

# Integrative and comparative genomic analyses of mammalian macrophage responses to intracellular mycobacterial pathogens

Thomas J. Hall<sup>a</sup>, Gillian P. McHugo<sup>a</sup>, Michael P. Mullen<sup>b</sup>, James A. Ward<sup>a</sup>, Kate E. Killick<sup>a,1</sup>, John A. Browne<sup>a</sup>, Stephen V. Gordon<sup>c,d</sup>, David E. MacHugh<sup>a,d,\*</sup>

<sup>a</sup> Animal Genomics Laboratory, UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland

<sup>b</sup> Bioscience Research Institute, Technological University of the Shannon, Athlone, Westmeath, N37 HD68, Ireland

<sup>c</sup> UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland

<sup>d</sup> UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland

## ARTICLE INFO

### Keywords:

GWAS  
Host-pathogen interaction  
Integrative genomics  
Macrophage  
Mycobacteria  
Tuberculosis

## ABSTRACT

*Mycobacterium tuberculosis*, the causative agent of human tuberculosis (hTB), is a close evolutionary relative of *Mycobacterium bovis*, which causes bovine tuberculosis (bTB), one of the most damaging infectious diseases to livestock agriculture. Previous studies have shown that the pathogenesis of bTB disease is comparable to hTB disease, and that the bovine and human alveolar macrophage (bAM and hAM, respectively) transcriptomes are extensively reprogrammed in response to infection with these intracellular mycobacterial pathogens. In this study, a multi-omics integrative approach was applied with functional genomics and GWAS data sets across the two primary hosts (*Bos taurus* and *Homo sapiens*) and both pathogens (*M. bovis* and *M. tuberculosis*). Four different experimental infection groups were used: 1) bAM infected with *M. bovis*, 2) bAM infected with *M. tuberculosis*, 3) hAM infected with *M. tuberculosis*, and 4) human monocyte-derived macrophages (hMDM) infected with *M. tuberculosis*. RNA-seq data from these experiments 24 h post-infection (24 hpi) was analysed using three computational pipelines: 1) differentially expressed genes, 2) differential gene expression interaction networks, and 3) combined pathway analysis. The results were integrated with high-resolution bovine and human GWAS data sets to detect novel quantitative trait loci (QTLs) for resistance to mycobacterial infection and resilience to disease. This revealed common and unique response macrophage pathways for both pathogens and identified 32 genes (12 bovine and 20 human) significantly enriched for SNPs associated with disease resistance, the majority of which encode key components of the NF- $\kappa$ B signalling pathway and that also drive formation of the granuloma.

## 1. Introduction

The phylum Actinobacteria represents one of the largest taxonomic groups among the 18 major lineages currently recognized within the domain Bacteria, including five subclasses and 14 suborders. It comprises gram-positive bacteria with a high genomic DNA GC content, ranging from 51 % in some *Corynebacterium* to more than 70 % in *Streptomyces* and *Frankia* [1]. Two important disease-causing species of Actinobacteria, *Mycobacterium tuberculosis* and *Mycobacterium bovis*, are members of the *M. tuberculosis* complex (MTBC) and are characterised by 99.95 % similarity at the nucleotide level [2]. Both species cause

tuberculosis (TB) in mammals, with *M. tuberculosis* displaying a specific host preference for humans, with this MTBC strain not environmentally maintained by animal reservoirs [3,4].

*M. tuberculosis* is one of the most deadly human pathogens, and until the COVID-19 pandemic was the leading cause of death from a single infectious agent, ranking above HIV/AIDS with approximately 1.6 million deaths in 2021 and is the thirteenth leading cause of death overall [5]. In 2021, there were also approximately 450,000 new cases that exhibited resistance to rifampicin (RR-TB), the most effective first-line drug, or that were multidrug-resistant (MDR-TB). Most deaths from hTB could be prevented with early diagnosis and appropriate

\* Corresponding author. Animal Genomics Laboratory, UCD School of Agriculture and Food Science, UCD College of Health and Agricultural Sciences, University College Dublin, Belfield, Dublin D04 V1W8, Ireland.

E-mail address: [david.machugh@ucd.ie](mailto:david.machugh@ucd.ie) (D.E. MacHugh).

<sup>1</sup> Current address: Novartis Ireland Ltd., Elm Park Business Campus, Dublin, D04A9N6, Ireland.

<https://doi.org/10.1016/j.tube.2023.102453>

Received 17 July 2023; Received in revised form 20 November 2023; Accepted 23 November 2023

Available online 2 December 2023

1472-9792/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

treatment.

*M. bovis*, the causative agent of bovine tuberculosis (bTB), is considered to be one of the most damaging infectious diseases to the global agricultural industry, with the global costs of disease conservatively estimated to cost €3 billion annually, imposing huge economic losses on farmers of infected herds [6–8]. *Mycobacterium bovis* can cause zoonotic TB (zTB) with serious implications for human health [9–11]. Due to the highly infectious nature of the pathogen, early detection and removal of infected animals is the most effective control measure [12]. In contrast to *M. tuberculosis*, *M. bovis* has a wide host range and can infect a broad spectrum of domestic and wild mammals including taurine and zebu cattle (*Bos taurus* and *B. indicus*, respectively), sheep (*Ovis aries*), goats (*Capra hircus*), llamas and alpacas (*Lama* spp.), pigs and wild boar (*Sus scrofa*), European badgers (*Meles meles*), brushtail possums (*Trichosurus vulpecula*), various species of deer (Cervidae), African and Asian elephants (*Loxodonta* spp. and *Elephas maximus*), African buffalo (*Syncerus caffer*), and many Felidae species including African lions (*Panthera leo*), leopards (*Panthera pardus*), and cheetahs (*Acinonyx jubatus*) [13–15].

Considered together, *M. tuberculosis* and *M. bovis* represent enormous burdens on global health systems and animal agriculture worldwide. Previous studies have shown that the pathogenesis of bTB disease in cattle is comparable to hTB disease and many aspects of *M. bovis* infection are also characteristic of *M. tuberculosis* infection [16–20]. Consequently, *M. bovis* infection of cattle and bTB disease are now recognized as a valuable model for understanding hTB caused by *M. tuberculosis* [21–25]. Transmission of both MTBC strains is via inhalation of contaminated aerosol droplets and the primary site of infection is the lungs [26,27], where these pathogens encounter resident alveolar macrophages (AM), the host's first line of defence, which normally phagocytize and destroy airborne bacteria [28,29]. However, MTBC mycobacteria such as *M. bovis* and *M. tuberculosis* can persist and replicate within alveolar macrophages via a plethora of evolved evasion mechanisms that subvert and interfere with host immune responses [30–33]. Some of these mechanisms encompass the early secretory antigenic target-6 (ESAT-6) secretion system-1 (ESX-1), which facilitates escape from the macrophage phagosome into the cytosol where it replicates [34]. The phagosomes that form when macrophages digest an invading microorganism normally go through rapid fusion with lysosomes; however, this is not the case with AM that phagocytize the pathogenic MTBC strains [35].

Previous studies, including those performed by our research group, have demonstrated that the bovine and human alveolar macrophage (bAM and hAM) transcriptomes are extensively reprogrammed in response to infection with *M. bovis* and *M. tuberculosis* [36–43]. These studies have also highlighted a complex system of gene expression regulation that drives host-pathogen interactions and innate immune response pathway execution, which are functionally associated with many macrophage processes that control or eliminate intracellular microbes. What remains unclear, however, is the innate immune response genes and pathways that are species-specific or common to both bAM and hAM infected with *M. bovis* and *M. tuberculosis*, respectively.

Even though *M. bovis* and *M. tuberculosis* share 99.95 % identity at the genome level, the innate immune responses to the pathogens can be characteristically different. Though the two pathogens encounter common immune response pathways in their preferred host, there are certain genes that are activated or repressed in response to a specific MTBC pathogen. For example, a study by our group investigated the responses of bAM to *M. bovis* and *M. tuberculosis* at 2, 6, 24 and 48 hpi [39]. RNA-seq data from the infected bAM indicated that host genes were differentially expressed between the cells infected with each MTBC strain, suggesting that different immune response pathways may be employed by the bovine host in response to *M. bovis* and *M. tuberculosis*. In this regard, it is important to note that cattle experimentally infected with *M. tuberculosis* displayed minimal pathology, even though diagnostic assays indicated a successful infection [44,45].

In the current study, we extend our macrophage infection approach to encompass comparative integrative analyses for both pathogens (*M. bovis* and *M. tuberculosis*) and their corresponding mammalian hosts (*B. taurus* and *H. sapiens*). To do this we used four different experimental infection groups, each with parallel non-infected control cells: 1) bAM infected with *M. bovis* (bAM-MB); 2) bAM infected with *M. tuberculosis* (bAM-MT); 3) hAM infected with *M. tuberculosis* (hAM-MT); and 4) human monocyte derived macrophages (hMDM) infected with *M. tuberculosis* (hMDM-MT). Transcriptomics data from these experiments were analysed, compared, and interpreted using three separate computational pipelines: 1) differentially expressed (DE) genes (DEG)—the standard approach to catalogue quantitative changes in gene expression levels between experimental groups; 2) differential gene expression interaction networks (DEN), which uses validated molecular interactions extracted from the scientific literature and other sources in combination with DE gene expression values to detect and identify functional gene subnetworks (modules); and 3) combined pathway analysis (CPA), where DE genes are subject to pathway enrichment across six different biological pathway resources. The outputs from these analyses were then integrated with two separate high-resolution bovine and human GWAS data sets with the aim of uncovering novel QTLs and obtaining new insights into the genomic architecture of the TB infection response in cattle and humans.

## 2. Materials and methods

### 2.1. Ethics statement

No animal or human procedures were performed for this computational genomics study, which used published and publicly available RNA-seq transcriptomics data.

### 2.2. General computational methods

All data-intensive computational procedures were performed on a 36-core/72-thread compute server (2 × Intel® Xeon® CPU E5- 2697 v4 processors, 2.30 GHz with 18 cores each), with 512 GB of RAM, 96 TB SAS storage (12 × 8 TB at 7200 rpm), 480 GB SSD storage, and with Ubuntu Linux OS (version 18.04 LTS). The complete computational and bioinformatics workflow is available with additional information as a public GitHub repository (github.com/ThomasHall1688/Bovine\_multi-omic\_integration). The individual components of the experimental and computational workflows are described below and these methodologies and procedures used are modified from those detailed previously by us [41].

### 2.3. Bovine genomic data acquisition

Genome-wide RNA-seq transcriptomics data we previously generated from a 48-h bAM time course challenge experiment using the sequenced *M. bovis* AF2122/97 and *M. tuberculosis* H37Rv strains was used (European Nucleotide Archive (ENA) accession: PRJEB23469). The complete laboratory methods used to isolate, culture and infect bAM with *M. bovis* AF2122/9 and *M. tuberculosis* H37Rv and generate strand-specific RNA-seq libraries using RNA harvested from these cells are described in detail elsewhere [36,39,46]. Briefly, these RNA-seq data were generated using bAM obtained by lung lavage of ten unrelated age-matched 7–12-week-old male Holstein-Friesian calves. These bAM were infected *in vitro* with 1) *M. bovis* AF2122/97, 2) *M. tuberculosis* H37Rv, or 3) incubated with media only. Following total RNA extraction from the two MTBC strain infected groups and the control non-infected bAM, strand-specific RNA-seq libraries were prepared. These comprised *M. bovis*-, *M. tuberculosis*- and non-infected samples from each post-infection time point (2, 6, 24 and 48 hpi across 10 animals (with the exception of one animal that did not yield sufficient alveolar macrophages for *in vitro* infection at 48 hpi). Raw sequence data for RNA-seq

analysis was generated as paired-end  $2 \times 90$  nucleotide reads using an Illumina® HiSeq™ 2000 apparatus. For the present study, the RNA-seq libraries derived from the 24 hpi timepoint for the *M. bovis*-infected bAM, *M. tuberculosis*-infected bAM, and control non-infected bAM were selected for transcriptomics analysis and downstream comparative data mining.

Bovine GWAS data sets for the present study were obtained from intra-breed imputed WGS-based GWAS analyses that used estimated breeding values (EBVs) derived from a bTB resistance phenotype, which were generated for 1502 Holstein-Friesian sires [47]. The bTB phenotype, the WGS-based imputed SNP data, and the quantitative genetics methods are described in detail elsewhere [47]; however, the following provides a brief summary. The bTB resistance phenotype was defined for every animal present during each herd-level bTB breakdown when a bTB reactor or an abattoir case was identified. Cattle that yielded a positive single intradermal comparative tuberculin test (SICTT), post-mortem lymph node lesion, or laboratory culture result/s were coded as “1” (bTB = 1) and all other cattle present in the herd during the bTB-breakdown were coded as “0” (bTB = 0). After phenotype data edits, bTB resistance EBVs were generated for 781,270 cattle and these were used to produce individual sire EBVs for each of three breed groups (Charolais, Limousin and Holstein-Friesian). For the present study we used GWAS data from the Holstein-Friesian breed group, which was generated using EBVs for bTB resistance, and genotype data from 1502 Holstein-Friesian male cattle. After SNP filtering using thresholds for minor allele frequency (MAF < 0.002) and deviation from Hardy-Weinberg equilibrium (HWE;  $P < 1 \times 10^{-6}$ ), there were 15,017, 692 autosomal SNPs for the Holstein-Friesian sire analysis. A single-SNP regression analysis was then performed using sire EBVs for bTB resistance and the nominal GWAS *P* values were used for downstream integrative genomics analyses.

#### 2.4. Human genomic data acquisition

RNA-seq transcriptomics data from a 72-h hAM and hMDM time course challenge experiment using the sequenced *M. tuberculosis* H37Rv strain was used for the human transcriptomics component of this study (GEO accession: GSE114371). The complete laboratory methods used to isolate, culture, and infect hAM and hMDM with *M. tuberculosis* H37Rv and generate strand-specific AmpliSeq™ RNA-seq libraries using RNA harvested from these cells are described in detail elsewhere [40]. Briefly, these RNA-seq data were generated using samples obtained from healthy human donors that tested negative for the tuberculin skin test (TST), under an approved IRB protocol at the Ohio State University Wexner Medical Centre [40]. Human AM and PBMC (used for culturing hMDM) were obtained from different donors and macrophages (hAM and hMDM) were infected with *M. tuberculosis* H37Rv using a multiplicity of infection (MOI) of 2:1. Infected and control non-infected hAM and hMDM were harvested in Trizol reagent after 2, 24 and 72 hpi and RNA was extracted from each experimental group and Ion Torrent sequencing libraries were prepared according to the AmpliSeq™ Library prep kit protocol [48].

The human GWAS data set for resistance to infection by *M. tuberculosis*, which was used for the human integrative genomics work described in this study was obtained from the UK Biobank GeneAtlas (<http://geneatlas.roslin.ed.ac.uk>). Detailed information about this public atlas of genetic associations for 118 non-binary and 660 binary traits catalogued in 452,264 UK Biobank participants of European ancestry has been published by Canela-Xandri et al. [49] and the UK Biobank deep phenotyping and genotyping project is described by Bycroft et al. [50]. The UK Biobank resource contains genotype data for 488,377 participants that was generated using custom high-density genome-wide SNP arrays containing more than 800,000 SNPs. These SNP data were then imputed up to more than 30 million genome-wide SNPs for the GeneAtlas GWAS resource [49]. The hTB GWAS data set used for the present study was generated from 2219 hTB cases and

450,045 disease-free controls. Autosomal SNP filtering criteria consisted of the following: a SNP call rate threshold > 0.98; HWE deviation threshold ( $P < 1 \times 10^{-50}$  on a subset of 344,057 unrelated White British individuals); MAF < 0.001; and an imputation score > 0.9 [49]. There were 9,113,113 SNPs remaining after these filtering steps. In common with the other traits, the binary hTB resistance trait GWAS data set was generated using a linear mixed model (LMM) and a genomic relationship matrix (GRM) using the DISSECT software tool [51]. Supplementary Fig. 1 shows a Q-Q plot obtained for the expected and observed SNP *P* values for the hTB resistance trait associations.

#### 2.5. Differential gene expression analysis of bovine and human RNA-seq data

The RNA-seq data was aligned to the bovine genome ARS-UCD1.2 [52] using the STAR aligner (version 2.7.1b) [53] with a 94 % average read mapping efficiency. Following this, differential gene expression analysis (experimental contrast: infected versus control) was performed using the DESeq2 package (version 1.24.0) [54] with a longitudinal time series design that accounted for time (hours post-infection, hpi) and treatment (control and infected). Multiple testing correction was performed using the Benjamini-Hochberg false discovery rate (FDR) method [55]. Consequently, an individual gene was considered to be differentially expressed (DE) if it exhibited an FDR-adjusted *P*-value less than 0.05 ( $P_{adj.} < 0.05$ ) and an absolute  $\log_2$  fold-change greater than one ( $|\log_2FC| > 1$ ).

The human differential gene expression analysis (experimental contrast: infected versus control) is fully described by Papp and colleagues [40]. Briefly, human AmpliSeq™ RNA-seq data was analysed using the Ion Torrent Mapping Alignment Program (TMAP) [56]. Differential expression analysis was performed with the R package edgeR [57]. For the purposes of the work described in this study, DE genes were selected based on  $FDR P_{adj.} < 0.05$  and  $|\log_2FC| > 1$ .

#### 2.6. Functional gene module identification using differential gene interaction networks

The GeneCards® ([www.genecards.org](http://www.genecards.org); version 4.12) gene compendium and knowledge database is a webtool that integrates multiple sources of biological information on all annotated and predicted human genes [58]. This database was used to identify a set of genes that are functionally associated with the host response to diseases caused by infection with mycobacteria. The search query used was *tuberculosis* OR *mycobacterium* OR *mycobacteria* OR *mycobacterial* and genes were ranked by a GeneCards® statistic—the *Relevance Score*—based on the *Elasticsearch* algorithm [59], which determines the strength of the relationships between genes and keyword terms. Gene IDs were converted to human Ensembl gene IDs [60] and retained for downstream analysis using the InnateDB knowledgebase and analysis platform for systems level analysis of the innate immune response ([www.innatedb.com](http://www.innatedb.com); version 5.4) [61].

A gene interaction network (GIN) was generated with the gene list output from GeneCards® using InnateDB with default settings and this network was visualised using Cytoscape [62]. The jActivesModules Cytoscape plugin (version 3.12.1) [63] was then used to superimpose the four bovine and human RNA-seq DE gene data sets and detect—through a greedy search algorithm—differentially active subnetworks (modules) of genes. This process generated four sets of locally coherent clusters that contain both DE genes and genes that are not DE but are members of the functional gene modules. These modules were identified using: the  $\log_2FC$  and  $P_{adj.}$  values of each DE gene; the overall connectivity of those genes with their immediate module co-members; and the comparison of that connectivity with a background comprised of randomly drawn networks using the same genes, but independent of the base network. Gene module identification revealed a set of modules for each of the four sets of DE genes (bAM infected with *M. bovis* – bAM-MB,



bAM infected with *M. tuberculosis* – bAM-MT, hAM infected with *M. tuberculosis* – hAM-MT, and hMDM infected with *M. tuberculosis* – hMDM-MT). Genes embedded in active modules that were detected as statistically significant for the four experimental contrasts were combined and annotated for downstream GWAS integration.

## 2.7. Combined pathway analyses

For the combined pathway analysis (CPA), the four DE gene data sets were uploaded to the InnateDB pathway analysis webtool [61] and bovine genes were converted to their human orthologs. Selection of genes for CPA was performed using  $P_{\text{adj.}} < 0.05$  and  $|\log_2\text{FC}| > 1$ , with all remaining genes acting as a background distribution by which to compare. The InnateDB overrepresentation analysis (ORA) tool was used, which performs a meta pathway analysis across multiple databases. The four different DE gene sets were queried against the following six pathway resource databases: 1) *Kyoto Encyclopaedia of Genes and Genomes* (KEGG – [www.genome.jp/kegg](http://www.genome.jp/kegg)) [64]; 2) *Integrating Network Objects with Hierarchies Pathway Database* (archived INOH – [dbarchive.biosciencedbc.jp/en/inoh/desc.html](http://dbarchive.biosciencedbc.jp/en/inoh/desc.html)) [65]; 3) *NCI-Nature Pathway Interaction Database* (archived NCI-PID – [www.ndexbio.org](http://www.ndexbio.org)) [66]; 4) *Reactome* (REACTOME – [reactome.org](http://reactome.org)) [67]; 5) *Biocarta Pathways* (archived BIOCARTA – <https://maayanlab.cloud/Harmonizome/datasheet/Biocarta+Pathways>) [68]; and 5) *NetPath* [69]. Following this, the top five pathways from each experimental contrast were identified using InnateDB pathway overrepresentation  $P_{\text{adj.}}$  values (B–H FDR) and curated for downstream analysis. For each of the five top pathways for the four experimental contrasts, all the genes within the pathway, regardless of whether differentially expressed, were extracted, tabulated and annotated for downstream GWAS integration.

## 2.8. Ingenuity® pathway analysis

Ingenuity® Pathway Analysis—IPA® (version 1.1, summer 2020 release; Qiagen, Redwood City, CA, USA) was used to perform a statistical enrichment analysis of DE gene sets for each experimental group [70]. This enabled identification of canonical pathways and functional processes of biological importance in these groups. Following best practice, the background gene set for pathway and functional process enrichment testing was the set of detectable genes for each experimental group [71]. To produce gene sets for the IPA Core Analysis within the recommended range for the number of input entities [70,72] and to include DE genes with small fold-change values, gene sets were filtered using only  $P_{\text{adj.}}$  thresholds of 0.05 for the four experimental groups. In addition, the target species selected was *Homo sapiens* and the cell type used was *Macrophage* (including *Microglia* OR *Bone marrow-derived macrophages* OR *Monocyte-derived macrophages* OR *Other macrophages* OR *Peritoneal macrophages* OR *Macrophages not otherwise specified*) with the *Experimentally Observed* and *High Predicted* confidence settings. For each of the ten top pathways for the four experimental contrasts, all the genes within the pathway, regardless of whether differentially expressed, were extracted, tabulated, and annotated for downstream GWAS integration.

## 2.9. Integration with bovine and human tuberculosis GWAS data sets

To facilitate integration of GWAS data with gene sets generated from functional genomics data analyses, the *gwintE*R software package was used ([github.com/ThomasHall1688/gwintE](https://github.com/ThomasHall1688/gwintE)) [41]. The *gwintE*R tool can be used to test the hypothesis that a specific set of genes is enriched for signal in a GWAS data set relative to the genomic background. For example, such a gene set could be an output from an active gene module network analysis of transcriptomics data from a tissue relevant to the GWAS phenotype. For the present study, *gwintE*R was used as follows: 1) a set of SNPs (the target SNP set) was collated across all genes in a specific gene set at increasing genomic intervals upstream and

downstream from each gene inclusive of the coding sequence (e.g.,  $\pm 0$  kb [intragenic only],  $\pm 10$  kb,  $\pm 20$  kb,  $\pm 30$  kb ...  $\pm 100$  kb); 2) for each genomic interval, a null distribution of 2500 SNP sets, each of which contains the same number of total combined SNPs as the target SNP set, was generated by resampling with replacement from the search space of the total population of SNPs in the GWAS data set; 3) the nominal (uncorrected) GWAS  $P$ -values for the target SNP set and the null distribution SNP sets were converted to local FDR-adjusted  $P$ -values ( $P_{\text{adj.}}$ ) using the *fdrtool* R package (current version 1.2.15) [73]; 4) a permuted  $P$ -value ( $P_{\text{perm.}}$ ) to the test the primary hypothesis for each observed genomic interval target SNP set was generated based on the proportion of permuted random SNP sets where the same or a larger number of SNPs exhibiting significant  $q$ -values (e.g.  $q < 0.05$  or  $q < 0.10$ ) are observed; 5) *gwintE*R generated data to plot  $P_{\text{perm.}}$  results by genomic interval class and obtain a graphical representation of the GWAS signal surrounding genes within the target gene set; 6) a summary output file of all SNPs in the observed target SNP set with genomic locations and  $q$ -values was generated for subsequent investigation.

For the integrative analyses of GWAS data sets with functional genomics outputs from infected bAM (bAM-MB and bAM-MT) and infected hAM and hMDM (hAM-MT and hMDM-MT), four different subsets of genes for each experimental contrast were used: 1) basic DE gene sets that were filtered to ensure manageable computational loads using stringent expression threshold criteria of  $|\log_2\text{FC}| > 2$  and  $P_{\text{adj.}} < 0.01$  for the bAM-MB, bAM-MT and hMDM-MT contrasts and less stringent expression criteria of  $|\log_2\text{FC}| > 1$  and  $P_{\text{adj.}} < 0.05$  for the hAM-MT contrast; 2) for the four contrasts individually (bAM-MB, bAM-MT, hAM-MT, and hMDM-MT), the genes embedded in active modules identified from the GIN using *jActiveModules*; 3) for the four contrasts individually, all genes, regardless of expression, that are members of the top five overrepresented pathways across the KEGG, INOH, NCI-PID, REACTOME, BIOCARTA and NETPATH databases; and 4) for the four contrasts individually, all genes, regardless of expression, that are members of the top 10 enriched pathways obtained using the IPA analyses.

## 3. Results

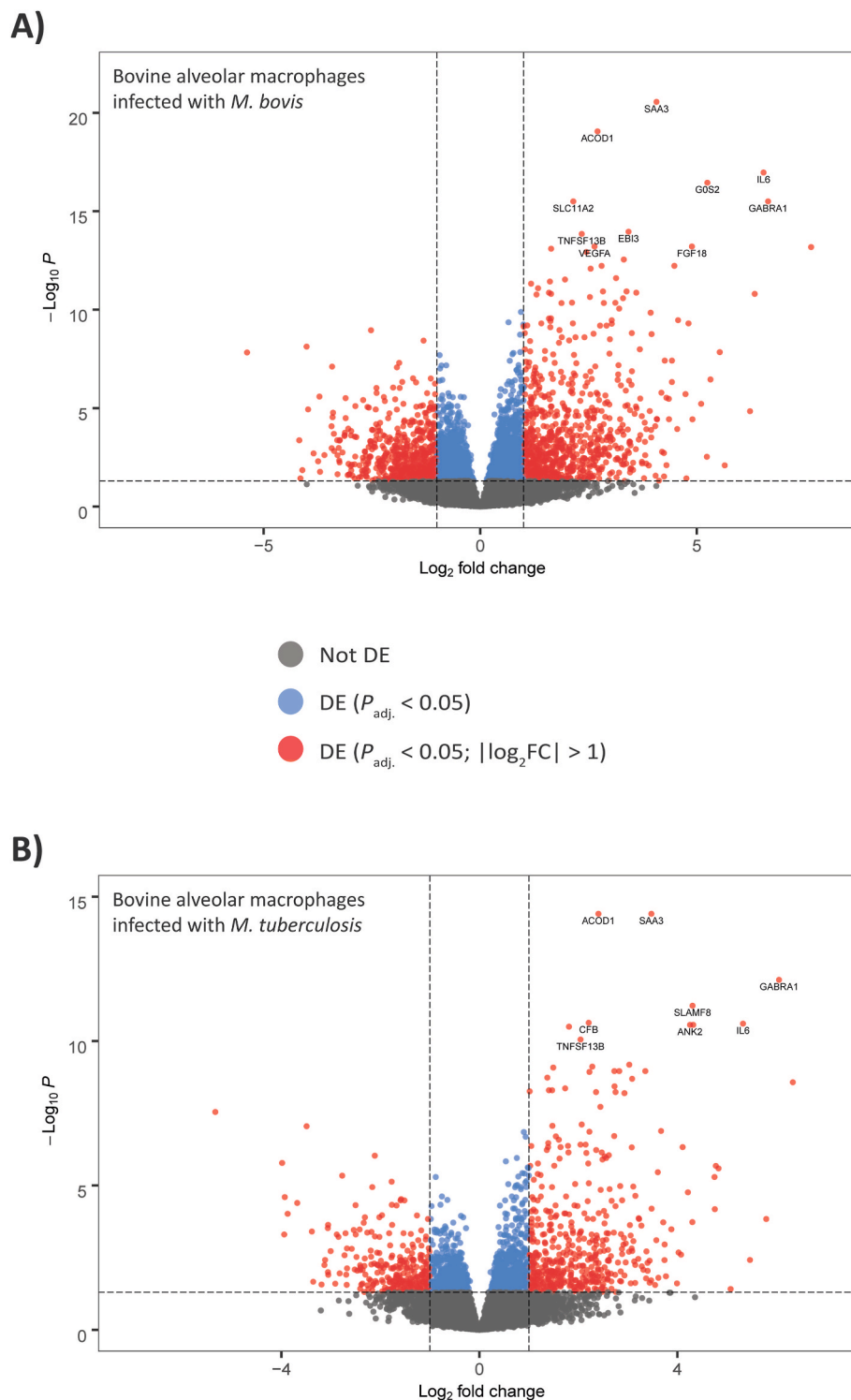
### 3.1. Differential gene expression of *M. bovis*- and *M. tuberculosis*-infected bovine alveolar macrophages

Quality filtering of RNA-seq read pairs yielded a mean of 22,347,042  $\pm$  2,433,115 reads per individual library ( $n = 29$  libraries). A mean of 19,290,873  $\pm$  2,166,803 read pairs (86.31 %) were uniquely mapped to locations in the ARS-UCD1.2 bovine genome assembly. Detailed filtering, mapping, and read count statistics are provided in **Supplementary Information File 1** (Worksheets 1 and 2) and multivariate PCA analysis of the individual animal sample expression data using DESeq2 revealed separation of the control non-infected AM versus the bAM-MB and the bAM-MT at the 24 hpi time point (**Supplementary Fig. 2**).

Using default criteria for differential expression ( $P_{\text{adj.}} < 0.05$ ) and considering the bAM-MB and bAM-MT relative to the control non-infected AM, 3591 DE genes were detected at 24 hpi in the bAM-MB contrast (1879 with increased expression and 1712 with decreased expression); and 1816 DE genes were detected at 24 hpi in the bAM-MT contrast (1039 increased and 777 decreased). **Fig. 1** shows volcano plots of DE genes for the bAM-MB and bAM-MT comparisons (see also **Supplementary Information File 1** – Worksheets 3 and 4).

To ensure manageable computational loads, the input lists of bAM-MB and bAM-MT DE gene sets for GWAS integration with the *gwintE*R tool (DEG-bAM-MB) and DEG-bAM-MT) were filtered with  $|\log_2\text{FC}| > 2$ , and  $P_{\text{adj.}} < 0.01$  and  $P_{\text{adj.}} < 0.05$  for the bAM-MB and bAM-MT contrasts, respectively. The less stringent  $P_{\text{adj.}}$  threshold used for the DEG-bAM-MT input gene set was a consequence of losing one sample after QC filtering ( $n = 9$ ). With these criteria, there were 378 and 284 genes for the DEG-





**Fig. 1.** Volcano plots showing differentially expressed genes for the two bAM experimental contrasts at 24 hpi. **A.** Bovine alveolar macrophages (bAM) infected with *M. bovis* versus control non-infected bAM. **B.** bAM infected with *M. bovis* versus control non-infected bAM. All genes with  $P_{\text{adj.}} < 0.05$  and  $|\log_2 \text{FC}| > 1$  are shown in red, with the top ten genes by  $P_{\text{adj.}}$  value labelled for each group. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

bAM-MB and DEG-bAM-MT input sets, respectively. The two DEG bAM *gwintE* input gene sets are fully detailed in **Supplementary Information File 1** (Worksheets 5 and 6). It is also important to note that 243 genes overlapped between the DEG-bAM-MB and DEG-bAM-MT gene sets.

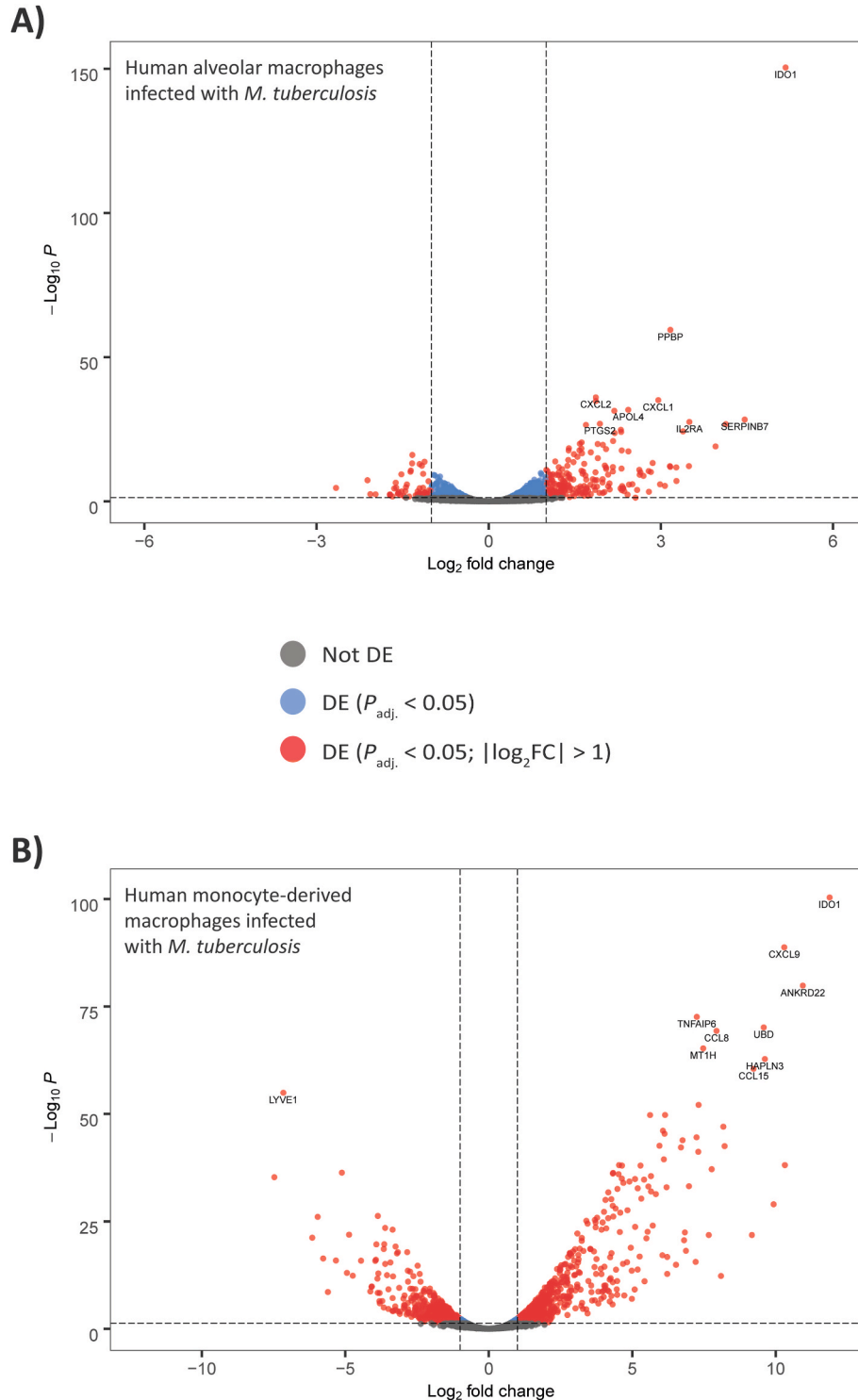
### 3.2. Differential gene expression of *M. tuberculosis*-infected human alveolar macrophages and human monocyte-derived macrophages

A full description of the methodology used to detect DE genes in *M. tuberculosis*-infected hAM and hMDM at 24 hpi is available in the source publication [40]. Briefly, a mean of  $8,557,929 \pm 2,22,249$  read

pairs (93.95 %) from 12 libraries (3 hAM control/3 hAM infected, 3 hMDM control/3 hMDM infected) were uniquely mapped to locations in the GRCh38 human genome assembly. Detailed mapping statistics and read count information on the DE genes are provided in **Supplementary Information File 2** (Worksheets 1 and 2).

Using default criteria for differential expression ( $P_{adj.} < 0.05$ ) and considering the *M. tuberculosis*-infected hAM (hAM-MT) and hMDM

(hMDM-MT) relative to the control non-infected hAM and hMDM, 899 DE genes were detected at 24 hpi for the hAM-MT contrast (567 increased and 332 decreased) and 1545 DE genes were detected at 24 hpi for the hMDM-MT contrast (796 increased and 749 decreased). **Fig. 2** shows volcano plots of DE genes for the hAM-MT and hMDM-MT comparisons (see also **Supplementary Information File 2** – Worksheets 3 and 4).



**Fig. 2.** Volcano plots showing differentially expressed genes for the hAM and hMDM experimental contrasts at 24 hpi. **A.** Human alveolar macrophages (hAM) infected with *M. tuberculosis* versus control non-infected hAM. **B.** Human monocyte-derived macrophages (hMDM) infected with *M. tuberculosis* versus control non-infected hMDM. All genes with  $P_{adj.} < 0.05$  and  $|\log_2 FC| > 1$  are shown in red, with the top ten genes by  $P_{adj.}$  value labelled for each group. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

As with the bovine data, the input lists of hAM-MT and hMDM-MT DE gene sets that were used for GWAS integration with the *gwniteR* tool (DEG-hAM-MT) and DEG-hMDM-MT) were filtered with  $|\log_2FC| > 1$ , and  $P_{adj.} < 0.05$  for the hAM-MT contrast and  $|\log_2FC| > 2$ , and  $P_{adj.} < 0.05$  for the hMDM-MT contrast. The less stringent  $|\log_2FC|$  cut-off used for the DEG-hAM-MT input gene set was a consequence of relatively low expression fold change values observed for the hAM-MT infection experiment, which was discussed by Papp and colleagues [40]. With these criteria, there were 277 and 415 genes for the DEG-hAM-MT and DEG-hMDM-MT input sets, respectively. The DEG-hAM-MB and DEG-hMDM-MT *gwniteR* input gene sets are fully detailed in **Supplementary Information File 2** (Worksheets 5 and 6). It is also important to note that 98 genes overlapped between the DEG-hAM-MT and DEG-hMDM-MT gene sets.

### 3.3. Common and group-specific differentially expressed genes across the four experimental contrasts

**Table 1** shows a breakdown of DE genes across the four infection groups for a range of statistical thresholds and fold-change cut-offs, including the default criteria (FDR  $P_{adj.} < 0.05$ ). The 118 up- and downregulated genes common to all experimental groups at  $P_{adj.} < 0.05$  are shown in **Fig. 3A** and detailed in **Supplementary Information File 2** (Worksheet 7). In addition, **Supplementary Figs. 3–6** show various between-group correlation plots for the 118 DE genes common to all experimental groups. Examples of shared upregulated genes include: *IL6*, which encodes a pro-inflammatory cytokine; *IL1B*, which also encodes an inflammatory protein produced by activated macrophages, *CCL20*, which encodes a chemokine that attracts dendritic cells; and *INHBA*, which encodes a member of the transforming growth factor beta superfamily. Examples of shared downregulated genes include: *PIK3IP1*, which encodes a negative regulator of *PIK3* activity, *CABLES1*, which encodes a regulator of p53/p73-induced apoptotic cell death; and *SORL1*, which encodes a transmembrane signalling receptor that is involved in control of phagocytosed mycobacteria [74]. The overlap of the DE genes across the four experimental contrasts is illustrated in **Fig. 3B**.

### 3.4. Detection of active gene subnetworks in bovine and human infected macrophages using a tuberculosis and mycobacterial infection gene interaction network

The GeneCards® search query generated a total of 2516 gene hits using the terms tuberculosis OR mycobacterium OR mycobacteria OR mycobacterial (**Supplementary Information Files 3 and 4** – Worksheet 1). To provide a computationally manageable number of genes for an InnateDB input data set, a GeneCards® relevance score (GCRS) threshold  $>2.5$  was used. This produced an input list of 260 functionally prioritised genes for generation of an InnateDB gene

**Table 1**

Differentially expressed genes detected in *M. bovis*-infected bovine AM (bAM-MB), *M. tuberculosis*-infected bovine AM (bAM-MT), *M. tuberculosis*-infected human AM (hAM-MT), *M. tuberculosis*-infected human MDM (hMDM-MT) relevant to controls.

Post-infection time point	$P_{adj.} < 0.05$ ; $ \log_2FC  > 0$ (increased/decreased)	$P_{adj.} < 0.05$ ; $ \log_2FC  > 1$ (increased/decreased)	$P_{adj.} < 0.01$ ; $ \log_2FC  > 0$ (increased/decreased)	$P_{adj.} < 0.01$ ; $ \log_2FC  > 1$ (increased/decreased)
bAM-MB	3620 (1898/1722)	1345 (764/581)	2059 (1168/891)	933 (577/356)
bAM-MT	1819 (1040/779)	734 (468/266)	805 (527/278)	422 (297/125)
hAM-MT	899 (567/332)	277 (211/66)	589 (405/184)	240 (194/46)
hMDM-MT	1545 (796/749)	1286 (665/621)	1063 (597/466)	1042 (584/458)

interaction network (GIN) and the top ten genes from this list ranked by GCRS were: interferon gamma receptor 1 (*IFNGR1*), toll like receptor 2 (*TLR2*), interleukin 12 receptor subunit beta 1 (*IL12RB1*), interleukin 12B (*IL12B*), solute carrier family 11 member 1 (*SLC11A1*), signal transducer and activator of transcription 1 (*STAT1*), interferon gamma receptor 2 (*IFNGR2*), cytochrome b-245 beta chain (*CYBB*), tumour necrosis factor (*TNF*), and interferon gamma (*IFNG*).

The large GIN produced by InnateDB starting with the input list of 260 functionally prioritised genes was visualised using Cytoscape and consisted of 6951 nodes (individual genes) and 21,653 edges (gene interactions) (**Fig. 4**). Following visualisation of the large GIN in Cytoscape, the jActivesModules Cytoscape plugin was used to detect statistically significant differentially activated subnetworks (modules) within the large GIN. This consisted of superimposing the DE genes from all for experimental groups onto the larger GIN, creating differentially expressed gene interaction networks for each of the four experimental contrasts. **Supplementary Information File 3** (Worksheets 2) provides information for all gene interactions and superimposed bovine DE genes represented in **Fig. 4A** and **B** and **Supplementary Information File 4** (Worksheets 2) provides information for all gene interactions and superimposed human DE genes represented in **Fig. 4C** and **D**. To illustrate active module subnetwork capture, the large GINs in **Fig. 4A–D** are accompanied by an example subnetwork of genes and gene interactions. The top five subnetworks from each experimental group were retained for downstream analyses.

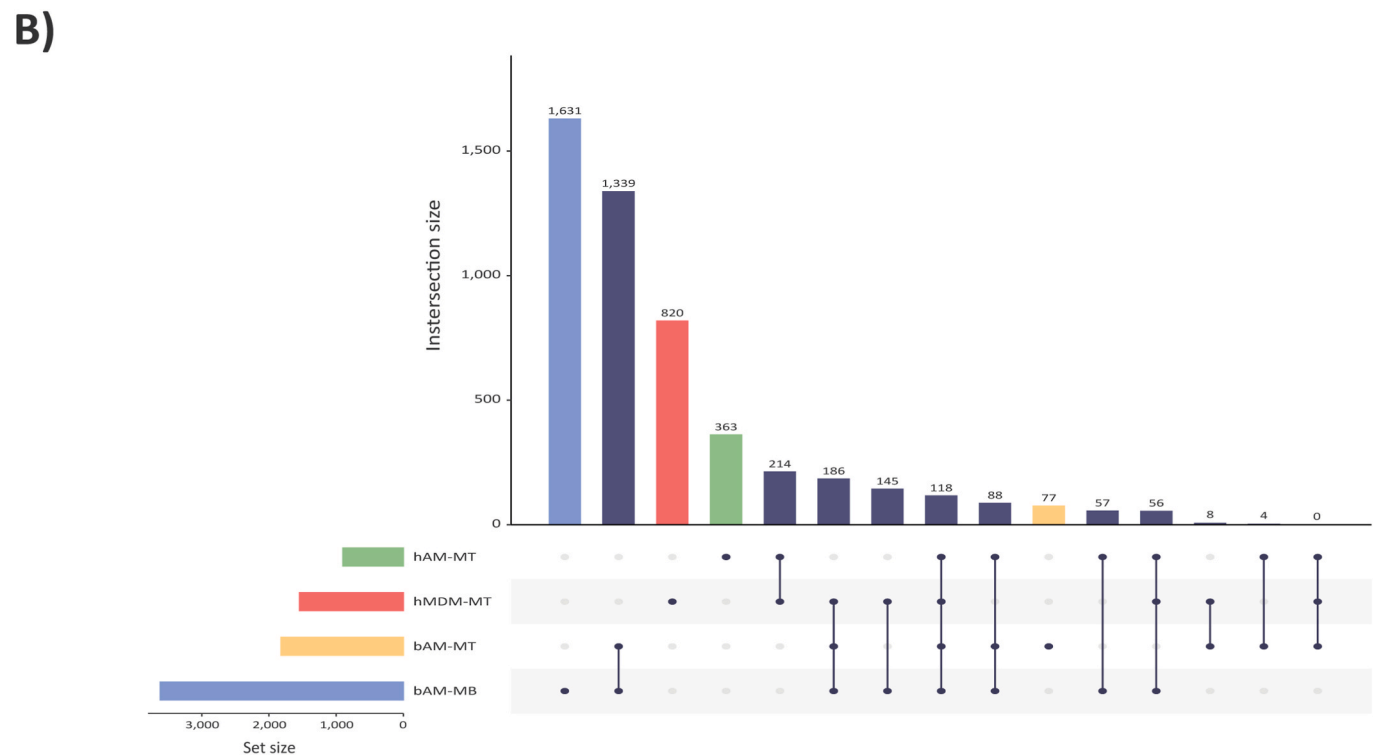
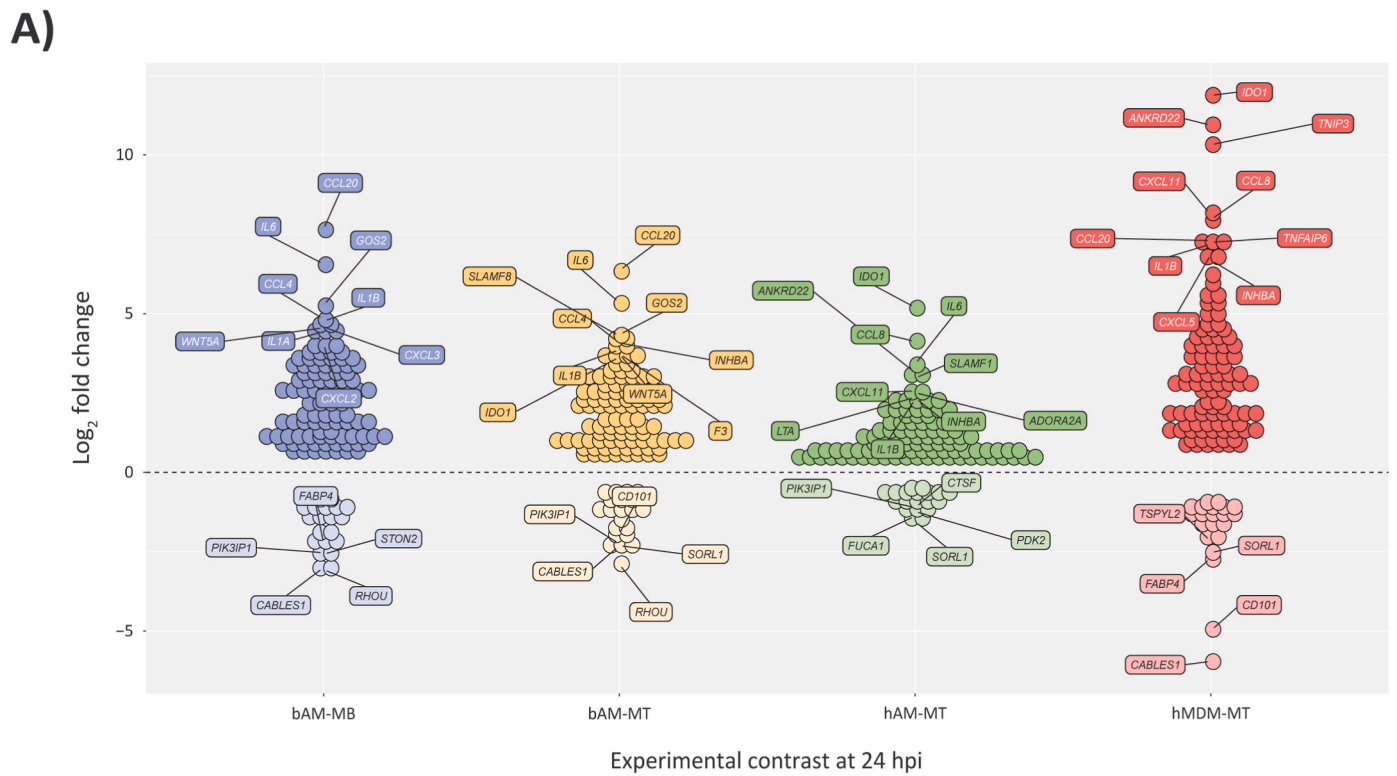
Module 1 from the bAM-MB group (M1-DEN-bAM-MB) contained 266 genes and there were 237 genes in M2-DEN-bAM-MB, 269 genes in M3-DEN-bAM-MB, 204 genes in M4-DEN-bAM-MB, and 235 genes in M5-DEN-bAM-MB (**Supplementary Information File 3** – Worksheet 3). Module 1 from the bAM-MT group (M1-DEN-bAM-MT) contained 148 genes and there were 158 genes in M2-DEN-bAM-MT, 160 genes in M3-DEN-bAM-MT, 137 genes in M4-DEN-bAM-MT, and 52 genes in M5-DEN-bAM-MT (**Supplementary Information File 3** – Worksheet 4). Module 1 from the hAM-MT group (M1-DEN-hAM-MT) contained 140 genes and there were 155 genes in M2-DEN-hAM-MT, 65 genes in M3-DEN-hAM-MT, 168 genes in M4-DEN-hAM-MT, and 180 genes in M5-DEN-hAM-MT (**Supplementary Information File 4** – Worksheet 3). Module 1 from the hMDM-MT group (M1-DEN-hMDM-MT) contained 139 genes and there were 137 genes in M2-DEN-hMDM-MT, 149 genes in M3-DEN-hMDM-MT, 93 genes in M4-DEN-hMDM-MT, and 182 genes in M5-DEN-hMDM-MT (**Supplementary Information File 4** – Worksheet 4).

After concatenation of each set of five modules and removal of duplicates, the input gene set derived from the bAM-MB group (DEN-bAM-MB) contained 398 genes and the DEN-bAM-MT input gene set contained 259 genes (**Supplementary Information File 3** – Worksheets 5 and 6). The DEN-hAM-MT input gene set contained 262 genes and the DEN-hMDM-MT input gene set contained 239 genes (**Supplementary Information File 4** – Worksheets 5 and 6).

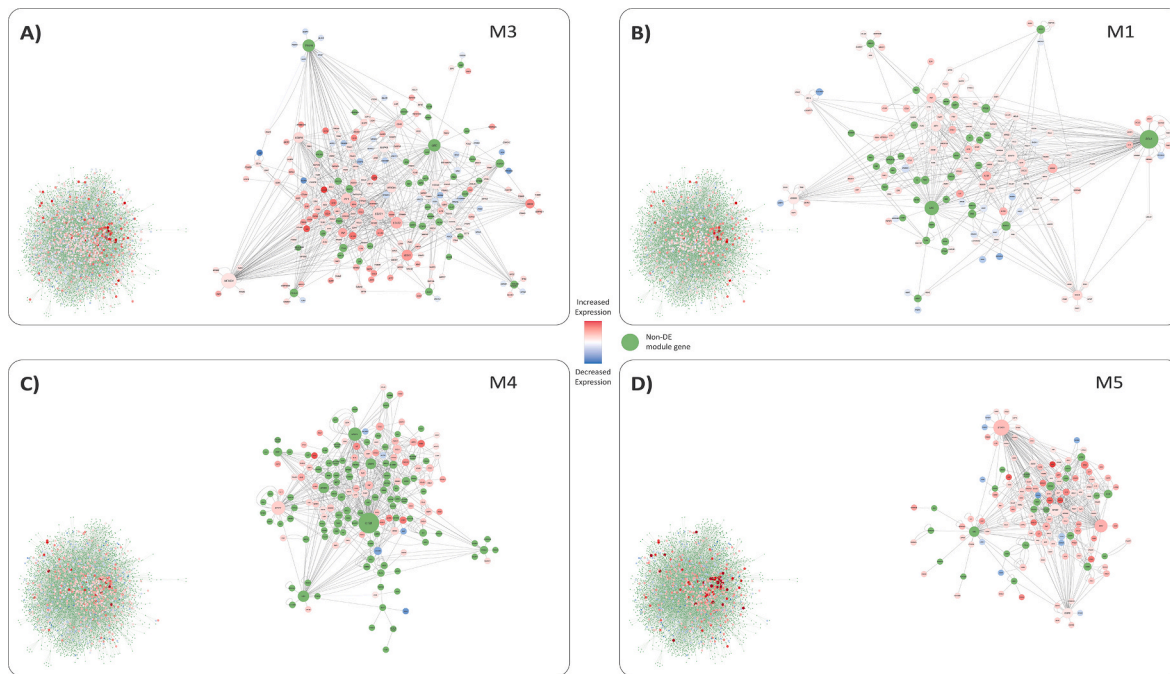
### 3.5. Combined open-source pathway analysis of differentially expressed genes

The DE genes from each of the four experimental contrasts was analysed individually for enriched pathways across the KEGG, INOH, NCI-PID, REACTOME, BIOCARTA and NETPATH pathway repositories using the InnateDB pathway ORA tool. Pathways with a B–H FDR  $P_{adj.} < 0.05$  were considered enriched. Analysis of the DE genes generated from the bAM-MB experimental contrast identified 46 upregulated pathways and one downregulated pathway. After filtering by the  $P_{adj.}$  values, the top five enriched pathways were *Cytokine-cytokine receptor interaction* (KEGG), *GPCR signalling* (INOH), *Jak-STAT signalling pathway* (KEGG), *RIG-I-like receptor signalling pathway* (KEGG), and *JAK STAT regulation* (INOH), which were all upregulated (see **Supplementary Information File 5** – Worksheet 1). Analysis of the DE genes generated from the bAM-MT experimental contrast identified 21 upregulated pathways and no





**Fig. 3.** Comparison of shared differentially expressed (DE) genes across the four experimental contrasts. A. Plot of 118 common shared up- and downregulated genes across the four experimental contrasts. The top 10 upregulated and the top 5 downregulated genes are shown for each group (by log<sub>2</sub>FC). B. UpSetR plot showing all overlapping DE genes ( $P_{adj.} < 0.05$ ) colour-coded by experimental contrast. The UpSetR plot was generated using the UpSetR package [125]. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 4.** Gene interaction network analysis and functional module identification using jActiveModules and differentially expressed genes across four experimental contrasts at 24 hpi. **A.** bovine alveolar macrophages (bAM) infected with *M. bovis*. **B.** bAM infected with *M. tuberculosis*. **C.** human alveolar macrophages (hAM) infected with *M. tuberculosis*. **D.** human monocyte-derived macrophages (hMDM) infected with *M. tuberculosis*. Example active modules are shown for each contrast.

downregulated pathways. The top five enriched pathways were *Cytokine-cytokine receptor interaction* (KEGG), *GPCR signalling* (INOH), *Jak-STAT signalling pathway* (KEGG), *HIF-1-alpha transcription factor network* (NCI-PID), and *IL23-mediated signalling events* (NCI-PID) (see **Supplementary Information File 5 – Worksheet 2**).

Analysis of the DE genes generated from the hAM-MT experimental contrast identified 31 upregulated pathways and no downregulated pathways. The top five enriched pathways were *Cytokine-cytokine receptor interaction* (KEGG), *Chemokine receptors bind chemokines* (REACTOME), *Peptide ligand-binding receptors* (REACTOME), *Class A/1 Rhodopsin-like receptors* (REACTOME), and *GPCR ligand binding* (REACTOME) (see **Supplementary Information File 6 – Worksheet 1**). Analysis of the DE genes generated from the hMDM-MT experimental contrast identified 111 upregulated pathways and three downregulated pathways. All three downregulated pathways were also significantly upregulated; however, they contained genes that were also significantly downregulated. All pathways with downregulated genes were involved in GPCR signalling. The top five enriched pathways were *Cytokine-cytokine receptor interaction* (KEGG), *Cytokine Signalling in Immune system* (REACTOME), *Interferon gamma signalling* (REACTOME), *Interferon Signalling* (REACTOME), and *Class A/1 (Rhodopsin-like receptors)* (REACTOME) (see **Supplementary Information File 6 – Worksheet 2**). **Supplementary Fig. 7** shows the top 16 overrepresented biological pathways for each of the four experimental contrasts.

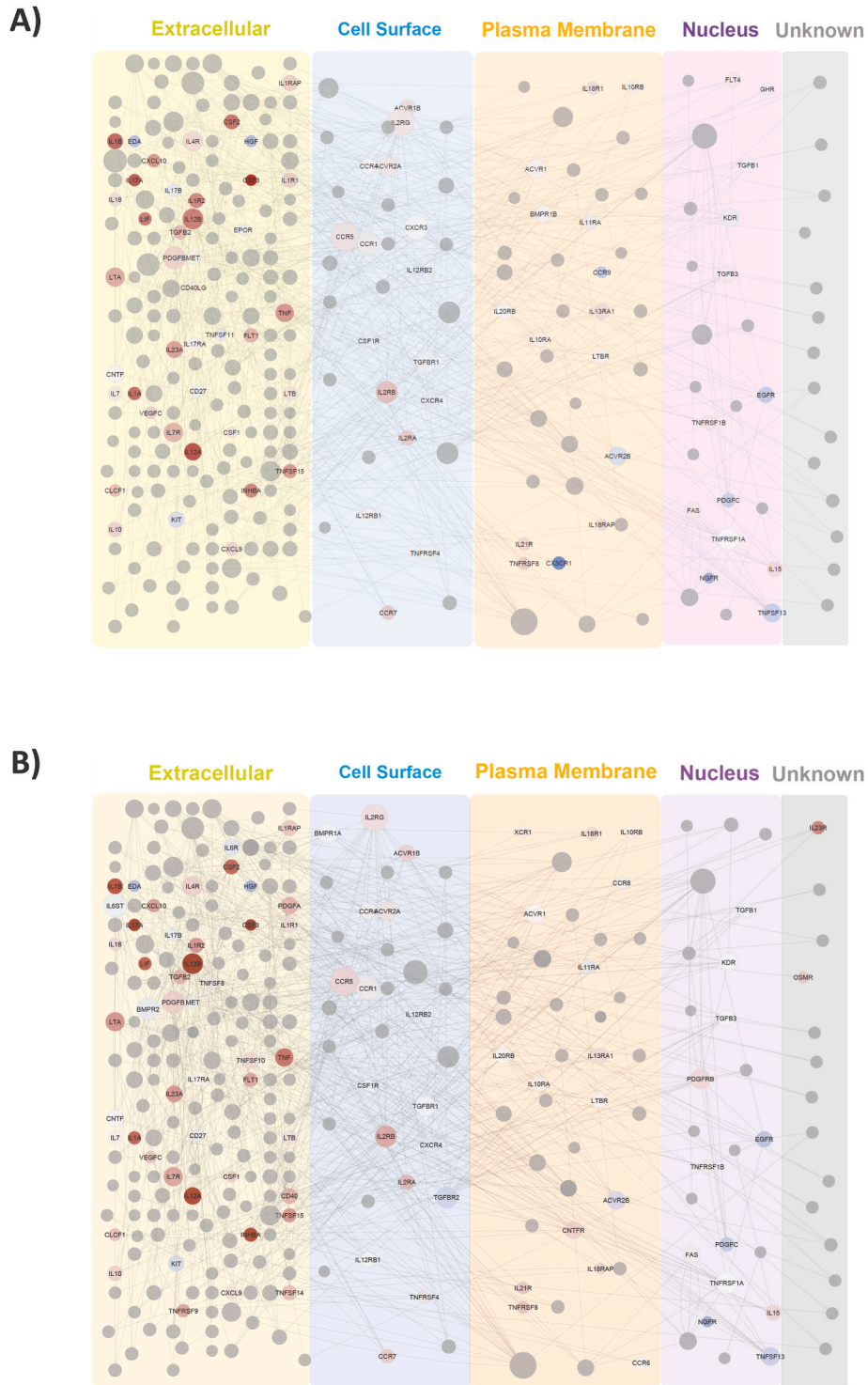
The top overrepresented pathway common to all four experimental contrasts was *Cytokine-cytokine receptor interaction* (KEGG) (see **Figs. 5 and 6**). This pathway was also the most significantly overrepresented pathway for all four experimental contrasts, though the DE genes enriched in this pathway differ between all four groups (**Figs. 5 and 6**). The genes from each of the top five pathways identified from the CPA for each experimental contrast, regardless of if the genes were DE or not, were combined, filtered for duplicates, and catalogued for GWAS integration. The CPA input gene set derived from the bAM-MB group (CPA-

bAM-MB) contained 557 genes and the CPA-bAM-MT input gene set contained 473 genes (**Supplementary Information File 5 – Worksheets 3 and 4**). The CPA-hAM-MT input gene set contained 640 genes and the CPA-hMDM-MT input gene set contained 737 genes (**Supplementary Information File 6 – Worksheets 3 and 4**).

### 3.6. Ingenuity® pathway analysis (IPA) of differentially expressed genes

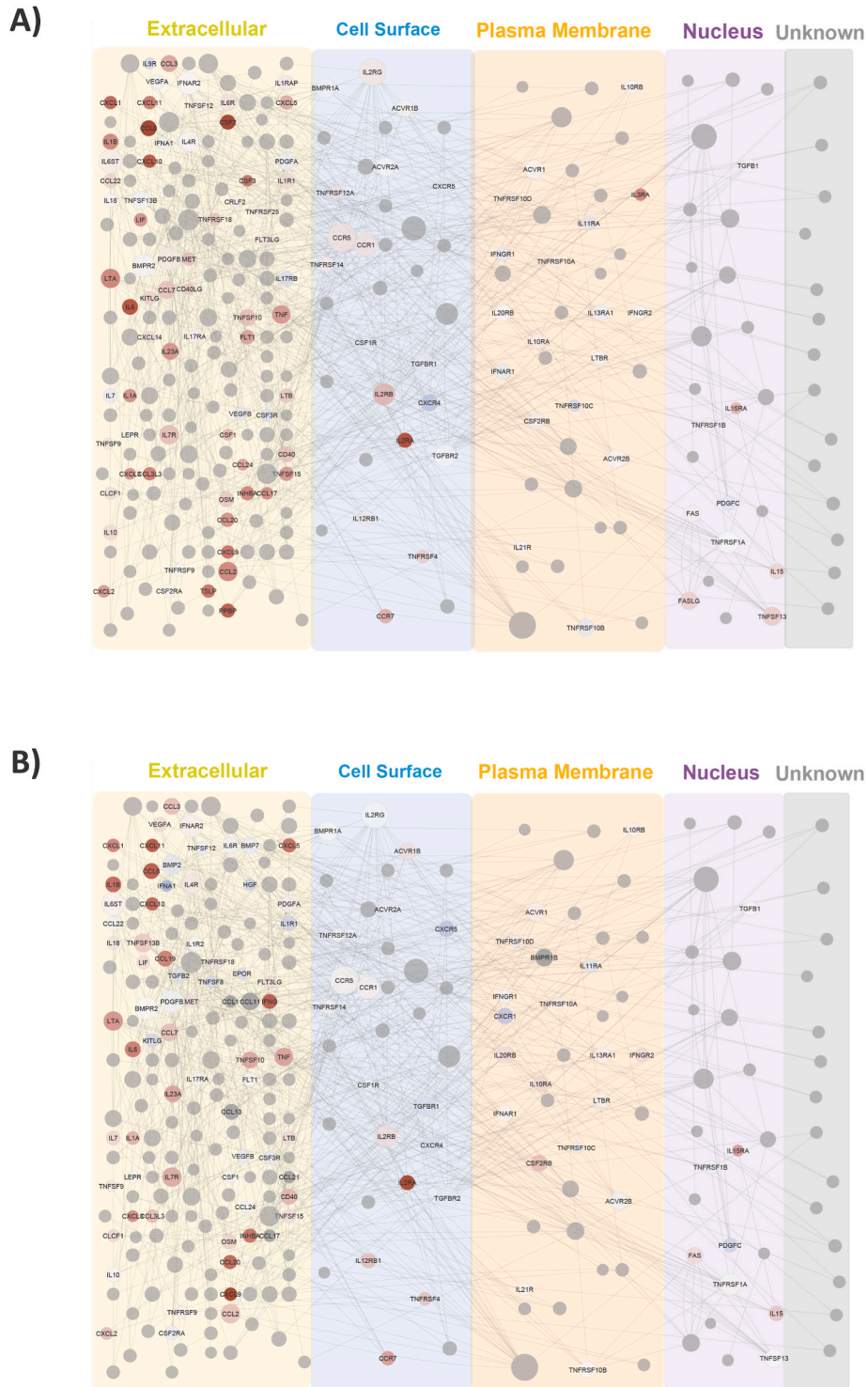
The DE genes from each experimental contrast were analysed for enriched pathways using IPA and the IPA Knowledge Base. This resulted in 1961 input genes (1073 upregulated and 888 downregulated) from a background detectable set of 7103 from bAM infected with *M. bovis* (bAM-MB), 978 input genes (571 upregulated and 407 downregulated) from a background detectable set of 6970 genes from bAM-MT, 581 input genes (354 upregulated and 227 downregulated) from a background detectable set of 6627 genes from hAM-MT and 941 input genes (509 upregulated and 432 downregulated) from a background detectable set of 6825 genes from hMDM-MT.

Using the B–H method for multiple test correction in IPA (FDR  $P_{adj}$  < 0.05), there were 68 statistically significant enriched IPA canonical pathways from the bAM-MB group, 201 from the bAM-MT group, 61 from the hAM-MT group, and 118 from the hMDM-MT group. The genes from each of the top 10 pathways for each of the four experimental groups (**Supplementary Information File 7 and 8 – Worksheets 1 and 2**), regardless of if the genes were DE or not, for each group from the IPA analysis were combined, duplicates removed, and catalogued for data integration using the bovine GWAS and human GWAS data sets (**Supplementary Information File 7 and 8 – Worksheets 3 and 4**). The IPA gene set derived from bAM infected with *M. bovis* (IPA-bAM-MB) contained 386 genes and the IPA-bAM-MT input set contained 276 genes (**Supplementary Information File 7 – Worksheets 3 and 4**). The IPA gene set derived from hAM infected with *M. tuberculosis* (IPA-hAM-MT) contained 207 genes and the IPA-hMDM-MT input set contained 274

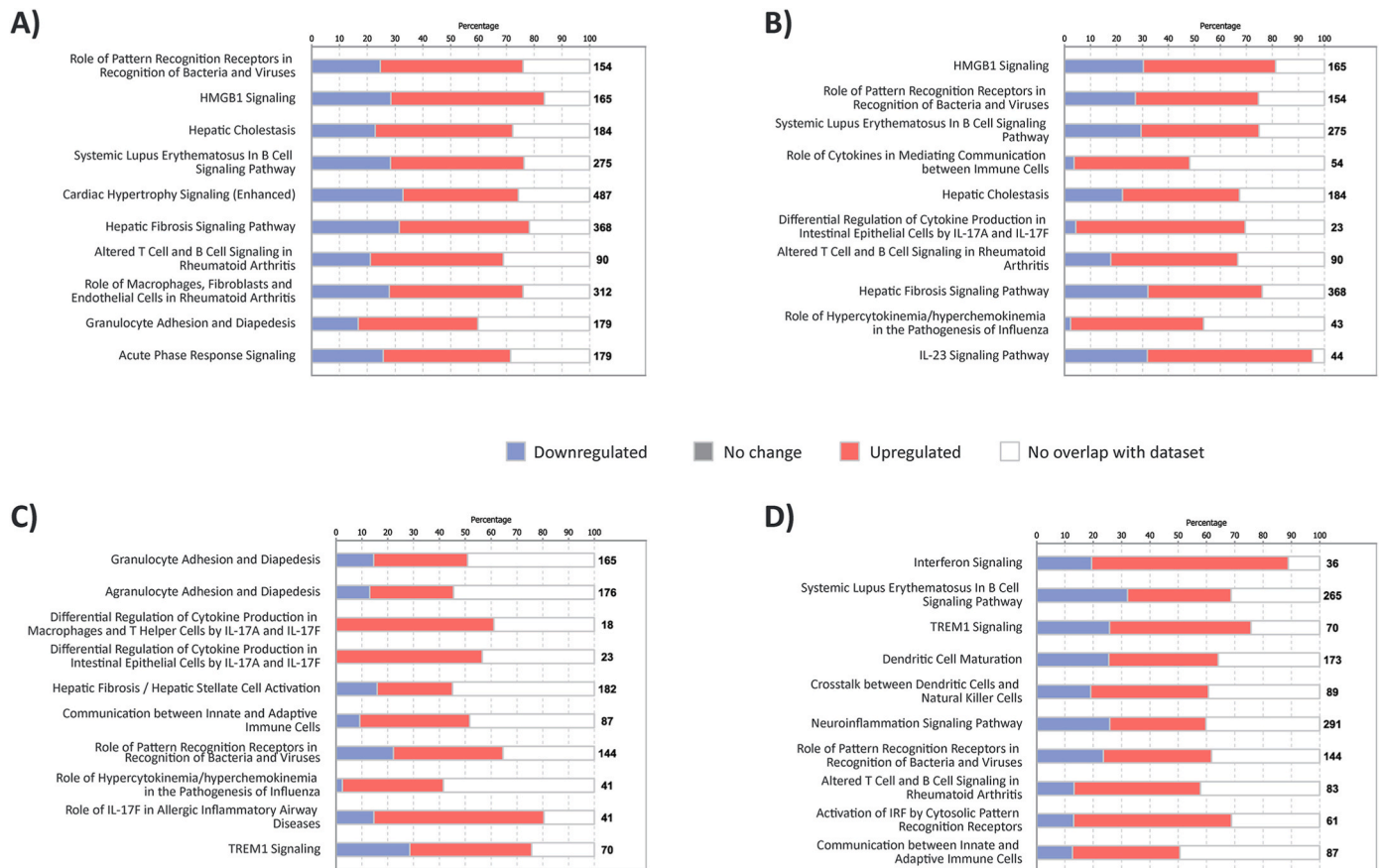


**Fig. 5.** Protein network schematics and cellular locations of the top overrepresented common biological pathway - *Cytokine-cytokine receptor signalling* (bovine). **A.** *M. bovis*-infected BAM. **B.** *M. tuberculosis*-infected BAM. Red and blue nodes indicate upregulation and downregulation, respectively. Grey nodes indicate no detectable expression change. The size of the nodes corresponds to the degree of connectivity (number of interactions). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)





**Fig. 6.** Protein network schematics and cellular locations of the top overrepresented common biological pathway – *Cytokine-cytokine receptor signalling* (human). **A.** *M. tuberculosis*-infected hAM. **B.** *M. tuberculosis*-infected hMDM. Red and blue nodes indicate upregulation and downregulation, respectively. Grey nodes indicate no detectable expression change. The size of the nodes corresponds to the degree of connectivity (number of interactions). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 7.** Top 10 enriched biological pathways for each experimental group from the Ingenuity® Pathway Analysis (IPA). Stacked horizontal bar charts showing the top twenty canonical pathways from IPA in order of descending  $-\log_{10} P_{adj}$ : **A.** bAM-MB, **B.** bAM-MT, **C.** hAM-MT, and **D.** hMDM-MT The numbers to the right of the bars show the number of genes in the pathway and the colours of the bars indicate the percentages of these genes that are downregulated, upregulated, show no change in expression, or have no overlap with the data set. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

**Table 2**  
The 16 different input bovine and human gene sets used for GWAS data integration.

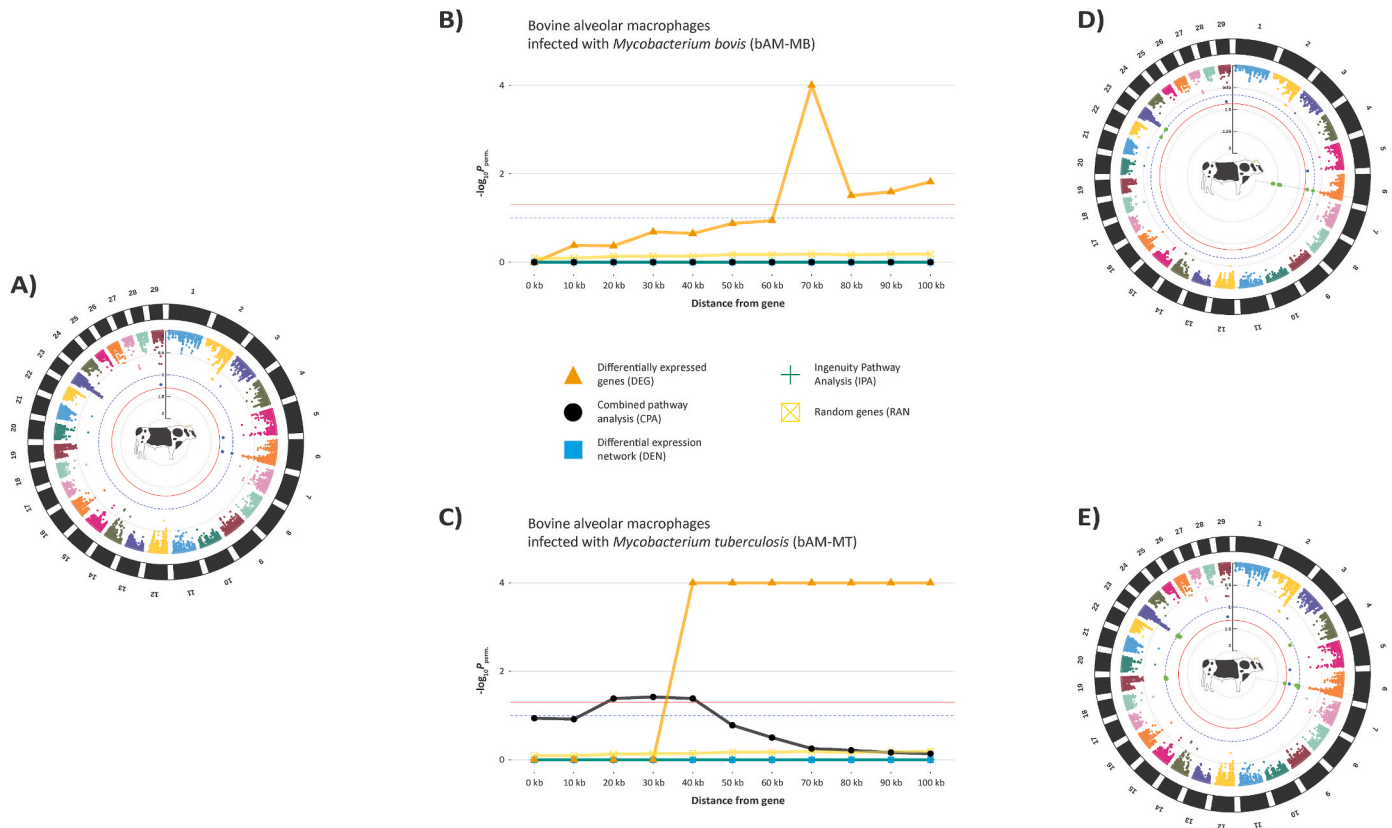
Infection group	Differentially expressed genes (DEG)		Differential expression network (DEN)		Combined Pathway Analysis (CPA)		Ingenuity Pathway Analysis (IPA)	
	Input gene set code	No. of genes	Input gene set code	No. of genes	Input gene set code	No. of genes	Input gene set code	No. of genes
Bovine alveolar macrophages infected with <i>M. bovis</i>	DEG-bAM-MB	378	DEN-bAM-MB	398	CPA-bAM-MB	557	IPA-bAM-MB	386
Bovine alveolar macrophages infected with <i>M. tuberculosis</i>	DEG-bAM-MT	390	DEN-bAM-MT	259	CPA-bAM-MT	473	IPA-bAM-MT	276
Human alveolar macrophages infected with <i>M. tuberculosis</i>	DEG-hAM-MT	277	DEN-hAM-MT	262	CPA-hAM-MT	473	IPA-hAM-MT	207
Human monocyte-derived macrophages infected with <i>M. tuberculosis</i>	DEG-hMDM-MT	415	DEN-hMDM-MT	239	CPA-hMDM-MT	737	IPA-hMDM-MT	274

gene (**Supplementary Information File 8** – Worksheets 3 and 4). **Fig. 7** summarises the top 10 IPA enriched pathways for each of the four experimental contrasts.

Some of the top 10 enriched pathways (**Fig. 7**) were common to two or more experimental groups (see **Supplementary Information File 8** – Worksheet 5) and include: *Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses*; *Activation of IRF by Cytosolic Pattern Recognition Receptors*; *TNFR2 Signalling*; *Dendritic Cell Maturation*; *Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis*; and *Hepatic Cholestasis*.

**3.7. Integration of functional genomics outputs with GWAS data sets and identification of additional species-specific SNP-trait associations**

The sixteen input gene sets generated from the four separate analyses of DE genes generated from the bAM-MB, bAM-MT, hAM-TB, and hMDM-TB infection challenge experiments are summarised in **Table 2** and further detailed in **Supplementary Information Files 1–8**. In addition to these sixteen putative functionally relevant gene sets, 100 sets of 250 genes randomly sampled from the bovine genome (BOV-RAN), and 100 sets of 250 genes randomly sampled from the human genome (HUM-RAN), were used for statistical context and comparison (data not shown). The results from the integrative analyses using the



**Fig. 8.** Integration of bAM functional genomics outputs and GWAS data for *M. bovis* infection resistance in Holstein-Friesian cattle. **A.** Circular Manhattan plots showing GWAS results pre-integration with blue and red highlighted data points indicating binned SNP clusters with  $FDR P_{adj.} < 0.10$  and  $< 0.05$ , respectively. **B./C.** Line plots of permuted  $P$  values ( $-\log_{10}P_{perm.}$ ) across different genomic intervals for SNPs from eight different input gene sets and random genes (**B.** bAM-MB and **C.** bAM-MT). **D./E.** Circular Manhattan plots showing GWAS results post-integration with blue and red data highlighted points indicating binned SNP clusters with  $FDR P_{adj.} < 0.10$  and  $< 0.05$ , respectively, and SNPs enriched by *gwinteR* coloured green (**D.** bAM-MB and **E.** bAM-MT). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

*gwinteR* tool with the DEG-bAM-MB, DEG-bAM-MT, DEN-bAM-MB, DEN-bAM-MT, CPA-bAM-MB, CPA-bAM-MT, IPA-bAM-MB, IPA-bAM-MT, and BOV-RAN input gene sets are summarised graphically in Fig. 8 and fully detailed in **Supplementary Information File 9** (Worksheets 1–4). Likewise, the results from the integrative analyses using the equivalent human data sets are summarised graphically in Fig. 9 and fully detailed in **Supplementary Information File 10** (Worksheets 1–4). Inspection of Fig. 8B and C shows that, in terms of SNP enrichment ( $P_{perm.} < 0.05$ ), the integrative analyses using *gwinteR* were most effective with the DEG-bAM-MB, DEG-bAM-MT, and CPA-bAM-MT data sets. A possible explanation for the inferior performance of the other data sets is that these methods rely heavily on gene orthology, i.e., all the bovine DE genes used to generate the DEN, CPA and IPA gene sets were first converted to human gene IDs. More generally, it is also important to note that the literature-supported relationships among molecules, diseases and biological processes in the IPA Knowledge Base are based on human, mouse (*Mus musculus*), and rat (*Rattus norvegicus*) data, which may place caveats on the use of IPA with other mammals [70]. In this regard, methods that do not rely on ortholog conversion, such as gene co-expression networks, seem to perform better. Fig. 9B and C demonstrate that the human data integration performed better with multiple techniques; statistically significant SNP enrichment was achieved with the DEG-hAM-MT, DEG-hMDM-MT, DEN-hAM-MT, and CPA-hAM-MT data sets. Fig. 10 provides information on the 32 genes that were proximal to SNPs significantly associated with bTB and hTB disease resilience, which were identified through integration of functional genomics outputs with the bovine and human GWAS data sets.

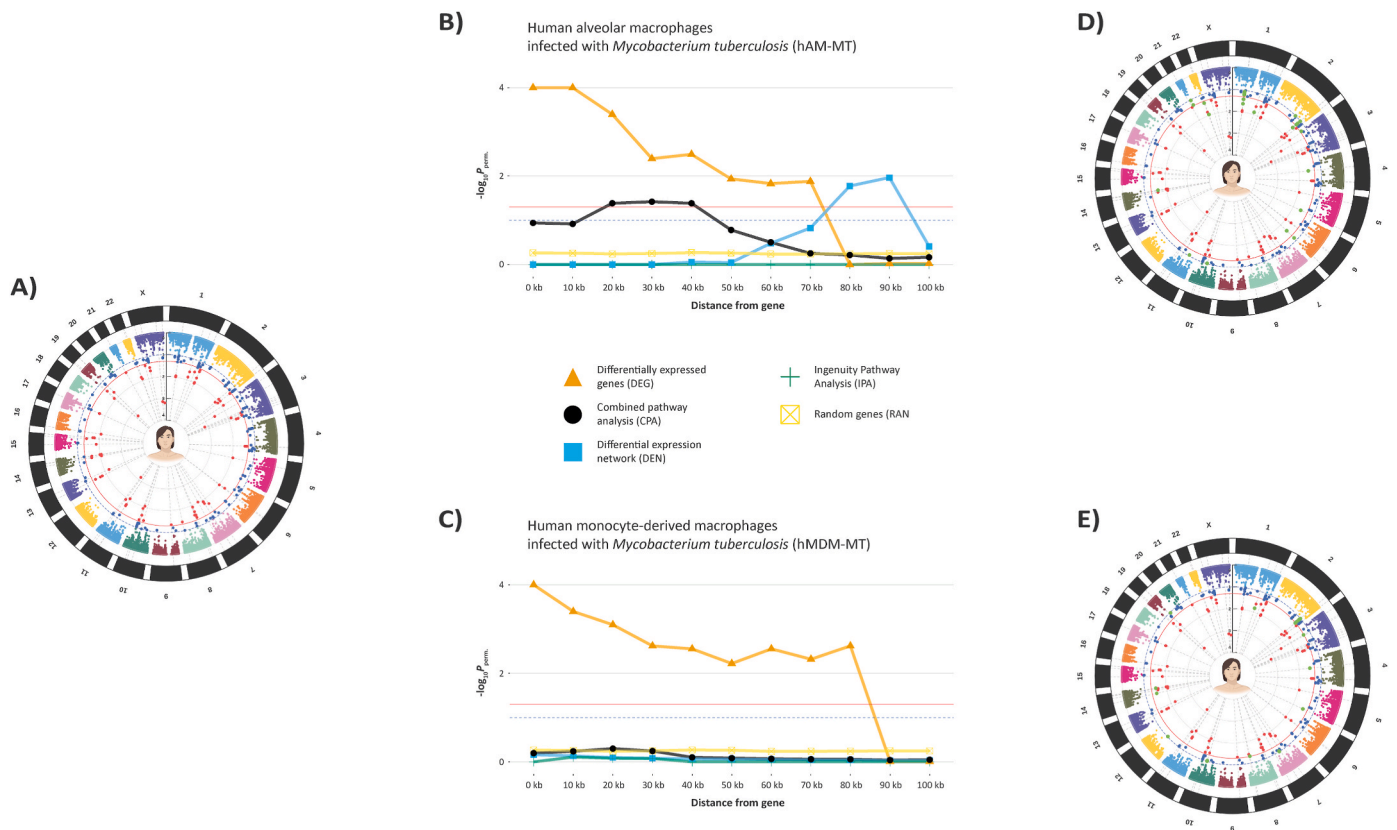
## 4. Discussion

### 4.1. Bovine and human macrophages share common immune response genes and pathways to mycobacterial infection

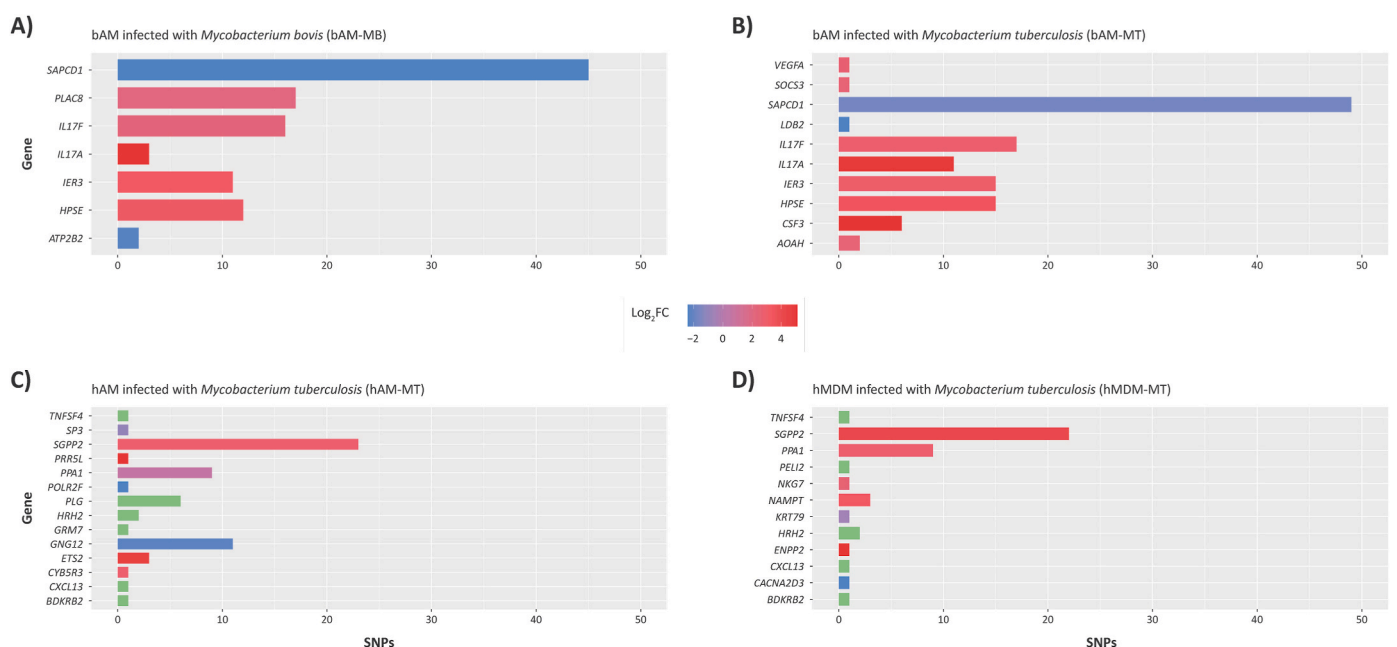
The work described in this study has generated new scientific information regarding host-pathogen interaction during the initial stages of infection for *M. bovis* and *M. tuberculosis*, which are the archetypal mycobacterial pathogens of cattle and humans, respectively. Using four separate multi-omics analysis workflows, we demonstrate that there are striking similarities in the patterns of differential expression and the cellular pathways that are perturbed by these infections in bovine and human macrophages. In terms of understanding host-pathogen interaction in an evolutionary context, it is now considered likely that the MTBC complex emerged relatively recently during the Neolithic period and probably within the last 5000 years [75–77]. Therefore, the recent shared evolutionary histories of *M. bovis* and *M. tuberculosis* may account for congruent host responses from the pathogen perspective. However, from the perspective of the mammalian host species, what is particularly remarkable—given the similarities in macrophage responses to infection—is that artiodactyls and primates last shared a common ancestor more than 80 million years ago [78].

As shown in Fig. 3 and **Supplementary Information File 2** (Worksheet 7), there were 118 DE genes in common across the four different experimental infection groups (bAM-MB, bAM-MT, hAM-MT, and hMDM-MT) 24 hpi. All these genes were significantly DE, and the  $\log_2$  fold change values were highly correlated for each pairwise comparison, with the highest correlation coefficient for these 118 genes





**Fig. 9.** Integration of hAM and hMDM functional genomics outputs and GWAS data for *M. tuberculosis* infection resistance in humans. **A.** Circular Manhattan plots showing GWAS results pre-integration with blue and red highlighted data points indicating binned SNP clusters with  $FDR P_{adj.} < 0.10$  and  $< 0.05$ , respectively. **B./C.** Line plots of permuted  $P$  values ( $-\log_{10}P_{perm.}$ ) across different genomic intervals for SNPs from eight different input gene sets and random genes (**B.** hAM-MT and **C.** hMDM-MT). **D./E.** Circular Manhattan plots showing GWAS results post-integration with blue and red data highlighted sets indicating binned SNP clusters with  $FDR P_{adj.} < 0.10$  and  $< 0.05$ , respectively, and SNPs enriched by *gWinteR* coloured green (**D.** hAM-MT and **E.** hMDM-MT). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



**Fig. 10.** Histograms of statistically significant SNP numbers and proximal genes enriched through integration of bovine and human functional genomics outputs with bTB and hTB GWAS data sets. **A.** bAM infected with *M. bovis* (bAM-MB). **B.** bAM infected with *M. tuberculosis* (bAM-MT). **C.** hAM infected with *M. tuberculosis* (hAM-MT). **D.** hMDM infected with *M. tuberculosis* (hMDM-MT). The colour of each bar indicates whether the gene was upregulated (red), downregulated (blue), or whether there was no significant differential expression for genes identified through the subnetwork (module) or pathway analyses. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

observed between the bAM-MB and bAM-MT contrasts ( $r = 0.9894$ ) (Supplementary Fig. 3). The correlation of expression differences for these genes between the hAM-MT and hMDM-MT infection groups was lower ( $r = 0.8088$ ) (Supplementary Fig. 4), reflecting cell type differences in the transcriptomes of these human macrophages infected with *M. tuberculosis*. The hAM-MT/bAM-MB and hAM-MT/bAM-MT infection group comparisons for these 118 shared DE genes produced correlation coefficients of  $r = 0.7122$  and  $r = 0.7170$ , respectively (see Supplementary Figs. 5 and 6); the interspecific comparison with the same MTBC pathogen (*M. tuberculosis*) being marginally more correlated.

Several groups of related genes were consistently upregulated for all four experimental infection contrasts, suggesting a core MTBC infection response gene repertoire intrinsic to macrophages from both species of mammal. Several of these genes encode different types of cytokines including interleukins (*IL1A*, *IL1B*, *IL23*, *IL6*, and *IL7R*); chemokines (*CCL2*, *CCL3*, *CCL4*, *CCL8*, *CCL20*, *CXCL2*, *CXCL3*, *CXCL5* and *CXCL11*); and members of the TNF family, such as *TNF*, *CD40*, *LTA* and *TNFAIP6*. In this regard, many of these genes were previously identified by our group as corresponding to a core innate immune response for bovine AM infected with both *M. bovis* and *M. tuberculosis* across a 48-h post-infection time course [39]. Interestingly, the group of cytokines that did exhibit a marked species-specific pattern of gene expression were interferons. Many interferon-induced genes were differentially expressed in each experimental infection contrast group (bAM-MB: 18, bAM-MT: 6, hAM-MT: 12, and hMDM-MT: 14); however, only *IFIT2* was DE across all four groups, suggesting that for MTBC infections of AM or MDM, interferon-induced proteins tend to function in a strain- or host species-specific fashion, as opposed to interleukins and chemokines, which were consistently DE across the experimental groups regardless of macrophage cell type, strain or host species.

Other genes that were DE across the four experimental contrasts included: *STAT1*, which was consistently upregulated and is a key modulator of the immune response to mycobacterial infection and a critical transcription factor in the JAK-STAT and interferon signalling pathway [79]; *PIK3IP1*, which was consistently downregulated and is a key negative regulator of the PI3K-AKT signalling pathway [42,80]; and *IDO1*, which was consistently upregulated and is involved with limitation and accumulation of tryptophan in immune cells, which can induce apoptosis, limit growth of intracellular pathogens, and control immunopathology resulting from unchecked immune responses [81]. Recent studies have indicated that the IDO1 protein may facilitate bacterial survival in macrophages infected with *M. tuberculosis*, and that downregulation of *IDO1* may lead to enhanced control of infection [82,83]. It is therefore noteworthy in the present study that *IDO1* is uniformly upregulated 24 hpi across all experimental groups (bAM-MB, linear fold change of +13.3; bAM-MT, 12.3-fold; hAM-MT, +36.1-fold; and hMDM-MT, +3765.6-fold) (see Supplementary File 2 – Worksheet 7). Finally, *SLAMF1*, which was consistently upregulated and induces expression of *IL12* and *TNF* in the presence of LPS [84], and regulates phagosome maturation and recruitment of the PI3K complex 2; for example, with *Escherichia coli* infection of macrophages this occurs through recognition of *OmpC* and *OmpF* on the bacterial cell surface [85,86]. Our group has also recently shown that the activity of *SLAMF1* is regulated at the epigenetic level, with H3K4me3 deposition at the transcriptional start site of *SLAMF1* modulating its expression in bovine AM infected with *M. bovis* [42].

In addition to common response genes, all four experimental infection contrasts had the same top response pathway, *Cytokine-cytokine receptor interaction* (Figs. 5 and 6). While the DE genes in each pathway varied among the four groups, this pathway was statistically highly enriched across all the groups (Supplementary Fig. 7), which again highlights the evolutionary conservation and functional importance of cytokine signalling during MTBC infections of mammalian macrophages.

As part of the differential expression network analysis (DEN), a large gene interaction network (GIN) was used with  $\log_2$  fold-change and FDR

$P_{adj}$  values as the key parameters to identify and extract smaller functional modules/subnetworks (see Fig. 4 and Supplementary Information Files 3 and 4). Five functional modules were extracted from the large GIN for each of the four experimental infection contrasts. Each of these subnetworks exhibits a scale-free topology [87,88]; most gene nodes within the network interact with one other gene (low degree), while a small subset interacts with substantially more (high degree). These nodes tend to be direct regulators of other genes such as transcription factors, or subunits of important proteins. Comparison of genes with high degree across all 20 modules (Supplementary Information File 4 – Worksheet 7), demonstrated that all four groups share genes of high degree, suggesting that the key regulators of gene modules that respond to intracellular mycobacterial infection are common across the two species and macrophage cell type and MTBC strain. Genes present in at least one functional module from all experimental infection contrasts included inflammation-related transcription factors such as *CEBPP*, *EGR1*, *IRF1*, *NFKB1*, *NFKBIA*, *STAT1*, *JAK2*, *UBC*, and *TNF*. In this regard, using a differential network approach to analyse microarray gene expression data from bovine MDM challenged with *M. bovis* and an attenuated *M. bovis* BCG vaccine strain, our group previously identified *NFKB1* and *EGR1* as key hub and bottleneck gene nodes, respectively [89].

Interestingly, although the *IRF1* gene was present in all experimental infection contrasts, and *IFNG*, *IFNG1*, *IFNG2*, *IFNG3* and *IFNG4* were present as high degree genes in some contrasts, only *IFNG* was present in more than one group and no interferon gene was present in the functional modules identified for the bAM-MT infection contrast. This observation indicates that interferon genes in these experimental contrasts exhibit specific functions depending on cell type, species and MTBC strain. There was also one other gene that did not appear in subnetworks identified for the bAM-MB infection contrast, but that was present in the remaining three *M. tuberculosis*-challenged groups (bAM-MT, hAM-MT, and hMDM-MT). This was the vitamin D receptor gene (*VDR*), which mediates the immunological function of vitamin D3, an activator of macrophages [27,90]. Vitamin D deficiency has been implicated in susceptibility to both bTB and hTB [91,92]; therefore, it is somewhat surprising that *VDR* was not a gene node in any of the modules from the bAM-MB infection contrast. Taking into consideration the inter- and intra-species robustness of the functional modules, further research is needed to improve identification and functional annotation of gene response modules extracted from knowledge-based gene interaction networks.

#### 4.2. Genes enriched for bTB and hTB GWAS SNPs in both species play central roles in NF- $\kappa$ B signalling and the formation of the granuloma

After all of the input gene sets were integrated with the bTB and hTB GWAS data sets using *gwintE*R, a total of 32 unique genes (12 bovine and 20 human genes; see Fig. 10) exhibited intragenic SNPs or SNPs within 100 kb up- and downstream that were enriched for associations with infection susceptibility/resistance. There was intraspecies overlap in enriched genes, such as bovine *IL17A* and human *CXCL13*, but there was no interspecies overlap. However, it is important to note that there was an enrichment observed for genes involved with initiation, formation, and regulation of the granuloma in both species.

Granulomas are densely compact, organized aggregates of immune cells consisting of epithelioid cells, blood-derived infected and uninfected macrophages, foamy macrophages, and multinucleated giant cells [93]. The enriched bovine SNPs lay in proximity to six genes related to granuloma biology. These are *CSF3*, *HPSE*, *IER3*, *IL17A*, *IL17F* and *VEGFA*. The HPSE protein (heparanase) has a role in inflammation and cell adhesion during granuloma formation and is also expressed in peripheral granulomas [94,95]. Expression of *IER3* has been observed in chronic lung granulomas [96], and it may act as an inhibitor or contributor to their formation. *VEGFA* expression in macrophages regulates granuloma formation in the non-angiogenic pathway during TB

disease and recruits immune cells to the granuloma [97].

The *CSF3* gene (aka *GCSF*) encodes a protein with many immunological roles, such as survival, proliferation, and polarization of macrophages [98,99]; it has also been detected in multiple macrophage infection studies using pathogenic MTBC strains [39,40,42]. Although it has been shown to be upregulated in sarcoidosis granulomas [100], the role of *CSF3* in TB-induced granuloma formation has not been extensively studied; however, induction of granulomatous tissue by *CSF3* has been documented in other conditions, such as granulomatous dermatitis and chronic granulomatous disease in immunocompromised patients [101,102]. However, it is important to note that colony stimulating factor proteins have recently been shown—in a membrane fusion-independent manner—to stimulate the formation of multinucleated giant cells (MGCs), which are key components of TB-induced granulomas. For example, *CSF1* in conjunction with a persistent ligand, such as TNF, can induce MGCs from bone marrow myeloid progenitors via polyploidy, typically associated with replication stress or DNA damage [103], which supports a larger role for CSF proteins in the formation of TB granulomas. In addition, our group has recently shown that the *CSF3* locus is epigenetically regulated through H3K4me3 deposition in bovine AM infected with *M. bovis* [42]. Consequently, based on the multi-omics results presented here that highlight *CSF3* in macrophage responses to mycobacterial infections, further research into the role of *CSF3* in bTB and hTB is warranted.

For the human functional genomics outputs and GWAS data integration results, the enriched SNPs identified were proximal to five genes associated with granuloma biology: *CACNA2D3*, *CXCL13*, *ENPP2*, *PLG*, and *TNFSF4*. Differential levels of *CACNA2D3*, a protein in the voltage-dependent calcium channel complex, had been demonstrated within granulomatous structures [104] and it has been proposed to play a role in the CRR5 pathway in macrophages, which regulates chemotaxis during inflammatory responses [105]. The *CXCL13* chemokine recruits B cells to sites of inflammation, including the granuloma [106,107]. Macrophages do not typically express this gene, which would explain the lack of differential expression observed in any of the four experimental infection contrasts. However, this gene was included as part of the GIN analysis, highlighting the importance of integrative analytical approaches that extend beyond simple catalogues of DE genes. The protein product of the *ENPP2* gene interacts with IL-13 during the IL-13-mediated sarcoidosis granulomatous response [108] and the product of *ENPP2* (aka autotaxin) phospholipase action, lysophosphatidic acid, converts recruited monocytes from bone marrow into macrophages [109]. Increased levels of *PLG* (plasminogen) have been shown to encourage granuloma formation; a recent study showed that *M. bovis* BCG bacilli coated in *PLG* leads to a decrease in phagocytosis, and an increase in granuloma formation [110]. *TNFSF4* has been proposed as an inhibitor to granuloma formation; polymorphisms in the *TNFSF4* gene are associated with Sjogren's syndrome, phenotypic effects of which include irregular granuloma formation [111,112]. Other human genes that were associated with enriched SNPs include *PLAC8*, *LDB2*, *PLOR2F*, *GNG12*, and *PRRS1*, which have roles in inflammation, pathway initiation and recruitment, but no direct link to granuloma formation for these genes has yet been established.

From the host's perspective, the ideal outcome of granuloma formation in TB disease is elimination or sterilisation of macrophages infected with pathogenic MTBC strains. However, in cases where the granuloma becomes necrotic, it is clearly beneficial to the pathogen and transmission of the infection [93]. As such, formation of the granuloma can be interpreted as a failure of alveolar macrophages to sufficiently control MTBC pathogens. Consequently, a better understanding of host genes and genomic variation that determines and influences formation and success of the granuloma could lead to novel therapeutics, diagnostics, prognostics, and in the case of livestock, potential targets for genome editing and genome-enabled breeding programmes for bTB disease resistance/resilience [113–115]. The results presented here relating to granuloma formation also emphasise the importance of

*M. bovis* infection in cattle and bTB disease as an animal model of hTB [20,21,116].

In addition to genes associated with granuloma formation, many of the 32 bovine and human genes identified through bTB and hTB GWAS integration also encode important components of the NF- $\kappa$ B signalling pathway. The nuclear factor- $\kappa$ B (NF- $\kappa$ B) complex constitutes a family of inducible transcription factors that regulate a large array of genes associated with innate and adaptive immunity and inflammatory responses [117]. Including both bovine and human genes, there were three genes encoding activators (*GNG12*, *IL17*, *PELI2*), two genes encoding inhibitors (*TNFSF4*, *IER3*, *SOC3*) and twelve genes encoding downstream targets (*GRM7*, *CACNA2D3*, *CSF3*, *CYB5R3*, *CXCL13*, *KRT79*, *NKG7*, *PRRS1*, *SOC3*, *SP3*, *VEGFA*, *HPSE*). Notable downstream target genes included *HPSE*, which encodes a lung injury serum protein that promotes neutrophil adherence and inflammation [118] and *GRM7*, which encodes a protein that modulates adaptive immunity and inflammation [119]. Many of the downstream targets are also regulators of granuloma initiation and formation, again highlighting the importance of this feature of TB disease pathogenesis. Because NF- $\kappa$ B-related genes were highlighted in the original GWAS integration, a supplementary analysis was conducted where 17 NF- $\kappa$ B subunit, activator and inhibitor genes (*NFKB1*, *NFKB2*, *RELA*, *RELB*, *REL*, *NFKBIA*, *NFKBIB*, *NKRF*, *NFKBIL1*, *NFKBIZ*, *NFKBIE*, *NDFIP1*, *NKAP*, *NKAPL*, *NKIRAS2*, *NFKBID*, and *NKAPP1*) were directly used as an input gene set for GWAS integration using the *gwintE*R tool (**Supplementary Information File 11**). This resulted in detection of 163 significantly enriched SNPs (6 human and 157 bovine) within or proximal to these NF- $\kappa$ B-associated genes.

It is important to note that the disparity between the human and bovine SNP enrichment for NF- $\kappa$ B genes may be a consequence of the phenotypes used in the bTB GWAS and the hTB GWAS. The bTB phenotype is a susceptibility/resistance sire EBV generated from *M. bovis* infection diagnostic epidemiology data [47]. Conversely, the hTB case-control GWAS phenotype is more directly related to disease resilience because a relatively large proportion of the hTB control cohort (>20%) may have been latently infected with *M. tuberculosis* [120,121]. Consequently, the overrepresentation of bovine genes with enriched SNPs in the NF- $\kappa$ B signalling pathway may reflect the importance of innate immune responses in the phenotype used for the bTB GWAS. The hTB case-control GWAS phenotype, on the other hand, may be more directly associated with adaptive immunity [122–124].

## 5. Conclusion

Although *M. bovis* and *M. tuberculosis* are 99.95% similar at the nucleotide level, they exhibit specific host tropism to sustain across different hosts yet generate comparable disease in their preferred host. In this regard, key to further understanding this host tropism will be elucidation of the shared and species-specific immunological mechanisms underpinning the bovine and human host responses to establishment of infection by *M. bovis* and *M. tuberculosis*, respectively. In addition, identification of critical mycobacterial infection response pathways shared between the two species underlines the importance of the bovine model for understanding human TB. Using four transcriptomics data sets that represent the responses to infection of bAM, hAM, and hMDM, common and distinct genes and gene response pathways across both host species have been identified. Using three different analysis pipelines, a key observation was the role played by the cytokine-cytokine receptor interaction pathway, which was the most enriched pathway in all four experimental groups across seven functional databases. After integration of the downstream functional genomics outputs with the GWAS data sets, 32 bovine and human genes contained or were proximal to SNPs significantly associated with resistance to infection or disease resilience. A striking feature of this result is that 11 of these genes, across all four groups, are directly involved with the formation of the granuloma, while 18 are involved with NF- $\kappa$ B



signalling. This indicates that the overall response pathways and gene regulatory networks (GRNs) are comparable across host species and mycobacterial strains. Our work therefore provides underpinning data with which to elucidate the molecular basis of host tropism across the animal- and human-adapted MTBC.

### Declaration of competing interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Data availability

The bovine RNA-seq data set was generated by the authors and can be obtained from the European Nucleotide Archive (ENA) with accession identifier: PRJEB23469. The human RNA-seq generated by Papp et al. can be obtained from the NCBI Gene Expression Omnibus (GEO) with accession identifier: GSE114371. Bovine GWAS summary statistics data were obtained from a published study that provides additional information about sequence and genotype data availability [47]. The human GWAS data set for resistance to infection by *M. tuberculosis* used for the integrative genomics work in this study was obtained from the UK Biobank *GeneATLAS* (<http://geneatlas.roslin.ed.ac.uk>) [49]. The computer code required to repeat and reproduce the analyses is available from a GitHub repository ([https://github.com/ThomasHall1688/Human\\_Bovine\\_comparison](https://github.com/ThomasHall1688/Human_Bovine_comparison)). The computer code used to integrate the functional genomics outputs and GWAS data sets is available from a GitHub repository (<https://github.com/ThomasHall1688/gwintE>) Bovine RNA-seq filtering and mapping statistics are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.83bk3j9q6>.

### Funding

This study was supported by Science Foundation Ireland (SFI) Investigator Programme Awards to DEM and SVG (grant nos. SFI/08/IN.1/B2038 and SFI/15/IA/3154); a Department of Agriculture, Food and the Marine (DAFM) project award to DEM (TARGET-TB; grant no. 17/RD/US-ROI/52); and a European Union Framework 7 project grant to DEM (no: KBBE-211602-MACROSYS).

### CRediT authorship contribution statement

**Thomas J. Hall:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Gillian P. McHugo:** Formal analysis, Software, Visualization, Writing – review & editing. **Michael P. Mullen:** Formal analysis, Software, Visualization, Writing – review & editing. **James A. Ward:** Formal analysis, Software, Visualization, Writing – review & editing. **Kate E. Killick:** Formal analysis, Software, Visualization, Writing – review & editing. **John A. Browne:** Investigation, Methodology, Writing – review & editing. **Stephen V. Gordon:** Conceptualization, Funding acquisition, Methodology, Resources, Writing – review & editing. **David E. MacHugh:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

### Acknowledgements

The authors would like to thank Donagh Berry and Siobhán Ring for provision of cattle GWAS data. We would also like to thank Carol Correia, Joseph Crispell, Eamonn Gormley, and Kerri Malone for useful discussions.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.tube.2023.102453>.

## References

- [1] Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, Chater KF, van Sinderen D. Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev* 2007;71:495–548. <https://doi.org/10.1128/mmb.00005-07>.
- [2] Garnier T, Eiglmeier K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler PR, Parkhill J, Barrell BG, Cole ST, Gordon SV, Hewinson RG. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* 2003;100:7877–82. <https://doi.org/10.1073/pnas.1130426100>.
- [3] Brites D, Gagneux S. Co-evolution of *Mycobacterium tuberculosis* and *Homo sapiens*. *Immunol Rev* 2015;264:6–24. <https://doi.org/10.1111/imr.12264>.
- [4] Gagneux S. Ecology and evolution of *Mycobacterium tuberculosis*. *Nat Rev Microbiol* 2018;16:202–13. <https://doi.org/10.1038/nrmicro.2018.8>.
- [5] World Health Organization. Global tuberculosis Report 2022. Geneva, Switzerland: World Health Organization; 2022. <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022>.
- [6] Steele JH. Introduction (Part 2 regional and country status reports). In: Thoen CO, Steele JH, editors. *Mycobacterium bovis* infection in animals and humans. Ames, IA, USA: Iowa State University Press; 1995. p. 169–72.
- [7] Schiller I, Oesch B, Vordermeier HM, Palmer MV, Harris BN, Orloski KA, Buddle BM, Thacker TC, Lyashchenko KP, Waters WR. Bovine tuberculosis: a review of current and emerging diagnostic techniques in view of their relevance for disease control and eradication. *Transbound Emerg Dis* 2010;57:205–20. <https://doi.org/10.1111/j.1865-1682.2010.01148.x>.
- [8] Waters WR, Palmer MV, Buddle BM, Vordermeier HM. Bovine tuberculosis vaccine research: historical perspectives and recent advances. *Vaccine* 2012;30:2611–22. <https://doi.org/10.1016/j.vaccine.2012.02.018>.
- [9] Olea-Popelka F, Muwonge A, Perera A, Dean AS, Mumford E, Erlacher-Vindel E, Forcella S, Silk BJ, Ditiu L, El Idrissi A, Raviglione M, Cosivi O, LoBue P, Fujiwara PI. Zoonotic tuberculosis in human beings caused by *Mycobacterium bovis* – a call for action. *Lancet Infect Dis* 2017;17:e21–5. [https://doi.org/10.1016/S1473-3099\(16\)30139-6](https://doi.org/10.1016/S1473-3099(16)30139-6).
- [10] Vayr F, Martin-Blondel G, Savall F, Soulat JM, Deffontaines G, Herin F. Occupational exposure to human *Mycobacterium bovis* infection: a systematic review. *PLoS Neglected Trop Dis* 2018;12:e0006208. <https://doi.org/10.1371/journal.pntd.0006208>.
- [11] Kock R, Michel AL, Yeboah-Manu D, Azhar EI, Torrelles JB, Cadmus SI, Brunton L, Chakaya JM, Marais B, Mboera L, Rahim Z, Haider N, Zumla A. Zoonotic tuberculosis - the changing landscape. *Int J Infect Dis* 2021;113(Suppl 1):S68–72. <https://doi.org/10.1016/j.ijid.2021.02.091>.
- [12] Clegg TA, Good M, Duignan A, Doyle R, Blake M, More SJ. Longer-term risk of *Mycobacterium bovis* in Irish cattle following an inconclusive diagnosis to the single intradermal comparative tuberculin test. *Prev Vet Med* 2011;100:147–54. <https://doi.org/10.1016/j.prevetmed.2011.02.015>.
- [13] Pesciaroli M, Alvarez J, Boniotti MB, Cagiola M, Di Marco V, Marianelli C, Pacciarini M, Pasquali P. Tuberculosis in domestic animal species. *Res Vet Sci* 2014. <https://doi.org/10.1016/j.rvsc.2014.05.015> (in press).
- [14] Gormley E, Corner LAL. Wild animal tuberculosis: stakeholder value systems and management of disease. *Front Vet Sci* 2018;5:327. <https://doi.org/10.3389/fvets.2018.00327>.
- [15] Malone KM, Gordon SV. *Mycobacterium tuberculosis* complex members adapted to wild and domestic animals. *Adv Exp Med Biol* 2017;1019:135–54. [https://doi.org/10.1007/978-3-319-64371-7\\_7](https://doi.org/10.1007/978-3-319-64371-7_7).
- [16] Neill SD, Bryson DG, Pollock JM. Pathogenesis of tuberculosis in cattle. *Tuberculosis* 2001;81:79–86.
- [17] Russell DG. Highlighting the parallels between human and bovine tuberculosis. *J Vet Med Educ* 2003;30:140–2.
- [18] Cassidy JP. The pathogenesis and pathology of bovine tuberculosis with insights from studies of tuberculosis in humans and laboratory animal models. *Vet Microbiol* 2006;112:151–61. <https://doi.org/10.1016/j.vetmic.2005.11.031>.
- [19] Pollock JM, Rodgers JD, Welsh MD, McNair J. Pathogenesis of bovine tuberculosis: the role of experimental models of infection. *Vet Microbiol* 2006;112:141–50. <https://doi.org/10.1016/j.vetmic.2005.11.032>.
- [20] Waters WR, Maggioli MF, McGill JL, Lyashchenko KP, Palmer MV. Relevance of bovine tuberculosis research to the understanding of human disease: historical perspectives, approaches, and immunologic mechanisms. *Vet Immunol Immunopathol* 2014;159:113–32. <https://doi.org/10.1016/j.vetimm.2014.02.009>.
- [21] Van Rhijn I, Godfroid J, Michel A, Rutten V. Bovine tuberculosis as a model for human tuberculosis: advantages over small animal models. *Microb Infect* 2008;10:711–5. <https://doi.org/10.1016/j.micinf.2008.04.005>.
- [22] Waters WR, Palmer MV. *Mycobacterium bovis* infection of cattle and white-tailed deer: translational research of relevance to human tuberculosis. *ILAR J* 2015;56:26–43. <https://doi.org/10.1093/ilar/ilv001>.
- [23] Waters WR, Palmer MV, Thacker TC, Davis WC, Sreevatsan S, Coussens P, Meade KG, Hope JC, Estes DM. Tuberculosis immunity: opportunities from

- studies with cattle. *Clin Dev Immunol* 2011;768542. <https://doi.org/10.1155/2011/768542>. 2011.
- [24] Williams A, Orme IM. Animal models of tuberculosis: an overview. *Microbiol Spectr* 2016;4. <https://doi.org/10.1128/microbiolspec.TB2-0004-2015>.
- [25] Gong W, Liang Y, Wu X. Animal models of tuberculosis vaccine research: an important component in the fight against tuberculosis. *BioMed Res Int* 2020;4263079. <https://doi.org/10.1155/2020/4263079>. 2020.
- [26] Dorhoi A, Kaufmann SH. Pathology and immune reactivity: understanding multidimensionality in pulmonary tuberculosis. *Semin Immunopathol* 2016;38:153–66. <https://doi.org/10.1007/s00281-015-0531-3>.
- [27] Weiss G, Schaible UE. Macrophage defense mechanisms against intracellular bacteria. *Immunol Rev* 2015;264:182–203. <https://doi.org/10.1111/immr.12266>.
- [28] O'Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ, Berry MP. The immune response in tuberculosis. *Annu Rev Immunol* 2013;31:475–527. <https://doi.org/10.1146/annurev-immunol-032712-095939>.
- [29] Cliff JM, Kaufmann SH, McShane H, van Helden P, O'Garra A. The human immune response to tuberculosis and its treatment: a view from the blood. *Immunol Rev* 2015;264:88–102. <https://doi.org/10.1111/immr.12269>.
- [30] Cambier CJ, Falkow S, Ramakrishnan L. Host evasion and exploitation schemes of *Mycobacterium tuberculosis*. *Cell* 2014;159:1497–509. <https://doi.org/10.1016/j.cell.2014.11.024>.
- [31] Schorey JS, Schlesinger LS. Innate immune responses to tuberculosis. *Microbiol Spectr* 2016;4. <https://doi.org/10.1128/microbiolspec.TB2-0010-2016>.
- [32] Awuh JA, Flo TH. Molecular basis of mycobacterial survival in macrophages. *Cell Mol Life Sci* 2017;74:1625–48. <https://doi.org/10.1007/s00118-016-2422-8>.
- [33] Chandra P, Grigsby SJ, Phillips JA. Immune evasion and provocation by *Mycobacterium tuberculosis*. *Nat Rev Microbiol* 2022;20:750–66. <https://doi.org/10.1038/s41579-022-00763-4>.
- [34] Goldberg MF, Saini NK, Porcelli SA. Evasion of innate and adaptive immunity by *Mycobacterium tuberculosis*. *Microbiol Spectr* 2014;2. <https://doi.org/10.1128/microbiolspec.MGM2-0005-2013>.
- [35] Hmama Z, Pena-Diaz S, Joseph S, Av-Gay Y. Immuno-evasion and immunosuppression of the macrophage by *Mycobacterium tuberculosis*. *Immunol Rev* 2015;264:220–32. <https://doi.org/10.1111/immr.12268>.
- [36] Nalpas NC, Magee DA, Conlon KM, Browne JA, Healy C, McLoughlin KE, Rue-Albrecht K, McGettigan PA, Killick KE, Gormley E, Gordon SV, MacHugh DE. RNA sequencing provides exquisite insight into the manipulation of the alveolar macrophage by tubercle bacilli. *Sci Rep* 2015;5:13629. <https://doi.org/10.1038/srep13629>.
- [37] Vegh P, Magee DA, Nalpas NC, Bryan K, McCabe MS, Browne JA, Conlon KM, Gordon SV, Bradley DG, MacHugh DE, Lynn DJ. MicroRNA profiling of the bovine alveolar macrophage response to *Mycobacterium bovis* infection suggests pathogen survival is enhanced by microRNA regulation of endocytosis and lysosome trafficking. *Tuberculosis* 2015;95:60–7. <https://doi.org/10.1016/j.tube.2014.10.011>.
- [38] Lavalett L, Rodriguez H, Ortega H, Sadee W, Schlesinger LS, Barrera LF. Alveolar macrophages from tuberculosis patients display an altered inflammatory gene expression profile. *Tuberculosis* 2017;107:156–67. <https://doi.org/10.1016/j.tube.2017.08.012>.
- [39] Malone KM, Rue-Albrecht K, Magee DA, Conlon K, Schubert OT, Nalpas NC, Browne JA, Smyth A, Gormley E, Aebersold R, MacHugh DE, Gordon SV. Comparative 'omics analyses differentiate *Mycobacterium tuberculosis* and *Mycobacterium bovis* and reveal distinct macrophage responses to infection with the human and bovine tubercle bacilli. *Microb Genom* 2018;4. <https://doi.org/10.1099/mgen.0.000163>.
- [40] Papp AC, Azad AK, Pietrzak M, Williams A, Handelman SK, Igo Jr RP, Stein CM, Hartmann K, Schlesinger LS, Sadee W. AmpliSeq transcriptome analysis of human alveolar and monocyte-derived macrophages over time in response to *Mycobacterium tuberculosis* infection. *PLoS One* 2018;13:e0198221. <https://doi.org/10.1371/journal.pone.0198221>.
- [41] Hall TJ, Mullen MP, McHugo GP, Killick KE, Ring SC, Berry DP, Correia CN, Browne JA, Gordon SV, MacHugh DE. Integrative genomics of the mammalian alveolar macrophage response to intracellular mycobacteria. *BMC Genom* 2021;22:343. <https://doi.org/10.1186/s12864-021-07643-w>.
- [42] Hall TJ, Vernimmen D, Browne JA, Mullen MP, Gordon SV, MacHugh DE, O'Doherty AM. Alveolar macrophage chromatin is modified to orchestrate host response to *Mycobacterium bovis* infection. *Front Genet* 2020;10:1386. <https://doi.org/10.3389/fgene.2019.01386>.
- [43] Mendonca LE, Pernet E, Khan N, Sanz J, Kaufmann E, Downey J, Grant A, Orlova M, Schurr E, Krawczyk C, Jones RG, Barreiro LB, Divangahi M. Human alveolar macrophage metabolism is compromised during *Mycobacterium tuberculosis* infection. *Front Immunol* 2023;13:1044592. <https://doi.org/10.3389/fimmu.2022.1044592>.
- [44] Whelan AO, Coad M, Cockle PJ, Hewinson G, Vordermeier M, Gordon SV. Revisiting host preference in the *Mycobacterium tuberculosis* complex: experimental infection shows *M. tuberculosis* H37Rv to be avirulent in cattle. *PLoS One* 2010;5:e8527. <https://doi.org/10.1371/journal.pone.0008527>.
- [45] Villarreal-Ramos B, Berg S, Whelan A, Holbert S, Carreras F, Salguero FJ, Khatri BL, Malone K, Rue-Albrecht K, Shaughnessy R, Smyth A, Ameni G, Aseffa A, Sarradin P, Winter N, Vordermeier M, Gordon SV. Experimental infection of cattle with *Mycobacterium tuberculosis* isolates shows the attenuation of the human tubercle bacillus for cattle. *Sci Rep* 2018;8:894. <https://doi.org/10.1038/s41598-017-18575-5>.
- [46] Magee DA, Conlon KM, Nalpas NC, Browne JA, Pirson C, Healy C, McLoughlin KE, Chen J, Vordermeier HM, Gormley E, MacHugh DE, Gordon SV. Innate cytokine profiling of bovine alveolar macrophages reveals commonalities and divergence in the response to *Mycobacterium bovis* and *Mycobacterium tuberculosis* infection. *Tuberculosis* 2014;94:441–50. <https://doi.org/10.1016/j.tube.2014.04.004>.
- [47] Ring SC, Purfield DC, Good M, Breslin P, Ryan E, Blom A, Evans RD, Doherty ML, Bradley DG, Berry DP. Variance components for bovine tuberculosis infection and multi-breed genome-wide association analysis using imputed whole genome sequence data. *PLoS One* 2019;14:e0212067. <https://doi.org/10.1371/journal.pone.0212067>.
- [48] Li W, Turner A, Aggarwal P, Matter A, Storvick E, Arnett DK, Broeckel U. Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole transcriptome RNA sequencing methodology for global gene expression analysis. *BMC Genom* 2015;16:1069. <https://doi.org/10.1186/s12864-015-2270-1>.
- [49] Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet* 2018;50:1593–9. <https://doi.org/10.1038/s41588-018-0248-z>.
- [50] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9. <https://doi.org/10.1038/s41586-018-0579-z>.
- [51] Canela-Xandri O, Law A, Gray A, Woolliams JA, Tenesa A. A new tool called DISSECT for analysing large genomic data sets using a Big Data approach. *Nat Commun* 2015;6:10162. <https://doi.org/10.1038/ncomms10162>.
- [52] Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsiek CG, Tseng E, Rowan TN, Low WY, Zimin A, Coudrey C, Hall R, Li W, Rhie A, Ghurye J, McKay SD, Thibaud-Nissen F, Hoffman J, Murdoch BM, Snelling WM, McDanel TG, Hammond JA, Schwartz JC, Nandolo W, Hagen DE, Dreischer C, Schultheiss SJ, Schroeder SG, Phillippy AM, Cole JB, Van Tassel CP, Liu G, Smith TPL, Medrano JF. De novo assembly of the cattle reference genome with single-molecule sequencing. *GigaScience* 2020;9. <https://doi.org/10.1093/gigascience/giaa021>.
- [53] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
- [54] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
- [55] Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Method* 1995;57:289–300.
- [56] Pietrzak M, Papp A, Curtis A, Handelman SK, Katakai M, Scharre DW, Rempala G, Sadee W. Gene expression profiling of brain samples from patients with Lewy body dementia. *Biochem Biophys Res Commun* 2016;479:875–80. <https://doi.org/10.1016/j.bbrc.2016.09.114>.
- [57] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
- [58] Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, Kaplan S, Dahary D, Warshawsky D, Guan-Golan Y, Kohn A, Rappaport N, Safran M, Lancet D. The GeneCards Suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinformatics* 2016;54(1). <https://doi.org/10.1002/cpbi.5>. 30.31–31.30.33.
- [59] Gormley C, Tong Z. Elasticsearch: the definitive guide. Sebastopol, CA: O'Reilly Media, Inc; 2015.
- [60] Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, Becker A, Bennett R, Berry A, Bhai J, Bhurji SK, Bicknell A, Boddu S, Branco Lins PR, Brooks L, Ramaraju SB, Charkhchi M, Coggurn A, Da Rin Fiorretto L, Davidson C, Dodiya K, Donaldson S, El Houdaigui B, El Naboulsi T, Fatima R, Giron CG, Genez T, Ghattaoraya GS, Martinez JG, Gujjarar C, Hardy M, Hollis Z, Hourlier T, Hunt T, Kay M, Kaykala V, Le T, Lemos D, Marques-Coelho D, Marugán JC, Merino GA, Mirabueno LP, Mushtaq A, Hossain SN, Ogeh DN, Sakthivel MP, Parker A, Perry M, Pilizota I, Prosovetskaia I, Pérez-Silva JG, Salam AIA, Saraiva-Agostinho N, Schuilenburg H, Sheppard D, Sinha S, Sips B, Stark W, Steed E, Sukumaran R, Sumathipala D, Suner MM, Surapaneni L, Sutinen K, Szpak M, Tricomi FF, Urbina-Gómez D, Veidenberg A, Walsh TA, Walts B, Wass E, Willhoft N, Allen J, Alvarez-Jarreta J, Chakiachvili M, Flint B, Giorgetti S, Haggerty L, Ilsley GR, Loveland JE, Moore B, Mudge JM, Tate J, Thybert D, Trevanion SJ, Winterbottom A, Frankish A, Hunt SE, Ruffier M, Cunningham F, Dyer S, Finn RD, Howe KL, Harrison PW, Yates AD, Flicek P. Ensembl 2023. *Nucleic Acids Res* 2023;51:D933–d941. <https://doi.org/10.1093/nar/gkac958>.
- [61] Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS, Lynn DJ. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Res* 2013;41:D1228–33. <https://doi.org/10.1093/nar/gks1147>.
- [62] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498–504. <https://doi.org/10.1101/gr.1239303>.
- [63] Ideker T, Ozier O, Schwikowski B, Siegel AF. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 2002;18(Suppl 1):S233–40. [https://doi.org/10.1093/bioinformatics/18.suppl\\_1.s233](https://doi.org/10.1093/bioinformatics/18.suppl_1.s233).
- [64] Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* 2023;51:D587–d592. <https://doi.org/10.1093/nar/gkac963>.

- [65] Yamamoto S, Sakai N, Nakamura H, Fukagawa H, Fukuda K, Takagi T. Database 2011. <https://doi.org/10.1093/database/bar052>. 2011:bar052.
- [66] Pillich RT, Chen J, Churas C, Fong D, Gyori BM, Ideker T, Karis K, Liu SN, Ono K, Pico A, Pratt D. NDEX IQuery: a multi-method network gene set analysis leveraging the Network Data Exchange. *Bioinformatics* 2023;39. <https://doi.org/10.1093/bioinformatics/btad118>.
- [67] Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, Griss J, Sevilla C, Matthews L, Gong C, Deng C, Varusai T, Ragueneau E, Haider Y, May B, Shamovsky V, Weiser J, Brunson T, Sanati N, Beckman L, Shao X, Fabregat A, Sidiropoulos K, Murillo J, Viteri G, Cook J, Shorsler S, Bader G, Demir E, Sander C, Haw R, Wu G, Stein L, Hermjakob H, D'Eustachio P. The reactome pathway knowledgebase 2022. *Nucleic Acids Res* 2022;50:D687–d692. <https://doi.org/10.1093/nar/gkab1028>.
- [68] Nishimura D. *BioCarta*. *Biotech Softw Internet Rep* 2001;2:117–20. <https://doi.org/10.1089/152791601750294344>.
- [69] Kandasamy K, Mohan SS, Raju R, Keerthikumar S, Kumar GS, Venugopal AK, Telikicherla D, Navarro JD, Mathivanan S, Pecquet C, Gollapudi SK, Tattikota SG, Mohan S, Padhukasahasram H, Subbannayya Y, Goel R, Jacob HK, Zhong J, Sekhar R, Nanjappa V, Balakrishnan L, Subbiah R, Ramachandra YL, Rahiman BA, Prasad TS, Lin JX, Houtman JC, Desiderio S, Renauld JC, Constantinescu SN, Ohara O, Hirano T, Kubo M, Singh S, Khatri P, Draghici S, Bader GD, Sander C, Leonard WJ, Pandey A. NetPath: a public resource of curated signal transduction pathways. *Genome Biol* 2010;11:R3. <https://doi.org/10.1186/gb-2010-11-1-r3>.
- [70] Krämer A, Green J, Pollard Jr J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 2014;30:523–30. <https://doi.org/10.1093/bioinformatics/btt703>.
- [71] Timmons JA, Szkop KJ, Gallagher IJ. Multiple sources of bias confound functional enrichment analysis of global -omics data. *Genome Biol* 2015;16:186. <https://doi.org/10.1186/s13059-015-0761-7>.
- [72] Qiagen. Qiagen ingenuity pathway analysis online manual. 2023. <https://qiagen.my.salesforce-sites.com/KnowledgeBase/KnowledgeNavigatorPage?categoryName=IPA>.
- [73] Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* 2008;24:1461–2. <https://doi.org/10.1093/bioinformatics/btn209>.
- [74] Vázquez CL, Rodgers A, Herbst S, Coade S, Gronow A, Guzman CA, Wilson MS, Kanzaki M, Nykjaer A, Gutierrez MG. The proneurotrophin receptor sortilin is required for *Mycobacterium tuberculosis* control by macrophages. *Sci Rep* 2016;6:29332. <https://doi.org/10.1038/srep29332>.
- [75] Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, Forrest SA, Bryant JM, Harris SR, Schuenemann VJ, Campbell TJ, Majander K, Wilbur AK, Guichon RA, Wolfe Steadman DL, Cook DC, Niemann S, Behr MA, Zumarraga M, Bastida R, Huson D, Nieselt K, Young D, Parkhill J, Buikstra JE, Gagneux S, Stone AC, Krause J. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* 2014;514:494–7. <https://doi.org/10.1038/nature13591>.
- [76] Kay GL, Sergeant MJ, Zhou Z, Chan JZ, Millard A, Quick J, Szikossy I, Pap I, Spigelman M, Loman NJ, Achtman M, Donoghue HD, Pallen MJ. Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* 2015;6:6717. <https://doi.org/10.1038/ncomms7717>.
- [77] Sabin S, Herbig A, Vågene ÅJ, Ahlström T, Bozovic G, Arcini C, Kühnert D, Bos KI. A seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic emergence of the *Mycobacterium tuberculosis* complex. *Genome Biol* 2020;21:201. <https://doi.org/10.1186/s13059-020-02112-1>.
- [78] Liu L, Zhang J, Rheindt FE, Lei F, Qu Y, Wang Y, Zhang Y, Sullivan C, Nie W, Wang J, Yang F, Chen J, Edwards SV, Meng J, Wu S. Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *Proc Natl Acad Sci U S A* 2017;114:E7282–e7290. <https://doi.org/10.1073/pnas.1616744114>.
- [79] Yi XH, Zhang B, Fu YR, Yi ZJ. STAT1 and its related molecules as potential biomarkers in *Mycobacterium tuberculosis* infection. *J Cell Mol Med* 2020;24:2866–78. <https://doi.org/10.1111/jcmm.14856>.
- [80] DeFrances MC, Debelius DR, Cheng J, Kane LP. Inhibition of T-cell activation by PIK3IP1. *Eur J Immunol* 2012;42:2754–9. <https://doi.org/10.1002/eji.201141653>.
- [81] van Baren N, Van den Eynde BJ. Tryptophan-degrading enzymes in tumoral immune resistance. *Front Immunol* 2015;6:34. <https://doi.org/10.3389/fimmu.2015.00034>.
- [82] Guo L, Dijkhuis A, Van Der Wel N, Lutter R. Indoleamine 2,3-dioxygenase-1 (Ido1) activity protects *M. tuberculosis* in macrophages from antibiotics. *Eur Respir J* 2019;54:PA4599. <https://doi.org/10.1183/13993003.congress-2019.PA4599>.
- [83] Singh B, Moodley C, Singh DK, Escobedo RA, Sharan R, Arora G, Ganatra SR, Shivanna V, Gonzalez O, Hall-Ursone S, Dick Jr EJ, Kaushal D, Alvarez X, Mehra S. Inhibition of indoleamine dioxygenase leads to better control of tuberculosis adjunctive to chemotherapy. *JCI Insight* 2023;8. <https://doi.org/10.1172/jci.insight.163101>.
- [84] Wang N, Satoskar A, Faubion W, Howie D, Okamoto S, Feske S, Gullo C, Clarke K, Sosa MR, Sharpe AH, Terhorst C. The cell surface receptor SLAM controls T cell and macrophage functions. *J Exp Med* 2004;199:1255–64. <https://doi.org/10.1084/jem.20031835>.
- [85] Ma C, Wang N, Detre C, Wang G, O'Keeffe M, Terhorst C. Receptor signaling lymphocyte-activation molecule family 1 (Slamf1) regulates membrane fusion and NADPH oxidase 2 (NOX2) activity by recruiting a Beclin-1/Vps34/ultraviolet radiation resistance-associated gene (UVRAG) complex. *J Biol Chem* 2012;287:18359–65. <https://doi.org/10.1074/jbc.M112.367060>.
- [86] Berger SB, Romero X, Ma C, Wang G, Faubion WA, Liao G, Compeer E, Keszei M, Rameh L, Wang N, Boes M, Regueiro JR, Reinecker HC, Terhorst C. SLAM is a microbial sensor that regulates bacterial phagosome functions in macrophages. *Nat Immunol* 2010;11:920–7. <https://doi.org/10.1038/ni.1931>.
- [87] Albert R. Scale-free networks in cell biology. *J Cell Sci* 2005;118:4947–57. <https://doi.org/10.1242/jcs.02714>.
- [88] Barabási AL, Albert R. Emergence of scaling in random networks. *Science* 1999;286:509–12.
- [89] Killick KE, Magee DA, Park SD, Taraktoglou M, Browne JA, Conlon KM, Nalpas NC, Gormley E, Gordon SV, MacHugh DE, Hokamp K. Key hub and bottleneck genes differentiate the macrophage response to virulent and attenuated *Mycobacterium bovis*. *Front Immunol* 2014;5:422. <https://doi.org/10.3389/fimmu.2014.00422>.
- [90] BoseDasgupta S, Pieters J. Macrophage-microbe interaction: lessons learned from the pathogen *Mycobacterium tuberculosis*. *Semin Immunopathol* 2018;40:577–91. <https://doi.org/10.1007/s00281-018-0710-0>.
- [91] Hu Q, Chen Z, Liang G, Mo F, Zhang H, Xu S, Wang Y, Kang L, Jin T. Vitamin D receptor gene associations with pulmonary tuberculosis in a Tibetan Chinese population. *BMC Infect Dis* 2016;16:469. <https://doi.org/10.1186/s12879-016-1699-4>.
- [92] Nelson CD, Reinhardt TA, Lippolis JD, Sacco RE, Nonnecke BJ. Vitamin D signaling in the bovine immune system: a model for understanding human vitamin D requirements. *Nutrients* 2012;4:181–96. <https://doi.org/10.3390/nu4030181>.
- [93] Guirado E, Schlesinger LS. Modeling the *Mycobacterium tuberculosis* granuloma - the critical battlefield in host immunity and disease. *Front Immunol* 2013;4:98. <https://doi.org/10.3389/fimmu.2013.00098>.
- [94] Elad S, Sherman Y, Palmon A, Vlodavsky I, Or R. Heparanase expression in peripical granulomas and radicular cysts. *Odontology* 2013;101:96–102. <https://doi.org/10.1007/s10266-011-0045-y>.
- [95] Irony-Tur-Sinai M, Vlodavsky I, Ben-Sasson SA, Pinto F, Sicsic C, Brenner T. A synthetic heparin-mimicking polyanionic compound inhibits central nervous system inflammation. *J Neurol Sci* 2003;206:49–57. [https://doi.org/10.1016/S0022-510X\(02\)00318-0](https://doi.org/10.1016/S0022-510X(02)00318-0).
- [96] Mehra S, Alvarez X, Didier PJ, Doyle LA, Blanchard JL, Lackner AA, Kaushal D. Granuloma correlates of protection against tuberculosis and mechanisms of immune modulation by *hard*. *J Infect Dis* 2013;207:1115–27. <https://doi.org/10.1093/infdis/jis778>.
- [97] Harding JS, Herbath M, Chen Y, Rayasam A, Ritter A, Csoka B, Hasko G, Michael IP, Fabry Z, Nagy A, Sandor M. VEGF-A from granuloma macrophages regulates granulomatous inflammation by a non-angiogenic pathway during mycobacterial infection. *Cell Rep* 2019;27:2119–2131.e2116. <https://doi.org/10.1016/j.celrep.2019.04.072>.
- [98] Hollmén M, Karaman S, Schwager S, Lisibach A, Christiansen AJ, Maksimow M, Varga Z, Jalkanen S, Detmar M. G-CSF regulates macrophage phenotype and associates with poor overall survival in human triple-negative breast cancer. *Oncotarget* 2016;5:e115177. <https://doi.org/10.1080/2162402x.2015.1115177>.
- [99] Wen Q, Kong Y, Zhao HY, Zhang YY, Han TT, Wang Y, Xu LP, Zhang XH, Huang XJ. G-CSF-induced macrophage polarization and mobilization may prevent acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant* 2019;54:1419–33. <https://doi.org/10.1038/s41409-019-0449-9>.
- [100] Casanova NG, Gonzalez-Garay ML, Sun B, Bime C, Sun X, Knox KS, Crouser ED, Sammani N, Gonzales T, Natt B, Chaudhary S, Lussier Y, Garcia JGN. Differential transcriptomics in sarcoidosis lung and lymph node granulomas with comparisons to pathogen-specific granulomas. *Respir Res* 2020;21:321. <https://doi.org/10.1186/s12931-020-01537-3>.
- [101] Ozaki S, Funasaka Y, Takubo M, Matayoshi T, Ueno T, Asayama T, Saeki H. Granulocyte colony-stimulating factor-induced granulomatous dermatitis with enlarged histiocytes clinically manifesting as painful edematous nodules with high fever similar to Sweet's syndrome. *J Dermatol* 2015;42:414–7. <https://doi.org/10.1111/1346-8138.12772>.
- [102] Sekhsaria S, Fleisher TA, Vowells S, Brown M, Miller J, Gordon I, Blaese RM, Dunbar CE, Leitman S, Malech HL. Granulocyte colony-stimulating factor recruitment of CD34+ progenitors to peripheral blood: impaired mobilization in chronic granulomatous disease and adenosine deaminase-deficient severe combined immunodeficiency disease patients. *Blood* 1996;88:1104–12.
- [103] Herrtwick L, Nanda I, Evangelou K, Nikolova T, Horn V, Sagar, Erny D, Stefanowski J, Rogell L, Klein C, Gharun K, Follo M, Seidl M, Kremer B, Münke N, Senges J, Fliegau M, Aschman T, Pfeifer D, Sarrazin S, Sieweke MH, Wagner D, Dierks C, Haaf T, Ness T, Zaiss MM, Voll RE, Deshmukh SD, Prinz M, Goldmann T, Hölscher C, Hauser AE, Lopez-Contreras AJ, Grün D, Gorgoulis V, Diefenbach A, Henneke P, Triantafyllopoulou A. DNA damage signaling instructs polyploid macrophage fate in geranylulomas. *Cell* 2018;174:1325–6. <https://doi.org/10.1016/j.cell.2018.08.015>.
- [104] Crouser ED, White P, Caceres EG, Julian MW, Papp AC, Locke LW, Sadee W, Schlesinger LS. A novel in vitro human granuloma model of sarcoidosis and latent tuberculosis infection. *Am J Respir Cell Mol Biol* 2017;57:487–98. <https://doi.org/10.1165/rcmb.2016-0321OC>.
- [105] Shaheen ZR, Christmann BS, Stafford JD, Moran JM, Buller RML, Corbett JA. CCR5 is a required signaling receptor for macrophage expression of inflammatory genes in response to viral double-stranded RNA. *Am J Physiol Regul Integr Comp Physiol* 2019;316:R525–r534. <https://doi.org/10.1152/ajpregu.00019.2019>.



- [106] Armas-González E, Domínguez-Luis MJ, Díaz-Martín A, Arce-Franco M, Castro-Hernández J, Danelon G, Hernández-Hernández V, Bustabad-Reyes S, Cantabrana A, Ugucioni M, Díaz-González F. Role of CXCL13 and CCL20 in the recruitment of B cells to inflammatory foci in chronic arthritis. *Arthritis Res Ther* 2018;20:114. <https://doi.org/10.1186/s13075-018-1611-2>.
- [107] Khader SA, Rangel-Moreno J, Fountain JJ, Martino CA, Reiley WW, Pearl JE, Winslow GM, Woodland DL, Randall TD, Cooper AM. In a murine tuberculosis model, the absence of homeostatic chemokines delays granuloma formation and protective immunity. *J Immunol* 2009;183:8004–14. <https://doi.org/10.4049/jimmunol.0901937>.
- [108] Locke LW, Crouser ED, White P, Julian MW, Caceres EG, Papp AC, Le VT, Sadee W, Schlesinger LS. IL-13-regulated macrophage polarization during granuloma formation in an in vitro human sarcoidosis model. *Am J Respir Cell Mol Biol* 2019;60:84–95. <https://doi.org/10.1165/rcmb.2018-0053OC>.
- [109] Ray R, Rai V. Lysophosphatidic acid converts monocytes into macrophages in both mice and humans. *Blood* 2017;129:1177–83. <https://doi.org/10.1182/blood-2016-10-743757>.
- [110] Echeverría-Valencia G, Silva-Miranda M, Ekaza E, Vallecillo AJ, Parada C, Sada-Ovalle I, Altare F, Espitia C. Interaction of mycobacteria with psimblasmin(ogen) affects phagocytosis and granuloma development. *Tuberculosis* 2019;117:36–44. <https://doi.org/10.1016/j.tube.2019.05.008>.
- [111] Nordmark G, Kristjansdóttir G, Theander E, Appel S, Eriksson P, Vasaitis L, Kvarnstrom M, Delaleu N, Lundmark P, Lundmark A, Sjowall C, Brun JG, Jonsson MV, Harboe E, Goransson LG, Johnsen SJ, Soderkvist P, Eloranta ML, Alm G, Baecklund E, Wahren-Herlenius M, Omdal R, Ronnblom L, Jonsson R, Syvanen AC. Association of *EBF1*, *FAM167A(C8orf13)-BLK* and *TNFSF4* gene variants with primary Sjogren's syndrome. *Gene Immun* 2011;12:100–9. <https://doi.org/10.1038/gene.2010.44>.
- [112] Sun F, Li P, Chen H, Wu Z, Xu J, Shen M, Leng X, Shi Q, Zhang W, Tian X, Li Y, Zhang F. Association studies of *TNFSF4*, *TNFAIP3* and *FAM167A-BLK* polymorphisms with primary Sjogren's syndrome in Han Chinese. *J Hum Genet* 2013;58:475–9. <https://doi.org/10.1038/jhg.2013.26>.
- [113] Bishop TF, Van Eenennaam AL. Genome editing approaches to augment livestock breeding programs. *J Exp Biol* 2020;223. <https://doi.org/10.1242/jeb.207159>.
- [114] Banos G, Winters M, Mrode R, Mitchell AP, Bishop SC, Woolliams JA, Coffey MP. Genetic evaluation for bovine tuberculosis resistance in dairy cattle. *J Dairy Sci* 2017;100:1272–81. <https://doi.org/10.3168/jds.2016-11897>.
- [115] Tsairidou S, Allen A, Banos G, Coffey M, Anacleto O, Byrne AW, Skuce RA, Glass EJ, Woolliams JA, Doeschl-Wilson AB. Can we breed cattle for lower bovine TB infectivity? *Front Vet Sci* 2018;5. <https://doi.org/10.3389/fvets.2018.00310>.
- [116] Buddle BM, Vordermeier HM, Hewinson RG. Experimental infection models of tuberculosis in domestic livestock. *Microbiol Spectr* 2016;4. <https://doi.org/10.1128/microbiolspec.TB2-0017-2016>.
- [117] Zhang Q, Lenardo MJ, Baltimore D. 30 years of NF-kappaB: a blossoming of relevance to human pathobiology. *Cell* 2017;168:37–57. <https://doi.org/10.1016/j.cell.2016.12.012>.
- [118] Kedia K, Wendler JP, Baker ES, Burnum-Johnson KE, Jarsberg LG, Stratton KG, Wright AT, Piehowski PD, Gritsenko MA, Lewinsohn DM, Sigal GB, Weiner MH, Smith RD, Jacobs JM, Nahid P. Application of multiplexed ion mobility spectrometry towards the identification of host protein signatures of treatment effect in pulmonary tuberculosis. *Tuberculosis* 2018;112:52–61. <https://doi.org/10.1016/j.tube.2018.07.005>.
- [119] Fallarino F, Volpi C, Fazio F, Notartomaso S, Vacca C, Busceti C, Biciato S, Battaglia G, Bruno V, Puccetti P, Fioretti MC, Nicoletti F, Grohmann U, Di Marco R. Metabotropic glutamate receptor-4 modulates adaptive immunity and restrains neuroinflammation. *Nat Med* 2010;16:897–902. <https://doi.org/10.1038/nm.2183>.
- [120] Houben RM, Dodd PJ. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. *PLoS Med* 2016;13:e1002152. <https://doi.org/10.1371/journal.pmed.1002152>.
- [121] Cohen A, Mathiasen VD, Schön T, Wejse C. The global prevalence of latent tuberculosis: a systematic review and meta-analysis. *Eur Respir J* 2019;54. <https://doi.org/10.1183/13993003.00655-2019>.
- [122] Bloom BR, Atun R, Cohen T, Dye C, Fraser H, Gomez GB, Knight G, Murray M, Nardell E, Rubin E, Salomon J, Vassall A, Volchenkov G, White R, Wilson D, Yadav P. Tuberculosis. In: Holmes KK, Bertozzi S, Bloom BR, Jha P, editors. *Major infectious diseases*. Washington (DC): the international bank for reconstruction and development. The World Bank; 2017. <https://www.ncbi.nlm.nih.gov/books/NBK525174/>.
- [123] Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A, Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M. Tuberculosis. *Nat Rev Dis Prim* 2016;2:16076. <https://doi.org/10.1038/nrdp.2016.76>.
- [124] Simmons JD, Stein CM, Seshadri C, Campo M, Alter G, Fortune S, Schurr E, Wallis RS, Churchyard G, Mayanja-Kizza H, Boom WH, Hawn TR. Immunological mechanisms of human resistance to persistent *lav* infection. *Nat Rev Immunol* 2018. <https://doi.org/10.1038/s41577-018-0025-3>.
- [125] Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 2017;33:2938–40. <https://doi.org/10.1093/bioinformatics/btx364>.