-omics entity-relationship model to integrate heterogeneous big-data for precise crop breeding, Frontiers in Bioscience, 27(5): 159.
https://doi.org/10.31083/j.fbl2705159

Han B., Lin S., Ye W., Chen A., Liu Y., and Sun D., 2025, COL6A1 promotes milk production and fat synthesis through the PI3K-Akt/insulin/AMPK/PPAR signaling pathways in dairy cattle, International Journal of Molecular Sciences, 26(5): 2255.
https://doi.org/10.3390/ijms26052255

Jin L., Xu L., Jin H., Zhao S., Jia Y., Li J., and Hua J., 2024, Accuracy of genomic predictions cross populations with different linkage disequilibrium patterns, Genes, 15(11): 1419.
https://doi.org/10.3390/genes15111419

Kim S., Lim B., Cho J., Lee S., Dang C., Jeon J., Kim J., and Lee J., 2021, Genome-wide identification of candidate genes for milk production traits in Korean holstein cattle, Animals, 11(5): 1392.
https://doi.org/10.3390/ani11051392

Križanac A., Reimer C., Heise J., Liu Z., Pryce J., Bennewitz J., Thaller G., Falker-Gieske C., and Tetens J., 2025, Sequence-based GWAS in 180 000 German Holstein cattle reveals new candidate variants for milk production traits, Genetics Selection Evolution, 57: 3.
https://doi.org/10.1186/s12711-025-00951-9

Li X., Wu Y., Yang X., Gao R., Lu Q., Lü X., and Chen Z., 2024, The cooperative regulatory effect of the miRNA-130 family on milk fat metabolism in dairy cows, Animal Bioscience, 37: 1289-1302.
https://doi.org/10.5713/ab.23.0485

Liang M., An B., Chang T., Deng T., Du L., Li K., Cao S., Du Y., Xu L., Zhang L., Gao X., Li J., and Gao H., 2022, Incorporating kernelized multi-omics data improves the accuracy of genomic prediction, Journal of Animal Science and Biotechnology, 13: 95.
https://doi.org/10.1186/s40104-022-00756-6

Lu X., Arbab A., Abdalla I., Liu D., Zhang Z., Xu T., Su G., and Yang Z., 2022, Genetic parameter estimation and genome-wide association study-based loci identification of milk-related traits in Chinese Holstein, Frontiers in Genetics, 13: 799664.
https://doi.org/10.3389/fgene.2021.799664

Mancin E., Proto G., Tuliozi B., Schiavo G., Bovo S., Fontanesi L., Sartori C., and Mantovani R., 2023, Uncovering genetic parameters and environmental influences on fertility milk production and quality in *Autochthonous reggiana* cattle, Journal of Dairy Science, 107(2): 956-977.
https://doi.org/10.3168/jds.2022-23035

Martins R., Filho C., Valloto A., Carvalheiro R., De Albuquerque L., De Almeida Teixeira R., and Dias L., 2025, Influence of different environmental challenges on the expression of productive traits in Holstein cattle in the southern region of Brazil, Tropical Animal Health and Production, 57(4): 182.
https://doi.org/10.1007/s11250-025-04436-1

Nasir T., Rehman A., Tariq M., Kubra K., Ashraf N., and Nasir F., 2024, Role of multi-omics in functional genomic transcriptomic proteomic and metabolic mechanisms for milk production growth fertility and health in livestock animals, Zeugma Biological Science, 6(3): 38-46.
https://doi.org/10.55549/zbs.1680746

Nguyen D., 2025, Transcriptomic insights into gene expression differences between high and low milk producing dairy cows, International Journal of Environmental Sciences, 11(13s): 1021-1033.
https://doi.org/10.64252/9kj25r03

Oloo R., Mrode R., Ekine-Dzivenu C., Ojango J., Bennewitz J., Gebreyohanes G., Okeyo A., and Chagunda M., 2025, Genetic relationships among resilience fertility and milk production traits in crossbred dairy cows performing in sub-saharan Africa, Journal of Animal Breeding and Genetics, 142(6): 630-642.
https://doi.org/10.1111/jbg.12933

Ravelo A., Ferm P., Guo Y., Omontese B., Morley P., Chen C., Noyes N., and Caixeta L., 2024, Using a multi-omics approach to explore potential associations with rumen content and serum of cows with different milk production levels based on genomic predicted transmitting ability for milk and phenotypic milk production, PLOS ONE, 19(7): e0305674.
https://doi.org/10.1371/journal.pone.0305674

Ristanić M., Zorc M., Glavinić U., Stevanović J., Blagojevic J., Maletić M., and Stanimirović Z., 2024, Genome-wide analysis of milk production traits and selection signatures in serbian holstein-friesian cattle, Animals, 14(5): 669.
https://doi.org/10.3390/ani14050669

Roudbari Z., Mokhtari M., Gorji A., Sadkowski T., Sadr A., and Shirali M., 2023, Identification of hub genes and target miRNAs crucial for milk production in holstein friesian dairy cattle, Genes, 14(11): 2105.
https://doi.org/10.3390/genes14112105

Shin W., Mun S., Kim S., Yu B., Srikulnath K., Ahn K., and Han K., 2025, Integrating miRNAs with QTL for identification of milk and reproduction related traits in dairy cattle, Genes and Genomics, 1: 11.
https://doi.org/10.1007/s13258-025-01650-x

Silpa M., König S., Sejian V., Malik P., Nair M., Fonsêca V., Maia A., and Bhatta R., 2021, Climate-resilient dairy cattle production: applications of genomic tools and statistical models, Frontiers in Veterinary Science, 8: 625189.
https://doi.org/10.3389/fvets.2021.625189

Tahir M., Mohyuddin S., Yao Y., Wang Y., Liang Y., Karrow N., and Mao Y., 2025, Association of NCOA6 gene polymorphism with milk production traits in Chinese holstein cows, Animals, 15(10): 1461.
https://doi.org/10.3390/ani15101461

Venkatesan R., Rani A., Umesh S., Sushil K., Kumar D., Sowmya P., Kesavan M., Singh R., Panchasara H., Kumar S., and Chhaya R., 2025, Genome-wide scan for SNPs and selective sweeps reveals candidate genes and QTLs for milk production and reproduction traits in Indian Kankrej cattle, 3 Biotech, 15(4): 90.
https://doi.org/10.1007/s13205-025-04263-z

Wadood A., Bordbar F., and Zhang X., 2025, Integrating omics approaches in livestock biotechnology: innovations in production and reproductive efficiency, Frontiers in Animal Science, 6: 1551244.
https://doi.org/10.3389/fanim.2025.1551244

Wang M., Yang N., Laterrière M., Gagné D., Omonijo F., and Ibeagha-Awemu E., 2024, Multi-omics integration identifies regulatory factors underlying bovine subclinical mastitis, Journal of Animal Science and Biotechnology, 15: 46.
https://doi.org/10.1186/s40104-024-00996-8

Xia L., Zhao Z., Yu X., Lu C., Jiang P., Yu H., Li X., Yu X., Liu J., Fang X., and Yang R., 2021, Integrative analysis of miRNAs and mRNAs revealed regulation of lipid metabolism in dairy cattle, Functional and Integrative Genomics, 21: 393-404.
https://doi.org/10.1007/s10142-021-00786-9

Xiong Z., Li L., Ouyang K., Qu M., and Qiu Q., 2025, Deciphering heat stress mechanisms and developing mitigation strategies in dairy cattle: a multi-omics perspective, Agriculture, 15(14): 1477.
https://doi.org/10.3390/agriculture15141477

Zhang C., Wang M., Liu H., Jiang X., Chen X., Liu T., Yin Q., Wang Y., Deng L., Yao J., and Wu S., 2023, Multi-omics reveals that the host-microbiome metabolism crosstalk of differential rumen bacterial enterotypes can regulate the milk protein synthesis of dairy cows, Journal of Animal Science and Biotechnology, 14: 63.
https://doi.org/10.1186/s40104-023-00862-z

Zhang X.Z., and Lin X.F., 2025, Single-cell multi-omics analysis of epigenetic and transcriptional regulatory mechanisms of goat skeletal muscle development, Bioscience Methods, 16(2): 83-99.
https://doi.org/10.5376/bm.2025.16.0009

Zheng W., Zhang J., Huang J., Han B., Liu Y., and Sun D., 2025, A pilot multi-omics study reveals genetic mechanisms regulating milk component traits in dairy cattle, Communications Biology, 8(1): 1150.
https://doi.org/10.1038/s42003-025-08615-6