Màster Universitari
Anàlisi de Dades Òmiques /
Omics Data Analysis
FACULTAT DE CIÈNCIES I TECNOLOGIA
UVIC | UVIC·UCC

Master of Science in Omics Data Analysis

Master Thesis

Host epigenetic changes to Salmonella Paratyphi Exposure

by

Sara Pischedda

Supervisor: Daniel O'Connor, PhD (Oxford Vaccine Group)

Academic tutor: Mireia Oliveilla

Biosciences Department

University of Vic – Central University of Catalonia

10 September 2023

Bioinformatics, YYYY, 0–0
doi: 10.1093/bioinformatics/xxxxx
Advance Access Publicationb Date: DD Month YYYY
Manuscript Category

Subject Section

Host epigenetic changes to Salmonella Paratyphi Exposure

Sara Pischedda[1], Deborah L. Cross[2,3], Antonio Salas[1,4], Federico Martinón-Torres[1], Andrew J. Pollard[2,3] & Daniel O'Connor[2*]

[1]Genetics, Vaccines, Infectious Diseases and Pediatrics (GENVIP) research group, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Hospital Clinico Universitario de Santiago de Compostela (CHUS), Santiago de Compostela, Spain

[2]Oxford Vaccine Group, Department of Paediatrics, University of Oxford, Oxford, United Kingdom.

[3]NHR Oxford Biomerical Research Centre and Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom.

[4]GenPoB Research Group, Instituto de Investigacinó Sanitaria (IDIS), Hospital Clínico Universitario de Santiago (SERGAS), Unidade de Xenética, Santiago de Compostela, Spain

*To whom correspondence should be addressed.

Abstract

Motivation: Typhoid fever, resulting from Salmonella enterica serovars Typhi and Paratyphi A, stands as a significant contributor to illness and mortality. Recent challenge studies on the host immunity to these Salmonella bacteria, indicate that both infection and vaccination lead to changes in gene expression within leukocytes, and these alterations could be associated with immune and inflammatory responses during the course of the disease. The aim of the present study is to explore the methylation changes occurring during and after the paratyphoid infection by comparing individuals in two cell population (cd14+ and cd56+ cells), at three different timepoints: before, during and after the infection.

Results: The analysis of the three timepoints performed separately in the two -cell leukocyte populations revealed that biggest differences in DNA methylation were observed in infected individuals in cd14+ monocytes cells. In this particularly cell types, highest differences in methylation in individuals before, during and after the paratyphoid infection, were observed in seven specific genes, with a global DNA hypomethylation for infected individuals than the other ones. The most significant CpGs were observed within the *SND1, PSAP* and *MNDA* genes. Gene set analysis revealed a significant enrichment for signaling pathways related to Golgi vesicle transport and MHC protein complex.

Contact: sara.pischedda01@gmail.com

1 Introduction

Typhoid fever, resulting from Salmonella enterica serovars Typhi and Paratyphi A, stands as a significant contributor to illness and mortality (Dougan & Baker, 2014). It constituted a pressing public health issue in economically disadvantaged regions worldwide, giving rise to roughly 10.9 million instances and causing about 100,000 dies each year (Typhoid & Paratyphoid, 2019).

Overcoming and managing typhoid infection present several challenges, in accurately diagnosing enteric fever due to the limited sensitivity of current tests, limited short-term effectiveness of available vaccines (World Health, 2019), and the emergence of antibiotic-resistant strains.

To address these concerns, infection biomarkers, more potent vaccines, and innovative treatments for enteric fever are imperative. The achievement of these objectives hinges upon acquiring a more comprehensive comprehension of the immune reaction to enteric fever, particularly discerning the roles of innate and adaptive cellular immunity.

The understanding of the underlying mechanisms and host responses to S. Paratyphi A infection is currently limited. Promising vaccine candidates, including live-attenuated strains and conjugate vaccines are in developmental stages (Jin et al., 2017; Jin et al., 2021). However, none of these candidates have yet undergone efficacy trials. The advancement of vaccines to counter paratyphoid infection is hindered by the lack of comprehensive knowledge regarding immune correlates of protection and the absence of a suitable animal model for testing.

Recent challenge studies on the host immunity to these Salmonella bacteria (Zhu et al., 2023), indicate that both infection and vaccination lead to changes in gene expression within leukocytes, and these alterations could be associated with immune and inflammatory responses during the course of the disease. One of the fundamental mechanisms driving these alterations involves epigenetics—a process that brings about changes in gene expression while preserving the integrity of the DNA sequence (Russo, Martienssen, & Riggs, 1996).

Epigenetic regulation plays a pivotal role in coordinating normal immune responses and upholding immunological memory. While the role of epigenetic processes in cancer and associated conditions has garnered substantial attention for many years (P. A. Jones & Baylin, 2002), in the recent years a surge of interest focused on how infections and vaccines can reshape the human epigenome, consequently influencing immune system activity.

Researchers have increasingly turned their focus towards comprehending how natural infections and vaccinations can induce changes in the epigenome, thereby influencing both the initial immune response and the long-term risk of diseases (Bannister, Messina, Novakovic, & Curtis, 2020). Evidence indicates that epigenetic modifications exert significant control over the promoters and enhancers of various regulators within the immune response (Zhang & Cao, 2019).

This susceptibility to epigenetic marks underscores the intricate interplay between epigenetic processes and the immune system, shedding light on the mechanisms by which infections and vaccinations can modulate immune activity.

Controlled human infection models (CHIMs) offer a unique opportunity to explore the immune response to pathogens at an epigenetic level within a tightly controlled environment. The application of different technologies, such as the study of the global DNA methylation in the genome, to specific leukocytes cells, such as the CD14+ and Cd56+ cells, obtained from participants in enteric fever CHIMs will make possible to characterize changes in gene expression occurring after typhoid exposure and during infection. In this way, researchers aim to achieve an unprecedented level of understanding about gene expression during both exposure and infection stages.

CD56 cells, also known as the neural cell adhesion molecule, is the archetypical phenotypic indicator of natural killer (NK) cells, though they are present in a broader range of immune cells beyond its association with alpha beta T cells, gamma delta T cells, dendritic cells, and monocytes. These cells have diverse roles in immunity, ranging from direct cytotoxicity to immunomodulation, making them important players in immune responses and potential targets for therapeutic interventions (Gianchecchi, Delfino, & Fierabracci, 2018; Poli et al., 2009).

CD14 was initially identified as a monocyte marker (Goyert et al., 1986), and functions as a pattern recognition receptor (PRR) that triggers intracellular responses upon recognizing various bacterial components (Zanoni & Granucci, 2012), it is also an important component of innate immunity, and a regulator of metabolism (Fernandez-Real et al., 2011).

Given these premises, the main aim of the project is to explore the methylation changes occurring during and after the paratyphoid infection by comparing individuals in two cell population (CD14+ and CD56+ cells), at the three different timepoints: before, during and after the infection.

2 Methods

Study design

The study design of the present project has been described previously (Dobinson et al., 2017). Briefly, a group of healthy volunteers enrolled for an observational, human challenge study of S. Paratyphi A infection were selected at three different time-points: a basal sample (D0), an infected sample (ED) — recruited 7-10 days after D0 — and a treated sample (D28) — recruited 28 days after the ED.

Peripheral blood mononuclear cells (PBMCs) were isolated from blood sample at the three-time points and CD14+ and CD56+ cells were extracted.

CD14 group was composted by 9 individuals at three time-point (D0, ED, D28), while CD56 group was represented by 7 individuals at the same three time points (D0, ED, D28).

***Tabla 1**. Clinical characteristics of the individuals in CD14+ and CD56+ cell populations.*

	CD14 (n = 9)	CD56 (n= 7)
Age	28.628 [±6.35]	28.674 [±6.31]
Sex		
Male	5 (55%)	5 (71%)
Female	4 (45%)	2 (29%)
Smoking		
Yes	4 (45%)	3 (43%)
No	5 (55%)	4 (57%)
Ethnicity		
Indian	1 (11%)	1 (14%)
White_other	1 (11%)	1 (14%)
White_british	6 (67%)	4 (57%)
White_irish	1 (11%)	1 (14%)
BMI	23.278 [±3.85]	23.779 [±3.81]
Dose		
700	1 (11%)	1 (14%)
825	1 (11%)	-
1000	3 (33%)	3 (43%)
1300	2 (22%)	2 (29%)
2400	1 (11%)	-
2750	1 (11%)	1 (14%)

Cell purification and sorting

The process of cell purification and sorting was performed as described previously (C. Jones et al., 2016; Toapanta et al., 2015)

DNA isolation and bisulfite treatment

Genomic DNA (gDNA) was extracted from sorted cell populations using the QIAamp DNA blood mini kit according to the manufacturer's instructions (QIAGEN, Chatsworth, CA). DNA concentration was quantified for all samples using absorbance (Thermo Fisher Scientific, Waltham, MA); samples with less than 250ng of total gDNA were excluded from further analysis. DNA quality assessment was conducted using the Agilent TapeStation (Agilent, Santa Clara, CA.).

After extraction, genome DNA underwent sodium bisulfite conversion as a preliminary step before proceeding with the Illumina microarray analysis. This method is widely recognized as the gold standard for investigating the methylation status of CpG sites across the entire genome since it provides a qualitative, quantitative and efficient approach to identify 5-methylcytosine at single base-pair resolution. Essentially, it involves converting unmethylated cytosine residues into uracil through sodium bisulfite-induced deamination, while leaving methylated cytosine (5-mC) unaffected. As a result, uracils are amplified as thymines in the subsequent PCR reaction, while 5-mC or 5-hmC residues remain as cytosines.

The EZ-96 DNA Methylation kit from Zymo Research Corp was employed in this study for bisulfite conversion. Subsequently, the treated DNA was hybridized onto the Illumina Infinium MethylationEPIC BeadChip, and the array was scanned using the Illumina iScan system. This process quantified the percentage methylation status of each CpG site across the entire study cohort.

DNA methylation analysis

Quality control of DNA methylation, along with its processing, standardization, and statistical assessments, was carried out utilizing the R statistical software (R Version 4.3.1). Distinct Bioconductor packages were employed, following the procedural framework detailed in Fortin's recent EPIC methylation analysis guidelines (Fortin, Triche, & Hansen, 2017). The original intensity files (IDAT), representing raw data, were imported into R. The IDAT preprocessing and transformation into β and M-values was carried out through the use of the minfi package (version 1.46.0) (Aryee et al., 2014). β-values denote the proportion of methylated signals relative to the overall signals (encompassing both methylated and unmethylated) for each genomic site. These values span from 0, signifying a lack of methylation across all cells, to 1, indicating total methylation at that particular site across all sample cells (Bibikova et al., 2006). M-values were derived as the logarithmic ratio (base 2) of the intensities between methylated and unmethylated probes, and assume positive and negative values. M values near 0 imply parity in intensity between the two probe types. Positive M-values suggest greater methylated than unmethylated sites, while negative values suggest the opposite (Du et al., 2010).

Within the Infinium 850K array, two categories of probes exist—Infinium I and Infinium II—yet a direct comparison between them is untenable. Therefore, an adjustment is required to mitigate design bias. The normalization routines were executed using the preprocessQuantile function found within the minfi package. Moreover, probes underwent several filtration steps, involving the elimination of probes with P-values exceeding 0.01 (when compared to the signal background), as well as those situated on the sex chromosomes. Further exclusions encompassed sites housing SNPs or with a minor allele frequency (MAF) below 0.05, as these could potentially disrupt probe binding due to genetic variations in binding regions. Finally, probes known to exhibit cross-reactions were also removed from the dataset.

Adjusting for individual differences in cellular heterogeneity in the blood sample from which genomic DNA was extracted can improve biological interpretation, given the highly specific nature of DNA methylation in relation to cell types (Ziller et al., 2013) i.e., variations in cell composition across different phenotypes can introduce confounding elements. In the present study, two specific cellular population were enriched from PBMCs; so, prior to investigating differentially methylated positions (DMPs), the FlowSorted.Blood.EPIC package and its adapted estimateCellCounts2 function was employed to assess the presence of various blood cell types inside these cellular population. This function, based on the Houseman algorithm (Houseman, Kelsey, Wiencke, & Marsit, 2015), permits estimation of relative proportions of distinct white blood cell subcategories such as CD4+ T-lymphocytes, CD8+ T-lymphocytes, natural killer (NK) cells, B-lymphocytes, monocytes, and neutrophils. Following the application of the Shapiro-Wilk test to assess the normality of cell type estimation, the Wilcoxon test was employed to compare the distributions of cell types across the various time points.

Downstream analysis consisting on the search for DMPs has been performed separately in CD56+ and CD14+ cells. For this purpose, the limma package was employed (Smyth, 2005), assuming a linear model where the M values of each probe were used as quantitative dependent variables in all analyses. The application of limma package makes possible to include covariate in the definition of the design matrix that represents the model to be fit. In both study groups, the comparison of different timepoints of the same individuals is the focus of the analysis, together with an adjustment of the model for age, sex and cell composition. The limma's duplicateCorrelation function, aims to estimate the correlation structure among these replicates, allowing for more accurate statistical analysis by appropriately modeling the dependencies between samples. It computes a correlation value that represents the strength of the association between replicates, which can then be incorporated into the linear modeling, to improve the precision of differential methylation analysis and avoiding an overfit and inflation of P-values, thus reducing the false positives.

Identified DMPs were annotated to genes using the IlluminaHumanMethylationEPICanno.ilm10b4.hg19 for EPIC BeadChip "annotation" package which includes information about the position of the methylation loci in the genome, about the genomic features they map to and the known SNPs they can overlap.

Following the Illumina's recommendations, the threshold used for statistical analysis in DMPs detection is P-value < 0.05 (adjusted for multiple testing using the Benjamini–Hochberg method (Benjamini & Hochberg, 1995)), and the absolute difference in the Delta Beta, considered as the difference of means of the Beta value of a single position between the group of interest > 0.1.

Principal Component Analysis (PCA) was carried out to compare DNA methylation of individuals before, during and after the infection. Receiver operating characteristic (ROC) curve analyses were conducted to evaluate the diagnostic efficacy of the most significant candidate DMPs, and the determined area under the curve (AUC) was used to assess prediction accuracy.

Gene set enrichment analysis

Once the analysis of DMPs between groups of samples has been completed, the long list of significant CpGs identified need to be interpreted, through the gene set enrichment analysis. Unlike gene expression data, DNA methylation can occur anywhere on the genome and the number of CpGs profiled and gene length is not of one-to-one correspondence. In fact, DNA methylation data result in multiple CpG association P-values per gene since several differential CpGs can belong to the same gene, as well as one single CpG position can be found in more than one gene. One of the most recent R packages able to avoid this problem by adjusting for the number of CpGs instead of gene length, is called methylGSA (Ren & Kuan, 2019). This package allows the specification of the array type used, 450K or EPIC; it supports pathways from "GO", "KEGG", and "Reactome and gives the possibility to include in the specificity of the functions, the option to consider the different genomic context when selecting the gene list like promoter regions (CpGs within genes belonging to "TSS1500", "TSS200", "1stExon", or "5'UTR"), gene body or all together. Here the threshold adopted to select pathways significantly enriched is false discovery rate (FDR) <0.05.

3 Results

Cell estimation

The analysis of cell composition revealed contrasting results for the two groups of cell population. In the CD56+ enriched cells, only monocytes cells differ among D0 and ED, while the remaining comparisons did not reveal any statistical significance (Figure 1A). However, in CD14+ cells, all the cell subtypes estimated were observed to significantly differ before, during and after the infection, particularly in the contrasts D0-ED and D28-ED (Figure 1A). The most statistically significant differences were detected for natural killer cells and neutrophils (P-value < $1x10^{-3}$) (Figure 1B).

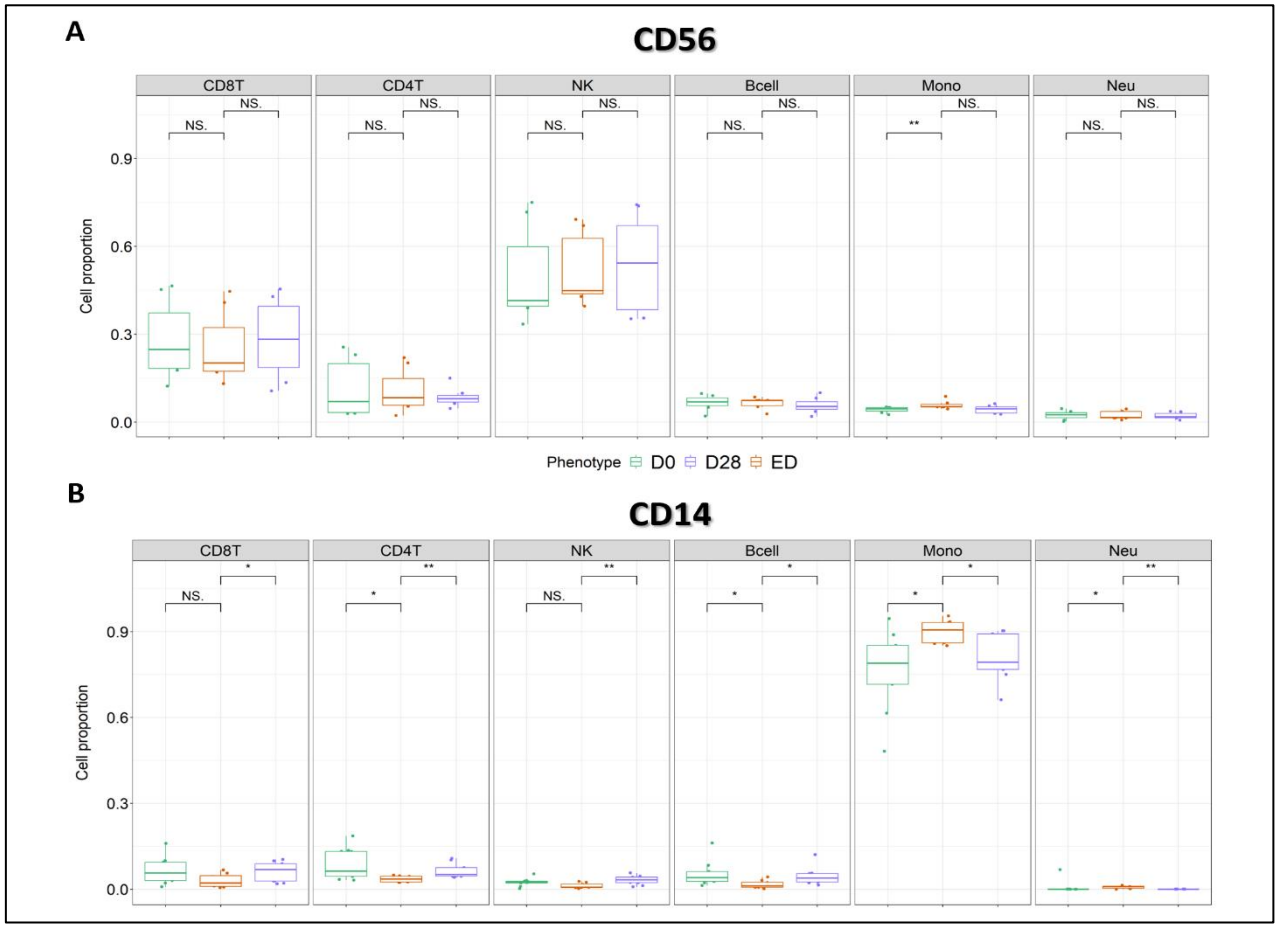

Fig. 1. *Boxplots showing the proportion of leukocyte cell type among the contrasts ED-D0 and ED-D28 in (A) CD56+ and (B) CD14+ cells. Asterix indicated significance levels (* for P-value < 0.01 and **for P-value < 0.001).*

Search for differentially methylated position (DMPs)

To search for DMPs among the three timepoints, the two cell subpopulations were analyzed individually. Before applying the linear model, the clinical characteristics of the patients were observed to assess for any possible cofounding factor. Cigarette smoking has previously been shown to influence and to have important effects on DNA methylation patterns (Lee & Pausova, 2013); in our cohort almost half of individuals were classified as smokers. However, when we evaluated the difference between smokers and no smokers in the basal timepoint, we did not detect any significant difference (Figure 2A and 2B). For this reason, the smoker status factor was not added to the linear regression model.

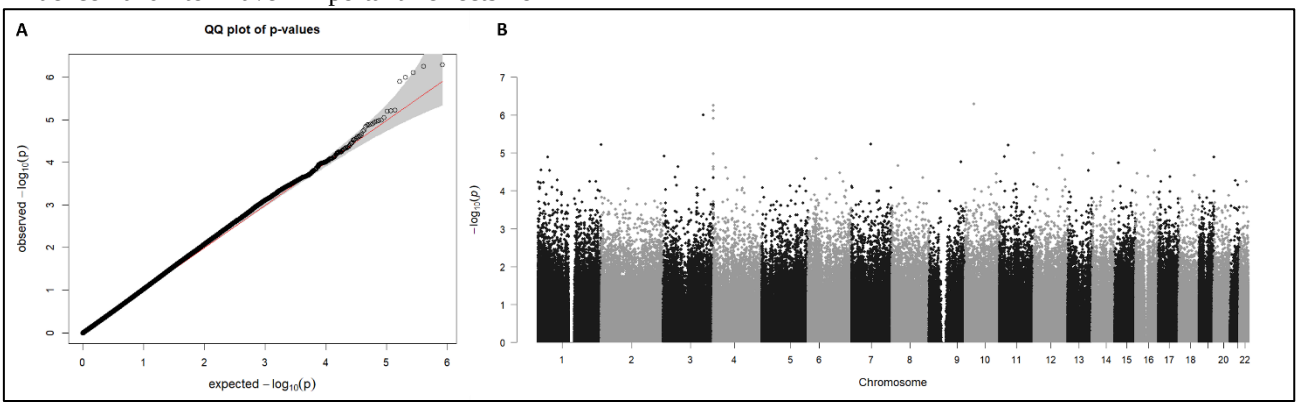

Fig. 2. *(A) QQ plot of the observed P-values against the expected P-values between smokers and non-smokers before the infectious challenge. (B) Manhattan plot of CpG sites associated with smoking. X-axis shows chromosomal positions. Y-axis shows -log10 P-values. No significant positions were observed between the two groups of subjects.*

CD56 cells

Only monocyte cell type was found to be significantly different among individuals before and during the infection, so this is the only cellular variable added, together with sex and age in the limma linear model.

The preprocessed filtering of CD56+ cells data, left for downstream analysis 813,220 CpGs, and 21 samples (7 individuals at three timepoints). The limma linear regression model adjusted for age, sex and monocyte composition did not give any significant results with an FDR <0.05. For this reason, we consider positions with a nominal P-value <0.001. In this way we were able to detect 718 CpGs that perfectly split the three subgroups in the PCA (Figure 3A).

The 718 DMPs identified were annotated to 297 unique genes. 316 CpGs (44%) were found to be hypomethylated in infected individuals when compared to individuals before infection, while 347 (48%) sites were hypomethylated in ED timepoints when compared to D28.

CD14

The six leukocyte subpopulations were found to be significantly different among individuals before, during and after infections in the CD14+ cells, so in the design model, all of them were added as confounding factors together with age and sex. Filter analysis of CD14+ cells data, left 814,342 CpGs, and 26 samples (9 at D0 and D28 and 8 at ED timepoints), one sample of the ED group was removed during preprocessing step for bad quality (sample with detection P-value > 0.05).

The limma model adjusted for age, sex and cell composition give 12 significant DMPs for the contrast ED-D0 and 10 DMPs for the contrast ED-D28. Eight DMPs were shared between the two contrasts. The total 14 significant CpGs split perfectly individuals before and after the infection from those during the infection (Figure 3B), along with the first principal component (PC1; accounting for 68.9% of the variance).

In both contrasts (D0-ED and D28-ED) a general hypomethylation toward the infected individuals was observed (Figure 4), with eleven positions out twelve in the first comparison and seven positions out ten in the second one.

Specifically, looking at the positions differentially methylated in the two contrasts (Table 2), we could observe that in both comparisons, the DMPs exhibited similar methylation patterns, being 11 of them (81%) hypomethylated in infected individuals when compared with those before and after infection. The CpGs were annotated to nine unique genes; for ten CpGs, the delta beta in both contrasts is higher than the threshold established (0.10), reflecting high methylation differences for these genes among the three timepoints.

The 6 annotated DMPs observed in CD14+ cells, with the highest difference in DNA methylation (Delta Beta >0.10) were selected to evaluate their ability, in terms of methylation levels, to discriminate infected individuals from those before and after the infection (Figure 5). In both contrasts, it is possible to see how the methylation pattern of D0 and D28 are very similar, suggesting that after the infection and with the treatment, the methylation level changes during infection return to the basal state post-treatment. To assess the potential predictive value of DMPs from these genes as biomarkers of paratyphoid infection, ROC analysis with AUC calculations was performed. The majority of ROC curves for the six genes in the contrast D0-ED yielded a mean AUC of 1.000, while the ROC curves for the six DMPs for the contrast D28-ED returned a mean AUC of 0.939.

Gene Set Analysis

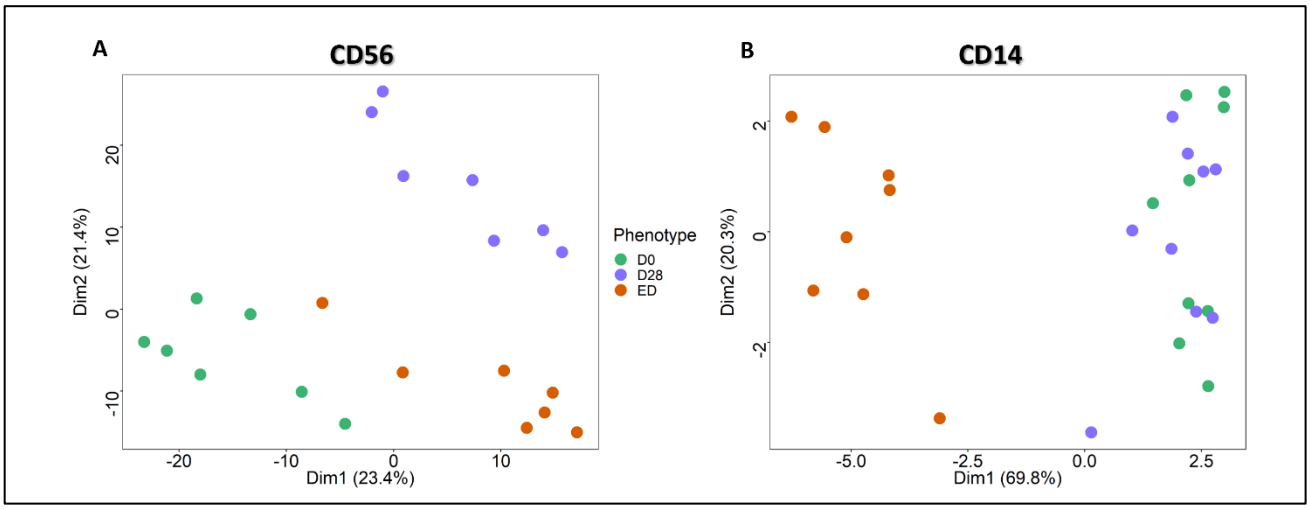

Fig. 3. *PCA of the significant DMPs found among the three timepoints in (A) CD56+ cells (P-value < 0.001) and (B) in CD14+ cells (FDR < 0.05).*

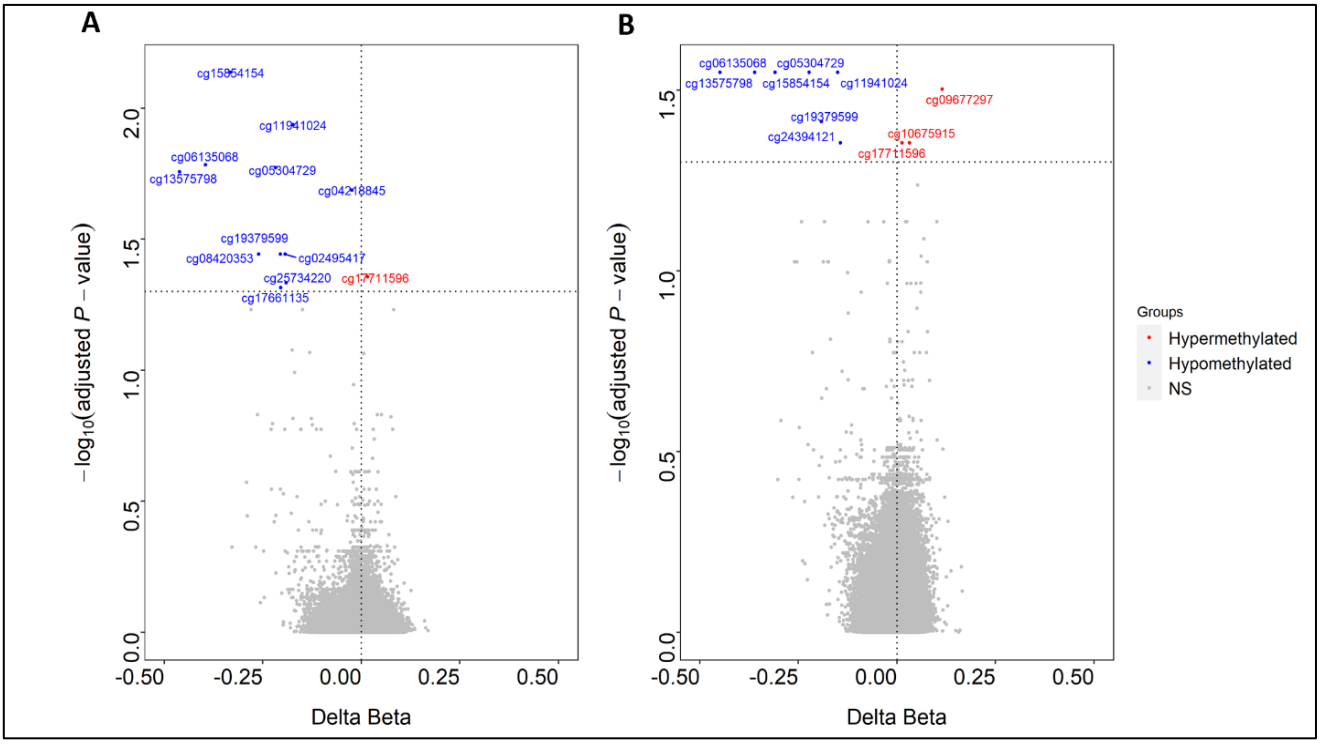

Fig. 4. *Volcano plot showing changes in DNA methylation in CD14+ cells in (A) individuals before infection compared with those infected and (B) in subjects after infections with those already treated. Blues and red labels indicated hypomethylated and hypermethylated positions, respectively, in ED timepoint when compared to D0 and D28.*

The analysis of pathways performed in CD56+ cell population reveal significant pathways for Reactome database and GO biological processes (BP), considering in the analysis all CpGs or only those within the promoter regions. The top BP gene sets in GO (Figure 6A) were found to be related with regulation of DNA replication and chromosome / chromatid segregation, as well as ATPase complex and regulation of response to DNA damage stimulus. Concerning the Reactome database, the top significant pathways observed (Figure 6B) were involved in the transcriptional regulation by TP53, cell cycle checkpoint, rRNA processing and DNA repair.

The same analysis performed in CD14+ cells, revealed different results, with statistical significance for GO BP (Figure 7) related to Golgi vesicle transport and MHC protein complex.

Table 2. *Differentially methylated positions observed in CD14+ cells*

CpG Name	Chromosome	Gene Name	Gene Group	FDR P-value	DeltaBeta ED-D0	DeltaBeta ED-D28
cg13575798	chr7	SND1	Body	0,01748	-0,46138	-0,44896
cg06135068	chr16			0,01642	-0,39513	-0,36122
cg15854154	chr10	PSAP	Body	0,00730	-0,33110	-0,30870
cg08420353	chr12			0,03600	-0,26076	-0,24170
cg05304729	chr1	MNDA	TSS1500	0,01686	-0,21773	-0,22241
cg19379599	chr21			0,03600	-0,20559	-0,19190
cg17661135	chr8	NRG1	Body	0,04830	-0,20421	-0,18618
cg02495417	chr14			0,03600	-0,19329	-0,18398
cg25734220	chr10	TCF7L2	Body	0,04634	-0,19098	-0,17371
cg11941024	chr15	VPS39	Body	0,01156	-0,17378	-0,14973
cg04218845	chr1	SSU72	Body	0,02053	-0,02441	-0,01870
cg10675915	chr11			0,04423	0,00609	0,03238
cg17711596	chr2	DGKD	TSS1500	0,04395	0,01576	0,01289
cg09677297	chr14	LINC01146	Body	0,03137	0,08261	0,11497

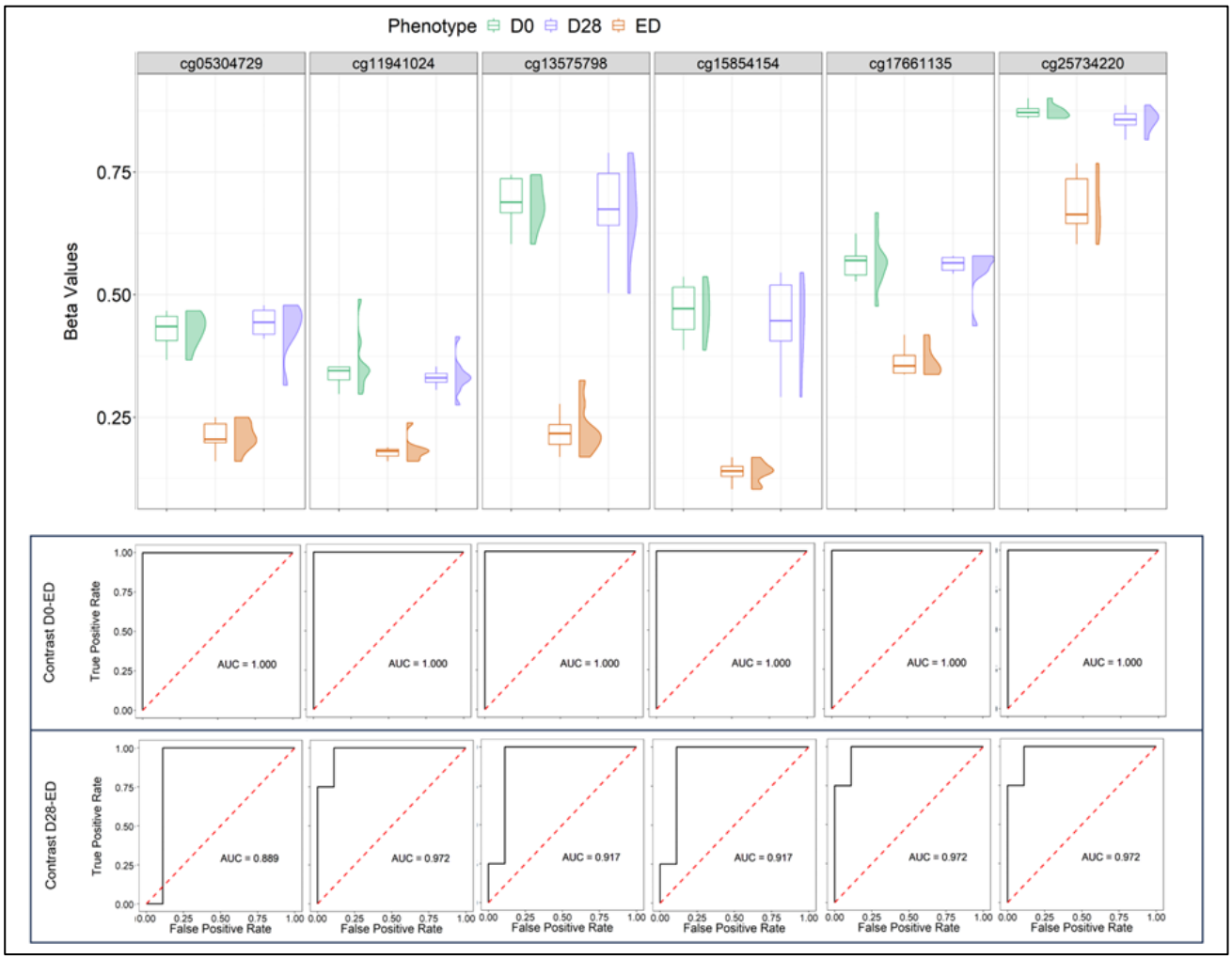

Fig. 5. *Boxplot and ROC curves of the DMPs with the highest difference in methylation between contrasts annotated to genes.*

4 Discussion

Enteric CHIMs provide a distinctive chance to investigate how the immune system responds to bacterial pathogens within a tightly controlled setting. Here, for the first time, we explored the immune cell methylation levels before, during and after paratyphoid infection, in two distinct cell populations, CD14+ and CD56+ cells. The analysis of the three timepoints performed separately in the two cell leukocyte populations revealed that biggest differences in DNA methylation were observed in infected individuals in CD14+ monocytes cells, where also a significant different composition of leukocyte cells, probably due to the purity of isolated cells, was observed among timepoints.

The innate immune system forms a barrier against pathogens, involving the participation of cells such as macrophages and dendritic cells (DCs). Circulating monocytes exhibit remarkable adaptability and are able to convert into different macrophage and DC types upon migrating to tissues or in response to specific cytokine signals (Peters, Ruppert, Gieseler, Najar, & Xu, 1991).

In this study, we observed more changes in the methylation patterns of some specific genes before, during and after the paratyphoid infection in CD14+ cells, than in CD56+. This may be due to the fact that CD14+ cells, being primarily associated with monocytes and macrophages, are the key players in the initial stage of the immune response and their activation, migration and differentiation can results in significant changes in response to pathogen-associated molecules.

In this particularly cell types (CD14+), highest differences in methylation in individuals before, during and after the paratyphoid infection, are observed in seven specific genes. In all of them, CpGs are found to be hypomethylated in infected individuals and the 6 of them settled in the gene body, while only one within the promoter. It is not clear how the methylation levels in gene body affects the expression of the gene. A recent review on cancer (Q. Wang et al., 2022) has revealed that DNA methylation is not limited to promoters but extends into gene bodies where methylation changes significantly impact gene expression regulation and are intricately tied to the development and

advancement of cancerous tumors. Even if the precise function of gene body methylation in gene transcription remains a subject of ongoing debate, it suppresses spurious gene transcription, regulate alternative splicing, and maintain stable and organized transcription processes.

The CpG site with the highest difference in DNA methylation between the three groups settled in the gene body is the SND1 gene, the staphylococcal nuclease domain-containing protein 1, that is a multifunctional protein, recently

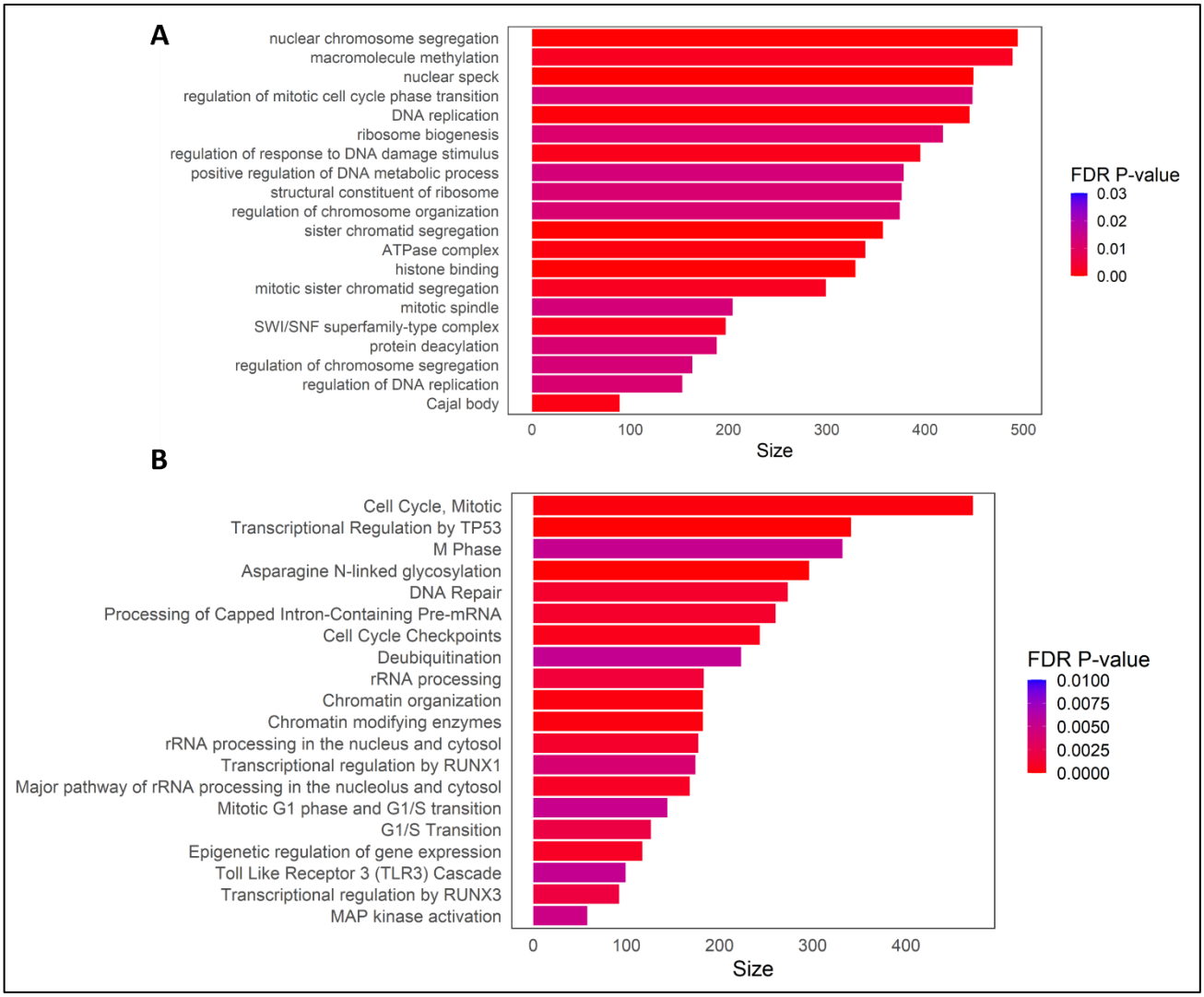

Fig. 6. *Barplot showing the results of the gene set analysis performed with DMPs found in CD56+ cells in (A) GO biological processes and (B) in Reactome database.*

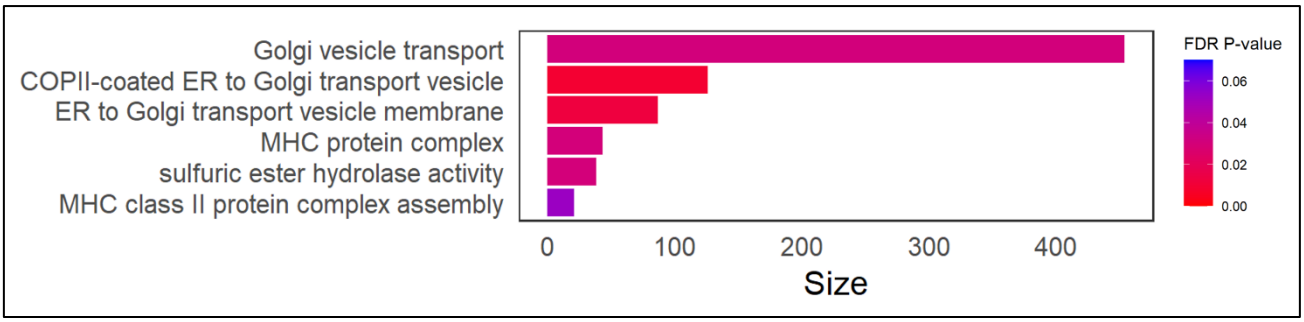

Fig.7. *Barplot showing the results of the gene set analysis performed with DMPs found in CD14+ cells in GO biological processes.*

associated with host defense against intracellular infections (X. Wang et al., 2021).

In this recent study, the gene was explored in the context of chlamydial infection and it was observed that the absence of SND1 resulted in a change in the characteristics and activity of dendritic cells (DCs), which is linked to an incapacity in fostering the creation of safeguarding Th1/17 immune responses. The second gene with the highest difference in methylation is the PSAP gene followed by the NRG1. PSAP gene, encoding for the lysosomal protein prosaposin is found to be present in higher expression levels in adult liver and body fluids (Kolter & Sandhoff, 2005); recently, it has been reported the implication of this gene in inflammation processes associated with cellular metabolism and mTOR signaling (van Leent et al., 2021). NRG1, instead, is the Neuregulin 1, and is a gene involved in cell interactions, and in the regulation of vital processes like cell proliferation, migration, and apoptosis across various systems. Recent findings propose NRG1's potential role in immune response modulation, possibly functioning as an anti-inflammatory or antioxidant agent (Alizadeh, Santhosh, Kataria, Gounni, & Karimi-Abdolrezaee, 2018).

The only one positions of the top seven, found within the promoter region is the cg05304729 annotated within the TSS1500 of the MNDA, a PYHIN protein, known as myeloid cell nuclear differentiation antigen, that is found to be primarily expressed in CD14+ monocytes, being significantly important in the induction of IFNα (Gu et al., 2022). Unlike other PYHIN proteins, the role of MNDA in this process isn't related to pathogen sensing; instead, it controls the expression of IRF7, a pivotal transcription factor required for IFNα induction. Intriguingly, MNDA itself becomes recruited to the IRF7 promoter following stimulation by type I interferons. These findings highlight MNDA's critical role in regulating the type I interferon cascade within human myeloid cells, shedding light on a novel function for human PYHIN proteins in the initiation of innate immune gene responses.

The gene set analysis revealed different results for contrasts in CD56+ and CD14+ cells. In the former, the majority of significantly enriched pathways were related to regulation of DNA replication, response to DNA damage stimulus, cell cycle checkpoint, and DNA repair.

While in the latter, the few significant GO biological processes were linked to Golgi vesicle transport and regulation of MHC class complex. These last results are in line with previous studies which report an interesting interaction between Salmonella Typhi and the Golgi vesicle transport system that contributes to the bacterium's pathogenicity. In fact, Salmonella species, have evolved mechanisms to manipulate host cells and create a suitable environment for their survival and replication (McGhie, Brawn, Hume, Humphreys, & Koronakis, 2009).

During infection, Salmonella Typhimurium can inject proteins known as effectors into host cells using a specialized structure called the type III secretion system (T3SS) (Dos Santos, Ferrari, & Conte-Junior, 2020). These effectors play a role in modulating various host cellular processes to the bacterium's advantage. In the context of Golgi vesicle transport, Salmonella has been found to manipulate this process to establish a replication niche called the Salmonella-containing vacuole (SCV) within host cells (Bakowski, Braun, & Brumell, 2008). This vacuole is derived from the host cell's membrane and provides a protected environment where Salmonella can replicate. The bacterium uses effectors to manipulate the SCV's interactions with the host cell's endomembrane system, including the Golgi apparatus. By disrupting Golgi vesicle transport and manipulating host cell processes, Salmonella can evade host defenses and establish a suitable environment for its survival and replication.

5 Conclusions

In conclusion in the present study, we detected methylation alterations occurring during paratyphoid infection in specific cell populations. The outcomes of this study might unveil novel epigenetic biomarkers of paratyphoid infection, shedding light on the biological mechanisms related to proteins and pathways that either facilitate or hinder paratyphoid infection. Such insights could be helpful in the development of innovative paratyphoid vaccines or enteric fever therapies.

6 Availability of data

R code used to analyse data is available at the following URL:
https://github.com/spischedda/TFM_OmicsData_Rcode/blob/main/TFM_OmicDataAnalysis_Rscript.R

7 References

Alizadeh, A., Santhosh, K. T., Kataria, H., Gounni, A. S., & Karimi-Abdolrezaee, S. (2018). Neuregulin-1 elicits a regulatory immune response following traumatic spinal cord injury. J Neuroinflammation, 15(1), 53. doi:10.1186/s12974-018-1093-9

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics, 30(10), 1363-1369. doi:10.1093/bioinformatics/btu049

Bakowski, M. A., Braun, V., & Brumell, J. H. (2008). Salmonella-containing vacuoles: directing traffic and nesting to grow. Traffic, 9(12), 2022-2031. doi:10.1111/j.1600-0854.2008.00827.x

Bannister, S., Messina, N. L., Novakovic, B., & Curtis, N. (2020). The emerging role of epigenetics in the immune response to vaccination and infection: a systematic review. Epigenetics, 15(6-7), 555-593. doi:10.1080/15592294.2020.1712814

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B, 57(1), 289-300.

Bibikova, M., Lin, Z., Zhou, L., Chudin, E., Garcia, E. W., Wu, B., . . . Fan, J. B. (2006). High-throughput DNA methylation profiling using universal bead arrays. Genome Res, 16(3), 383-393. doi:10.1101/gr.4410706

Dobinson, H. C., Gibani, M. M., Jones, C., Thomaides-Brears, H. B., Voysey, M., Darton, T. C., . . . Pollard, A. J. (2017). Evaluation of the Clinical and Microbiological Response to Salmonella Paratyphi A Infection in the First Paratyphoid

Human Challenge Model. Clin Infect Dis, 64(8), 1066-1073. doi:10.1093/cid/cix042

Dos Santos, A. M. P., Ferrari, R. G., & Conte-Junior, C. A. (2020). Type three secretion system in Salmonella Typhimurium: the key to infection. Genes Genomics, 42(5), 495-506. doi:10.1007/s13258-020-00918-8

Dougan, G., & Baker, S. (2014). Salmonella enterica serovar Typhi and the pathogenesis of typhoid fever. Annu Rev Microbiol, 68, 317-336. doi:10.1146/annurev-micro-091313-103739

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. J. B. b. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics, 11(1), 587.

Fernandez-Real, J. M., Perez del Pulgar, S., Luche, E., Moreno-Navarrete, J. M., Waget, A., Serino, M., . . . Zorzano, A. (2011). CD14 modulates inflammation-driven insulin resistance. Diabetes, 60(8), 2179-2186. doi:10.2337/db10-1210

Fortin, J. P., Triche, T. J., Jr., & Hansen, K. D. (2017). Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics, 33(4), 558-560. doi:10.1093/bioinformatics/btw691

Gianchecchi, E., Delfino, D. V., & Fierabracci, A. (2018). NK cells in autoimmune diseases: Linking innate and adaptive immune responses. Autoimmun Rev, 17(2), 142-154. doi:10.1016/j.autrev.2017.11.018

Goyert, S. M., Ferrero, E. M., Seremetis, S. V., Winchester, R. J., Silver, J., & Mattison, A. C. (1986). Biochemistry and expression of myelomonocytic antigens. J Immunol, 137(12), 3909-3914.

Gu, L., Casserly, D., Brady, G., Carpenter, S., Bracken, A. P., Fitzgerald, K. A., . . . Bowie, A. G. (2022). Myeloid cell nuclear differentiation antigen controls the pathogen-stimulated type I interferon cascade in human monocytes by transcriptional regulation of IRF7. Nat Commun, 13(1), 14. doi:10.1038/s41467-021-27701-x

Houseman, E. A., Kelsey, K. T., Wiencke, J. K., & Marsit, C. J. (2015). Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective. BMC Bioinformatics, 16, 95. doi:10.1186/s12859-015-0527-y

Jin, C., Gibani, M. M., Moore, M., Juel, H. B., Jones, E., Meiring, J., . . . Pollard, A. J. (2017). Efficacy and immunogenicity of a Vi-tetanus toxoid conjugate vaccine in the prevention of typhoid fever using a controlled human infection model of Salmonella Typhi: a randomised controlled, phase 2b trial. Lancet, 390(10111), 2472-2480. doi:10.1016/S0140-6736(17)32149-9

Jin, C., Hill, J., Gunn, B. M., Yu, W. H., Dahora, L. C., Jones, E., . . . Pollard, A. J. (2021). Vi-specific serological correlates of protection for typhoid fever. J Exp Med, 218(2). doi:10.1084/jem.20201116

Jones, C., Sadarangani, M., Lewis, S., Payne, I., Saleem, M., Derrick, J. P., & Pollard, A. J. (2016). Characterisation of the Immunomodulatory Effects of Meningococcal Opa Proteins on Human Peripheral Blood Mononuclear Cells and CD4+ T Cells. PLoS One, 11(4), e0154153. doi:10.1371/journal.pone.0154153

Jones, P. A., & Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. Nat Rev Genet, 3(6), 415-428.

Kolter, T., & Sandhoff, K. (2005). Principles of lysosomal membrane digestion: stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids. Annu Rev Cell Dev Biol, 21, 81-103. doi:10.1146/annurev.cellbio.21.122303.120013

Lee, K. W., & Pausova, Z. (2013). Cigarette smoking and DNA methylation. Front Genet, 4, 132. doi:10.3389/fgene.2013.00132

McGhie, E. J., Brawn, L. C., Hume, P. J., Humphreys, D., & Koronakis, V. (2009). Salmonella takes control: effector-driven manipulation of the host. Curr Opin Microbiol, 12(1), 117-124. doi:10.1016/j.mib.2008.12.001

Peters, J. H., Ruppert, J., Gieseler, R. K., Najar, H. M., & Xu, H. (1991). Differentiation of human monocytes into CD14 negative accessory cells: do dendritic cells derive from the monocytic lineage? Pathobiology, 59(3), 122-126. doi:10.1159/000163628

Poli, A., Michel, T., Theresine, M., Andres, E., Hentges, F., & Zimmer, J. (2009). CD56bright natural killer (NK) cells: an important NK cell subset. Immunology, 126(4), 458-465. doi:10.1111/j.1365-2567.2008.03027.x

Ren, X., & Kuan, P. F. (2019). methylGSA: a Bioconductor package and Shiny app for DNA methylation data length bias adjustment in gene set testing. Bioinformatics, 35(11), 1958-1959. doi:10.1093/bioinformatics/bty892

Russo, V. E., Martienssen, R. A., & Riggs, A. D. (1996). Epigenetic mechanisms of gene regulation: Cold Spring Harbor Laboratory Press.

Smyth, G. K. (2005). Limma: linear models for microarray data. In Bioinformatics and computational biology solutions using R and Bioconductor (pp. 397-420): Springer.

Toapanta, F. R., Bernal, P. J., Fresnay, S., Darton, T. C., Jones, C., Waddington, C. S., . . . Sztein, M. B. (2015). Oral Wild-Type Salmonella Typhi Challenge Induces Activation of Circulating Monocytes and Dendritic Cells in Individuals Who Develop Typhoid Disease. PLoS Negl Trop Dis, 9(6), e0003837. doi:10.1371/journal.pntd.0003837

Typhoid, G. B. D., & Paratyphoid, C. (2019). The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis, 19(4), 369-381. doi:10.1016/S1473-3099(18)30685-6

van Leent, M. M. T., Beldman, T. J., Toner, Y. C., Lameijer, M. A., Rother, N., Bekkering, S., . . . Duivenvoorden, R. (2021). Prosaposin mediates inflammation in atherosclerosis. Sci Transl Med, 13(584). doi:10.1126/scitranslmed.abe1433

Wang, Q., Xiong, F., Wu, G., Liu, W., Chen, J., Wang, B., & Chen, Y. (2022). Gene body methylation in cancer: molecular mechanisms and clinical applications. Clin Epigenetics, 14(1), 154. doi:10.1186/s13148-022-01382-9

Wang, X., Zhang, C., Wang, S., Rashu, R., Thomas, R., Yang, J., & Yang, X. (2021). SND1 promotes Th1/17 immunity against chlamydial lung infection through enhancing dendritic cell function. PLoS Pathog, 17(2), e1009295. doi:10.1371/journal.ppat.1009295

World Health, O. (2019). Typhoid vaccines: WHO position paper, March 2018 - Recommendations. Vaccine, 37(2), 214-216. doi:10.1016/j.vaccine.2018.04.022

Zanoni, I., & Granucci, F. (2012). Regulation and dysregulation of innate immunity by NFAT signaling downstream of pattern recognition receptors (PRRs). Eur J Immunol, 42(8), 1924-1931. doi:10.1002/eji.201242580

Zhang, Q., & Cao, X. (2019). Epigenetic regulation of the innate immune response to infection. Nat Rev Immunol, 19(7), 417-432.

Zhu, H., Chelysheva, I., Cross, D. L., Blackwell, L., Jin, C., Gibani, M. M., . . . O'Connor, D. (2023). Molecular correlates of vaccine-induced protection against typhoid fever. J Clin Invest, 133(16). doi:10.1172/JCI169676

Ziller, M. J., Gu, H., Muller, F., Donaghey, J., Tsai, L. T., Kohlbacher, O., . . . Meissner, A. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature, 500(7463), 477-481. doi:10.1038/nature12433