Evolutionary Applications

WILEY

Evolutionary Applications

ORIGINAL ARTICLE OPEN ACCESS

Integrative Genomics Identifies Candidate Genes Underlying Trypanotolerance in Hybrid African Cattle

Gillian P. McHugo[1] | James A. Ward[1] | Said Ismael Ng'ang'a[2,3] | Laurent A. F. Frantz[2,3] | John A. Browne[1] | Michael Salter-Townshend[4] | Grace M. O'Gorman[5] | Kieran G. Meade[1,6,7] | Emmeline W. Hill[1] | Thomas J. Hall[1] | David E. MacHugh[1,6,7]

[1]UCD School of Agriculture and Food Science, University College Dublin, Dublin, Ireland | [2]Palaeogenomics Group, Department of Veterinary Sciences, Ludwig Maximilian University, Munich, Germany | [3]School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK | [4]UCD School of Mathematics and Statistics, University College Dublin, Dublin, Ireland | [5]UK Agri-Tech Centre, Innovation Centre, York, UK | [6]UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland | [7]UCD One Health Centre, University College Dublin, Dublin, Ireland

Correspondence: David E. MacHugh (david.machugh@ucd.ie)

Received: 31 July 2025 | **Revised:** 2 May 2026 | **Accepted:** 18 May 2026

ABSTRACT

Integrative genomics combines data from different omics sources to link genotypes and phenotypes with the aim of unravelling biological networks and pathways that undergird complex traits, particularly with respect to disease. In this respect, integrative genomics, leveraging population and functional genomic data, can be employed to understand the evolutionary processes that have shaped adaptation to infectious diseases in domestic cattle. This approach can be particularly informative for African cattle, which exhibit a complex mosaic of genomic ancestry from *Bos taurus* (taurine) and *Bos indicus* (indicine) populations. Some African taurine populations have an important evolutionary adaptation known as trypanotolerance, a genetically determined tolerance to infection by trypanosome parasites (*Trypanosoma* spp.) that cause African animal trypanosomiasis (AAT). AAT is one of the largest constraints to livestock production in sub-Saharan Africa and causes a financial burden of approximately $4.5 billion annually. In this study, we identified putative candidate genes underlying trypanotolerance by integrating local ancestry inference (LAI) from genome-wide SNP data across multiple trypanotolerant and trypanosusceptible hybrid cattle populations with RNA-seq and expression microarray transcriptomic data from multiple tissues collected across time course trypanosome infection experiments. These candidate genes included *AGO2*, *CBL*, *CNOT1*, *EDN1*, *IL1B*, *NFKB1*, *RIPK1* and *TRAF2*. Functional analysis of the gene set outputs from this work highlighted GO terms associated with the immune system (including the major histocompatibility complex—MHC) and cell signalling processes. These results signpost future work to elucidate the cellular networks and pathways that drive trypanotolerance.

1 | Introduction

Integrative genomics, the integration of data from different omics sources, is an important strategy for obtaining novel scientific insights from several different but complementary omics technologies. It can be used to link phenotypic information with multiple types of biomolecular data, with the aim of unravelling the biological networks and pathways that underpin complex traits in multicellular organisms (Porcu et al. 2021; Subramanian et al. 2020). In this regard, an integrative approach encompassing multi-omics analyses and network biology has been acknowledged in recent years as a powerful strategy for understanding the vertebrate immune system and immune responses to pathogens and parasites (Hao et al. 2020; Urbanski et al. 2019; Wang et al. 2023; Wong et al. 2021). This is because each individual omics technology is unable to capture the full biological complexity of host-pathogen interaction for many infectious diseases,

and the integration of multiple omics sources can provide more detailed and comprehensive insights (Cano-Gamez and Trynka 2020; Karczewski and Snyder 2018). In addition, it has become increasingly cost-effective as the generation of genome-scale data has become more economical (Gamazon et al. 2013; Koestler et al. 2014).

Integrative and functional population genomics techniques have shown that the immune system of modern humans has been shaped by adaptation to pathogens and parasites, regional human ancestry contributions, gene flow from archaic hominins such as Neanderthals and Denisovans, and the emergence of agriculture, sedentarism and complex societies (Karlsson et al. 2014; Quintana-Murci 2019; Randolph et al. 2024). For example, European and African ancestry components in modern human populations have different effects on the activity of innate immune pathways associated with responses to viral infections (Randolph et al. 2021). Also, analyses of paleogenomic and GWAS data have shown that the likelihood of avoiding severe COVID-19 caused by SARS-CoV-2 infection is positively influenced by Neanderthal haplotypes on human chromosome 12 (HSA12) and negatively by those on chromosome 3 (HSA3) (Zeberg and Paabo 2020, 2021). In addition, genetic adaptation to pathogens since the development of farming, particularly over the last three thousand years after the Bronze Age, has contributed to the risk of inflammatory diseases in modern European populations (Kerner et al. 2023). Integrative genomics approaches, whereas less widely used on non-human species, also have the potential to unravel the complex evolutionary history of cattle including production traits and their response to infectious diseases (de Lima et al. 2020; Ghoreishifar et al. 2026; Gonzalez-Prendes et al. 2022; Hall et al. 2024, 2021; Kadarmideen and Mazzoni 2018; O'Grady et al. 2025; Suravajhala et al. 2016).

Domestic cattle can be categorised into *Bos taurus* (taurine), *Bos indicus* (indicine), and various grades of hybrid taurine-indicine breeds or populations (Bovine HapMap Consortium et al. 2009; Decker et al. 2014). The divergence between the taurine and indicine lineages substantially predates the development of animal agriculture, with high-resolution genome-scale analyses providing estimates of at least 300 kya for the split between the groups (Chen et al. 2018; Hou et al. 2024; Wang et al. 2018; Wu et al. 2018). These diverse ancestries, subsequent admixture, and selection for resistance to infectious diseases caused by viruses, bacteria and parasites have shaped the evolution of the domestic cattle genome in Africa. Examples include rinderpest disease caused by rinderpest virus (RPV) (Flori et al. 2014), tuberculosis caused by *Mycobacterium bovis* (Callaby et al. 2020; Kassahun et al. 2015), and most notably, trypanosomiasis due to bloodstream parasites of the *Trypanosoma* genus (Bahbahani et al. 2018; Tijjani et al. 2019).

Patterns of genomic variation in African cattle form clines of *B. indicus* and *B. taurus* ancestry primarily across the east–west axis of the continent (Hanotte et al. 2002). However, at a more granular regional level, nomadic pastoralism, livestock trade, and organised crossbreeding have resulted in a rich diversity of taurine, indicine and admixed populations (Freeman et al. 2004; Kim, Hanotte, et al. 2017; Kim et al. 2020; MacHugh et al. 1997). Importantly, retention of substantial African taurine ancestry in West and Central Africa is largely due to a genetically determined tolerance to trypanosomes that cause trypanosomiasis, a major impediment to livestock production in sub-Saharan Africa (Steverding 2008; Yaro et al. 2016). These trypanotolerant cattle exhibit a greater ability to control parasitaemia and anaemia, which makes them more productive in areas infested with trypanosomes than *B. indicus* cattle or other *B. taurus* breeds (Dayo et al. 2012; Murray et al. 1984; Yaro et al. 2016). These trypanotolerant breeds are therefore an important genetic resource as they are uniquely adapted to livestock production in the humid tropical and semitropical regions of Africa (Dayo et al. 2012; Food and Agriculture Organization 2023). The genes and genomic regulatory elements (GREs) containing sequence variations that contribute to the polygenic trypanotolerance trait remain poorly understood, although some candidate genes have been proposed (Alvarez et al. 2016; Noyes et al. 2011; Yaro et al. 2016). Improved knowledge of the genes and genomic variants responsible for trypanotolerance could inform genome-enabled breeding or genome editing programmes to increase the productivity of livestock populations in sub-Saharan Africa (Yaro et al. 2016).

N'Dama cattle are the most well-studied trypanotolerant population and control parasitaemia and anaemia significantly better than trypanosusceptible cattle, and have increased trypanotolerance compared to hybrid populations (Achukwi et al. 1997; Berthier et al. 2015; Meade et al. 2009; Murray et al. 1984; Noyes et al. 2011; O'Gorman et al. 2009, 2006; Paling et al. 1991). Lagune and Baoulé cattle were found to exhibit trypanotolerance levels similar to those of the N'Dama population (Berthier et al. 2015). The endangered Muturu population is not as well characterised as other trypanotolerant populations; however, minimal gene flow was observed between the N'Dama and Muturu populations, illustrating their genetic separation and the importance of their conservation as a trypanotolerant population (Ibeagha-Awemu and Erhardt 2005; Tijjani et al. 2019). Although the trypanotolerant African hybrid populations analysed in this study are classed as trypanotolerant, their response to trypanosome infection is intermediate to that of the trypanotolerant populations with higher levels of African *B. taurus* ancestry and trypanosusceptible populations (Berthier et al. 2015; Food and Agriculture Organization 2023). This means that they exhibit a reduced ability to control both parasitaemia and anaemia during trypanosome infection than other trypanotolerant populations, likely because they are hybrid populations, and the degree of trypanotolerance has been shown to correlate with levels of African *B. taurus* ancestry (Berthier et al. 2015). The trypanosusceptible African hybrid populations analysed in this study show no evidence of trypanotolerance despite their long history of exposure to trypanosome parasites (Berthier et al. 2015; Food and Agriculture Organization 2023; Meade et al. 2009; Noyes et al. 2011; O'Gorman et al. 2009, 2006). Similarly, the *B. indicus* and European *B. taurus* populations have no documented trypanotolerance despite their introduction into trypanosome-infested areas (Berthier et al. 2015; Food and Agriculture Organization 2023). Finally, for the residually admixed European populations, because trypanosomiasis in cattle is rare in Europe, the low level of African *B. taurus* ancestry in the breeds means that trypanotolerance is unlikely, as the trait is known to correlate with increased African *B. taurus* ancestry (Barbato et al. 2020; Berthier et al. 2015; Food and Agriculture Organization 2023).

In this study, we investigated the genomic architecture of the trypanotolerance trait. We performed local ancestry inference (LAI) using genome-wide SNP data for multiple trypanotolerant and trypanosusceptible hybrid cattle populations. These results were then integrated with gene expression data from some of the same animals, obtained as both RNA-seq and microarray data from time course trypanosome infection experiments. Overall, 1154 individual genome-wide SNP samples from 26 different cattle populations were analysed. A total of 220 microarray samples, collected across 12 time points and four tissues from 50 of those animals, and 120 RNA-seq sample datasets obtained across four time points from an additional 30 animals in trypanosome infection experiments were also analysed. The datasets were integrated by testing, using interval enrichment analysis, whether genomic intervals containing local ancestry peaks and troughs were significantly enriched for genes identified in the functional genomics analyses.

2 | Materials and Methods

2.1 | Data Sources

2.1.1 | SNP Data

As previously described (McHugo, Ward, Ng'ang'a, et al. 2025), Illumina BovineHD 777K BeadChip SNP datasets were generated for 39 African cattle (23 Somba, 8 N'Dama and 8 Boran). The N'Dama and Boran animals were obtained from DNA samples collected during a trypanosome infection time course experiment performed in 2003 (Meade et al. 2009; O'Gorman et al. 2009, 2006). These SNP datasets were merged with publicly available Illumina BovineHD 777K BeadChip genome-wide high-density SNP datasets for additional animals that were assembled from published studies and databases (Bahbahani et al. 2017; Barbato et al. 2020; Sempere et al. 2015; Upadhyay et al. 2017; Verdugo et al. 2019; Ward et al. 2022; Wragg et al. 2022), which, after filters were applied, resulted in a total sample set of 750 animals. The publicly available samples were 6 Alentejana, 26 Angus, 25 Ankole, 82 Boran, 50 Borgou, 19 Chianina, 111 East African Shorthorn Zebu, 28 Gir, 36 Holstein Friesian, 23 Jersey, 16 Karamojong, 22 Keteku, 5 Lagune, 13 Marchigiana, 5 Maremmana, 19 Muturu, 47 N'Dama Guinea, 33 N'Dama hybrid, 38 Nelore, 27 Nganda, 51 Romagnola, 16 Sheko and 13 Tharparkar animals. The high-density BovineHD 777K SNP dataset was then downsampled to the subset of SNPs in common with the Illumina Bovine SNP50 BeadChip.

Low-density genome-wide SNP array datasets (Illumina Bovine SNP50 BeadChip) were also obtained for 39 cattle that were part of a similar trypanosomiasis time course infection study using RNA-seq data (Berthier et al. 2015; Peylhard et al. 2023). Several of the populations examined in this study were not included in the initial dataset used for the previous local ancestry analysis (McHugo, Ward, Ng'ang'a, et al. 2025); therefore, additional low-density genome-wide SNP array datasets (Illumina Bovine SNP50 BeadChip) were obtained from published studies via the Web-Interfaced Next Generation Database (WIDDE) resource for cattle and sheep SNP array genotype data (Decker et al. 2014; Flori et al. 2014; Gautier et al. 2009; Sempere et al. 2015). In total, 26 different populations were represented (Table 1): three European *B. taurus* populations (Holstein-Friesian, Angus and Jersey); four African *B. taurus* populations (Muturu, Lagune, Guinean N'Dama and Baoulé); three *B. indicus* populations (Tharparkar, Gir and Nelore); five residually admixed European populations (Romagnola, Chianina, Marchigiana, Maremmana and Alentejana); five trypanotolerant African hybrid populations (hybrid N'Dama, Borgou, Somba, Keteku and Sheko) and six trypanosusceptible African hybrid populations (Fulani Zebu, Ankole, Nganda, East African Shorthorn Zebu, Karamojong and Boran).

2.1.2 | Gene Expression Data

As previously described (McHugo, Ward, Browne, et al. 2025), Affymetrix GeneChip Bovine Genome Array datasets were obtained from published studies with parallel trypanosome infection time course experimental designs (Meade et al. 2009; Noyes et al. 2011; O'Gorman et al. 2009, 2006) in which trypanotolerant African hybrid N'Dama and trypanosusceptible African hybrid Boran cattle were experimentally infected with the *Trypanosoma congolense* clone IL1180 (Geigy and Kauffmann 1973; Nantulya et al. 1984) delivered via the bites of infected tsetse flies (*Glossina morsitans morsitans*) (Akol and Murray 1982; Dwinger et al. 1987). Samples were collected before infection and at various days post-infection (dpi), and included peripheral blood mononuclear cells (PBMC) isolated from blood (BL), liver (LI), lymph node (LN) and spleen (SP) samples. This resulted in a total of 220 samples from 50 animals (25 trypanotolerant N'Dama and 25 trypanosusceptible Boran) collected across 12 time points and four tissues before filtering (Table 2). Figure 1 illustrates the experimental design and study workflow. The computer code required to repeat and reproduce the analyses is available at http://doi.org/10.5281/zenodo.11517978.

A total of 120 RNA-seq sample datasets were obtained from a similar trypanosome infection time course experiment in which 30 animals (six trypanotolerant African *B. taurus* Lagune, six trypanotolerant African *B. taurus* Baoulé, six trypanotolerant African hybrid N'Dama, six trypanotolerant African hybrid Borgou, and six trypanosusceptible African hybrid Fulani Zebu) were infected via intravenous inoculation with 10^5 trypanosomes of the same *T. congolense* IL1180 strain (Berthier et al. 2015; Geigy and Kauffmann 1973; Nantulya et al. 1984; Peylhard et al. 2023). Whole blood (WB) samples were taken before infection and at 20, 30 and 40 dpi (four time points) and total RNA was extracted for sequencing (Berthier et al. 2015; Peylhard et al. 2023).

2.2 | Analysis of Genomic Data

2.2.1 | Preparation of Genomic Data

The additional publicly available SNP samples to be merged with our previously analysed dataset (McHugo, Ward, Ng'ang'a, et al. 2025) were converted to binary PLINK files with Illumina allele coding for the FORWARD strand using PLINK (v. 1.90 beta 6.25) (Chang et al. 2015), and the iConvert.py script from SNPchiMp (v. 3) (Nicolazzi et al. 2015),

TABLE 1 | Population code, population name, group, country of origin, number of samples pre- and post-filtering, and sources of SNP data.

Code	Population	Group	Country of origin	No. pre-filter	No. post-filter	Source
ALEN	Alentejana	Residually admixed European	Portugal	6	6	a, b
ANGU	Angus	European *Bos taurus*	United Kingdom	26	26	c
ANKO	Ankole	Trypanosusceptible African hybrid	Uganda	25	25	d
BAOU	Baoulé	African *B. taurus*	Burkina Faso	37	37	e, f
BORA	Boran	Trypanosusceptible African hybrid	Kenya	90	90	g, h, i
BORG	Borgou	Trypanotolerant African hybrid	Benin	261	261	j, k, e, f
CHIA	Chianina	Residually admixed European	Italy	19	19	l, b
EASZ	East African Shorthorn Zebu	Trypanosusceptible African hybrid	Kenya	111	111	d, k
FULA	Fulani Zebu	Trypanosusceptible African hybrid	Benin, Burkina Faso	51	51	e, f
GIR	Gir	*Bos indicus*	India	28	28	c, d
HOLS	Holstein Friesian	European *B. taurus*	Netherlands	36	36	c
JERS	Jersey	European *B. taurus*	United Kingdom	23	23	c
KARA	Karamojong	Trypanosusceptible African hybrid	Uganda	16	16	d
KETE	Keteku	Trypanotolerant African hybrid	Nigeria	22	22	k
LAGU	Lagune	African *B. taurus*	Benin	56	51	e, f, c
MARC	Marchigiana	Residually admixed European	Italy	13	13	l
MARE	Maremmana	Residually admixed European	Italy	5	5	b
MUTU	Muturu	African *B. taurus*	Nigeria	19	19	k
NDAG	N'Dama Guinea	African *B. taurus*	Guinea	47	47	k, c, d
NDAM	N'Dama hybrid	Trypanotolerant African hybrid	Burkina Faso, Unspecified, Togo, Mali, Kenya	80	80	e, l, a, f, i
NELO	Nelore	*B. indicus*	Brazil	38	38	c, b, d
NGAN	Nganda	Trypanosusceptible African hybrid	Uganda	27	27	d, k
ROMA	Romagnola	Residually admixed European	Italy	51	51	l, c
SHEK	Sheko	Trypanotolerant African hybrid	Ethiopia	16	16	c
SOMB	Somba	Trypanotolerant African hybrid	Benin	23	23	i
THAR	Tharparkar	*B. indicus*	Pakistan	13	13	l
			Total	1159	1154	

Note: a: Verdugo et al. (2019), b: Upadhyay et al. (2017), c: Web-Interfaced Next Generation Database (WIDDE; Sempere et al. (2015)), d: Bahbahani et al. (2017), e: Gautier et al. (2009), f: Berthier et al. (2015), g: Wragg et al. (2022), h: Decker et al. (2014), i: McHugo, Ward, Ng'ang'a, et al. (2025), j: Flori et al. (2014), k: Ward et al. (2022) and l: Barbato et al. (2020).

since some samples were downloaded with TOP strand allele coding. The Illumina Bovine SNP50 BeadChip SNP locations were updated from UMD3.1 to the newer bovine genome assembly ARS-UCD1.2 (Rosen et al. 2020) using coordinates from the NAGRP Data Repository (Schnabel 2018) and the PLINK --update-chr and --update-map functions. The additional SNP samples were downsampled to the subset of the 46,713 SNPs in common with the BovineHD 777K BeadChip using PLINK with a list of the SNPs from the previously downsampled low-density SNP dataset generated using dplyr (v. 1.1.2) (Wickham, François, et al. 2023) and readr (v. 2.1.4) (Wickham, Hester, and Bryan 2023) with R (v. 4.3.2) (R Core Team 2023). The samples were then merged with the previously downsampled low-density SNP dataset with PLINK. The merged low-density SNP data were filtered as described in our previous study (McHugo, Ward, Ng'ang'a, et al. 2025). Briefly, this included removing samples with call rates < 0.95, as well as samples duplicated across multiple data sources that had an identity-by-state value ≥ 0.99 with PLINK. The dataset was then filtered using PLINK to retain autosomal SNPs with a minimum call rate of 95% and a minor allele frequency (MAF) of at least 5%.

2.2.2 | Population Genomic Analyses

The results of the population genomic analyses of high- and low-density SNP datasets were obtained from our previous study, and the same analyses of the merged low-density SNP dataset were performed as described therein (McHugo, Ward, Ng'ang'a, et al. 2025). Briefly, an inbreeding analysis was performed with PLINK, and principal component analysis (PCA) was performed using smartpca after file conversion with convertf, both part of the EIGENSOFT package (v. 7.1.2) (Patterson et al. 2006). Genetic structure analysis was performed using structure_threader (v. 1.3.4) (Pina-Martins et al. 2017) with fastStructure (v. 1.0) (Raj et al. 2014), with the model complexity or number of populations (K) set from 2 to 27. The chooseK function was used to evaluate the outputs and identify a range of K values that best accounted for the structure in the data (Raj et al. 2014).

2.2.3 | Local Ancestry Analysis and Functional Enrichment of Introgressed Regions

After conversion of the binary PLINK files into BIMBAM format with PLINK, local ancestry inference (LAI) was performed separately for each bovine autosome in the populations, with additional low-density SNP samples added to the previously analysed dataset (McHugo, Ward, Ng'ang'a, et al. 2025) using 30 expectation-maximisation (EM) steps, three upper clusters, 15 lower clusters, and 200 mixing generations using the Efficient Local Ancestry Inference (ELAI) software (v. 1.0) (Guan 2014). Based on the previous local ancestry results (McHugo, Ward, Ng'ang'a, et al. 2025), the donor populations for the LAI were the Angus (European *B. taurus*), Guinean N'Dama (African *B. taurus*) and Gir (*B. indicus*). Mean ancestry scores across the individual hybrid animals and a genome-wide z-score for each of the three ancestry components were estimated for each hybrid population.

Functional enrichment of the introgressed regions was performed using gprofiler2 (v. 0.2.2) (Kolberg et al. 2023) with R. The background set was the set of genes within genomic intervals of 1 Mb up- and downstream from a SNP in the dataset. The query sets were the genes within 1 Mb upstream and downstream of the SNPs with a z-score ≥ 2.0 for each ancestry. An interval size range for this analysis of ±1 Mb around focal LAI segments was selected for consistency with our previous cattle local ancestry study (McHugo, Ward, Ng'ang'a, et al. 2025) and based on functional population genomics studies of gene flow and admixture in archaic and modern human populations, which have a similar evolutionary divergence to the *B. taurus* and *B. indicus* cattle lineages (Chen et al. 2018; Colbran et al. 2019; McQuillan et al. 2022; Wang et al. 2018; Wu et al. 2018).

2.3 | Analysis of Gene Expression Data

2.3.1 | Differential Expression Analysis

The results of differential expression analysis of the microarray data were obtained from our previous study (McHugo, Ward, Browne, et al. 2025). Differential expression analysis of the RNA-seq data was performed using edgeR (v. 4.0.16) (Robinson et al. 2010) with R. The scripts published with the intermediate tables and raw data from Peylhard et al. (2023) were modified to analyse response (RESP) contrasts to identify changes in expression over time in the trypanotolerant samples (LAGU, BAOU, NDAM and BORG) relative to the trypanosusceptible FULA ($\text{RESP}_i = [T_i - T_0] - [S_i - S_0]$, where T represents the trypanotolerant populations, and S represents the trypanosusceptible populations) (Figure 1) (Noyes et al. 2011). This was done to facilitate comparison with the previous analysis of the microarray data (McHugo, Ward, Browne, et al. 2025) and because the original analysis of the RNA-seq data did not include these contrasts (Peylhard et al. 2023). As in the analysis conducted by Peylhard et al. (2023), generalized linear model likelihood ratio tests were used to calculate $\log_2$ fold change values and Benjamini-Hochberg (B-H) corrected p-values (Benjamini and Hochberg 1995; McCarthy et al. 2012). Genes with an adjusted p-value of ≤ 0.05 (B-H $p_{adj.} \leq 0.05$) were considered to be significantly differentially expressed.

2.3.2 | Gene Interaction Network Analysis

The GeneCards database (https://www.genecards.org/, v. 5.19) (Stelzer et al. 2016), which integrates data from almost 200 different biological databases, was searched for genes relating to the search term 'trypano*'. The results were exported, filtered to select genes with a relevance score ≥ 1.75 and converted from gene symbols to Ensembl IDs to prepare a computationally manageable number of genes based on the methodology described by Hall et al. (2021) and using dplyr, gprofiler2, readr and stringr (v. 1.5.0) (Wickham 2023) with R. The resulting file was then used as input for network analysis to identify interactions among these genes and other genes in the bovine interactome using InnateDB (https://www.innatedb.com/, v. 5.4) (Breuer et al. 2013) including interactions predicted by orthology. The node IDs of the resulting network were prepared using dplyr, readr and stringr with R. The network was then imported into Cytoscape (v. 3.8.0) (Shannon et al. 2003). The differential gene expression results obtained using the microarray data (McHugo, Ward, Browne, et al. 2025) and the RNA-seq data described above were imported as node tables.

Expression-activated subnetworks or modules within the base network were identified using jActiveModules (v. 3.2.1) (Ideker et al. 2002) for the B-H corrected p-values for each of the eight final contrasts in the microarray and RNA-seq datasets, as the final time point produced the highest number of significantly differentially expressed genes (DEGs) in both datasets. The highest scoring module for each contrast was selected, analysed with Cytoscape and yFiles Layout Algorithms (v. 1.1.3) (Becker and Rojas 2001) was used to remove overlaps. The modules were outputted as network files and images and analysed with R.

2.3.3 | Functional Enrichment Analysis of Gene Expression Data

Functional enrichment was performed for the complete set of significant DEGs identified from the RNA-seq data for the response contrasts, and for the modules identified from both the microarray and RNA-seq data, using gprofiler2 with R. The

TABLE 2 | Days post-infection (dpi), population code, tissue, data type, number of samples pre- and post-filtering, and sources of gene expression data.

Dpi	Population	Tissue	Data type	No. pre-filtering	No. post-filtering	Source
0	BAOU	WB	RNA-seq	6	5	a
0	BORA	BL	Microarray	5	5	b
0	BORA	LI	Microarray	20	18	c
0	BORA	LN	Microarray	5	4	c
0	BORA	SP	Microarray	5	4	c
0	BORG	WB	RNA-seq	6	5	a
0	FULA	WB	RNA-seq	6	5	a
0	LAGU	WB	RNA-seq	6	6	a
0	NDAM	BL	Microarray	5	5	b
0	NDAM	LI	Microarray	20	19	c
0	NDAM	LN	Microarray	5	5	c
0	NDAM	SP	Microarray	5	5	c
0	NDAM	WB	RNA-seq	6	6	a
12	BORA	LI	Microarray	5	5	c
12	NDAM	LI	Microarray	5	5	c
14	BORA	BL	Microarray	5	5	b
14	NDAM	BL	Microarray	5	5	b
15	BORA	LI	Microarray	5	5	c
15	NDAM	LI	Microarray	5	5	c
18	BORA	LI	Microarray	5	5	c
18	NDAM	LI	Microarray	5	5	c
20	BAOU	WB	RNA-seq	6	5	a
20	BORG	WB	RNA-seq	6	5	a
20	FULA	WB	RNA-seq	6	5	a
20	LAGU	WB	RNA-seq	6	6	a
20	NDAM	WB	RNA-seq	6	6	a
21	BORA	LI	Microarray	5	5	c
21	BORA	LN	Microarray	5	5	c
21	BORA	SP	Microarray	5	5	c
21	NDAM	LI	Microarray	5	5	c
21	NDAM	LN	Microarray	5	4	c
21	NDAM	SP	Microarray	5	5	c
25	BORA	BL	Microarray	5	5	b
25	NDAM	BL	Microarray	5	5	b
26	BORA	LI	Microarray	5	5	c
26	NDAM	LI	Microarray	5	3	c
29	BORA	LI	Microarray	5	5	c

(Continues)

TABLE 2 | (Continued)

Dpi	Population	Tissue	Data type	No. pre-filtering	No. post-filtering	Source
29	NDAM	LI	Microarray	5	5	c
30	BAOU	WB	RNA-seq	6	5	a
30	BORG	WB	RNA-seq	6	5	a
30	FULA	WB	RNA-seq	6	5	a
30	LAGU	WB	RNA-seq	6	6	a
30	NDAM	WB	RNA-seq	6	6	a
32	BORA	LI	Microarray	5	5	c
32	NDAM	LI	Microarray	5	5	c
34	BORA	BL	Microarray	5	5	b
34	NDAM	BL	Microarray	5	5	b
35	BORA	LI	Microarray	5	5	c
35	BORA	LN	Microarray	5	5	c
35	BORA	SP	Microarray	5	5	c
35	NDAM	LI	Microarray	5	5	c
35	NDAM	LN	Microarray	5	4	c
35	NDAM	SP	Microarray	5	5	c
40	BAOU	WB	RNA-seq	6	5	a
40	BORG	WB	RNA-seq	6	5	a
40	FULA	WB	RNA-seq	6	5	a
40	LAGU	WB	RNA-seq	6	6	a
40	NDAM	WB	RNA-seq	6	6	a
		Total	Microarray	220	211	
			RNA-seq	120	108	

Note: a: Peylhard et al. (2023), b: McHugo, Ward, Browne, et al. (2025) and c: Noyes et al. (2011).

background set for the significant DEGs was the set of detectable genes identified by Peylhard et al. (2023), whereas the background set for the modules was the base network. The analyses were restricted to Gene Ontology (GO) terms, and the driver GO terms were highlighted.

2.4 | Integration of Genomic Local Ancestry Inference Results and Gene Expression Data

The results of the LAI (McHugo, Ward, Ng'ang'a, et al. 2025) and expression analyses (McHugo, Ward, Browne, et al. 2025), including the new results described in this study, were integrated with INRICH (INterval enRICHment analysis, v. 1.1) (Lee et al. 2012). A reference gene file of all bovine genes was prepared using biomaRt (v. 2.58.2) (Durinck et al. 2009) and readr with R. Reference SNP files were prepared from the bim files for the associated datasets using dplyr and readr with R. A target gene set file containing the genes in the first module identified for each of the eight final contrasts in the microarray and RNA-seq datasets converted to Ensembl IDs was prepared using dplyr, gprofiler2, readr, stringr and tidyr (v. 1.3.0) (Wickham, Vaughan, and Girlich 2023) with R. Associated genomic interval files were prepared to include regions 1 Mb up- and downstream from SNPs with a z-score ≥ 2.0 for each ancestry component of the LAI results (McHugo, Ward, Ng'ang'a, et al. 2025), excluding the results from the low-density SNP dataset analysed with MOSAIC due to the low resolution obtained. This procedure involved preparing files with z-scores converted to p-values using dplyr, readr and stringr with R. These data were used as input for linkage disequilibrium (LD) clumping of SNPs 1 Mb up- and downstream from those with converted p-values ≤ 0.02275 (equivalent to z-scores ≥ 2.0) using the associated SNP data files with PLINK. The resulting clumped files were then converted to interval files for use with INRICH with the intervals extending 1 Mb up- and downstream from the central SNP in each clump using dplyr, readr and stringr with R. Interval enrichment analyses

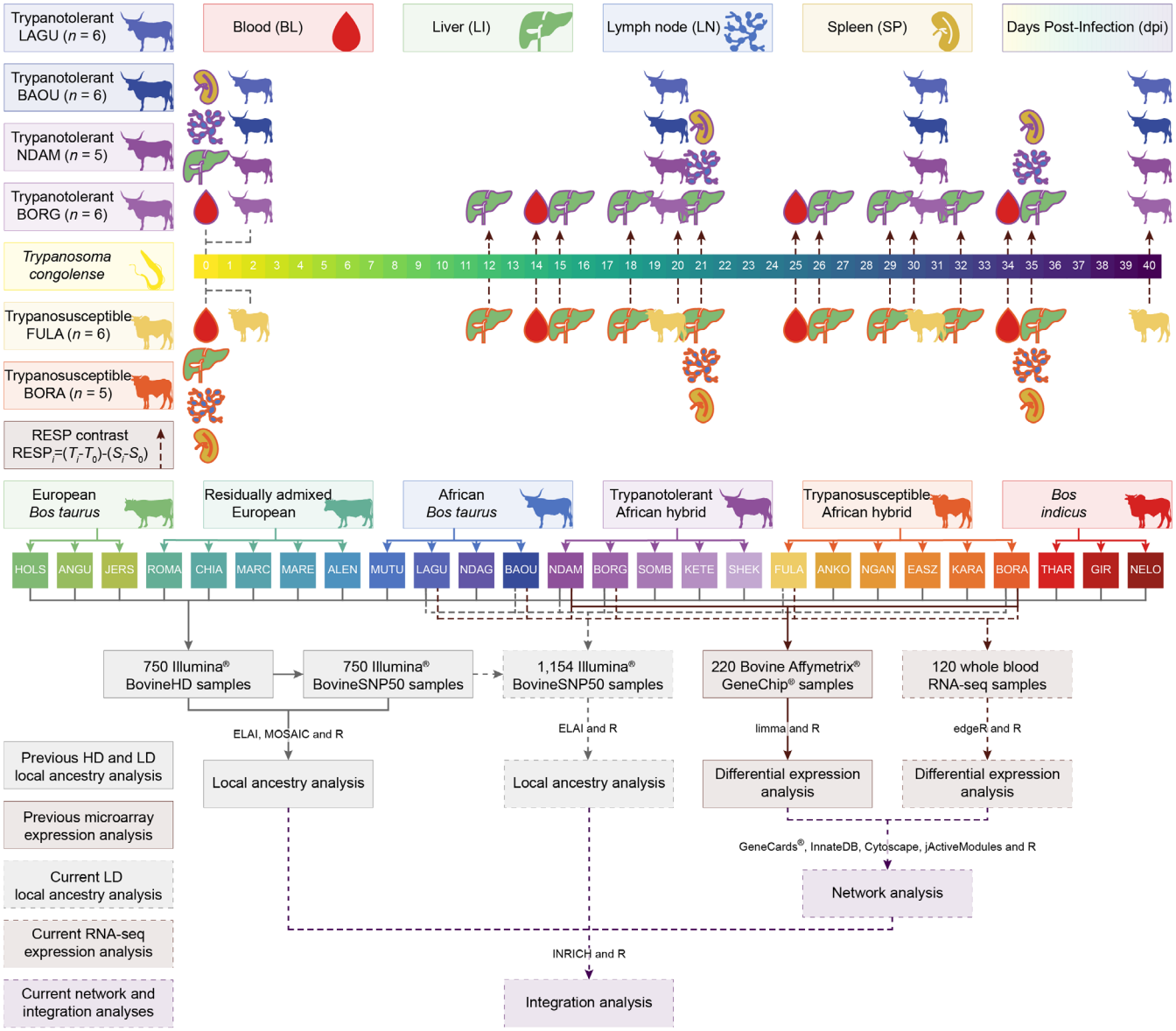

FIGURE 1 | Diagram showing the experimental design and study workflow. Trypanosome image by Matus Valach and cattle images by Tracy A. Heath, T. Michael Keesey and Steven Traver via https://www.phylopic.org/ and tissue images via https://healthicons.org/. Colours from the khroma (v. 1.10.0) (Frerebeau 2023) and viridis (v. 0.6.3) (Garnier et al. 2023) R packages.

were then performed for the target gene set file containing the modules and the interval files from the LAI results with the reference gene file and appropriate reference SNP files using INRICH with a target size filter of 2500 to include all the modules and specifying non-human data.

The results of the analyses were visualised using ComplexUpset (v. 1.3.3) (Krassowski 2020), dplyr, ggh4x (v. 0.2.4) (van den Brand 2023), ggplot2 (v. 3.4.2) (Wickham 2009), ggrepel (v. 0.9.3) (Slowikowski 2023), ggtext (v. 0.1.2) (Wilke and Wiernik 2022), magick (v. 2.8.1) (Ooms 2023) with ImageMagick (v. 6.9.12.96) (ImageMagick Studio LLC 2023), magrittr (v. 2.0.3) (Bache and Wickham 2022), parallel (v. 4.3.2) (R Core Team 2023), patchwork (v. 1.1.2) (Pedersen 2023), purrr (v. 1.0.1) (Wickham and Henry 2023), readr, rlang (v. 1.1.1) (Henry and Wickham 2024), scales (v. 1.2.1) (Wickham, Pedersen, and Seidel 2023), stringr, tibble (v. 3.2.1) (Müller and Wickham 2023), tidyr with R. Colours were generated from khroma (v. 1.10.0) (Frerebeau 2023) and viridis (v. 0.6.3) (Garnier et al. 2023).

3 | Results

3.1 | Population Genomics Results Reiterate Known Evolutionary Relationships

3.1.1 | Post-Filtering SNP Data

After filtering for missing genotypes and identity-by-state (Figure S1), there were 1154 animals in the merged low-density dataset (Table 1). The inbreeding results did not indicate additional filtering was required (Figure S2). Filtering for autosomal SNPs with a minimum call rate of 95% and MAF of at least 5% retained 29,869 SNPs with a total genotyping rate of 99.47%.

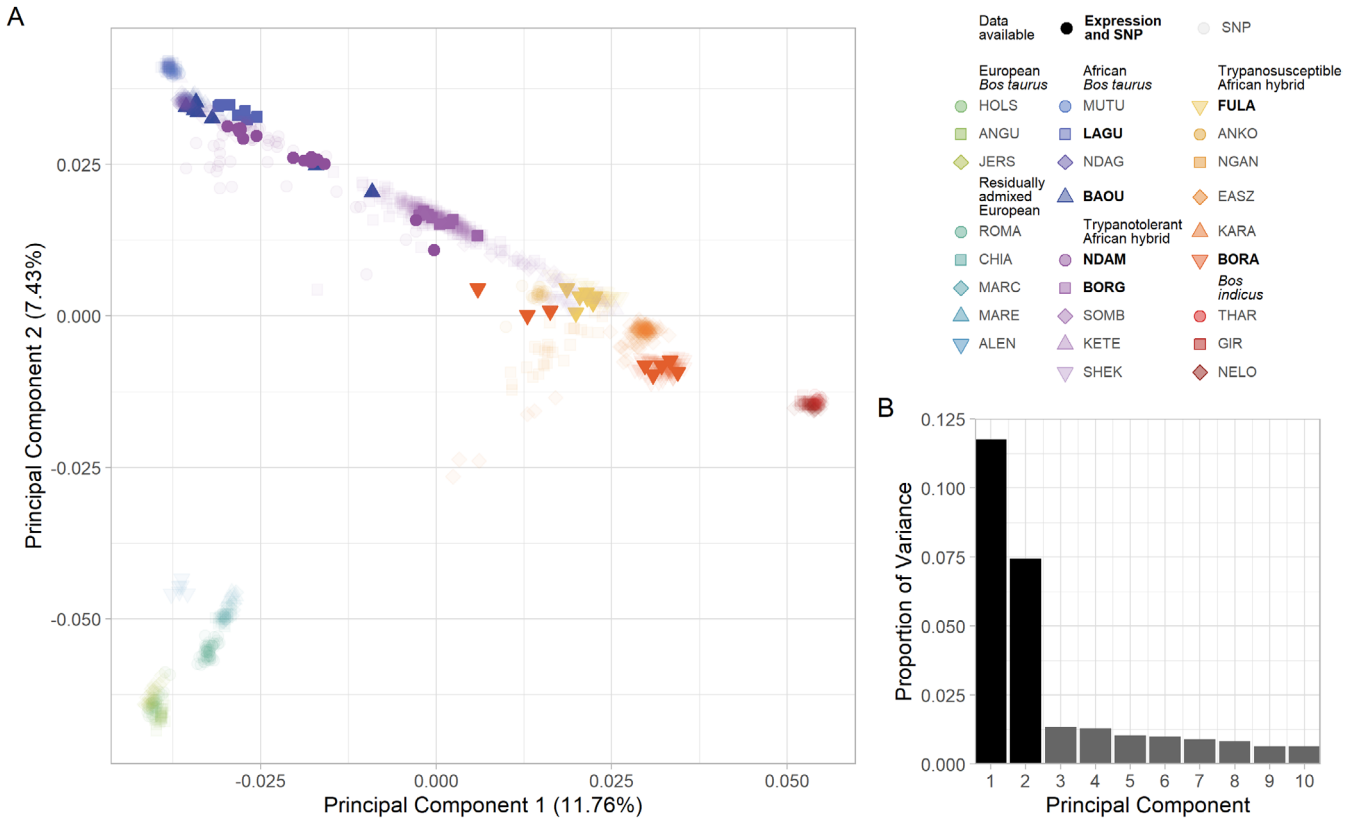

FIGURE 2 | (A) Principal component analysis of low-density SNP data for the cattle samples coloured according to population, showing the first two principal components and (B) bar chart of proportion of variance of the top 10 principal components. The transparency indicates the availability of gene expression data for the sample. A version of this figure without transparency is shown in Figure S3.

3.1.2 | Principal Component Analysis Recapitulates the Biogeography of Domestic Cattle

The first principal component (PC1) explained 11.76% of the total variation in the SNP data and separated the *B. taurus* and *B. indicus* lineages (Figure 2, Figure S3). The second principal component (PC2) explained an additional 7.43% of the total variation and separated the European *B. taurus* and African *B. taurus* groups (Figure 2). The hybrid and residually admixed animals emerged between the reference populations, with the residually admixed European animals clustering close to the European *B. taurus* populations and the African hybrid animals mostly dispersed between the African *B. taurus* and *B. indicus* populations (Figure 2). The trypanotolerant African hybrid animals are closest to the African *B. taurus* populations, whereas the trypanosusceptible African hybrid animals are closest to the *B. indicus* populations (Figure 2).

3.1.3 | Genetic Structure Analysis Highlights the Multiple Ancestries of African Cattle

The genetic structure results for the number of assumed populations (K) set to three partitioned the European and African *B. taurus* and *B. indicus* ancestries in the dataset (Figures S6–S9). The model complexity that maximised the marginal likelihood was 17, and the number of model components used to explain the structure in the data was 21.

3.1.4 | Local Ancestry Inference Revealed Peaks Across Admixed African Cattle Populations

Mean LAI values were estimated for each of the three ancestry components (European *B. taurus*, African *B. taurus* and *B. indicus*) for the six populations with available gene expression data (LAGU, BAOU, NDAM, BORG, FULA and BORA) (Figure 3). Genome-wide z-scores for the mean LAI results were used to identify SNPs with z-scores ≥ 2.0 for each population (Table S2). There were no SNPs that passed the $z \geq 2.0$ threshold for the African *B. taurus* ancestry component in the African *B. taurus* populations (LAGU and BAOU), whereas the trypanotolerant (NDAM and BORG) and trypanosusceptible (FULA and BORA) African hybrid populations had the lowest number of SNPs passing the $z \geq 2.0$ threshold for the African *B. taurus* and *B. indicus* ancestry components, respectively (Table S2). The European *B. taurus* ancestry components had the highest number of SNPs passing the $z \geq 2.0$ threshold in all six populations (Table S2).

3.1.5 | Introgressed Genomic Regions Showed Enrichment of Biologically Relevant Functional Categories

The proportions of the numbers of genes found within 1 Mb up- and downstream from each SNP with a z-score ≥ 2.0 are similar to those of the numbers of SNPs found for each ancestry

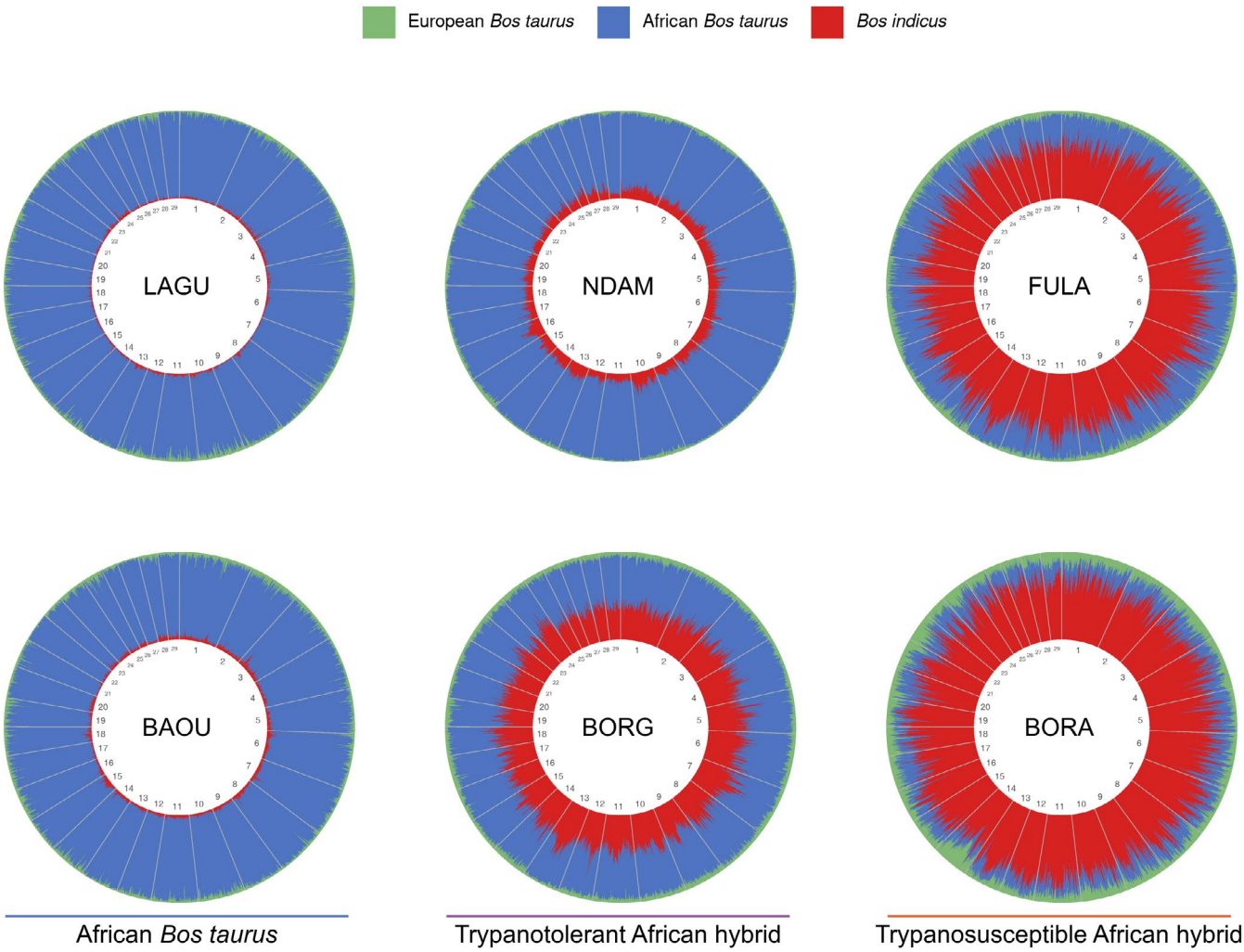

FIGURE 3 | Local ancestry plots showing mean European *Bos taurus*, African *B. taurus* and *Bos indicus* ancestry components using low-density SNP data for the six populations with gene expression data available across all autosomes. Each vertical line on the round genome plots represents a SNP and is coloured according to the ancestry results.

component in the populations studied (Tables S1 and S2). The top driver GO terms enriched for the European *B. taurus* genes in the African *B. taurus* populations with available gene expression data (LAGU, BAOU) included those related to alcohol dehydrogenase activity (*GO:0006069 ethanol oxidation* and *GO:0004024 alcohol dehydrogenase activity, zinc-dependent*) and aspartic-type endopeptidase activity (*GO:0004190 aspartic-type endopeptidase activity*). The top driver GO terms enriched for the *B. indicus* genes include those related to catabolic processes and catalytic activity (*GO:0030574 collagen catabolic process*, *GO:0006032 chitin catabolic process* and *GO:0004190 aspartic-type endopeptidase activity*); cell signalling (*GO:0051606 detection of stimulus*, *GO:0007186 protein-coupled receptor signalling pathway* and *GO:0004930 G protein-coupled receptor activity*); haemoglobin complex (*GO:0005833 haemoglobin complex*) and olfactory receptor activity (*GO:0004984 olfactory receptor activity*) (Figure S10). As there were no African *B. taurus* SNPs that passed the $z \geq 2$ threshold, there were no African *B. taurus* genes for functional enrichment in these populations (Tables S1 and S2).

The trypanotolerant African hybrid populations that had gene expression data available (NDAM and BORG) also had top driver GO terms enriched for the European *B. taurus* genes which related to alcohol dehydrogenase activity (*GO:0006069 ethanol oxidation* and *GO:0004024 alcohol dehydrogenase activity, zinc-dependent*) as well as chemotaxis (*GO:0050918 positive chemotaxis*, *GO:0031731 CCR6 chemokine receptor binding* and *GO:0042056 chemoattractant activity*); spliceosomal complex assembly (*GO:0000348 mRNA branch site recognition*, *GO:0005686 U2 snRNP* and *GO:0045131 pre-mRNA branch point binding*); cellular component organisation (*GO:0043933 protein-containing complex organisation* and *GO:0043229 intracellular organelle*); and collagen catabolic process (*GO:0030574 collagen catabolic process*) (Figure S11).

The top driver GO terms enriched for the *B. indicus* genes include those related to the skin and keratin (*GO:0031424 keratinisation*, *GO:0045109 intermediate filament organisation*, *GO:0045095 keratin filament*, *GO:0030280 structural constituent of skin epidermis* and *GO:0045103 intermediate filament-based process*); L-amino acid transmembrane transport (*GO:0097638 L-arginine import across plasma membrane*, *GO:1903352 L-ornithine transmembrane transport* and *GO:0000064 L-ornithine transmembrane transporter activity*); oxidoreductase activity

(*GO:0004499 N,N-dimethylaniline monooxygenase activity*, *GO:0047822 hypotaurine dehydrogenase activity*, *GO:0047023 androsterone dehydrogenase activity* and *GO:0047086 ketosteroid monooxygenase activity*) and aminoglycoside antibiotic metabolic process (*GO:0030647 aminoglycoside antibiotic metabolic process*) (Figure S11). As with the African *B. taurus* populations, there were too few African *B. taurus* genes for these populations for functional enrichment (Tables S1 and S2).

The trypanosusceptible African hybrid populations with gene expression data (FULA and BORA) had driver GO terms enriched for African *B. taurus* genes that related to the immune system, particularly the major histocompatibility complex (MHC) (*GO:0002396 MHC protein complex assembly*, *GO:0042613 MHC class II protein complex*, *GO:0023023 MHC protein complex binding*, *GO:0050829 defense response to Gram-negative bacterium*, *GO:0007156 homophilic cell adhesion* via *plasma membrane adhesion molecules*, *GO:0050870 positive regulation of T cell activation*, *GO:0050830 defense response to Gram-positive bacterium*, *GO:0002250 adaptive immune response* and *GO:0042605 peptide antigen binding*) (Figure S12). Other driver GO terms include those related to lysozyme activity (*GO:0003796 lysozyme activity* and *GO:0031902 late endosome membrane*); cell signalling (*GO:0007186 G protein-coupled receptor signalling pathway*) and olfactory receptor activity (*GO:0004984 olfactory receptor activity*) (Figure S12). There were no GO terms enriched for European *B. taurus* genes and only one driver GO term enriched for *B. indicus* genes for these populations (*GO:0007156 homophilic cell adhesion* via *plasma membrane adhesion molecules*) (Figure S12).

3.2 | Gene Expression Analysis Results and the Immunobiology of Trypanosomiasis

3.2.1 | Differentially Expressed Genes Included Immune Genes

The differential expression analysis of the RNA-seq data identified very few significant DEGs with only 45 genes, including 43 unique genes, across all response contrasts (Figure S13, Table S3). The majority of these were for the LAGU population at 40 days post-infection, which had 25 significantly DEGs (Figures S13 and S14, Table S3). The 40 dpi time point was also the only time point that had significant DEGs for all four populations (Figures S13–S17, Table S3). The duplicated genes among the significant DEGs included *NEIL2*, which had significantly decreased expression in the LAGU, NDAM and BORG samples at 40 dpi (Table S3). In addition, *LZTS3* had significantly decreased expression in the NDAM samples at 30 and 40 dpi, whereas *SLC11A1* had significantly increased expression in the NDAM samples at 20 and 30 dpi (Table S3).

3.2.2 | Network Analysis Identified Functionally Relevant Modules

The search of the GeneCards database (Stelzer et al. 2016) for genes relating to the term 'trypano*' generated a list of 1036 genes. The application of a filter to select genes with a relevance score ≥1.75 left 417 genes with valid bovine Ensembl IDs. The gene interaction network (GIN) generated using InnateDB (Breuer et al. 2013) with this list of genes contained 5666 genes (nodes) and 15,651 interactions (edges) (Figure S18). Modules were identified using jActiveModules within this base network for each of the final response contrasts for the microarray and RNA-seq data (Figure 4, Figures S19–S26). The smallest module identified was for the microarray blood samples at 34 days post-infection (MICRO BL 34) with 474 genes and 1166 interactions, whereas the largest module identified was for the microarray spleen samples at 35 days post-infection (MICRO SP 35) with 2446 genes and 6326 interactions (Figure 5). The MICRO SP 35 module also had the highest number of unique genes (Figure 5, Figure S27). There were eight genes with valid gene symbols that were present in all eight modules, these were *AGO2* (argonaute RISC catalytic component 2), *CBL* (Cbl proto-oncogene), *CNOT1* (CCR4-NOT transcription complex subunit 1), *EDN1* (endothelin 1), *IL1B* (interleukin 1 beta), *NFKB1* (nuclear factor kappa B subunit 1), *RIPK1* (receptor interacting serine/threonine kinase 1) and *TRAF2* (TNF receptor associated factor 2) (Figure 5, Table S4).

3.2.3 | Functional Enrichment of Gene Expression Data Highlighted the Immune System

There were no GO terms significantly enriched for the combined set of significant DEGs across all response contrasts in the RNA-seq dataset, relative to the background set of detectable genes. However, a single GO term (*GO:0030667 secretory granule membrane*) was significantly enriched when no background set was used (Figure S28). The GO terms significantly enriched for the genes in the modules identified for each of the final contrasts in the microarray and RNA-seq data against the background set of the base network are shown in Figure 6.

The driver GO terms significantly enriched for the genes in the MICRO BL 34 module include those related to the immune system (*GO:0002376 immune system process*, *GO:0034097 response to cytokine* and *GO:0050729 positive regulation of inflammatory response*) and cellular response to chemical stimulus (*GO:0070887 cellular response to chemical stimulus*) (Figure 6). The driver GO terms significantly enriched for the genes in the MICRO LN 35 module include those related to the cell cycle and cellular organisation (*GO:1903047 mitotic cell cycle process*, *GO:0043067 regulation of programmed cell death*, *GO:0006261 DNA-templated DNA replication*, *GO:0051240 positive regulation of multicellular organismal process*, *GO:2000147 positive regulation of cell motility*, *GO:0140513 nuclear protein-containing complex* and *GO:0005694 chromosome*) (Figure 6). Although no GO terms were significantly enriched for the genes in the MICRO LI 35 module against the background set of the base network, there were many GO terms significantly enriched when no background set was used (Figure S29). In addition, 491 of the 505 GO terms significantly enriched for the MICRO LI 35 module genes were also among the 2044 GO terms significantly enriched for the genes in the base network (Figure S30). The driver GO terms significantly enriched for the genes in the MICRO SP 35 module include those related to cellular responses and binding (*GO:0050794 regulation of cellular process*, *GO:0071478 cellular response to radiation*, *GO:0010033 response to organic substance*, *GO:0005634 nucleus*, *GO:0140096 catalytic activity,*

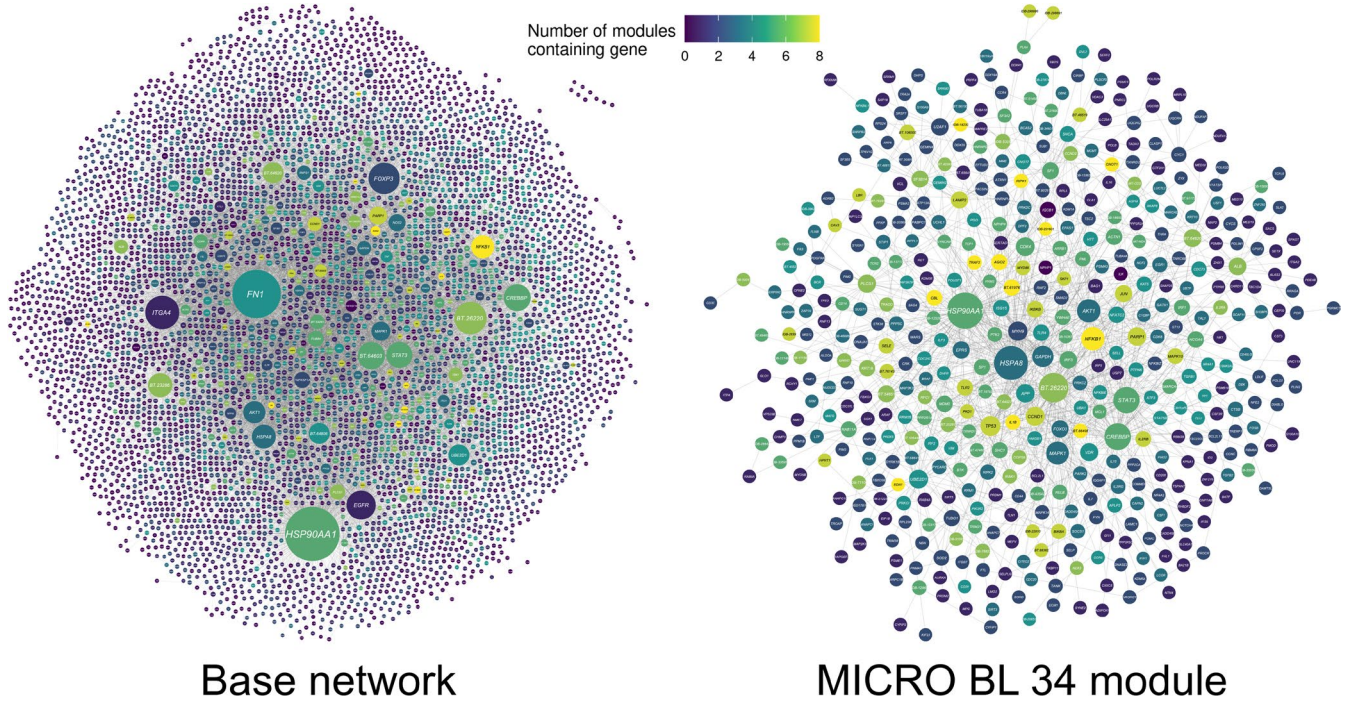

FIGURE 4 | Base network generated using InnateDB with the top results of a search of the GeneCards for genes relating to the term 'trypano*' and functional module identified using jActiveModules and the differential expression results for the MICRO BL 34 contrast. Each node in the networks represents a gene coloured according to the number of functional modules that contain that gene. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). The base network contains 5666 genes (nodes) and 15,651 interactions (edges), whereas the MICRO BL 34 module contains 474 genes and 1166 interactions.

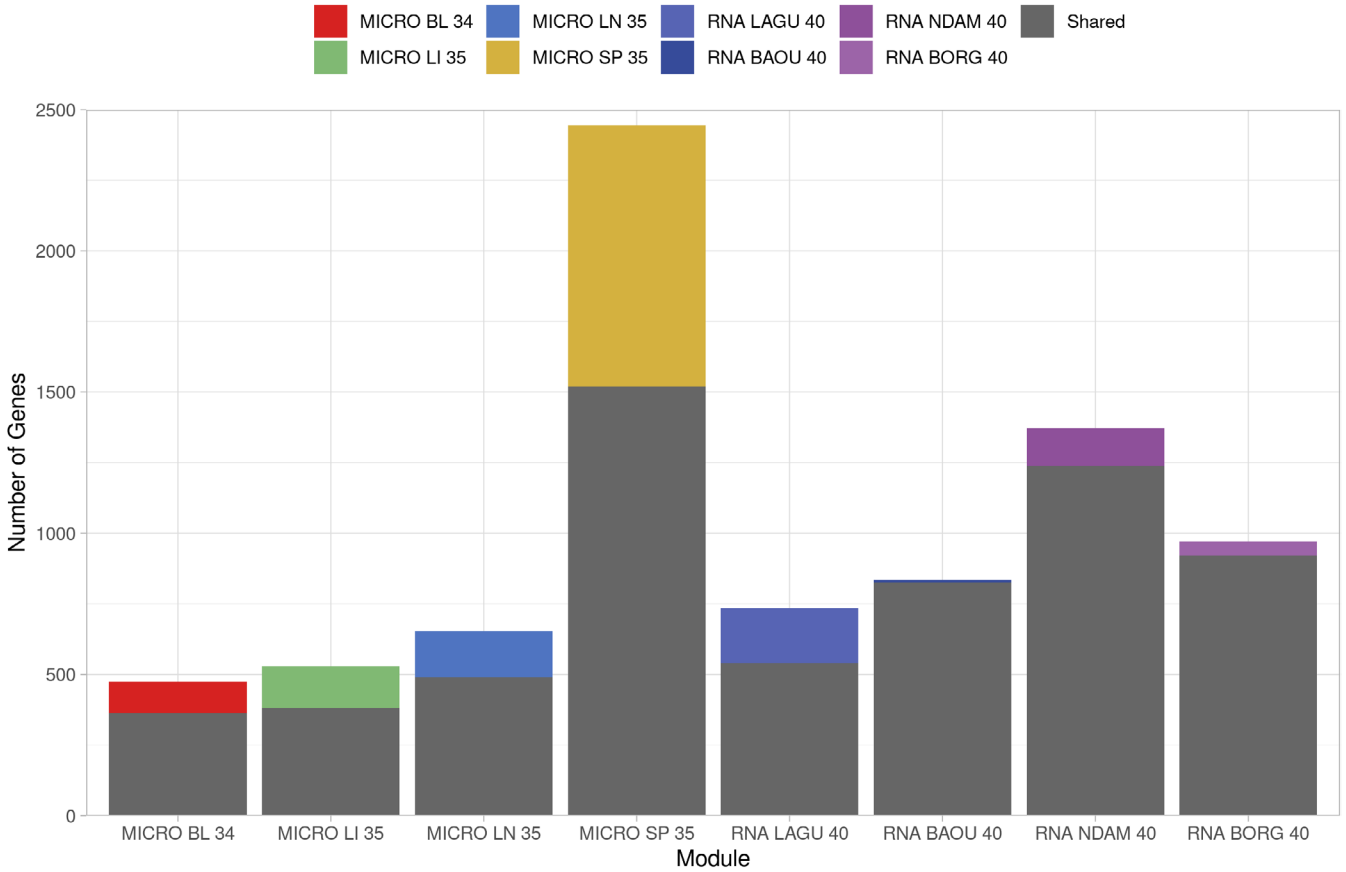

FIGURE 5 | Bar chart showing the number of genes in the functional modules identified using jActiveModules and the differential expression results for each of the final response contrasts for the microarray and RNA-seq data. The colours indicate the number of genes unique to each module or shared between multiple modules.

acting on a protein, *GO:0003677 DNA binding*, *GO:0000981 DNA-binding transcription factor activity, RNA polymerase II-specific*, *GO:0005515 protein binding* and *GO:0005524 ATP binding*) (Figure 6).

The driver GO terms significantly enriched for the genes in the RNA LAGU 40 module include those related to the immune system and cell signalling (*GO:0002682 regulation of immune system process*, *GO:0038061 non-canonical NF-kappaB signal transduction*, *GO:0006915 apoptotic process*, *GO:0048519 negative regulation of biological process*, *GO:0019079 viral genome replication*, *GO:0044271 cellular nitrogen compound biosynthetic process* and *GO:0002221 pattern recognition receptor signalling pathway*) (Figure 6). The driver GO terms significantly enriched for the genes in the RNA BAOU 40 module include those related to cell regulation, cell signalling and cytokine activity (*GO:0048518 positive regulation of biological process*, *GO:0005615 extracellular space*, *GO:0005102 signalling receptor binding*, *GO:0098772 molecular function regulator activity* and *GO:0005125 cytokine activity*) (Figure 6). The driver GO terms significantly enriched for the genes in the RNA NDAM 40 module include those related to cell signalling and tissue and organ development (*GO:0007154 cell communication*, *GO:0048513 animal organ development*, *GO:0009888 tissue development*, *GO:0001934 positive regulation of protein phosphorylation*, *GO:0005615 extracellular space* and *GO:0005102 signalling receptor binding*) (Figure 6). The driver GO terms significantly enriched for the genes in the RNA BORG 40 module include those related to cell signalling and regulation (*GO:0007166 cell surface receptor signalling pathway*, *GO:0030154 cell differentiation*, *GO:0141124 intracellular signalling cassette*, *GO:0019220 regulation of phosphate metabolic process*, *GO:0010628 positive regulation of gene expression*, *GO:0005102 signalling receptor binding*, *GO:0004888 transmembrane signalling receptor activity* and *GO:0140677 molecular function activator activity*) (Figure 6).

3.3 | Results From the Integration of Genomic and Gene Expression Data Showed Significant Functional Overlaps

A summary of the input data for the integration of the genomic and gene expression data are shown in Table S5. This integration generated several significant overlaps between the intervals within 1 Mb up- and downstream from genes with a mean local ancestry z-score ≥ 2.0 and the genes contained in the modules identified using the gene expression data (Table 3). These overlaps include the *B. indicus* ancestry of the trypanosusceptible FULA population, which was significantly enriched for genes in both the RNA BAOU 40 and RNA BORG 40 modules (Table 3). Similarly, the *B. indicus* ancestry of the trypanosusceptible populations in the original LAI analyses (McHugo, Ward, Ng'ang'a, et al. 2025) were significantly enriched for the genes in the MICRO LI 35 module (Table 3). This was true for all combinations of LAI software and SNP data density examined (Table 3) and indicates our integrative methodology can detect genomic loci associated with interpopulation differences in susceptibility to trypanosomiasis. The *B. indicus* ancestry of the selected residually admixed European populations analysed with ELAI using low-density SNP data (McHugo, Ward, Ng'ang'a, et al. 2025) was significantly enriched for genes in both the MICRO LN 35 and RNA LAGU 40 modules (Table 3). The African *B. taurus* ancestry of the selected trypanotolerant African hybrids analysed with ELAI using low-density SNP data (McHugo, Ward, Ng'ang'a, et al. 2025) was significantly enriched for genes in the MICRO SP 35 module (Table 3), again suggesting this approach can help dissect the genetic architecture of trypanotolerance. Finally, the European *B. taurus* ancestry of the trypanotolerant LAGU population was significantly enriched for genes in the MICRO LN 35 module (Table 3).

4 | Discussion

4.1 | Population Genomics Analyses Confirm the Recent Evolutionary History of Domestic Cattle

The results of the population genomics analyses of the SNP data were consistent with the results from our previous analysis (McHugo, Ward, Ng'ang'a, et al. 2025) and multiple published studies of hybrid cattle populations (Barbato et al. 2020; Decker et al. 2014; Hanotte et al. 2002; Kim et al. 2020; Ward et al. 2022). Visualisation of the PCA results by plotting PC1 and PC2 recovered the classic '*Bos* triangle' with the first two PCs explaining a high proportion of the total variation within the data and separating the reference European *B. taurus*, African *B. taurus* and *B. indicus* populations (Figure 2, Figures S3–S5). Also, as we previously observed (McHugo, Ward, Ng'ang'a, et al. 2025), the locations of the hybrid populations—nearer to the reference populations they share the most ancestry with—are in agreement with multiple published studies (Figure 2) (Bahbahani et al. 2017; Barbato et al. 2020; Berthier et al. 2015; Decker et al. 2014; Flori et al. 2014; Gautier et al. 2009; Sempere et al. 2015; Upadhyay et al. 2017; Verdugo et al. 2019; Ward et al. 2022; Wragg et al. 2022).

The results of the genetic structure analysis for $K = 3$ mirrored those of the PCA, with the separation of the African and European *B. taurus* and *B. indicus* populations, which is also consistent with our previous results (McHugo, Ward, Ng'ang'a, et al. 2025) (Figures S6–S9). Again, the hybrid populations clearly showed evidence of global admixture proportions that were in agreement with both their positions on the PCA and previous studies (Figure 2) (Bahbahani et al. 2017; Barbato et al. 2020; Berthier et al. 2015; Decker et al. 2014; Flori et al. 2014; Gautier et al. 2009; Sempere et al. 2015; Upadhyay et al. 2017; Verdugo et al. 2019; Ward et al. 2022; Wragg et al. 2022). As in our previous study (McHugo, Ward, Ng'ang'a, et al. 2025), the high modelled K values required to best explain the variation among the populations indicated that some of the populations are closely related to the point that they may not be genetically distinct populations (Raj et al. 2014).

The results of the LAI analysis demonstrated similar patterns of peaks and troughs dispersed across the genome for each population as those found in our previous analysis (McHugo, Ward, Ng'ang'a, et al. 2025) using ELAI with high- and low-density SNP data and using MOSAIC with high-density SNP data (Figures 3 and 7). The African *B. taurus* populations (LAGU and BAOU) exhibited similar low peaks of European *B. taurus* and *B. indicus* ancestry, whereas the African trypanotolerant hybrid NDAM population had slightly higher *B. indicus* peaks

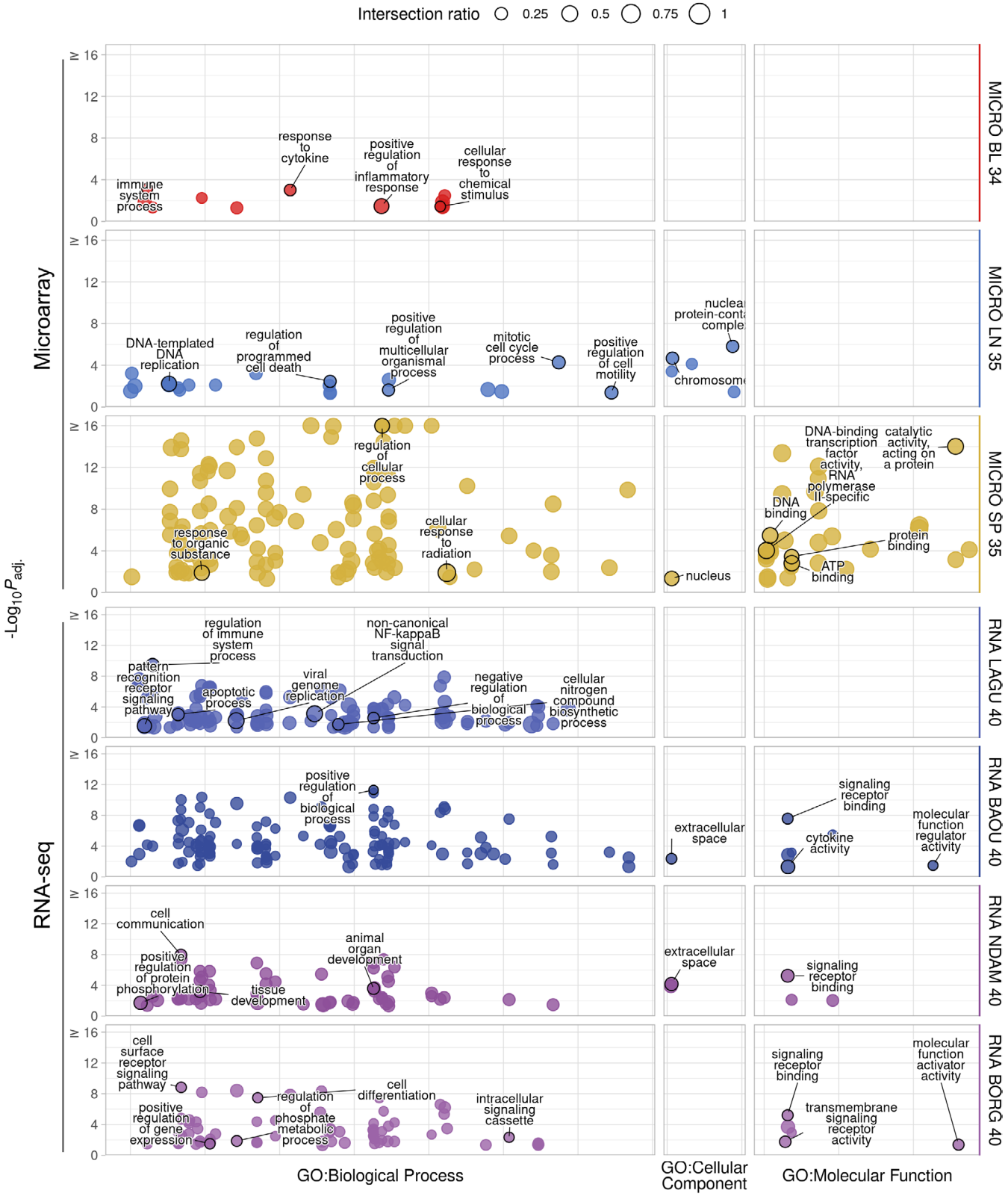

FIGURE 6 | g:Profiler functional enrichment of the genes in the functional modules identified using jActiveModules and the differential expression results for each of the final response contrasts for the microarray and RNA-seq data. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The *y*-axis shows the $-\log_{10}p_{adj.}$ value up to a maximum of 16, and the panels along the *y*-axis and colours indicate the module. The panels along the *x*-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of 10 are indicated with a black outline and label.

TABLE 3 | Corrected *p*-values for the statistically significant overlaps between the intervals within 1 Mb up- and downstream from genes with a mean local ancestry *z*-score ≥ 2.0 and the genes contained in the modules identified using the gene expression data.

Local ancestry						
Analysis	**Software**	**Data**	**Group/population**	**Ancestry**	**Module**	**Corrected *p***
a	ELAI	LD	FULA	*Bos indicus*	RNA BAOU 40	0.003
b	ELAI	LD	Selected residually admixed European	*Bos indicus*	MICRO LN 35	0.005
b	ELAI	HD	Trypanosusceptible African hybrids	*Bos indicus*	MICRO LI 35	0.011
b	ELAI	LD	Trypanosusceptible African hybrids	*Bos indicus*	MICRO LI 35	0.012
b	ELAI	LD	Selected residually admixed European	*Bos indicus*	RNA LAGU 40	0.013
b	MOSAIC	HD	Trypanosusceptible African hybrids	*Bos indicus*	MICRO LI 35	0.036
a	ELAI	LD	FULA	*Bos indicus*	RNA BORG 40	0.039
b	ELAI	LD	Selected trypanotolerant African hybrids	African *Bos taurus*	MICRO SP 35	0.043
a	ELAI	LD	LAGU	European *Bos taurus*	MICRO LN 35	0.050

Note: The local ancestry results are broken down into analysis, software used, SNP data density, group/population, and ancestry components. a: This study, b: (McHugo, Ward, Ng'ang'a, et al. 2025).

(Figure 3). This is consistent with previous studies of the origins of these populations as well as the results of the population genomics analyses (Figure 2, Figures S3–S9) (Barbato et al. 2020; Berthier et al. 2015; Gautier et al. 2009; Verdugo et al. 2019). The African trypanotolerant hybrid BORG population showed even higher peaks of *B. indicus* ancestry, which was markedly similar to the mean LAI results obtained previously for the selected trypanotolerant hybrid group (McHugo, Ward, Ng'ang'a, et al. 2025) (Figure 3). This was not unexpected, as this group included some of the same samples of the BORG population. In a similar manner, both the LAI results of the trypanosusceptible African hybrid populations (FULA and BORA) resembled the mean LAI results for the trypanosusceptible African hybrid group from our previous results (McHugo, Ward, Ng'ang'a, et al. 2025); again, this group included some of the same BORA samples (Figure 3). Consequently, the proportions of the numbers of SNPs passing the genome-wide *z*-score ≥ 2.0 threshold, as well as the numbers of genes within 1 Mb up- and downstream from these SNPs, were similar to our previous results (McHugo, Ward, Ng'ang'a, et al. 2025) (Tables S1 and S2). The lack of SNPs, and therefore genes, that passed the threshold for the African *B. taurus* ancestry component in the African *B. taurus* populations was likely due to the high and relatively uniform proportion of the African *B. taurus* ancestry component across the genome. This would result in no SNPs exceeding the threshold of two standard deviations from the mean (Figure 3, Tables S1 and S2). Similarly, the lower numbers of SNPs that exceeded the threshold for the African *B. taurus* and *B. indicus* ancestry components for the trypanotolerant and trypanosusceptible African hybrid populations, respectively, were likely due to the higher proportions of the reference population ancestries to which each hybrid group is most closely related (Figure 3, Tables S1 and S2).

The introgressed genomic regions also showed similar patterns in terms of functional enrichment as our previous LAI results (McHugo, Ward, Ng'ang'a, et al. 2025), with fewer or even no GO terms enriched for the ancestry component most similar to the population (Figures S10–S12). Significant driver GO terms enriched for genes near the peaks of the other ancestry contributions included those relating to the MHC and other components of the immune system, which was also in agreement with previous studies of local ancestry in cattle (Figures S10–S12) (Buggiotti et al. 2021; Chen et al. 2020; Guan et al. 2022; Li et al. 2023). It is notable that immune system functions were well represented in the top functional enrichment categories for the introgressed genomic regions since there are well documented differences among European *B. taurus*, African *B. taurus* and *B. indicus* cattle populations in terms of susceptibilities to various infectious diseases including bovine tuberculosis caused by *M. bovis* (Allen et al. 2010; Lee et al. 2024); East Coast fever and tropical theileriosis caused by *Theileria parva* and *Theileria annulate*, respectively (Bahbahani and Hanotte 2015) and African animal trypanosomiasis (AAT) caused by *Trypanosoma* spp. (Yaro et al. 2016). Similarly, it is notable that, as for the previous LAI analysis (McHugo, Ward, Ng'ang'a, et al. 2025), significant driver GO terms relating to olfaction were identified and genes related to olfaction have been identified by previous local ancestry studies of both cattle and humans and by selection signature studies in cattle (Figures S10–S12) (Melo et al. 2017; Pan et al. 2022; Sun et al. 2023; Yang et al. 2014). Although this may simply be due to the large number of olfactory receptor

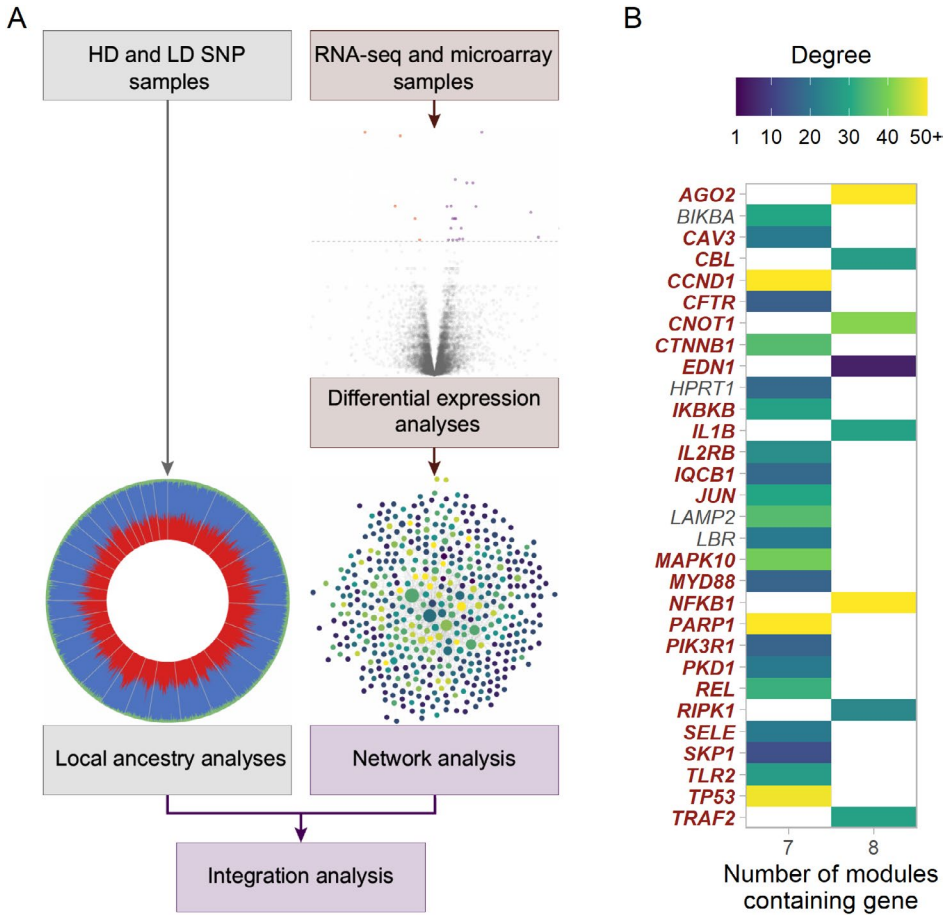

FIGURE 7 | (A) Diagram showing the study workflow and (B) heatmap showing the degree or number of interactions in the base network for the top 30 genes with valid gene symbols that were found in seven or more of the eight functional modules. Genes in a red bold font were also found within the peaks of the local ancestry analyses.

genes dispersed throughout the cattle genome (Lee et al. 2013; Niimura and Nei 2007), recent studies have suggested that up to 580 olfactory receptors may be expressed by macrophages, immune cells involved in the detection and phagocytosis of pathogens (Orecchioni et al. 2022). In addition, macrophages are the host's first line of defence to mycobacterial infection, with evasion and reprogramming of host macrophages being a key component of intracellular mycobacterial infections that cause tuberculosis disease (Hall et al. 2024), and variation in olfactory receptor genes was found to be associated with bovine tuberculosis (Ring et al. 2019).

Other significant driver GO terms detected in the present study that were common with the previous LAI analysis (McHugo, Ward, Ng'ang'a, et al. 2025) included those relating to cell signalling, L-amino acid transmembrane transport, oxidoreductase activity, metabolic processes, cellular component organisation, haemoglobin complex and homophilic cell adhesion via plasma membrane adhesion molecules (Figures S10–S12). A notable difference from our previous results was the enrichment of GO terms related to the skin and keratin among genes near peaks of *B. indicus* ancestry in the trypanotolerant hybrid populations (NDAM and BORG) (Figure S11). This may be a result of selection for heat tolerance traits such as the 'slick' phenotype associated with mutation in the prolactin receptor gene (*PRLR*) in Latin American Criollo cattle, which derive some of their ancestry from African cattle (Gebeyehu et al. 2025; Porto-Neto et al. 2018; Ward et al. 2024).

4.2 | Differential Expression Analysis Highlighted Immune Genes

As with the microarray data (McHugo, Ward, Browne, et al. 2025), and as expected due to their design, the use of response contrasts to identify changes in expression over time in the trypanotolerant population samples (LAGU, BAOU, NDAM and BORG) relative to the trypanosusceptible FULA in the differential expression analysis of the RNA-seq data found lower numbers of significant DEGs than the direct contrasts between or within the populations (Peylhard et al. 2023; Rue-Albrecht et al. 2014). However, despite the low numbers, the significant DEGs followed the established pattern also observed for the microarray data and direct RNA-seq contrasts, with increasing numbers of significant DEGs over the course of the infection, leading to the highest numbers of significant DEGs observed for the final time point (O'Gorman et al. 2009; Peylhard et al. 2023). In addition, it is known that trypanotolerant breeds with high levels of *B. taurus* ancestry have enhanced abilities to control anaemia, whereas hybrid animals exhibit intermediate levels of control when compared to susceptible *B. indicus* breeds (Berthier et al. 2015). As the LAGU, BAOU and

NDAM populations are known to have higher levels of *B. taurus* ancestry, which has also been shown by the population genomics analyses, it is therefore unsurprising that these populations exhibited more significant DEGs than the more admixed BORG population (Bahbahani et al. 2017; Barbato et al. 2020; Berthier et al. 2015; Flori et al. 2014; Gautier 2015; Verdugo et al. 2019; Ward et al. 2022).

For the RNA-seq results, as with the microarray analysis (McHugo, Ward, Browne, et al. 2025), there were overlaps in the significant DEGs between the contrast types. These include the duplicated genes *NEIL2* (nei like DNA glycosylase 2), which is involved in the immune system, the tumour suppressor *LZTS3* (leucine zipper tumour suppressor family member 3) and *SLC11A1* (solute carrier family 11 member 1), which encodes a regulator of iron homeostasis in macrophages (Archer et al. 2015; Gu et al. 2023; Tapryal et al. 2023). In particular, variants of *SLC11A1* are associated with susceptibility to infectious diseases including tuberculosis in cattle (Archer et al. 2015; Holder et al. 2020). Of the 43 unique genes identified as significantly differentially expressed by the response contrasts, only 10 were not identified as significantly differentially expressed by any of the direct contrasts used by Peylhard et al. (2023). These genes include *ANXA6* (annexin A6), *NT5C2* (5′-nucleotidase, cytosolic II), *PIGR* (polymeric immunoglobulin receptor) and *TET2* (tet methylcytosine dioxygenase 2), which were identified as significantly differentially expressed for various contrasts using a gene expression microarray (McHugo, Ward, Browne, et al. 2025) and are known to play roles in the immune system and haematological disorders (Cong et al. 2021; Jordheim 2018; Mercher et al. 2012; Rashidi et al. 2023; Sphyris and Mani 2011). Notably, anaemia is the main cause of death due to trypanosomiasis, and the ability to control this anaemia is thought to be critical to trypanotolerance in cattle (Naessens 2006; Stijlemans et al. 2015).

Other genes, which were not detectable with microarray data, included *GPATCH2L* (g-patch domain containing 2 like), *NFAT5* (nuclear factor of activated T cells 5), *PICALM* (phosphatidylinositol binding clathrin assembly protein), *PTPN4* (protein tyrosine phosphatase non-receptor type 4), *SLC38A4* (solute carrier family 38 member 4) and *TTBK2* (tau tubulin kinase 2). Interestingly, *TTBK2* has been identified as a candidate gene underlying trypanotolerance in the Sheko cattle breed, and mutations in this gene have been hypothesised to be associated with response to the presence of trypanosome parasites in the brain white matter, cerebral fluid, thyroid and parathyroid glands (Mekonnen et al. 2019). In addition, *NFAT5* encodes a member of a group of transcription factors shown to mediate cytokine production during trypanosome infection (Kayama et al. 2009), whereas PICALM interacts with phosphatidylinositol 4-kinases that are thought to be drug targets for human trypanosomiasis (Li et al. 2024). The *PTPN4* gene is involved in the immune system (Huai et al. 2015), and *GPATCH2L* has been highlighted in a study of haemorrhagic fever (Redwan et al. 2019). Finally, *SLC38A4* is related to both *SLC40A1* and *SLC11A1*, which were observed to be differentially expressed in the microarray data (McHugo, Ward, Browne, et al. 2025) and direct RNA-seq contrasts, respectively (Peylhard et al. 2023). Finally, the lack of enriched GO terms for the significant DEGs from the RNA-seq data was likely due to the low number of genes.

4.3 | Network Analysis Identified Modules of Genes Involved in Immunobiology

The base GIN network for genes associated with trypanosome infection and trypanosomiasis contained a similar number of genes and interactions as previous GINs generated by Hall et al. (2024) and Hall et al. (2021) for integrative genomics studies of *M. bovis* infection in cattle. In common with many GINs obtained using transcriptomics data, we observed a scale-free topology for functional modules (Yang 2020). This indicates that most genes within each functional module interact with only one other gene, whereas a small subset of genes interact with substantially more (Albert 2005; Barabasi and Albert 1999; Hall et al. 2024). However, the functional modules identified for each of the final contrasts in the microarray and RNA-seq data in the present study were considerably larger than those identified by Hall et al. (2024) and Hall et al. (2021) (Figures 5 and 7). When it is unable to adequately detect larger subnetworks, the jActiveModules algorithm can still retrieve subnetworks consisting of a small number of genes, or even single genes; therefore, the large modules identified in the present study indicate that the microarray and RNA-seq datasets are robust (Ideker et al. 2002). In addition, modules identified from RNA-seq data tend to be larger than those identified from microarray data (Hatem et al. 2012), and this was the case for all the modules, apart from that identified using the final spleen microarray contrast. The large number of genes shared between the modules is also unsurprising, as they were drawn from the same base network. Also, the significant DEGs identified using the microarray data had a large number of genes in common between the tissues (McHugo, Ward, Browne, et al. 2025). Our results are also in agreement with previous studies that observed high numbers of genes in common between modules identified from a shared base network (Hall et al. 2024, 2021).

The eight genes shared between all the modules include genes related to the immune system that have already been highlighted by previous studies of response to trypanosome infection in cattle such as *IL1B* (O'Gorman et al. 2006) and *NFKB1* (O'Gorman et al. 2006; Peylhard et al. 2023) (Figures 5 and 7, Table S4). Other genes have been similarly highlighted by studies of trypanosome infection in mice, including *EDN1* (Corral et al. 2013) and *TRAF2* (Santamaría et al. 2023), as well as by studies of other bovine infections, such as *CBL*, which was highlighted by a study of bovine anaplasmosis caused by *Anaplasma marginale* (Ahlawat et al. 2023) and *CNOT1*, which was highlighted by a study of Johne's disease caused by *Mycobacterium avium paratuberculosis* (MAP) (Kleinwort et al. 2019). Finally, *RIPK1* is known to play a role in inflammation (Newton 2015), whereas *AGO2* has recently been found to regulate immune responses (Wang et al. 2025). The driver GO terms significantly enriched for the genes in the microarray modules were similar to those found to be enriched for the significant DEGs in the microarray results outputs used to generate the modules (Figure 6) (McHugo, Ward, Browne, et al. 2025). Therefore, the presence of driver GO terms related to the immune system, and cytokines in particular, among those significantly enriched for genes in the MICRO BL 34 is expected as these genes were significantly differentially expressed for both the microarray

data (McHugo, Ward, Browne, et al. 2025), the RNA-seq data and other published studies (Noyes et al. 2011; O'Gorman et al. 2009, 2006; Peylhard et al. 2023; Uzonna et al. 1999). In addition, the anaemia caused by trypanosome infection can be considered an immune response, which may be driven by cytokines (Stijlemans et al. 2015, 2008, 2010). The enrichment of driver GO terms related to the cell cycle and cellular organisation for genes in the MICRO LN 35 module was also consistent with the results of the analysis of the microarray data (Noyes et al. 2011) (Figure 6). The lack of GO terms significantly enriched for genes in the MICRO LI 35 module against the background set of the base network was likely due to the similarity between the module and the base network, as evidenced by the almost total overlap in significantly enriched GO terms with the base network when no background set was used (Figures S29 and S30). The driver GO terms significantly enriched for the genes in the MICRO SP 35 module included those related to cellular responses and binding (Figure 6). This aligns with the small number of GO terms enriched for the small number of significant DEGs in the spleen samples in the previous microarray results (McHugo, Ward, Browne, et al. 2025) and highlights the utility of the jActiveModules method for incorporating genes that are highly interconnected with the significant DEGs but that were otherwise not detected by the differential expression analysis (Hall et al. 2024, 2021). The power of a network-based approach using jActiveModules was also evident from the modules identified from the RNA-seq data, where fewer significant DEGs were detected than in the microarray data, indicating that incorporating highly interconnected genes was more beneficial.

The driver GO terms significantly enriched for the genes in the RNA LAGU 40 module included those related to the immune system and cell signalling, whereas the driver GO terms significantly enriched for the genes in the RNA BAOU 40 module included those related to cell regulation, cell signalling and cytokine activity (Figure 6). The driver GO terms significantly enriched for the genes in the RNA NDAM 40 module also included those related to cell signalling, in addition to tissue and organ development. The driver GO terms significantly enriched for the genes in the RNA BORG 40 module also included those related to cell signalling and regulation (Figure 6). That driver GO terms related to cell signalling and regulation were common among the terms enriched for the genes in the modules is logical, as the modules were identified from a base network made of interacting genes that are frequently linked by cell signalling and regulatory pathways. Similarly, the presence of driver GO terms related to the immune system among those significantly enriched for the genes in the RNA-seq modules is to be expected, as the modules were identified using the differential expression results from an infection experiment that identified several immune genes as significantly differentially expressed.

4.4 | Integration Revealed Significant Overlap in the Results From the Genomic and Gene Expression Data

The presence of genes related to the immune system in both the modules identified with the microarray and RNA-seq data and the LAI results may explain the significant enrichment found in the intervals within 1 Mb up- and downstream from genes with a mean local ancestry z-score ≥ 2.0 and the genes contained in the modules identified using the gene expression data (Figure 7, Table 3). Notably, immune genes were located in genomic regions around the peaks of the LAI results for hybrid cattle populations, and these regions were also enriched for genes identified by functional modules of response contrasts for the same hybrid populations, using both RNA-seq and microarray gene expression data across multiple tissues and populations.

We hypothesised that the genes within 1 Mb of the top African *B. taurus* ancestry peaks in the trypanotolerant populations could underpin the complex trypanotolerance trait, particularly if these regions showed significant enrichment for the genes identified during the differential expression and network analyses. However, there was only one such significant overlap: the African *B. taurus* ancestry of the selected trypanotolerant African hybrids analysed with ELAI using low-density SNP data (McHugo, Ward, Ng'ang'a, et al. 2025) was significantly enriched for genes in the MICRO SP 35 module (Table 3). These genes may therefore be considered as candidate genes for trypanotolerance. Although it must also be noted that the MICRO SP 35 module was the largest, and the number of genes contained in this module may have increased the probability of overlap with the local ancestry peaks, despite the p-value correction to account for the number of genes examined (Lee et al. 2012).

Interestingly, the European *B. taurus* ancestry of the trypanotolerant LAGU population was enriched for genes in the MICRO LN 35 module ($p_{adj.} = 0.05$; Table 3). Although it is possible that these genes may also be candidate genes for trypanotolerance as they have been found within peaks of *B. taurus* ancestry, which can also be considered troughs of trypanosusceptible *B. indicus* ancestry, this is unlikely. The pure reference European and African *B. taurus* populations were clearly separated by the various population genomics analyses, and it is likely that the LAI algorithms could distinguish between the trypanosusceptible European *B. taurus* and trypanotolerant African *B. taurus* ancestry. In a similar manner, the peaks of *B. indicus* ancestry in the selected residually admixed European populations analysed with ELAI using low-density SNP data (McHugo, Ward, Ng'ang'a, et al. 2025) were significantly enriched for genes in both the MICRO LN 35 and RNA LAGU 40 modules (Table 3). A possible explanation for this is that the immune genes in these modules overlapped with those in the regions around the local ancestry peaks of *B. indicus* ancestry, as these regions were enriched for GO terms relating to the MHC. This makes sense as increased diversity in the MHC region is associated with disease resistance, and MHC genes are under balancing selection in cattle (Codner et al. 2012; Ellis 2004; Ellis and Hammond 2014).

It is also notable that the peaks of *B. indicus* ancestry in several trypanosusceptible populations were significantly enriched for genes in the functional modules. These overlaps include the trypanosusceptible FULA population, which was significantly enriched for genes in both the RNA BAOU 40 and RNA BORG 40 modules (Table 3). Similarly, the peaks of *B. indicus* ancestry in the trypanosusceptible populations in the original LAI analyses (McHugo, Ward, Ng'ang'a, et al. 2025) were

significantly enriched for the genes in the MICRO LI 35 module for all combinations of LAI software and SNP data density examined (Table 3). It is also possible that the genes in these additional functional modules can be considered candidate genes for trypanotolerance, as they are found in peaks of trypanosusceptible *B. indicus* ancestry across multiple trypanosusceptible hybrid populations. This implies that these regions of increased trypanosusceptible *B. indicus* introgression in multiple trypanosusceptible hybrid populations were significantly enriched for genes identified as part of downstream network analyses of differential expression comparing the responses of trypanosusceptible and trypanotolerant cattle to trypanosome infection using multiple populations, tissues and data types. Therefore, because these genomic regions were found to be significantly enriched for multiple functional modules, it suggests that the genes shared by multiple modules may be prioritised as candidate genes for trypanotolerance. The similarity of these genomic regions is highlighted by the fact that the driver GO term *GO:0007156 homophilic cell adhesion* via *plasma membrane adhesion molecules*, and the upstream GO term *GO:0098742 cell–cell adhesion* via *plasma-membrane adhesion molecules* were enriched for the genes located in the regions around the peaks of *B. indicus* ancestry in the FULA population and the trypanosusceptible African hybrid populations in the LAI analysis using high-density SNP data with MOSAIC and low-density SNP data with ELAI (McHugo, Ward, Ng'ang'a, et al. 2025). The *GO:0098742* term was also found in another study to be enriched for genes in introgressed African taurine genomic regions in hybrid African cattle populations (Friedrich et al. 2023).

4.5 | Limitations of This Study

The complex nature of the trypanotolerance trait, coupled with the many selective forces exerted by other pathogens on the immune systems of hybrid African cattle, means these results must be interpreted with caution. For example, trypanotolerant breeds are also less susceptible to other infectious diseases, such as helminthiasis and tick-borne diseases, and genes associated with the immune response have been found to be under selection in West African cattle populations (Gautier et al. 2009; Kim, Ka, et al. 2017). This suggests an alternative hypothesis such that the immune genes found in the peaks of African *B. taurus* ancestry in the trypanotolerant populations may be associated with resistance to another infectious disease. In addition, an alternative hypothesis can be proposed that the genes enriched for *B. indicus* ancestry in the trypanosusceptible populations may confer resistance to bovine tuberculosis (bTB), rather than trypanotolerance; *B. indicus* cattle have long been known to have lower susceptibility to bTB compared to *B. taurus* cattle (Allen et al. 2010; Liston and Soparkar 1917). *Bos indicus* and *B. taurus* cattle exhibit differing immune responses to infection with *M. bovis*, the causative agent of bTB, and the genes underlying this polygenic disease resistance trait remain unknown (Kumar et al. 2023; Vordermeier et al. 2012). Notably, integration of data from different experiments involving infectious diseases in cattle, including trypanosomiasis and tuberculosis, found many of the same genes were differentially expressed, suggesting common immune mechanisms in response to these infections (Beiki et al. 2016, 2018). Pleiotropy of immune genes, as well as selection acting on other physiological processes, such as adaptation to African environments (e.g., aridity and heat tolerance), should also be considered when interpreting the results. There is also the potential for false positives and spurious links of selection studies to gene functions that must be considered, despite the suggestive results apparent from the results of both the local ancestry and gene expression analyses. Another limitation of this study is the relatively low number of SNPs (29,869) used in the genomic analyses, which may cause problems due to low SNP density for populations with older admixture histories. However, due to linkage disequilibrium (LD) and the low effective population sizes (N_e) of the populations, this is unlikely to be a major issue. Future work will expand the information available for these and additional populations to include whole-genome sequence data, and analysis of this data may generate further insights. Although previous studies may have had difficulty identifying candidate genes in hybrid populations due to the mixture of different ancestries affecting genome signatures, local ancestry analysis offers an elegant and powerful approach by detecting selectively retained haplotypes or ancestral segments within hybrid populations.

In conclusion, integration of genomic data in the form of high- and low-density SNP data from a range of trypanotolerant and trypanosusceptible cattle populations with RNA-seq and microarray transcriptomics data from the same populations has provided a new approach for identification of trypanotolerance candidate genes.

Acknowledgements

We thank Morris Agaba, Olivier Hanotte, Stephen J. Kemp, Daniel G. Bradley and Stephen V. Gordon for assistance with sample resources and for useful scientific discussion. This research work was funded by Science Foundation Ireland (SFI) under Investigator Programme Awards (grant nos: SFI/01/F.1/B028 and SFI/15/IA/3154). J.A.W. was supported by the Centre for Research Training in Genomics Data Science (grant no. SFI/18/CRT/6214).

Funding

This research work was funded by Science Foundation Ireland (SFI; sfi.ie) under Investigator Programme Awards (grant nos: SFI/01/F.1/B028 and SFI/15/IA/3154).

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

No new data were generated for this study. The computer code required to repeat and reproduce the analyses is available at http://doi.org/10.5281/zenodo.11517978.

References

Achukwi, M. D., V. N. Tanya, E. W. Hill, et al. 1997. "Susceptibility of the Namchi and Kapsiki Cattle of Cameroon to Trypanosome Infection." *Tropical Animal Health and Production* 29, no. 4: 219–226. https://doi.org/10.1007/BF02632308.

Ahlawat, S., V. Choudhary, R. Kaur, et al. 2023. "Unraveling the Genetic Mechanisms Governing the Host Response to Bovine Anaplasmosis." *Gene* 877: 147532. https://doi.org/10.1016/j.gene.2023.147532.

Akol, G. W., and M. Murray. 1982. "Early Events Following Challenge of Cattle With Tsetse Infected With *Trypanosoma congolense*: Development of the Local Skin Reaction." *Veterinary Record* 110, no. 13: 295–302. https://doi.org/10.1136/vr.110.13.295.

Albert, R. 2005. "Scale-Free Networks in Cell Biology." *Journal of Cell Science* 118, no. Pt 21: 4947–4957. https://doi.org/10.1242/jcs.02714.

Allen, A. R., G. Minozzi, E. J. Glass, et al. 2010. "Bovine Tuberculosis: The Genetic Basis of Host Susceptibility." *Proceedings of the Royal Society B: Biological Sciences* 277, no. 1695: 2737–2745. https://doi.org/10.1098/rspb.2010.0830.

Alvarez, I., L. Perez-Pardal, A. Traore, I. Fernandez, and F. Goyache. 2016. "Lack of Haplotype Structuring for Two Candidate Genes for Trypanotolerance in Cattle." *Journal of Animal Breeding and Genetics* 133, no. 2: 105–114. https://doi.org/10.1111/jbg.12181.

Archer, N. S., N. T. Nassif, and B. A. O'Brien. 2015. "Genetic Variants of *SLC11A1* Are Associated With Both Autoimmune and Infectious Diseases: Systematic Review and Meta-Analysis." *Genes and Immunity* 16, no. 4: 275–283. https://doi.org/10.1038/gene.2015.8.

Bache, S. M., and H. Wickham. 2022. "magrittr: A Forward-Pipe Operator for R." https://magrittr.tidyverse.org.

Bahbahani, H., A. Afana, and D. Wragg. 2018. "Genomic Signatures of Adaptive Introgression and Environmental Adaptation in the Sheko Cattle of Southwest Ethiopia." *PLoS One* 13, no. 8: e0202479. https://doi.org/10.1371/journal.pone.0202479.

Bahbahani, H., and O. Hanotte. 2015. "Genetic Resistance: Tolerance to Vector-Borne Diseases and the Prospects and Challenges of Genomics." *Revue Scientifique et Technique, Office International des Epizooties* 34, no. 1: 185–197. https://doi.org/10.20506/rst.34.1.2353.

Bahbahani, H., A. Tijjani, C. Mukasa, et al. 2017. "Signatures of Selection for Environmental Adaptation and Zebu × Taurine Hybrid Fitness in East African Shorthorn Zebu." *Frontiers in Genetics* 8: 68. https://doi.org/10.3389/fgene.2017.00068.

Barabasi, A. L., and R. Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286, no. 5439: 509–512. https://doi.org/10.1126/science.286.5439.509.

Barbato, M., F. Hailer, M. Upadhyay, et al. 2020. "Adaptive Introgression From Indicine Cattle Into White Cattle Breeds From Central Italy." *Scientific Reports* 10, no. 1: 1279. https://doi.org/10.1038/s41598-020-57880-4.

Becker, M. Y., and I. Rojas. 2001. "A Graph Layout Algorithm for Drawing Metabolic Pathways." *Bioinformatics* 17, no. 5: 461–467. https://doi.org/10.1093/bioinformatics/17.5.461.

Beiki, H., A. Nejati-Javaremi, A. Pakdel, A. Masoudi-Nejad, Z. L. Hu, and J. M. Reecy. 2016. "Large-Scale Gene Co-Expression Network as a Source of Functional Annotation for Cattle Genes." *BMC Genomics* 17, no. 1: 846. https://doi.org/10.1186/s12864-016-3176-2.

Beiki, H., A. Pakdel, A. N. Javaremi, A. Masoudi-Nejad, and J. M. Reecy. 2018. "Cattle Infection Response Network and Its Functional Modules." *BMC Immunology* 19, no. 1: 2. https://doi.org/10.1186/s12865-017-0238-4.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B: Methodological* 57, no. 1: 289–300.

Berthier, D., M. Peylhard, G. K. Dayo, et al. 2015. "A Comparison of Phenotypic Traits Related to Trypanotolerance in Five West African Cattle Breeds Highlights the Value of Shorthorn Taurine Breeds." *PLoS One* 10, no. 5: e0126498. https://doi.org/10.1371/journal.pone.0126498.

Bovine HapMap Consortium, R. A. Gibbs, J. F. Taylor, et al. 2009. "Genome-Wide Survey of SNP Variation Uncovers the Genetic Structure of Cattle Breeds." *Science* 324, no. 5926: 528–532. https://doi.org/10.1126/science.1167936.

Breuer, K., A. K. Foroushani, M. R. Laird, et al. 2013. "InnateDB: Systems Biology of Innate Immunity and Beyond—Recent Updates and Continuing Curation." *Nucleic Acids Research* 41: D1228–D1233. https://doi.org/10.1093/nar/gks1147.

Buggiotti, L., A. A. Yurchenko, N. S. Yudin, et al. 2021. "Demographic History, Adaptation, and NRAP Convergent Evolution at Amino Acid Residue 100 in the World Northernmost Cattle From Siberia." *Molecular Biology and Evolution* 38, no. 8: 3093–3110. https://doi.org/10.1093/molbev/msab078.

Callaby, R., R. Kelly, S. Mazeri, et al. 2020. "Genetic Diversity of Cameroon Cattle and a Putative Genomic Map for Resistance to Bovine Tuberculosis." *Frontiers in Genetics* 11: 550215. https://doi.org/10.3389/fgene.2020.550215.

Cano-Gamez, E., and G. Trynka. 2020. "From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases." *Frontiers in Genetics* 11: 424. https://doi.org/10.3389/fgene.2020.00424.

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4: 7. https://doi.org/10.1186/s13742-015-0047-8.

Chen, N., Y. Cai, Q. Chen, et al. 2018. "Whole-Genome Resequencing Reveals World-Wide Ancestry and Adaptive Introgression Events of Domesticated Cattle in East Asia." *Nature Communications* 9, no. 1: 2337. https://doi.org/10.1038/s41467-018-04737-0.

Chen, Q., J. Zhan, J. Shen, et al. 2020. "Whole-Genome Resequencing Reveals Diversity, Global and Local Ancestry Proportions in Yunling Cattle." *Journal of Animal Breeding and Genetics* 137, no. 6: 641–650. https://doi.org/10.1111/jbg.12479.

Codner, G. F., M. J. Stear, R. Reeve, L. Matthews, and S. A. Ellis. 2012. "Selective Forces Shaping Diversity in the Class I Region of the Major Histocompatibility Complex in Dairy Cattle." *Animal Genetics* 43, no. 3: 239–249. https://doi.org/10.1111/j.1365-2052.2011.02239.x.

Colbran, L. L., E. R. Gamazon, D. Zhou, P. Evans, N. J. Cox, and J. A. Capra. 2019. "Inferred Divergent Gene Regulation in Archaic Hominins Reveals Potential Phenotypic Differences." *Nature Ecology & Evolution* 3, no. 11: 1598–1606. https://doi.org/10.1038/s41559-019-0996-x.

Cong, B., Q. Zhang, and X. Cao. 2021. "The Function and Regulation of TET2 in Innate Immunity and Inflammation." *Protein & Cell* 12, no. 3: 165–173. https://doi.org/10.1007/s13238-020-00796-6.

Corral, R. S., N. A. Guerrero, H. Cuervo, N. Gironès, and M. Fresno. 2013. "Trypanosoma Cruzi Infection and Endothelin-1 Cooperatively Activate Pathogenic Inflammatory Pathways in Cardiomyocytes." *PLoS Neglected Tropical Diseases* 7, no. 2: e2034. https://doi.org/10.1371/journal.pntd.0002034.

Dayo, G. K., M. Gautier, D. Berthier, et al. 2012. "Association Studies in QTL Regions Linked to Bovine Trypanotolerance in a West African Crossbred Population." *Animal Genetics* 43, no. 2: 123–132. https://doi.org/10.1111/j.1365-2052.2011.02227.x.

de Lima, A. O., J. E. Koltes, W. J. S. Diniz, et al. 2020. "Potential Biomarkers for Feed Efficiency-Related Traits in Nelore Cattle Identified by Co-Expression Network and Integrative Genomics Analyses." *Frontiers in Genetics* 11: 189. https://doi.org/10.3389/fgene.2020.00189.

Decker, J. E., S. D. McKay, M. M. Rolf, et al. 2014. "Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle." *PLoS Genetics* 10, no. 3: e1004254. https://doi.org/10.1371/journal.pgen.1004254.

Durinck, S., P. T. Spellman, E. Birney, and W. Huber. 2009. "Mapping Identifiers for the Integration of Genomic Datasets With the R/Bioconductor Package biomaRt." *Nature Protocols* 4, no. 8: 1184–1191. https://doi.org/10.1038/nprot.2009.97.

Dwinger, R. H., M. Murray, and S. K. Moloo. 1987. "Potential Value of Localized Skin Reactions (Chancres) Induced by *Trypanosoma congolense* Transmitted by *Glossina morsitans* Centralis for the Analysis of Metacyclic Trypanosome Populations." *Parasite Immunology* 9, no. 3: 353–362. https://doi.org/10.1111/j.1365-3024.1987.tb00513.x.

Ellis, S. 2004. "The Cattle Major Histocompatibility Complex: Is It Unique?" *Veterinary Immunology and Immunopathology* 102, no. 1–2: 1–8. https://doi.org/10.1016/j.vetimm.2004.06.007.

Ellis, S. A., and J. A. Hammond. 2014. "The Functional Significance of Cattle Major Histocompatibility Complex Class I Genetic Diversity." *Annual Review of Animal Biosciences* 2: 285–306. https://doi.org/10.1146/annurev-animal-022513-114234.

Flori, L., S. Thevenon, G. K. Dayo, et al. 2014. "Adaptive Admixture in the West African Bovine Hybrid Zone: Insight From the Borgou Population." *Molecular Ecology* 23, no. 13: 3241–3257. https://doi.org/10.1111/mec.12816.

Food and Agriculture Organization. 2023. *Domestic Animal Diversity Information System (DAD-IS)*. Food and Agriculture Organization of the United Nations. https://www.fao.org/dad-is.

Freeman, A. R., C. M. Meghen, D. E. MacHugh, et al. 2004. "Admixture and Diversity in West African Cattle Populations." *Molecular Ecology* 13, no. 11: 3477–3487. https://doi.org/10.1111/j.1365-294X.2004.02311.x.

Frerebeau, N. 2023. "khroma: Colour Schemes for Scientific Data Visualization." https://packages.tesselle.org/khroma.

Friedrich, J., R. I. Bailey, A. Talenti, et al. 2023. "Mapping Restricted Introgression Across the Genomes of Admixed Indigenous African Cattle Breeds." *Genetics, Selection, Evolution: GSE* 55, no. 1: 91. https://doi.org/10.1186/s12711-023-00861-8.

Gamazon, E. R., R. S. Huang, M. E. Dolan, N. J. Cox, and H. K. Im. 2013. "Integrative Genomics: Quantifying Significance of Phenotype-Genotype Relationships From Multiple Sources of High-Throughput Data." *Frontiers in Genetics* 3: 202. https://doi.org/10.3389/fgene.2012.00202.

Garnier, S., N. Ross, R. Rudis, A. P. Camargo, M. Sciaini, and C. Scherer. 2023. "viridis(Lite)—Colorblind-Friendly Color Maps for R." https://sjmgarnier.github.io/viridis.

Gautier, M. 2015. "Genome-Wide Scan for Adaptive Divergence and Association With Population-Specific Covariates." *Genetics* 201, no. 4: 1555–1579. https://doi.org/10.1534/genetics.115.181453.

Gautier, M., L. Flori, A. Riebler, et al. 2009. "A Whole Genome Bayesian Scan for Adaptive Genetic Divergence in West African Cattle." *BMC Genomics* 10: 550. https://doi.org/10.1186/1471-2164-10-550.

Gebeyehu, S., B. Heins, T. Sonstegard, et al. 2025. "Analysis of SLICK Allele in African Taurine and Zebu Cattle Breeds." *Animal Genetics* 56, no. 1: e13499. https://doi.org/10.1111/age.13499.

Geigy, R., and M. Kauffmann. 1973. "Sleeping Sickness Survey in the Serengeti Area (Tanzania) 1971. I. Examination of Large Mammals for Trypanosomes." *Acta Tropica* 30, no. 1: 12–23.

Ghoreishifar, M., I. M. Macleod, T. Nguyen, et al. 2026. "Bridging GWAS to Genes: An Integrative Multi-Omics Approach Using Cattle Data." *BMC Genomics* 27, no. 1: 171. https://doi.org/10.1186/s12864-026-12525-0.

Gonzalez-Prendes, R., C. Ginja, J. Kantanen, et al. 2022. "Integrative QTL Mapping and Selection Signatures in Groningen White Headed Cattle Inferred From Whole-Genome Sequences." *PLoS One* 17, no. 10: e0276309. https://doi.org/10.1371/journal.pone.0276309.

Gu, X., S. Chen, Z. Wang, Q. Bu, and S. An. 2023. "LZTS3/TAGLN Suppresses Cancer Progression in Human Colorectal Adenocarcinoma Through Regulating Cell Proliferation, Migration, and Actin Cytoskeleton." *Archives of Medical Research* 54, no. 7: 102894. https://doi.org/10.1016/j.arcmed.2023.102894.

Guan, X., S. Zhao, W. Xiang, et al. 2022. "Genetic Diversity and Selective Signature in Dabieshan Cattle Revealed by Whole-Genome Resequencing." *Biology-Basel* 11, no. 9: 1327. https://doi.org/10.3390/biology11091327.

Guan, Y. 2014. "Detecting Structure of Haplotypes and Local Ancestry." *Genetics* 196, no. 3: 625–642. https://doi.org/10.1534/genetics.113.160697.

Hall, T. J., G. P. McHugo, M. P. Mullen, et al. 2024. "Integrative and Comparative Genomic Analyses of Mammalian Macrophage Responses to Intracellular Mycobacterial Pathogens." *Tuberculosis (Edinburgh, Scotland)* 147: 102453. https://doi.org/10.1016/j.tube.2023.102453.

Hall, T. J., M. P. Mullen, G. P. McHugo, K. E. Killick, S. C. Ring, and D. P. Berry. 2021. "Integrative Genomics of the Mammalian Alveolar Macrophage Response to Intracellular Mycobacteria." *BMC Genomics* 22, no. 1: 343. https://doi.org/10.1186/s12864-021-07643-w.

Hanotte, O., D. G. Bradley, J. W. Ochieng, Y. Verjee, E. W. Hill, and J. E. Rege. 2002. "African Pastoralism: Genetic Imprints of Origins and Migrations." *Science* 296, no. 5566: 336–339. https://doi.org/10.1126/science.1069878.

Hao, S., K. K. Yan, L. Ding, C. Qian, H. Chi, and J. Yu. 2020. "Network Approaches for Dissecting the Immune System." *iScience* 23, no. 8: 101354. https://doi.org/10.1016/j.isci.2020.101354.

Hatem, A., K. Kaya, and Ü. V. Çatalyürek. 2012. "Microarray vs. RNA-Seq: A Comparison for Active Subnetwork Discovery." Paper presented at the Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, Orlando, Florida. https://doi.org/10.1145/2382936.2382994.

Henry, L., and H. Wickham. 2024. "rlang: Functions for Base Types and Core R and 'tidyverse' Features." https://rlang.r-lib.org.

Holder, A., R. Garty, C. Elder, et al. 2020. "Analysis of Genetic Variation in the Bovine *SLC11A1* Gene, Its Influence on the Expression of *NRAMP1* and Potential Association With Resistance to Bovine Tuberculosis." *Frontiers in Microbiology* 11: 1420. https://doi.org/10.3389/fmicb.2020.01420.

Hou, J., X. Guan, X. Xia, et al. 2024. "Evolution and Legacy of East Asian Aurochs." *Science Bulletin (Beijing)* 69, no. 21: 3425–3433. https://doi.org/10.1016/j.scib.2024.09.016.

Huai, W., H. Song, L. Wang, et al. 2015. "Phosphatase PTPN4 Preferentially Inhibits TRIF-Dependent TLR4 Pathway by Dephosphorylating TRAM." *Journal of Immunology* 194, no. 9: 4458–4465. https://doi.org/10.4049/jimmunol.1402183.

Ibeagha-Awemu, E. M., and G. Erhardt. 2005. "Genetic Structure and Differentiation of 12 African *Bos indicus* and *Bos taurus* Cattle Breeds, Inferred From Protein and Microsatellite Polymorphisms." *Journal of Animal Breeding and Genetics* 122, no. 1: 12–20. https://doi.org/10.1111/j.1439-0388.2004.00478.x.

Ideker, T., O. Ozier, B. Schwikowski, and A. F. Siegel. 2002. "Discovering Regulatory and Signalling Circuits in Molecular Interaction Networks." *Bioinformatics* 18, no. Suppl 1: S233–S240. https://doi.org/10.1093/bioinformatics/18.suppl_1.s233.

ImageMagick Studio LLC. 2023. "ImageMagick." https://imagemagick.org.

Jordheim, L. P. 2018. "Expanding the Clinical Relevance of the 5′-Nucleotidase cN-II/NT5C2." *Purinergic Signal* 14, no. 4: 321–329. https://doi.org/10.1007/s11302-018-9627-2.

Kadarmideen, H. N., and G. Mazzoni. 2018. "Transcriptomics-Genomics Data Integration and Expression Quantitative Trait Loci Analyses in Oocyte Donors and Embryo Recipients for Improving In Vitro Production of Dairy Cattle Embryos." *Reproduction, Fertility, and Development* 31, no. 1: 55–67. https://doi.org/10.1071/RD18338.

Karczewski, K. J., and M. P. Snyder. 2018. "Integrative Omics for Health and Disease." *Nature Reviews. Genetics* 19, no. 5: 299–310. https://doi.org/10.1038/nrg.2018.4.

Karlsson, E. K., D. P. Kwiatkowski, and P. C. Sabeti. 2014. "Natural Selection and Infectious Disease in Human Populations." *Nature Reviews. Genetics* 15, no. 6: 379–393. https://doi.org/10.1038/nrg3734.

Kassahun, Y., V. Mattiangeli, G. Ameni, et al. 2015. "Admixture Mapping of Tuberculosis and Pigmentation-Related Traits in an African-European Hybrid Cattle Population." *Frontiers in Genetics* 6: 210. https://doi.org/10.3389/fgene.2015.00210.

Kayama, H., R. Koga, K. Atarashi, et al. 2009. "NFATc1 Mediates Toll-Like Receptor-Independent Innate Immune Responses During *Trypanosoma cruzi* Infection." *PLoS Pathogens* 5, no. 7: e1000514. https://doi.org/10.1371/journal.ppat.1000514.

Kerner, G., A. L. Neehus, Q. Philippot, et al. 2023. "Genetic Adaptation to Pathogens and Increased Risk of Inflammatory Disorders in Post-Neolithic Europe." *Cell Genomics* 3, no. 2: 100248. https://doi.org/10.1016/j.xgen.2022.100248.

Kim, J., O. Hanotte, O. A. Mwai, et al. 2017. "The Genome Landscape of Indigenous African Cattle." *Genome Biology* 18, no. 1: 34. https://doi.org/10.1186/s13059-017-1153-y.

Kim, K., T. Kwon, T. Dessie, et al. 2020. "The Mosaic Genome of Indigenous African Cattle as a Unique Genetic Resource for African Pastoralism." *Nature Genetics* 52, no. 10: 1099–1110. https://doi.org/10.1038/s41588-020-0694-2.

Kim, S. J., S. Ka, J. W. Ha, et al. 2017. "Cattle Genome-Wide Analysis Reveals Genetic Signatures in Trypanotolerant N'Dama." *BMC Genomics* 18, no. 1: 371. https://doi.org/10.1186/s12864-017-3742-2.

Kleinwort, K. J. H., S. M. Hauck, R. L. Degroote, et al. 2019. "Peripheral Blood Bovine Lymphocytes and MAP Show Distinctly Different Proteome Changes and Immune Pathways in Host-Pathogen Interaction." *PeerJ* 7: e8130. https://doi.org/10.7717/peerj.8130.

Koestler, D. C., M. J. Jones, and M. S. Kobor. 2014. "The Era of Integrative Genomics: More Data or Better Methods?" *Epigenomics* 6, no. 5: 463–467. https://doi.org/10.2217/epi.14.44.

Kolberg, L., U. Raudvere, I. Kuzmin, P. Adler, J. Vilo, and H. Peterson. 2023. "G:Profiler-Interoperable Web Service for Functional Enrichment Analysis and Gene Identifier Mapping (2023 Update)." *Nucleic Acids Research* 51, no. W1: W207–W212. https://doi.org/10.1093/nar/gkad347.

Krassowski, M. 2020. "ComplexUpset."

Kumar, R., S. Gandham, A. Rana, H. K. Maity, U. Sarkar, and B. Dey. 2023. "Divergent Proinflammatory Immune Responses Associated With the Differential Susceptibility of Cattle Breeds to Tuberculosis." *Frontiers in Immunology* 14: 1199092. https://doi.org/10.3389/fimmu.2023.1199092.

Lee, K., D. T. Nguyen, M. Choi, et al. 2013. "Analysis of Cattle Olfactory Subgenome: The First Detail Study on the Characteristics of the Complete Olfactory Receptor Repertoire of a Ruminant." *BMC Genomics* 14: 596. https://doi.org/10.1186/1471-2164-14-596.

Lee, P. H., C. O'Dushlaine, B. Thomas, and S. M. Purcell. 2012. "INRICH: Interval-Based Enrichment Analysis for Genome-Wide Association Studies." *Bioinformatics* 28, no. 13: 1797–1799. https://doi.org/10.1093/bioinformatics/bts191.

Lee, S., C. Clementine, and H. Kim. 2024. "Exploring the Genetic Factors Behind the Discrepancy in Resistance to Bovine Tuberculosis Between African Zebu Cattle and European Taurine Cattle." *Scientific Reports* 14, no. 1: 2370. https://doi.org/10.1038/s41598-024-52606-2.

Li, G., Y. Wu, Y. Zhang, et al. 2024. "Research Progress on Phosphatidylinositol 4-Kinase Inhibitors." *Biochemical Pharmacology* 220: 115993. https://doi.org/10.1016/j.bcp.2023.115993.

Li, Z., J. He, F. Yang, et al. 2023. "A Look Under the Hood of Genomic-Estimated Breed Compositions for Brangus Cattle: What Have We Learned?" *Frontiers in Genetics* 14: 1080279. https://doi.org/10.3389/fgene.2023.1080279.

Liston, W. G., and M. B. Soparkar. 1917. "The Susceptibility of Indian Milch Cattle to Tuberculosis." *Indian Journal of Medical Research* 5, no. 1: 17–18.

MacHugh, D. E., M. D. Shriver, R. T. Loftus, P. Cunningham, and D. G. Bradley. 1997. "Microsatellite DNA Variation and the Evolution, Domestication and Phylogeography of Taurine and Zebu Cattle (*Bos taurus* and *Bos indicus*)." *Genetics* 146, no. 3: 1071–1086. https://doi.org/10.1093/genetics/146.3.1071.

McCarthy, D. J., Y. Chen, and G. K. Smyth. 2012. "Differential Expression Analysis of Multifactor RNA-Seq Experiments With Respect to Biological Variation." *Nucleic Acids Research* 40, no. 10: 4288–4297. https://doi.org/10.1093/nar/gks042.

McHugo, G. P., J. A. Ward, J. A. Browne, et al. 2025. "Functional Genomics of Trypanotolerant and Trypanosusceptible Cattle Infected With Trypanosoma congolense Across Multiple Time Points and Tissues." *PLoS Neglected Tropical Diseases* 19, no. 8: e0012882. https://doi.org/10.1371/journal.pntd.0012882.

McHugo, G. P., J. A. Ward, S. I. Ng'ang'a, et al. 2025. "Genome-Wide Local Ancestry and the Functional Consequences of Admixture in African and European Cattle Populations." *Heredity (Edinb)* 134, no. 1: 49–63. https://doi.org/10.1038/s41437-024-00734-w.

McQuillan, M. A., A. Ranciaro, M. E. B. Hansen, et al. 2022. "Signatures of Convergent Evolution and Natural Selection at the Alcohol Dehydrogenase Gene Region Are Correlated With Agriculture in Ethnically Diverse Africans." *Molecular Biology and Evolution* 39, no. 10: msac183. https://doi.org/10.1093/molbev/msac183.

Meade, K. G., G. M. O'Gorman, E. W. Hill, et al. 2009. "Divergent Antimicrobial Peptide (AMP) and Acute Phase Protein (APP) Responses to *Trypanosoma congolense* Infection in Trypanotolerant and Trypanosusceptible Cattle." *Molecular Immunology* 47, no. 2–3: 196–204. https://doi.org/10.1016/j.molimm.2009.09.042.

Mekonnen, Y. A., M. Gultas, K. Effa, O. Hanotte, and A. O. Schmitt. 2019. "Identification of Candidate Signature Genes and Key Regulators Associated With Trypanotolerance in the Sheko Breed." *Frontiers in Genetics* 10: 1095. https://doi.org/10.3389/fgene.2019.01095.

Melo, T. P., G. M. F. de Camargo, L. G. de Albuquerque, and R. Carvalheiro. 2017. "Genome-Wide Association Study Provides Strong Evidence of Genes Affecting the Reproductive Performance of Nellore Beef Cows." *PLoS One* 12, no. 5: e0178551. https://doi.org/10.1371/journal.pone.0178551.

Mercher, T., C. Quivoron, L. Couronne, C. Bastard, W. Vainchenker, and O. A. Bernard. 2012. "TET2, a Tumor Suppressor in Hematological Disorders." *Biochimica et Biophysica Acta* 1825, no. 2: 173–177. https://doi.org/10.1016/j.bbcan.2011.12.002.

Müller, K., and H. Wickham. 2023. "tibble: Simple Data Frames." https://tibble.tidyverse.org.

Murray, M., J. C. Trail, C. E. Davis, and S. J. Black. 1984. "Genetic Resistance to African Trypanosomiasis." *Journal of Infectious Diseases* 149, no. 3: 311–319. https://doi.org/10.1093/infdis/149.3.311.

Naessens, J. 2006. "Bovine Trypanotolerance: A Natural Ability to Prevent Severe Anaemia and Haemophagocytic Syndrome?" *International Journal for Parasitology* 36, no. 5: 521–528. https://doi.org/10.1016/j.ijpara.2006.02.012.

Nantulya, V. M., A. J. Musoke, F. R. Rurangirwa, and S. K. Moloo. 1984. "Resistance of Cattle to Tsetse-Transmitted Challenge With *Trypanosoma brucei* or *Trypanosoma congolense* After Spontaneous Recovery From Syringe-Passaged Infections." *Infection and Immunity* 43, no. 2: 735–738. https://doi.org/10.1128/iai.43.2.735-738.1984.

Newton, K. 2015. "RIPK1 and RIPK3: Critical Regulators of Inflammation and Cell Death." *Trends in Cell Biology* 25, no. 6: 347–353. https://doi.org/10.1016/j.tcb.2015.01.001.

Nicolazzi, E. L., A. Caprera, N. Nazzicari, et al. 2015. "SNPchiMp v.3: Integrating and Standardizing Single Nucleotide Polymorphism Data for Livestock Species." *BMC Genomics* 16, no. 1: 283. https://doi.org/10.1186/s12864-015-1497-1.

Niimura, Y., and M. Nei. 2007. "Extensive Gains and Losses of Olfactory Receptor Genes in Mammalian Evolution." *PLoS One* 2, no. 8: e708. https://doi.org/10.1371/journal.pone.0000708.

Noyes, H., A. Brass, I. Obara, et al. 2011. "Genetic and Expression Analysis of Cattle Identifies Candidate Genes in Pathways Responding to *Trypanosoma congolense* Infection." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 22: 9304–9309. https://doi.org/10.1073/pnas.1013486108.

O'Gorman, G. M., S. D. Park, E. W. Hill, et al. 2006. "Cytokine mRNA Profiling of Peripheral Blood Mononuclear Cells From Trypanotolerant and Trypanosusceptible Cattle Infected With *Trypanosoma congolense*." *Physiological Genomics* 28, no. 1: 53–61. https://doi.org/10.1152/physiolgenomics.00100.2006.

O'Gorman, G. M., S. D. Park, E. W. Hill, et al. 2009. "Transcriptional Profiling of Cattle Infected With *Trypanosoma congolense* Highlights Gene Expression Signatures Underlying Trypanotolerance and Trypanosusceptibility." *BMC Genomics* 10: 207. https://doi.org/10.1186/1471-2164-10-207.

O'Grady, J. F., G. P. McHugo, J. A. Ward, et al. 2025. "Integrative Genomics Sheds Light on the Immunogenetics of Tuberculosis in Cattle." *Communications Biology* 8, no. 1: 479. https://doi.org/10.1038/s42003-025-07846-x.

Ooms, J. 2023. "magick: Advanced Graphics and Image-Processing in R." https://docs.ropensci.org/magick.

Orecchioni, M., H. Matsunami, and K. Ley. 2022. "Olfactory Receptors in Macrophages and Inflammation." *Frontiers in Immunology* 13: 1029244. https://doi.org/10.3389/fimmu.2022.1029244.

Paling, R. W., S. K. Moloo, J. R. Scott, G. Gettinby, F. A. McOdimba, and M. Murray. 1991. "Susceptibility of N'dama and Boran Cattle to Sequential Challenges With Tsetse-Transmitted Clones of *Trypanosoma congolense*." *Parasite Immunology* 13, no. 4: 427–445. https://doi.org/10.1111/j.1365-3024.1991.tb00295.x.

Pan, Y., C. Zhang, Y. Lu, et al. 2022. "Genomic Diversity and Post-Admixture Adaptation in the Uyghurs." *National Science Review* 9, no. 3: nwab124. https://doi.org/10.1093/nsr/nwab124.

Patterson, N., A. L. Price, and D. Reich. 2006. "Population Structure and Eigenanalysis." *PLoS Genetics* 2, no. 12: e190. https://doi.org/10.1371/journal.pgen.0020190.

Pedersen, T. L. 2023. "patchwork: The Composer of Plots." https://patchwork.data-imaginist.com.

Peylhard, M., D. Berthier, G.-K. Dayo, et al. 2023. "Whole Blood Transcriptome Profiles of Trypanotolerant and Trypanosusceptible Cattle Highlight a Differential Modulation of Metabolism and Immune Response During Infection by *Trypanosoma congolense*." *Peer Community Journal* 3: e17. https://doi.org/10.24072/pcjournal.239.

Pina-Martins, F., D. N. Silva, J. Fino, and O. S. Paulo. 2017. "Structure_threader: An Improved Method for Automation and Parallelization of Programs Structure, fastStructure and MavericK on Multicore CPU Systems." *Molecular Ecology Resources* 17, no. 6: e268–e274. https://doi.org/10.1111/1755-0998.12702.

Porcu, E., J. Sjaarda, K. Lepik, et al. 2021. "Causal Inference Methods to Integrate Omics and Complex Traits." *Cold Spring Harbor Perspectives in Medicine* 11, no. 5: a040493. https://doi.org/10.1101/cshperspect.a040493.

Porto-Neto, L. R., D. M. Bickhart, A. J. Landaeta-Hernandez, et al. 2018. "Convergent Evolution of Slick Coat in Cattle Through Truncation Mutations in the Prolactin Receptor." *Frontiers in Genetics* 9: 57. https://doi.org/10.3389/fgene.2018.00057.

Quintana-Murci, L. 2019. "Human Immunology Through the Lens of Evolutionary Genetics." *Cell* 177, no. 1: 184–199. https://doi.org/10.1016/j.cell.2019.02.033.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org.

Raj, A., M. Stephens, and J. K. Pritchard. 2014. "fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets." *Genetics* 197, no. 2: 573–589. https://doi.org/10.1534/genetics.114.164350.

Randolph, H. E., K. A. Aracena, Y. L. Lin, Z. Mu, and L. B. Barreiro. 2024. "Shaping Immunity: The Influence of Natural Selection on Population Immune Diversity." *Immunological Reviews* 323, no. 1: 227–240. https://doi.org/10.1111/imr.13329.

Randolph, H. E., J. K. Fiege, B. K. Thielen, et al. 2021. "Genetic Ancestry Effects on the Response to Viral Infection Are Pervasive but Cell Type Specific." *Science* 374, no. 6571: 1127–1133. https://doi.org/10.1126/science.abg0928.

Rashidi, S., R. Mansouri, M. Ali-Hassanzadeh, A. Muro, P. Nguewa, and R. Manzano-Roman. 2023. "The Most Prominent Modulated Annexins During Parasitic Infections." *Acta Tropica* 243: 106942. https://doi.org/10.1016/j.actatropica.2023.106942.

Redwan, E. M., A. A. AlJaddawi, and V. N. Uversky. 2019. "Structural Disorder in the Proteome and Interactome of Alkhurma Virus (ALKV)." *Cellular and Molecular Life Sciences* 76, no. 3: 577–608. https://doi.org/10.1007/s00018-018-2968-8.

Ring, S. C., D. C. Purfield, M. Good, et al. 2019. "Variance Components for Bovine Tuberculosis Infection and Multi-Breed Genome-Wide Association Analysis Using Imputed Whole Genome Sequence Data." *PLoS One* 14, no. 2: e0212067. https://doi.org/10.1371/journal.pone.0212067.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. "edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data." *Bioinformatics* 26, no. 1: 139–140. https://doi.org/10.1093/bioinformatics/btp616.

Rosen, B. D., D. M. Bickhart, R. D. Schnabel, et al. 2020. "De Novo Assembly of the Cattle Reference Genome With Single-Molecule Sequencing." *GigaScience* 9, no. 3: giaa021. https://doi.org/10.1093/gigascience/giaa021.

Rue-Albrecht, K., D. A. Magee, K. E. Killick, N. C. Nalpas, S. V. Gordon, and D. E. MacHugh. 2014. "Comparative Functional Genomics and the Bovine Macrophage Response to Strains of the *Mycobacterium* Genus." *Frontiers in Immunology* 5: 536. https://doi.org/10.3389/fimmu.2014.00536.

Santamaría, M. H., L. D. Ríos, and R. S. Corral. 2023. "Chronic Trypanosoma Cruzi Infection Activates the TWEAK/Fn14 Axis in Cardiac Myocytes and Fibroblasts Driving Structural and Functional Changes That Affect the Heart." *Experimental Parasitology* 248: 108491. https://doi.org/10.1016/j.exppara.2023.108491.

Schnabel, R. D. 2018. "ARS-UCD1.2 Cow Genome Assembly: Mapping of All Existing Variants." https://www.animalgenome.org/repository/cattle/UMC_bovine_coordinates.

Sempere, G., K. Moazami-Goudarzi, A. Eggen, D. Laloe, M. Gautier, and L. Flori. 2015. "WIDDE: A Web-Interfaced Next Generation Database for Genetic Diversity Exploration, With a First Application

in Cattle." *BMC Genomics* 16: 940. https://doi.org/10.1186/s12864-015-2181-1.

Shannon, P., A. Markiel, O. Ozier, et al. 2003. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome Research* 13, no. 11: 2498–2504. https://doi.org/10.1101/gr.1239303.

Slowikowski, K. 2023. "ggrepel: Automatically Position Non-Overlapping Text Labels With 'ggplot2'." https://ggrepel.slowkow.com.

Sphyris, N., and S. A. Mani. 2011. "pIgR: Frenemy of Inflammation, EMT, and HCC Progression." *Journal of the National Cancer Institute* 103, no. 22: 1644–1645. https://doi.org/10.1093/jnci/djr421.

Stelzer, G., N. Rosen, I. Plaschkes, et al. 2016. "The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses." *Current Protocols in Bioinformatics* 54: 1.30.1–1.30.33. https://doi.org/10.1002/cpbi.5.

Steverding, D. 2008. "The History of African Trypanosomiasis." *Parasites and Vectors* 1, no. 1: 3. https://doi.org/10.1186/1756-3305-1-3.

Stijlemans, B., A. Beschin, S. Magez, J. A. Van Ginderachter, and P. De Baetselier. 2015. "Iron Homeostasis and *Trypanosoma brucei* Associated Immunopathogenicity Development: A Battle/Quest for Iron." *BioMed Research International* 2015: 819389. https://doi.org/10.1155/2015/819389.

Stijlemans, B., A. Vankrunkelsven, L. Brys, S. Magez, and P. De Baetselier. 2008. "Role of Iron Homeostasis in Trypanosomiasis-Associated Anemia." *Immunobiology* 213, no. 9–10: 823–835. https://doi.org/10.1016/j.imbio.2008.07.023.

Stijlemans, B., A. Vankrunkelsven, G. Caljon, et al. 2010. "The Central Role of Macrophages in Trypanosomiasis-Associated Anemia: Rationale for Therapeutical Approaches." *Endocrine, Metabolic & Immune Disorders Drug Targets* 10, no. 1: 71–82. https://doi.org/10.2174/187153010790827966.

Subramanian, I., S. Verma, S. Kumar, A. Jere, and K. Anamika. 2020. "Multi-Omics Data Integration, Interpretation, and Its Application." *Bioinformatics and Biology Insights* 14: 1177932219899051. https://doi.org/10.1177/1177932219899051.

Sun, L., K. Qu, Y. Liu, et al. 2023. "Assessing Genomic Diversity and Selective Pressures in Bashan Cattle by Whole-Genome Sequencing Data." *Animal Biotechnology* 34, no. 4: 835–846. https://doi.org/10.1080/10495398.2021.1998094.

Suravajhala, P., L. J. Kogelman, and H. N. Kadarmideen. 2016. "Multi-Omic Data Integration and Analysis Using Systems Genomics Approaches: Methods and Applications in Animal Production, Health and Welfare." *Genetics, Selection, Evolution: GSE* 48, no. 1: 38. https://doi.org/10.1186/s12711-016-0217-x.

Tapryal, N., A. Chakraborty, K. Saha, et al. 2023. "The DNA Glycosylase NEIL2 Is Protective During SARS-CoV-2 Infection." *Nature Communications* 14, no. 1: 8169. https://doi.org/10.1038/s41467-023-43938-0.

Tijjani, A., Y. T. Utsunomiya, A. G. Ezekwe, O. Nashiru, and O. Hanotte. 2019. "Genome Sequence Analysis Reveals Selection Signatures in Endangered Trypanotolerant West African Muturu Cattle." *Frontiers in Genetics* 10: 442. https://doi.org/10.3389/fgene.2019.00442.

Upadhyay, M. R., W. Chen, J. A. Lenstra, et al. 2017. "Genetic Origin, Admixture and Population History of Aurochs (*Bos primigenius*) and Primitive European Cattle." *Heredity (Edinb)* 118, no. 2: 169–176. https://doi.org/10.1038/hdy.2016.79.

Urbanski, A. H., J. D. Araujo, R. Creighton, and H. I. Nakaya. 2019. "Integrative Biology Approaches Applied to Human Diseases." In *Computational Biology*, edited by H. Husi. Codon Publications.

Uzonna, J. E., R. S. Kaushik, J. R. Gordon, and H. Tabel. 1999. "Cytokines and Antibody Responses During *Trypanosoma congolense* Infections in Two Inbred Mouse Strains That Differ in Resistance." *Parasite Immunology* 21, no. 2: 57–71. https://doi.org/10.1046/j.1365-3024.1999.00202.x.

van den Brand, T. 2023. "ggh4x: Hacks for 'ggplot2'." https://teunbrand.github.io/ggh4x/.

Verdugo, M. P., V. E. Mullin, A. Scheu, et al. 2019. "Ancient Cattle Genomics, Origins, and Rapid Turnover in the Fertile Crescent." *Science* 365, no. 6449: 173–176. https://doi.org/10.1126/science.aav1002.

Vordermeier, M., G. Ameni, S. Berg, et al. 2012. "The Influence of Cattle Breed on Susceptibility to Bovine Tuberculosis in Ethiopia." *Comparative Immunology, Microbiology and Infectious Diseases* 35, no. 3: 227–232. https://doi.org/10.1016/j.cimid.2012.01.003.

Wang, K., J. A. Lenstra, L. Liu, et al. 2018. "Incomplete Lineage Sorting Rather Than Hybridization Explains the Inconsistent Phylogeny of the Wisent." *Communications Biology* 1, no. 1: 169. https://doi.org/10.1038/s42003-018-0176-6.

Wang, X., D. Fan, Y. Yang, R. C. Gimple, and S. Zhou. 2023. "Integrative Multi-Omics Approaches to Explore Immune Cell Functions: Challenges and Opportunities." *iScience* 26, no. 4: 106359. https://doi.org/10.1016/j.isci.2023.106359.

Wang, Y., Z. Chen, K. Liang, et al. 2025. "AGO2 Mediates Immunotherapy Failure via Suppressing Tumor IFN-Gamma Response-Dependent $CD8^+$ T Cell Immunity." *Cell Reports* 44, no. 4: 115445. https://doi.org/10.1016/j.celrep.2025.115445.

Ward, J. A., G. P. McHugo, M. J. Dover, et al. 2022. "Genome-Wide Local Ancestry and Evidence for Mitonuclear Coadaptation in African Hybrid Cattle Populations." *iScience* 25, no. 7: 104672. https://doi.org/10.1016/j.isci.2022.104672.

Ward, J. A., S. I. Ng'ang'a, I. A. S. Randhawa, et al. 2024. "Genomic Insights Into the Population History and Adaptive Traits of Latin American Criollo Cattle." *Royal Society Open Science* 11, no. 3: 231388. https://doi.org/10.1098/rsos.231388.

Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wickham, H. 2023. "stringr: Simple, Consistent Wrappers for Common String Operations." https://stringr.tidyverse.org.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan. 2023. "dplyr: A Grammar of Data Manipulation." https://dplyr.tidyverse.org.

Wickham, H., and L. Henry. 2023. "purrr: Functional Programming Tools." https://purrr.tidyverse.org.

Wickham, H., J. Hester, and J. Bryan. 2023. "readr: Read Rectangular Text Data." https://readr.tidyverse.org.

Wickham, H., T. L. Pedersen, and D. Seidel. 2023. "scales: Scale Functions for Visualization." https://scales.r-lib.org.

Wickham, H., D. Vaughan, and M. Girlich. 2023. "tidyr: Tidy Messy Data." https://tidyr.tidyverse.org.

Wilke, C. O., and B. M. Wiernik. 2022. "ggtext: Improved Text Rendering Support for 'ggplot2'." https://wilkelab.org/ggtext.

Wong, A. K., R. S. G. Sealfon, C. L. Theesfeld, and O. G. Troyanskaya. 2021. "Decoding Disease: From Genomes to Networks to Phenotypes." *Nature Reviews. Genetics* 22, no. 12: 774–790. https://doi.org/10.1038/s41576-021-00389-x.

Wragg, D., E. A. J. Cook, P. Latre de Late, et al. 2022. "A Locus Conferring Tolerance to *Theileria* Infection in African Cattle." *PLoS Genetics* 18, no. 4: e1010099. https://doi.org/10.1371/journal.pgen.1010099.

Wu, D. D., X. D. Ding, S. Wang, et al. 2018. "Pervasive Introgression Facilitated Domestication and Adaptation in the *Bos* Species Complex." *Nature Ecology & Evolution* 2, no. 7: 1139–1145. https://doi.org/10.1038/s41559-018-0562-y.

Yang, X. 2020. "Multitissue Multiomics Systems Biology to Dissect Complex Diseases." *Trends in Molecular Medicine* 26, no. 8: 718–728. https://doi.org/10.1016/j.molmed.2020.04.006.

Yang, X., S. Al-Bustan, Q. Feng, et al. 2014. "The Influence of Admixture and Consanguinity on Population Genetic Diversity in Middle East." *Journal of Human Genetics* 59, no. 11: 615–622. https://doi.org/10.1038/jhg.2014.81.

Yaro, M., K. A. Munyard, M. J. Stear, and D. M. Groth. 2016. "Combatting African Animal Trypanosomiasis (AAT) in Livestock: The Potential Role of Trypanotolerance." *Veterinary Parasitology* 225: 43–52. https://doi.org/10.1016/j.vetpar.2016.05.003.

Zeberg, H., and S. Paabo. 2020. "The Major Genetic Risk Factor for Severe COVID-19 Is Inherited From Neanderthals." *Nature* 587, no. 7835: 610–612. https://doi.org/10.1038/s41586-020-2818-3.

Zeberg, H., and S. Paabo. 2021. "A Genomic Region Associated With Protection Against Severe COVID-19 Is Inherited From Neandertals." *Proceedings of the National Academy of Sciences of the United States of America* 118, no. 9: e2026309118. https://doi.org/10.1073/pnas.2026309118.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Figure S1:** Heatmap of mean identity by state values for SNP data in European, African and Asian cattle populations. **Figure S2:** Tukey box plots showing the distribution of inbreeding values (F) for SNP data for each population of European, African and Asian cattle. Outliers are indicated with a black outline. **Figure S3:** (A) Principal component analysis (PCA) of SNP data for cattle coloured according to population showing the first two principal components and (B) bar chart of proportion of variance of the top 10 principal components. **Figure S4:** (A) Principal component analysis (PCA) of the high-density SNP data for the cattle samples from McHugo, Ward, Ng'ang'a, et al. (2025) coloured according to population showing the first two principal components and (B) bar chart of proportion of variance of the top 10 principal components. The transparency indicates the availability of microarray gene expression data for the sample. **Figure S5:** (A) Principal component analysis (PCA) of the low-density SNP data for the cattle samples from McHugo, Ward, Ng'ang'a, et al. (2025) coloured according to population showing the first two principal components and (B) bar chart of proportion of variance of the top 10 principal components. The transparency indicates the availability of microarray gene expression data for the sample. **Figure S6:** Hierarchical clustering of the SNP data for European, African and Asian cattle populations. Results are shown for an assumed value of the number of ancestral populations, $K = 3$. The transparency indicates the availability of gene expression data for the sample. **Figure S7:** Hierarchical clustering of the SNP data for European, African and Asian cattle populations. Results are shown for an assumed value of the number of ancestral populations, $K = 3$. **Figure S8:** Hierarchical clustering of the high-density SNP data from McHugo, Ward, Ng'ang'a et al. (2025) for European, African and Asian cattle populations. Results are shown for an assumed value of the number of ancestral populations, $K = 3$. The transparency indicates the availability of microarray gene expression data for the sample. **Figure S9:** Hierarchical clustering of the low-density SNP data from McHugo, Ward, Ng'ang'a, et al. (2025) for European, African and Asian cattle populations. Results are shown for an assumed value of the number of ancestral populations, $K = 3$. The transparency indicates the availability of microarray gene expression data for the sample. **Figure S10:** g:Profiler functional enrichment of introgressed regions in the African *Bos taurus* cattle populations with gene expression data available according to local ancestry analysis of SNP data. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The y-axis shows the $-\log_{10}p_{adj}$. value up to a maximum of 16, and the panels along the y-axis and colours indicate the ancestry component. The panels along the x-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of ten are indicated with a black outline and label. **Figure S11:** g:Profiler functional enrichment of introgressed regions in the trypanotolerant African hybrid cattle populations with gene expression data available according to local ancestry analysis of SNP data. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The y-axis shows the $-\log_{10}p_{adj}$. value up to a maximum of 16, and the panels along the y-axis and colours indicate the ancestry component. The panels along the x-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of ten are indicated with a black outline and label. **Figure S12:** g:Profiler functional enrichment of introgressed regions in the trypanosusceptible African hybrid cattle populations with gene expression data available according to local ancestry analysis of SNP data. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The y-axis shows the $-\log_{10}p_{adj}$. value up to a maximum of 16, and the panels along the y-axis and colours indicate the ancestry component. The panels along the x-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of 10 are indicated with a black outline and label. **Figure S13:** Bar chart showing the numbers of significantly differentially expressed genes for the response contrasts of the RNA-seq data. The extent of the bar above and below 0 on the y-axis indicates the numbers of significantly differentially expressed genes with increased and decreased expression, respectively. The position on the x-axis indicates the number of days post-infection, and the colour and shapes within the bars represent the population. **Figure S14:** Volcano plot showing the results of the response contrast for the RNA-seq data from the LAGU population at 40 days post-infection. Each dot represents a gene with the position on the x- and y-axes indicating the $\log_2$ fold change and $-\log_{10}p_{adj}$., respectively. Genes above the horizontal dashed line are significantly differentially expressed, with the colours representing the change in expression. The top 10 most significant genes for increased and decreased expression with gene symbols are labelled. **Figure S15:** Volcano plot showing the results of the response contrast for the RNA-seq data from the BAOU population at 40 days post-infection. Each dot represents a gene with the position on the x- and y-axes indicating the $\log_2$ fold change and $-\log_{10}p_{adj}$., respectively. Genes above the horizontal dashed line are significantly differentially expressed, with the colours representing the change in expression. The top 10 most significant genes for increased and decreased expression with gene symbols are labelled. **Figure S16:** Volcano plot showing the results of the response contrast for the RNA-seq data from the NDAM population at 40 days post-infection. Each dot represents a gene with the position on the x- and y-axes indicating the $\log_2$ fold change and $-\log_{10}p_{adj}$., respectively. Genes above the horizontal dashed line are significantly differentially expressed, with the colours representing the change in expression. The top 10 most significant genes for increased and decreased expression with gene symbols are labelled. **Figure S17:** Volcano plot showing the results of the response contrast for the RNA-seq data from the BORG population at 40 days post-infection. Each dot represents a gene with the position on the x- and y-axes indicating the $\log_2$ fold change and $-\log_{10}p_{adj}$., respectively. Genes above the horizontal dashed line are significantly differentially expressed, with the colours representing the change in expression. The top 10 most significant genes for increased and decreased expression with gene symbols are labelled. **Figure S18:** Base network generated using InnateDB with the top results of a search of the GeneCards database for genes relating to the term 'trypano*'. Each node in the network represents a gene, whereas the edges connecting the nodes represent gene interactions. The nodes are sized according to their degree or number of interactions. **Figure S19:** Functional module identified using jActiveModules and differential expression results for the MICRO BL 34 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S20:** Functional module identified using jActiveModules and differential expression results for the MICRO

LI 35 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S21:** Functional module identified using jActiveModules and differential expression results for the MICRO LN 35 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S22:** Functional module identified using jActiveModules and differential expression results for the MICRO SP 35 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S23:** Functional module identified using jActiveModules and differential expression results for the RNA LAGU 40 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S24:** Functional module identified using jActiveModules and differential expression results for the RNA BAOU 40 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S25:** Functional module identified using jActiveModules and differential expression results for the RNA NDAM 40 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S26:** Functional module identified using jActiveModules and differential expression results for the RNA BORG 40 contrast. Each node in the network represents a gene coloured according to expression, with significant differential expression indicated by the outline. The edges connecting the nodes represent gene interactions, and the nodes are sized according to their degree (number of interactions). **Figure S27:** Upset plot showing the top 20 intersections between the genes in the functional modules identified using jActiveModules and differential expression results for each of the final response contrasts for the microarray and RNA-seq data. The horizontal bars indicate the total number of genes in each module, whereas the vertical bars indicate the number of genes in common between the modules annotated with black dots connected by lines in the intersection matrix. The background colour of the stripes in the intersection matrix and the colour of the horizontal and vertical bars represent the module. Black bars indicate an overlap between different modules. **Figure S28:** g:Profiler functional enrichment of significantly differentially expressed genes in the RNA-seq response contrasts with no background dataset specified. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The *y*-axis shows the $-\log_{10} p_{adj}$. value up to a maximum of 16, and the panels along the *y*-axis and colours indicate the module. The panels along the *x*-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of 10 are indicated with a black outline and label. **Figure S29:** g:Profiler functional enrichment of the genes in the MICRO LI 35 functional module with no background dataset specified. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The *y*-axis shows the $-\log_{10} p_{adj}$. value up to a maximum of 16, and the panels along the *y*-axis and colours indicate the module. The panels along the *x*-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of 10 are indicated with a black outline and label. **Figure S30:** g:Profiler functional enrichment of the genes in the base network with no background dataset specified. Each dot represents a significantly enriched GO term, with the size indicating the ratio of the intersection between the term and the introgressed genes. The *y*-axis shows the $-\log_{10} p_{adj}$. value up to a maximum of 16, and the panels along the *y*-axis and colours indicate the module. The panels along the *x*-axis indicate the source of the term and the position within the panels groups terms from the same GO subtree. The top driver GO terms up to a maximum of 10 are indicated with a black outline and label. **Table S1:** Numbers of SNPs with *z*-score ≥ 2.0 for mean European *Bos taurus*, African *B. taurus* and *Bos indicus* ancestry components for the six populations with gene expression data available across all autosomes. The numbers in brackets indicate the percentage of the total 29,869 SNPs in the dataset. **Table S2:** Numbers of genes within 1 Mb up- and downstream of SNPs with *z*-score ≥ 2.0 for mean European *Bos taurus*, African *B. taurus* and *Bos indicus* ancestry components for the six populations with gene expression data available across all autosomes. The numbers in brackets indicate the percentage of the total 34,080 genes in the dataset. **Table S3:** Population, days post-infection, and the significantly differentially expressed genes with increased and decreased expression, with gene symbols for the response contrasts. **Table S4:** Gene symbol and modules containing each of the 243 genes with valid gene symbols that were found in four or more of the eight functional modules. **Table S5:** Numbers of intervals within 1 Mb up- and downstream of SNPs with *z*-score ≥ 2.0 for mean European *Bos taurus*, African *B. taurus* and *B. indicus* ancestry components for the six populations with gene expression data available across all autosomes and groups of populations from the original local ancestry analysis.