Transcriptome profiling to understand host-bacteria interactions: Past, present and future

Sheila Nathan

Department of Biological Sciences and Biotechnology, Faculty of Science and Technology,
Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

e-mail: sheila@ukm.edu.my

ABSTRACT: Despite advances in antimicrobials, vaccination and public health measures, bacterial infectious diseases remain a leading cause of morbidity and mortality worldwide. With the increase in antimicrobial resistance and the emergence of new bacterial pathogens, there remains a need for better understanding of the host response to infection. This would lead to new paths of basic research and the identification of potential diagnostic biomarkers and new drug targets to overcome issues of antibiotic resistance. Bacterial pathogens have evolved strategies to promote their survival by significantly over-riding the transcriptional profile of the host cells they infect. In particular, pathogen-encoded effector molecules modulate host cells through different mechanisms. Transcriptomics is a powerful way to gauge these changes in mechanism of either bacterial or eukaryotic cells under a given condition. Microarray technology and more recently RNA sequencing are well established tools to unravel host-pathogen interactions which is of central importance to understand the host response to a particular bacterial infection or the mechanisms employed by a pathogenic bacteria to subvert host defenses. The establishment of dual RNA-sequencing has enabled the profiling of gene expression changes simultaneously in an infecting bacterium and its infected host. This technology is able to provide greater discriminative power to study the pathogen and host simultaneously throughout the infection process. This review provides an overview of the transcriptome-based tools utilised and new knowledge obtained from dissecting host-pathogen interactions.

KEYWORDS: transcriptome, dual RNA-sequencing, single cell RNA sequencing, *Burkholderia pseudomallei*

INTRODUCTION

Historically, infectious diseases have adopted a reductionist approach to study *in vitro* and *in vivo* infections by assessing the interaction of a particular microbe with a single host cell type. This has enabled important discoveries and advanced our understanding of mechanisms that underlie infection and disease. Pathogens, as we know, encompass a huge range of organisms such as bacteria, viruses, parasites and fungi, and interaction studies have utilised both traditional as well as 'omics-based approaches to dissect the host-pathogen interaction. This review will focus on bacterial-host interaction studies using transcriptome-based approaches.

Pathogenic bacteria need to rapidly adjust their virulence and fitness program to prevent eradication by the host as illustrated in the damage-response framework [1]. All bacterial pathogens encounter rapidly changing environmental conditions and adverse host reactions during the course of infection. Accordingly, a dynamic cascade of events is initiated that triggers a global alteration of gene expression patterns in both interacting organisms to adapt for survival. The outcome of the interaction between microbes and their host is usually a function of the immunological status of the host [1]. Generally, most hosts respond to molecular signs of a microbe invasion by initiating local defense mechanisms comprised of innate and acquired immune surveillance systems. Aggressive efforts to dissect the infected host response have utilised various infection model hosts. Mice and cells in culture are traditionally the preferred host model, however, cells do not represent the true physiology of the infected host whilst the number of mice in any one experiment can be quite restrictive. Hosts such as *Caenorhabditis elegans* and *Galleria melonella* larvae are more amenable to large scale experiments and data are easily translatable to the human host due to the conservation of the primary immune response.

Monitoring infection-linked transcriptome alterations represents a powerful approach to identify virulence-related factors and regulatory pro-
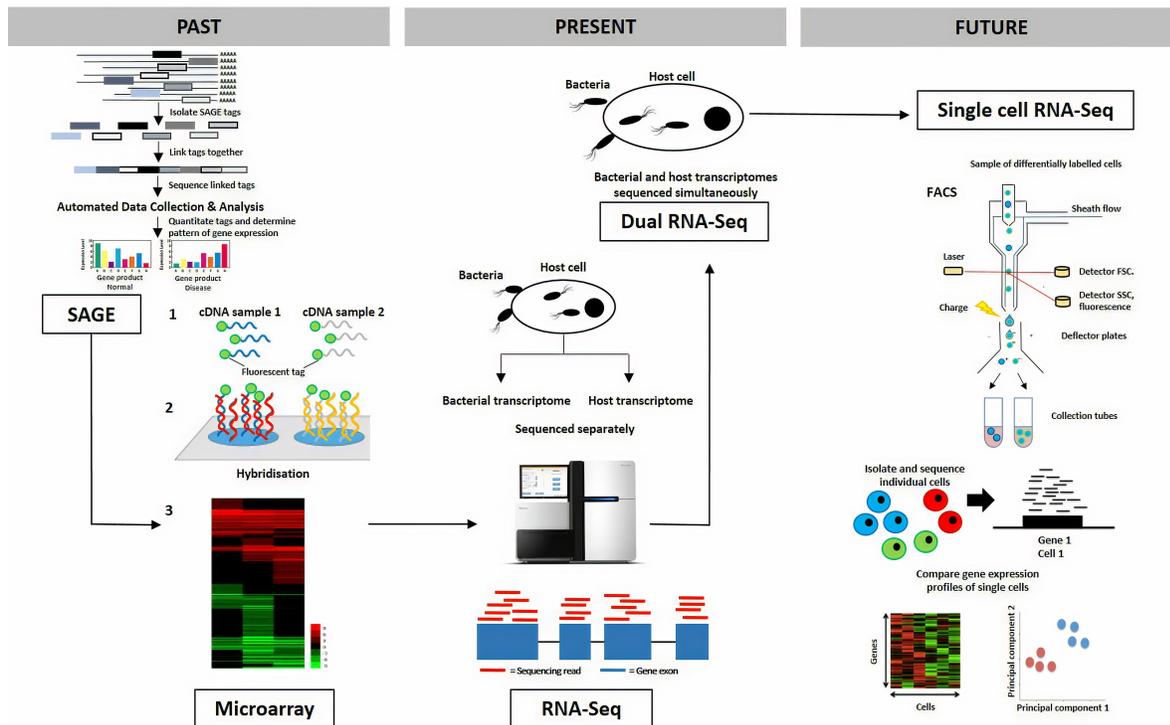
Fig. 1 The development of platform technologies for investigating host-pathogen interactions of pathogenic bacteria by transcriptome profiling. Historically, whole genome gene expression changes were profiled with SAGE and CAGE followed by microarray technology. The limitations of these platforms were overcome with the introduction of RNA-Seq while dual RNA-Seq has enabled the simultaneous mapping and comparative gene expression analysis of both host and bacteria. Moving on, platforms such as single cell (sc) RNA-Seq and sc dual RNA-Seq will increase the sensitivity of capturing bacterial transcripts towards greater power in comparative analysis of changes in pathogen gene expression.

cesses that drive bacterial pathogenesis. One of the earliest approaches was Serial Analysis of Gene Expression (SAGE) which incorporates Sanger sequencing to generate and sequence short 3′ tags to quantify transcript abundance. Improvements to SAGE have included Massively Parallel Signature Sequencing (MPSS) and Cap Analysis of Gene Expression (CAGE) [2] which is similar in principle to SAGE but targets 5′ transcription start sites. The arrival of microarray technology transformed gene expression studies but was eventually superseded by RNA sequencing (RNA-Seq). This high-throughput sequencing approach provides significant advantages with increased sensitivity and enables novel transcript discovery as well as non-coding RNA and alternative splicing analysis. More recently, the development of dual RNA-Seq allows simultaneous analysis of host and pathogen transcriptomes. Lately, single-cell (sc) RNA-Seq and sc-dual RNA-Seq have revolutionised the analyses of transcriptomes by enabling precise and sensitive transcrip-

tional profiling at single cell and single-nucleotide resolution. This timeline in the development of transcriptome platforms to interrogate host-bacteria interactions is summarised in Fig. 1.

SAGE

Traditional methods to measure the cellular expression of individual transcripts were based on gene-specific DNA probes used for detection by hybridis-ation (e.g. Northern blot) or by amplification (e.g. quantitative real-time PCR (qRT-PCR)). However, such approaches were limited to the analysis of a small number of pre-selected target genes. The timeline for host-pathogen interaction studies that focused on expression profiling started in the late 20th century using techniques such as SAGE and microarray. SAGE [3] produces a snapshot of the messenger RNA (mRNA) population in a sample of interest whereby gene fragments are tagged and the number of individual tags is quantified to represent mRNA transcript numbers (Fig. 1). One of

the first successful SAGE analysis of host-bacteria interactions was the infection of epithelial cells by *Pseudomonas aeruginosa* [4]. By integrating both bacterial and cell transcriptional changes, the investigators were able to demonstrate that overexpression of *P. aeruginosa* lipoproteins activated epithelial toll-like receptors (TLR) 2 and 5 as well as the transcription factor NF$_{Kappa}$B which in turn, increased expression of cytokines, chemokines, proteinase inhibitors, and components necessary to strengthen the host physical barrier.

Microarray

The microarray platform uses the concept of signal intensity to determine the quantitative value of gene expression. DNA probes are mobilised onto glass slides or membranes [5] and fluorescently-labelled host or bacterial RNA or cDNA samples are hybridised onto the chips. Changes in transcript abundance are reflected by fluorescence signal intensity of the respective gene probe (Fig. 1). Microarrays were used to interrogate expression changes of several important bacterial pathogens such as *Vibrio cholerae* [6], *Helicobacter pylori* [7], *Clostridium pneumoniae* [8] and *Listeria monocytogenes* [9]. A tiling microarray (which can reveal antisense RNA expression and other non-coding RNA transcripts) transcriptome study of *Bacillus subtilis* exposed to a wide range of nutritional and environmental conditions established one of the most comprehensive repertoires of transcription units in a prokaryote [10].

Burkholderia pseudomallei is the causative agent of melioidosis, an important disease in Malaysia, Thailand, Australia, and other tropical regions [11] and there is still a dearth of information on the bacteria's pathogenic mechanisms as well as how it adapts to different environments [12]. We adopted the microarray approach to investigate the host response to infection by this tropical pathogen. We established the *B. pseudomallei–C. elegans* infection model and performed microarray analysis of the infected worms. Expression profiling of infected worms identified an increasing number of host genes that were suppressed significantly over time and within the genes that were modulated, a large fraction of them were genes under the control of the worm GATA transcription factor [13] which was shown to be a key transcription factor for activation of the host immune response during a *P. aeruginosa* infection [14]. Our further analysis demonstrated that the down-regulation of GATA targets was not due to the absence of the GATA transcription fac-

tor but rather, *B. pseudomallei* secretes an effector molecule through its Type-3 Secretion System that induces the host ubiquitin proteosomal system to degrade its own GATA transcription factor, thereby repressing the activation of the immune response. We conducted a parallel study using a *B. pseudomallei*-mouse infection model where we showed that the infected host pro-inflammatory response was overexpressed which explained the hallmark symptoms of melioidosis such as fever, formation of abscesses [15]. Our analysis demonstrated that the bacteria disrupts the balance between pro- and anti-inflammatory responses of the host which contributes to tissue and cellular injury. *B. pseudomallei* also inhibited the complement cascade, preventing the formation of the membrane attack complex that is required for bacterial lysis [15].

To extend our findings on suppression of the infected host, we also developed *B. pseudomallei* microarray chips to interrogate the bacterial transcriptome when adapting to the stressful conditions presented by the host milieu. To achieve this, macrophage cells were infected with *B. pseudomallei*, bacteria were recouped from the infected cells for bacterial RNA preparation and hybridisation onto *B. pseudomallei* arrays [16]. A large fraction of the bacterial genome was suppressed over the period of infection. For example, suppression of flagella genes expression helped the bacteria to avoid being sensed by the TLRs which in turn delayed activation of the immune response signalling cascade. This provided a window of opportunity for *B. pseudomallei* to adapt to the new host environment and subsequently, start to replicate and secrete toxic materials through the secretion systems to damage the host (Fig. 2).

An abundance of new knowledge on host-bacteria interactions was made available from studies utilising microarray technology. Nonetheless, the limitations of microarrays to dissect changes in either host or pathogen gene expression include high background noise resulting from non-specific signals [17], concerns of cross hybridisation due to the use of non-specific probes that are homologous to both host and pathogen sequences, and the preparation of the chip which was dependent on the availability of a full genome sequence of the organism studied. Other restrictions include failure to capture information on novel transcripts and low dynamic range of fluorescence intensities [18]. These restrictions guided the application of next generation sequencing technology to investigate host-bacteria interactions.
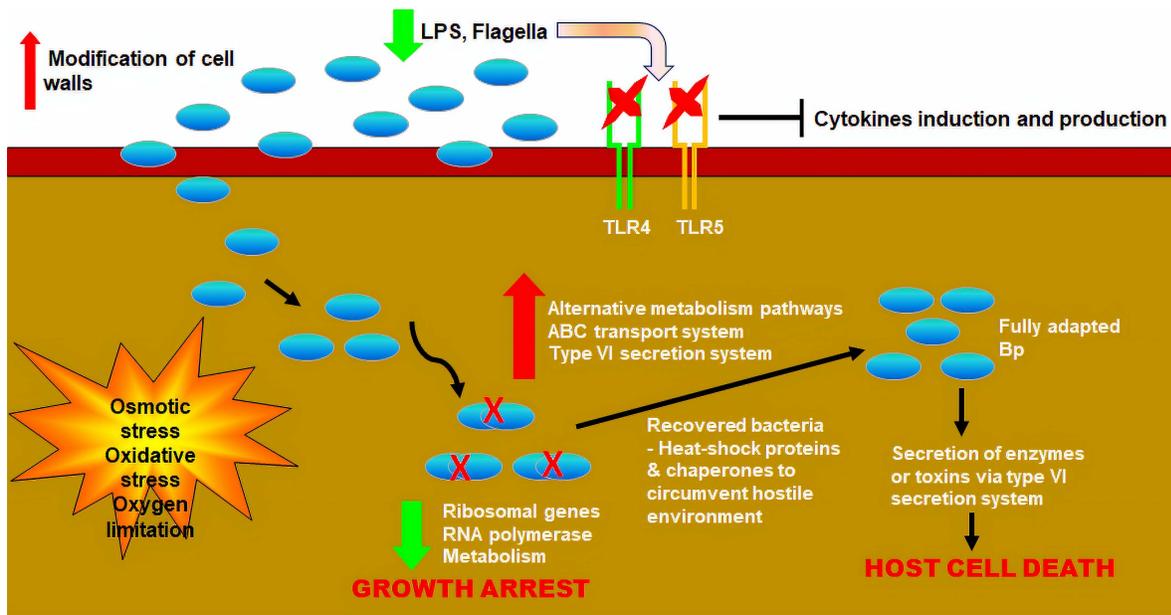
Fig. 2 *B. pseudomallei* adaptation during invasion of macrophage cells. During invasion of macrophage cells, genes for lipopolysaccharide biosynthesis and flagellar assembly were repressed. The repression of these genes caused the bacteria to avoid recognition by TLR4 and TLR5, thus inhibiting cytokine production and the inflammatory response. Once the bacteria successfully invaded the host cells, the bacteria are exposed to various stresses (e.g. osmotic and oxidative stress) in the intracellular compartment. Initially, the bacteria adopt a dormant state whereby genes encoding for ribosomes, RNA polymerase and metabolic proteins are down-regulated resulting in growth arrest. Once the bacteria have adapted, they express the ATP-binding cassette (ABC) transport system genes to obtain nutrients and also secrete various enzymes or toxins via the type VI secretion system, causing host cell lysis and death, enabling dissemination of the bacteria.

RNA sequencing (RNA-Seq)

Expression profiling of host and pathogen now routinely adopts the RNA-Seq approach that enables a more digital readout compared to the microarray platform [19]. RNA-Seq, or deep sequencing of cDNA libraries circumvents many of the problems associated with microarray profiling or tag-based sequencing. RNA-Seq can comprehensively and systematically define the transcriptome of an organism with minimal bias [20], across different experimental conditions or cell types [21] and without probe design or cross-hybridisation problems. In addition, RNA-Seq can detect transcribed intronic and intergenic regions, as well as post-transcriptional regulatory events such as alternative splicing and differential isoform expression [22, 23]. RNA-Seq data are consistent with microarray results [24] but are more sensitive, with essentially an ability to detect a significantly larger number of differentially expressed genes compared to microarrays [25]. RNA-Seq is annotation independent, allowing novel transcript discovery without being reliant on array

design or pre-existing genome annotation [26].

Some examples of RNA-Seq analysis on host-bacteria interactions include *Salmonella enterica* serovar Typhimurium (*S. typhimurium*) adaptation in macrophages [27], *V. cholerae* transcriptional changes during infection [28] and how *Streptococcus suis* adapts to different host environments [29]. Conejero et al [30] performed whole genome transcriptional analysis of pulmonary tissue and blood of mice challenged with *B. pseudomallei* and compared the expression profiles to that of blood transcriptional profiles of acute melioidosis patients from northeast Thailand. In this elegant study, the investigators demonstrated that the blood transcriptome of *B. pseudomallei* infected mice accurately mirrors the transcriptional signature of disease severity in human melioidosis. They went on to propose that although comparisons of human and animal disease models demonstrate poor correlation in expression profiles of different models of disease, identification of key genes or pathways shared in the response may expose known and novel infection-related immune
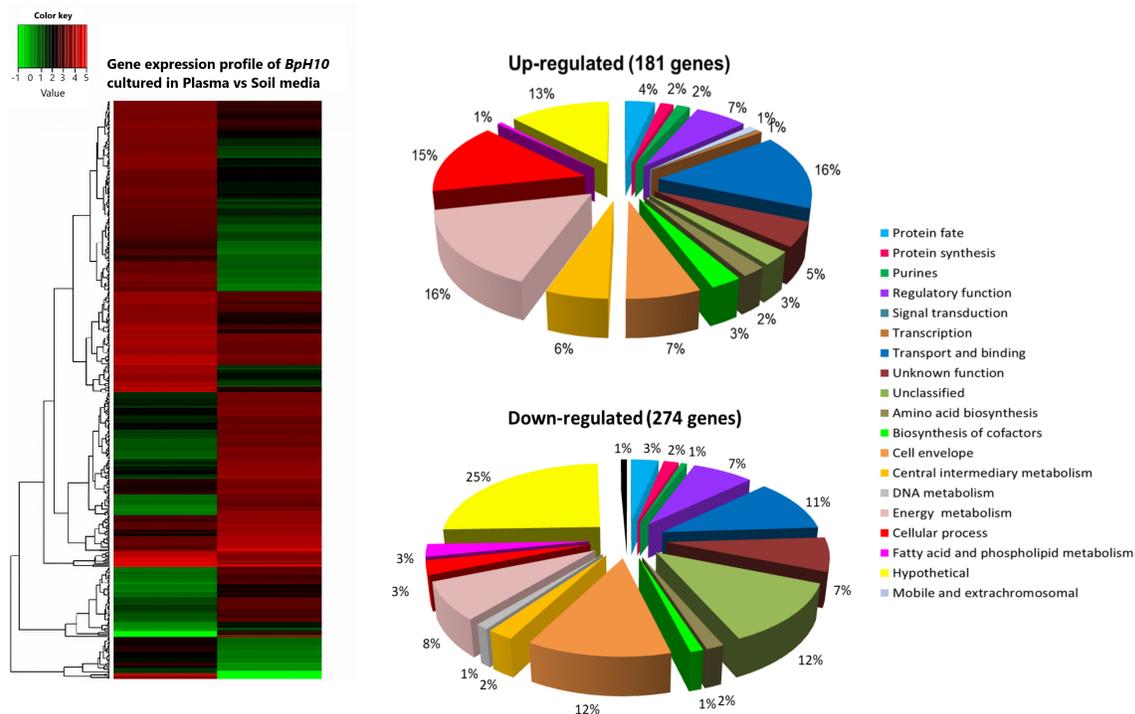
Fig. 3 Heat map and functional classification of genes that were significantly regulated in *B. pseudomallei* grown in plasma compared to soil. Coloured scales in the heat map represent log2 fold change values with green and red for down-regulated and up-regulated genes, respectively.

response mechanisms and assist in the identification of biomarkers pertinent to the host-pathogen relationship. Hirose et al [31] analysed the transcriptome profile of *Streptococcus pyogenes* using a mouse model of necrotising fasciitis. Their findings pointed to *S. pyogenes* altering its metabolism, reducing cell proliferation whilst up-regulating the expression of major toxins in the subcutaneous soft-tissue infection in an attempt to acquire nutrients from destroyed tissues.

RNA-Seq on a *B. pseudomallei* strain H10 grown in soil-extract media [32] and blood plasma was undertaken to catalogue transcriptional changes when the soil-dwelling pathogen transits to the host environment and adapts to its secondary role as a human pathogen [33]. Genes that were significantly regulated in *B. pseudomallei* grown in plasma compared to soil were functionally classified based on Comprehensive Microbial Resources (CMR) annotations. Most of these genes encode core functions such as energy metabolism, transport and binding, cellular process, regulatory proteins, cell envelope and central intermediary metabolism (Fig. 3). Of

note were a significant number of biofilm-associated genes and genes encoding for known virulence factors which were overexpressed under human-mimic conditions. This complemented our previous expression profiling data which noted that biofilm-related genes (surface-associated motility, surface composition and cell wall biogenesis) were overexpressed as were a number of virulence-associated genes and genes related to two-component signal transduction systems (TCS) [34]. TCS related proteins are responsible in regulating biofilm formation in a number of pathogenic bacteria whereby the bacteria sense environmental signals via the TCS and adapt to these changes by transcribing genes that are not regulated in non-biofilm producers [35].

We proposed that *B. pseudomallei* produces biofilm in response to the surrounding host environment to protect itself once removed from its natural environment as well as rerouting its focus to the secretion of virulence factors and secondary energy sourcing. An obvious transition to further expand on these host-mimicking conditions was to look at changes in the bacterial and host gene expression in

a live infection model. To achieve this purpose, we adopted the new technique of dual-RNA sequencing.

Dual RNA sequencing

Dual RNA sequencing (dual RNA-Seq) leverages on established next-generation sequencing-enabled RNA-Seq approaches that enables one to simultaneously capture genome-wide transcriptional changes of both infecting bacteria and host cells [36]. By simultaneously investigating both organisms from the same biological sample, dual RNA-Seq can provide unique insight into bacterial infection processes and reciprocal host responses at once under the same conditions at the same time point. This is not possible with conventional RNA-Seq where it is necessary to separate the prokaryote RNA from the eukaryote RNA and sequence each separately, which presents several challenges. For example, bacterial RNAs can constitute <1% of the total RNA in an infected cell. Moreover, up to 98% of total RNA in an infected cell is eukaryotic ribosomal RNA (rRNA), requiring rRNA depletion [37] or mRNA enrichment strategies to ensure sufficient number of transcripts can be sequenced at high coverage but with reasonable cost [38, 39]. Additionally, traditional cell lysis techniques are often not suitable for both eukaryotic and prokaryotic organisms at once. Finally, the trade-off between the multiplicity of infection (MOI) used and sequencing depth (cost) can lead to a mixed population of infected and uninfected cells, which may bias results. Dual RNA sequencing was successfully established by Westermann et al [40] who used GFP-labelled *Salmonella* to infect cells in culture after which infected cells were isolated by fluorescence-activated cell sorting (FACS) [41] and sequenced. Reads were mapped to both *Salmonella* and human genome sequences that enabled the expression profiling of *Salmonella*, host and host mitochondrial genomes. They then went on to propose how *Salmonella* pathogenicity islands control the induction of the host immune response. Studies undertaking dual RNA-Seq on *in vivo* samples have been performed on tissues rich in extracellular bacteria, such as *Yersinia pseudotuberculosis*-infected lymph nodes [42], *P. aeruginosa*-infected lung tissue [43], *S. pneumoniae* in lung and pleural space [44], and *Mycobacterium leprae*-infected human tissue lesions [45]. A new pipeline called Pathseq was recently used to recover the *M. tuberculosis* transcriptome from alveolar macrophages isolated from the murine lung [46]. Low et al. [47] were one of the first to undertake comparative transcriptome analysis of both the bacterium and the host in

an *in vivo* murine-*Clostridium perfringens* infection model. Their analysis led to the demonstration that *C. perfringens* up-regulated genes were those possibly involved in replication, virulence, subversion of host immune systems, and adaptation to host conditions. Concurrently, in the infected host, genes with modulated expression were involved in innate immunity and in response to changes in bacterial gene products.

As noted above, we wanted to simultaneously analyse the host and *B. pseudomallei* transcriptome changes during an active infection of a whole animal live infection model (Fig. 4). Mice were infected with the bacteria and at the height of infection, total mouse and bacterial RNA was obtained, following which, we enriched the bacterial RNA content by removing mouse and bacterial rRNA as well as mouse poly-A RNA [48]. In the infected mouse, a large number of genes were overexpressed, and these were delineated into functions related to metabolic and biological processes as well as the immune response. In the infected mouse at 5 days post-infection, we noted that the expression of cytokines is high leading to induction of the T_h1 and T_h2 response (Fig. 5A). On the other hand, the complement and coagulation cascade to promote bacterial lysis, is generally lowly expressed (Fig. 5B). Meanwhile, for *B. pseudomallei*, a relatively large fraction of the genome was down-regulated, in particular proteins related to cellular, metabolic, and biological processes suggesting that the bacteria was focusing its resources and energy on surviving in the hostile host environment. This is indeed reflected by the overexpression of genes functionally related to stress response, pathogenicity and motility.

Pisu et al [41] explored the *in vivo* molecular dynamics of *M. tuberculosis* (MTb) infection by performing dual RNA-Seq on MTb-infected, ontogenetically distinct macrophage lineages isolated directly from murine lungs. Their study uncovered a divergent transcriptional response of MTb between alveolar macrophages and interstitial macrophages. Alveolar macrophages supported MTb growth through improved access to iron and fatty acids while interstitial macrophages restricted Mtb growth through iron requisitioning and higher levels of nitric oxide. The first description of an interaction network between a bacterium and the human host at a site of nfection using the dual RNA-Seq approach demonstrated that *Haemophilus ducreyi* survives in an abscess by utilising L-ascorbate as an alternative carbon source, possibly taking advantage of host ascorbic acid recycling, and the pathogen also
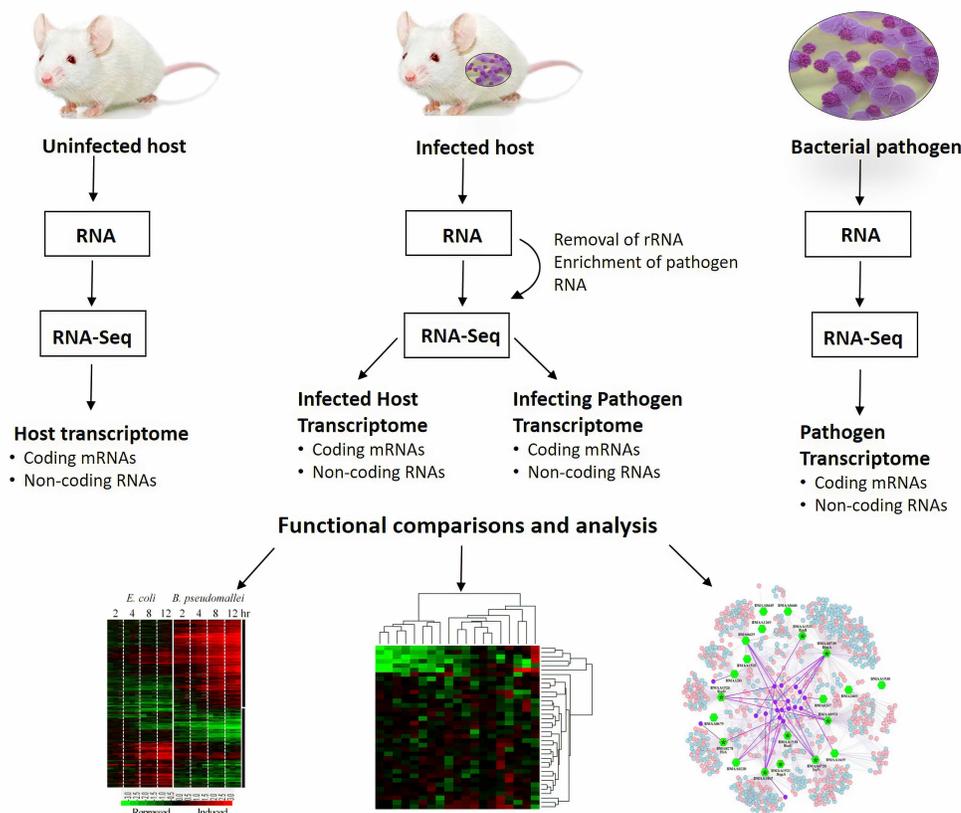
Fig. 4 A typical approach taken to simultaneously investigate host-bacteria interactions *in vivo* using the dual RNA-Seq approach. Transcripts from infected animal total RNA and bacteria-enriched mRNA are sequenced and mapped on to respective reference genomes. Functional comparisons of coding and non-coding RNAs as well as interactome maps provide a comprehensive picture on changes in gene expression of host and bacteria captured at the same time point or under similar conditions.
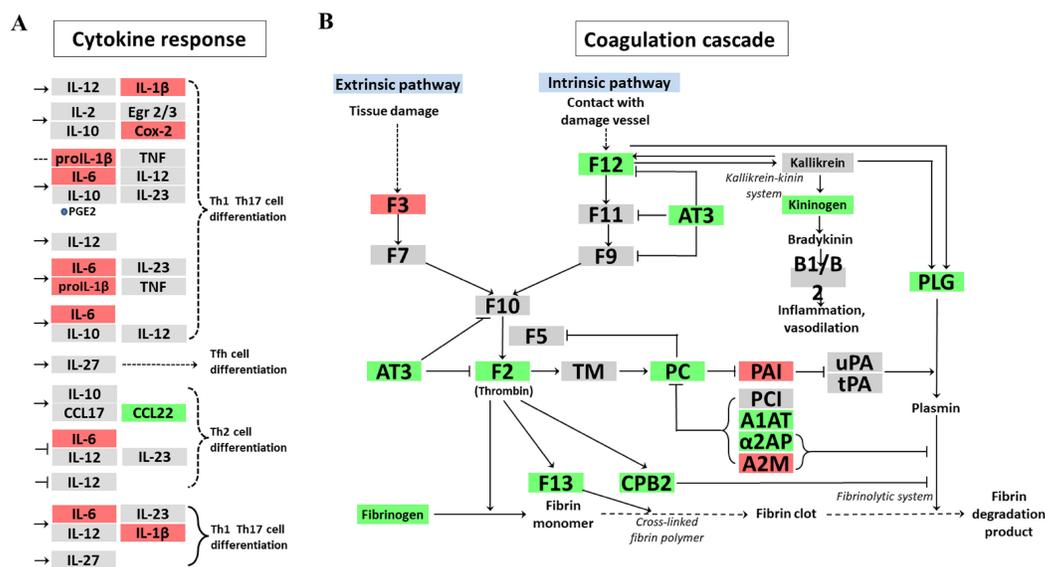

Fig. 5 Changes in gene expression of *B. pseudomallei*-infected mice mapped onto representative KEGG pathways. (A) The mouse cytokine response. (B) The mouse complement and coagulation cascade response.

adapts by up-regulating genes involved in anaerobic metabolism and inorganic ion and nutrient transport [49]. Rienksma et al [39] looked at changes in the transcriptome of intracellular *Mycobacterium bovis* in macrophages and noted that *M. bovis* acquired cholesterol from the host which prompted up-regulation of host *de novo* cholesterol synthesis to compensate for the depletion of intracellular cholesterol levels.

Dual RNA-Seq of an *in vivo* infection provides important insights into host-pathogen interactions but presents a number of experimental challenges. Transcriptome sequencing is performed on total RNA extracted from infected cells or tissues. The RNA sample is a heterogeneous mixture of host and bacterial RNA with host rRNA and mRNA dominating the transcriptional profile. On the other hand, bacterial mRNA is typically a minor fraction of an infected cell or animal tissue, even under optimised *in vitro* conditions, and especially during early infection of the host when bacterial numbers are low. As noted above, removal of rRNA and enrichment of bacterial RNA using kits such as the RiboZero Gold system [50] as well as the separation of infected from non-infected cells will increase the population of bacterial transcripts, nonetheless, these steps require extensive optimisation for different cell and tissue types as well as taking into account the MOI of the pathogen being investigated.

Moving forward, the question on what is next for host-pathogen interaction studies needs to be addressed. It is safe to say the 'omics-platforms such as RNA-Seq and dual RNA-Seq are well established and provide an unbiased insight into all transcripts but does not address the issue of heterogeneity in infections. Single cell RNA and single dual RNA sequencing platforms provide finer resolution of the interactions, which allow analysis of host immune cell composition and their responses to infection, however, the ability of current single-cell RNA-Seq protocols to sample the bacterial transcriptome has not been fully optimised.

Single cell RNA sequencing

The interaction between a pathogen and a host is a highly dynamic process in which both orchestrate intricate gene regulatory pathways. Understanding the pathophysiology of infection is critical to the rational design of prophylactic and therapeutic strategies to tackle infectious diseases. The outcome of infection resulting from the encounter of pathogens and host cells is often measured as population-averaged results, leaving the important host and bacteria cell-to-cell heterogeneity out of the picture. For example, the pathogen *S. typhimurium* infects different cell types, stimulating the host tissue inflammatory response or altering gene expression of macrophages for optimal *S. typhimurium* survival [51]. Generally, during a bacterial infection of a host cell, multiple outcomes are observed including bacterial clearance, bacterial survival and persistence, or host cell death. This implies heterogeneous cellular behavior [51, 52], which makes a single-cell approach crucial for the dissection of the factors contributing to the different infection outcomes [53]. Recent advances in single-cell analysis provide an attractive approach to probe the cellular population diversity and characterise infection pathophysiology at single-cell resolution. For example, single-cell analysis of bacteria infected animal lung tissues was undertaken by first isolating immune and non-immune single cells for massively parallel single-cell RNA sequencing of host and bacterial mRNA. In depth analysis led to the identification of infected cells without the presence of contaminating bystander cells, determination of bacterial loads and gene expression profiles of both host and pathogen [54].

Advances in 'third-generation' sequencing methods have further enhanced multi-transcriptomics approaches. One example is nanopore sequencing, a unique, scalable technology that enables direct, real-time analysis of long DNA or RNA fragments [55]. Improvements in nanopore sequencing potentially enable direct RNA sequencing (i.e. circumventing the cDNA intermediate) and without read length limitation. This allows unequivocal characterisation and quantification of transcript isoforms which provides a true reflection of gene expression. A further advantage is the ability to do gene expression analysis even with low amounts of RNA such as from single cells [56]. Admittedly, a major limitation in single cell RNA-Seq is the dependence on oligo-dT priming, which has restricted the examination to only individual host cells since this approach does not provide information on the state of the individual infecting pathogen. This can be overcome with the use of single-cell dual RNA sequencing (scDual RNA-Seq). Nonetheless, scDual RNA-Seq requires a high MOI that could mask certain infection stages and phenotypes due to a potential variable number of bacteria in the infected host cells [57].

CONCLUSION

The ever increasing new knowledge on molecular changes resulting from the intersection of host, pathogenic bacteria and their local environment has provided a better understanding of infectious disease mechanisms and host antimicrobial strategies. This understanding has been achieved through continued advances in high-throughput transcriptomics platforms and the resulting findings should contribute to the development of novel therapeutics and vaccines.

REFERENCES

1. Achkar JM, Casadevall A (2013) Antibody-mediated immunity against tuberculosis: implications for vaccine development. *Cell Host Microbe* **13**, 250–262.
2. Kodzius R, Kojima M, Nishiyori H, Nakamura M, Fukuda S, Tagami M, Sasaki D, Imamura K, et al (2006) CAGE: cap analysis of gene expression. *Nat Methods* **3**, 211–222.
3. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. *Science* **270**, 484–487.
4. Vos JB, Datson NA, Rabe KF, Hiemstra PS (2007) Exploring host-pathogen interactions at the epithelial surface: Application of transcriptomics in lung biology. *Am J Physiol Lung Cell Mol Physiol* **292**, L367–L377.
5. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.
6. Merrell DS, Butler SM, Qadri F, Dolganov NA, Alam A, Cohen MB, Calderwood SB, Schoolnik GK, et al (2002). Host-induced epidemic spread of the cholera bacterium. *Nature* **417**, 642–645.
7. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, et al (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**, 250–255.
8. Stabler R, Dawson L, Wren B (2010) *Clostridium difficile* using DNA microarrays. *Methods Mol Biol* **646**, 149–162.
9. Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, Loh E, Gripenland J, et al (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**, 950–956.
10. Nicolas P, Mäder U, Dervyn E, Rochat T, Leduc A, Pigeonneau N, Bidnenko E, Marchadier E, et al (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**, 1103–1106.
11. Nathan S, Chieng S, Kingsley PV, Mohan A, Podin Y, Ooi MH, Mariappan V, Vellasamy KM, et al (2018) Melioidosis in Malaysia: Incidence, clinical challenges, and advances in understanding pathogenesis. *Trop Med Infect Dis* **3**, ID 25.
12. Yip CH, Ghazali AK, Nathan S (2020) *Burkholderia pseudomallei* pathogenesis and survival in different niches. *Biochem Soc Trans* **48**, 569–579.
13. Lee SH, Wong RR, Chin CY, Lim TY, Eng SA, Kong C, Ijap NA, Lau MS, et al (2013) *Burkholderia pseudomallei* suppresses *Caenorhabditis elegans* immunity by specific degradation of a GATA transcription factor. *Proc Natl Acad Sci USA* **110**, 15067–15072.
14. Shapira M, Hamlin BJ, Rong J, Chen K, Ronen M, Tan MW (2006) A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc Natl Acad Sci USA* **103**, 14086–14091.
15. Chin CY, Monack DM, Nathan S (2010) Genome wide transcriptome profiling of a murine acute melioidosis model reveals new insights into how *Burkholderia pseudomallei* overcomes host innate immunity. *BMC Genomics* **11**, ID 672.
16. Chieng S, Carreto L, Nathan S (2012) *Burkholderia pseudomallei* transcriptional adaptation in macrophages. *BMC Genomics* **13**, ID 328.
17. Cappelli G, Volpe E, Grassi M, Liseo B, Colizzi V, Mariani F (2006) Profiling of *Mycobacterium tuberculosis* gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function. *Res Microbiol* **157**, 445–455.
18. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* **10**, 57–63.
19. Creecy JP, Conway T (2015) Quantitative bacterial transcriptomics with RNA-Seq. *Curr Opin Microbiol* **23**, 133–140.
20. Nookaew I, Papini M, Pornputtapong N, Scalcinati G, Fagerberg L, Uhlen M, Nielsen J (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: A case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **40**, 10084–10097.
21. Li S, Tighe SW, Nicolet CM, Grove D, Levy S, Farmerie W, Viale A, Wright C, et al (2014) Multi-platform assessment of transcriptome profiling using RNA-Seq in the ABRF next-generation sequencing study. *Nat Biotechnol* **32**, 915–925.
22. Mortazavi A, Williams B, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621–628.
23. Kalam H, Fontana MF, Kumar D (2017) Alternate

splicing of transcripts shape macrophage response to *Mycobacterium tuberculosis* infection. *PLoS Pathog* **13**, e1006236.

24. Trost B, Moir CA, Gillespie ZE, Kusalik A, Mitchell JA, Eskiw CH (2015) Concordance between RNA-Sequencing data and DNA microarray data in transcriptome analysis of proliferative and quiescent fibroblasts. *R Soc Open Sci* **2**, ID 150402.

25. Xu X, Zhang Y, Williams J, Antoniou E, McCombie WR, Wu S, Zhu W, Davidson NO, et al (2013) Parallel comparison of Illumina RNA-Seq and Affymetrix microarray platforms on transcriptomic profiles generated from 5-aza-deoxy-cytidine treated HT-29 colon cancer cells and simulated datasets. *BMC Bioinformatics* **14**, S1.

26. Croucher NJ, Thomson NR (2010) Studying bacterial transcriptome using RNA-Seq. *Curr Opin Microbiol* **13**, 619–624.

27. Srikumar S, Kröger C, Hebrard M, Colgan A, Owen SV, Sivasankaran SK, Cameron ADS, Hokamp K, et al (2015) RNA-Seq brings new insights to the intra-macrophage transcriptome of *Salmonella* Typhimurium. *PLoS Pathog* **11**, e1005262.

28. Mandlik A, Livny J, Robins WP, Ritchie JM, Mekalanos JJ, Waldor MK (2011) RNA-Seq-based monitoring of infection-linked changes in *Vibrio cholerae* gene expression. *Cell Host Microbe* **10**, 165–174.

29. Wu Z, Wu C, Shao J, Zhu Z, Wang W, Zhang W, Tang M, Pei N, et al (2014) The *Streptococcus suis* transcriptional landscape reveals adaptation mechanisms in pig blood and cerebrospinal fluid. *RNA* **20**, 882–898.

30. Conejero L, Potempa K, Graham CM, Spink N, Blankley S, Salguero FJ, Pankla-Sranujit R, Khaenam P, et al (2015) The blood transcriptome of experimental melioidosis reflects disease severity and shows considerable similarity with the human disease. *J Immunol* **195**, 3248–3261.

31. Hirose Y, Yamaguchi M, Okuzaki D, Motooka D, Hamamoto H, Hanada T, Sumitomo T, Nakata M, et al 2019 *Streptococcus pyogenes* transcriptome changes in the inflammatory environment of necrotizing fasciitis. *Appl Environ Microbiol* **85**, e0142819.

32. Yoder-Himes DR, Chain PSG, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, Sorek R (2009) Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing. *Proc Natl Acad Sci USA* **106**, 3976–3981.

33. Kong C, Ghazali AK, Hara Y, Aziz F, Nathan S (2020) Transcriptional landscape of *Burkholderia pseudomallei* cultured under environmental and clinical conditions. (in preparation).

34. Chin CY, Hara Y, Ghazali AK, Yap SJ, Kong C, Wong YC, Rozali N, Koh SF, et al (2015) Global transcriptional analysis of *Burkholderia pseudomallei* high and low biofilm producers reveals insights into biofilm production and virulence. *BMC Genomics* **16**, ID 471.

35. Ferrières L, Clarke DJ (2003) The RcsC sensor kinase is required for normal biofilm formation in *Escherichia coli* K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol Microbiol* **50**, 1665–1682.

36. Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-Seq of pathogen and host. *Nat Rev Microbiol* **10**, 618–630.

37. Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, Assefa SA, He M, et al (2009) A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella Typhi*. *PLoS Genet* **5**, e1000569.

38. Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, Livny J, Earl AM, et al (2012) Efficient and robust RNA-Seq process for cultured bacteria and complex community transcriptomes. *Genome Biol* **13**, R23.

39. Rienksma RA, Suarez-Diez M, Mollenkopf HJ, Dolganov GM, Dorhoi A, Schoolnik GK, Martin dos Santos VAP, Kaufmann SHE, et al (2015) Comprehensive insights into transcriptional adaptation of intracellular mycobacteria by microbe-enriched dual RNA sequencing. *BMC Genomics* **16**, 1–15.

40. Westermann AJ, Forstner KU, Amman F, Barquist L, Chao Y, Schulte LN, Muller L, Reinhardt R, et al (2016) Dual RNA-Seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* **529**, 496–501.

41. Pisu D, Huang L, Grenier JK, Russell (2020) Dual RNA-Seq of Mtb-infected macrophages *in vivo* reveals ontologically distinct host-pathogen interactions. *Cell Rep* **30**, 335–350.

42. Nuss AM, Beckstette M, Pimenova M, Schmuhl C, Opitz W, Pisano F, Heroven AK, Dersch P (2017) Tissue dual RNA-Seq allows fast discovery of infection-specific functions and riboregulators shaping host-pathogen transcriptomes. *Proc Natl Acad Sci USA* **114**, E791–E800.

43. Damron F, Oglesby-Sherrouse A, Wilks A, Barbier M (2016) Dual-seq transcriptomics reveals the battle for iron during *Pseudomonas aeruginosa* acute murine pneumonia. *Sci Rep* **6**, ID 39172.

44. Ritchie ND, Evans TJ (2019) Dual RNAseq in *Streptococcus pneumoniae* infection reveals compartmentalized neutrophil responses in lung and pleural space. *mSystems* **4**, e0021619.

45. Montoya DJ, Andrade P, Silva BJA, Teles RMB, Ma F, Bryson B, Sadanand S, Noel T, et al (2019) Dual RNA-Seq of human leprosy lesions identifies bacterial determinants linked to host immune response. *Cell Rep* **26**, 3574–3585.

46. Peterson EJR, Bailo R, Rothchild AC, Arrieta-Ortiz ML, Kaur A, Pan M, Mai D, Abidi AA, et al (2019) Path-seq identifies an essential mycolate remodeling program for mycobacterial host adaptation. *Mol Syst*

Biol **15**, e8584.

47. Low L, Harrison PF, Gould J, Powell DR, Choo JM, Forster SC, Chapman R, Gearing LJ, et al (2018) Concurrent host-pathogen transcriptional responses in a *Clostridium perfringens* murine myonecrosis infection. *mBio* **9**, e0047318.

48. Ghazali AK, Ng SM, Chua SB, Yip CH, Hoh CC, Nathan S (2020) *Burkholderia pseudomallei* transcriptional profile in different niches to understand pathogenicity. (in preparation).

49. Griesenauer B, Tran TM, Fortney KR, Janowicz DM, Johnson P, Gao H, Barnes S, Wilson LS, et al (2019) Determination of an interaction network between an extracellular bacterial pathogen and the human host. *mBio* **10**, e0119319.

50. Benes V, Blake J, Doyle K (2011) Ribo-Zero Gold Kit: Improved RNA-Seq results after removal of cytoplasmic and mitochondrial ribosomal RNA. *Nat Methods* **8**, iii–iv.

51. Avraham R, Haseley N, Brown D, Penaranda C, Jijon HB, Trombetta JJ, Satija R, Shalek AK, et al (2015) Pathogen cell-to-cell variability drives heterogeneity in host immune responses. *Cell* **162**, 1309–1321.

52. Saliba AE, Li L, Westermann AJ, Appenzeller S, Stapels DA, Schulte LN, Helaine S, Vogel J (2016) Single-cell RNA-Seq ties macrophage polarization to growth rate of intracellular *Salmonella*. *Nat Microbiol* **2**, ID 16206.

53. Buckner MMC, Finlay BB (2011) Host-microbe interaction: Innate immunity cues virulence. *Nature* **472**, 179–180.

54. Lin WN, Tay MZ, Lu R, Liu Y, Chen CH, Cheow LF (2020) The role of single-cell technology in the study and control of infectious diseases. *Cells* **9**, 1440–1469.

55. Loman NJ, Quinlan AR. (2014) Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics* **30**, 3399–3401.

56. Ayub M, Hardwick SW, Luisi BF, Bayley H (2013) Nanopore-based identification of individual nucleotides for direct RNA sequencing. *Nano Lett* **13**, 6144–6150.

57. Avital G, Avraham R, Fan A, Hashimshony T, Hung DT, Yanai I (2017) scDual-Seq: mapping the gene regulatory program of *Salmonella* infection by host and pathogen single-cell RNA-Sequencing. *Genome Biol* **18**, ID 200.