Early Transcriptional Landscapes of *Chlamydia trachomatis*-Infected Epithelial Cells at Single Cell Resolution

Regan J. Hayward[1], James W. Marsh[2], Michael S. Humphrys[3], Wilhelmina M. Huston[4] and Garry S. A. Myers[1,4]*

[1] Faculty of Science, School of Life Sciences, The ithree Institute, University of Technology Sydney, Ultimo, NSW, Australia, [2] Department of Microbiome Science, Max Planck Institute for Developmental Biology, Tübingen, Germany, [3] Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, United States, [4] Faculty of Science, School of Life Sciences, University of Technology Sydney, Ultimo, NSW, Australia

Chlamydia are Gram-negative obligate intracellular bacterial pathogens responsible for a variety of disease in humans and animals worldwide. *Chlamydia trachomatis* causes trachoma in disadvantaged populations, and is the most common bacterial sexually transmitted infection in humans, causing reproductive tract disease. Antibiotic therapy successfully treats diagnosed chlamydial infections, however asymptomatic infections are common. High-throughput transcriptomic approaches have explored chlamydial gene expression and infected host cell gene expression. However, these were performed on large cell populations, averaging gene expression profiles across all cells sampled and potentially obscuring biologically relevant subsets of cells. We generated a pilot dataset, applying single cell RNA-Seq (scRNA-Seq) to *C. trachomatis* infected and mock-infected epithelial cells to assess the utility, pitfalls and challenges of single cell approaches applied to chlamydial biology, and to potentially identify early host cell biomarkers of chlamydial infection. Two hundred sixty-four time-matched *C. trachomatis*-infected and mock-infected HEp-2 cells were collected and subjected to scRNA-Seq. After quality control, 200 cells were retained for analysis. Two distinct clusters distinguished 3-h cells from 6- and 12-h. Pseudotime analysis identified a possible infection-specific cellular trajectory for *Chlamydia*-infected cells, while differential expression analyses found temporal expression of metallothioneins and genes involved with cell cycle regulation, innate immune responses, cytoskeletal components, lipid biosynthesis and cellular stress. We find that changes to the host cell transcriptome at early times of *C. trachomatis* infection are readily discernible by scRNA-Seq, supporting the utility of single cell approaches to identify host cell biomarkers of chlamydial infection, and to further deconvolute the complex host response to infection.

Keywords: *Chlamydia (Chlamydia trachomatis)*, infection, single cell, transcriptomics, bioinformatics

INTRODUCTION

Chlamydia are Gram-negative obligate intracellular bacterial pathogens that cause disease in humans and a wide variety of animals. In humans, *Chlamydia trachomatis* typically infects cells within the ocular and genital mucosa, causing the most prevalent bacterial sexually transmitted infections (STI; Reyburn, 2016), inducing acute and chronic reproductive tract diseases that impact all socioeconomic groups, and trachoma in disadvantaged populations (Burton and Mabey, 2009). Disease outcomes arise from complex inflammatory cascades and immune-mediated host processes that can lead to tissue damage and fibrotic scarring in the upper genital tract or the conjunctiva (Taylor et al., 2014; Menon et al., 2015). Reproductive tract disease outcomes include pelvic inflammatory disease (PID), preterm delivery, ectopic pregnancy, hydrosalpinx, tubal factor infertility (TFI), and chronic pelvic pain in women, as well as epididymitis, testicular pain and infertility in men. Antibiotic therapy with azithromycin or doxycycline successfully treats diagnosed infections, however asymptomatic infections are common (Hafner et al., 2014; Ali et al., 2015). Without overt symptoms that lead individuals to seek primary health care, antibiotic interventions are not able to be employed. Asymptomatic infection rates are estimated to exceed diagnosed infection rates by at least 4.3-fold (Ali et al., 2015).

Chlamydia have a unique biphasic developmental cycle with distinct morphological forms. The cycle begins with attachment and entry of the infectious elementary bodies (EBs) into host cells, typically mucosal epithelial cells. After entry, EBs reside within membrane-bound vacuoles that escape phagolysosomal fusion (Scidmore et al., 1996). Differentiation into the replicating reticulate bodies (RBs) occurs within the first 2–3 h, followed by continued growth of the inclusion accommodating the increased number of RBs. Over the course of infection, *Chlamydia* parasitises and modifies the host cell by deploying type III effectors and other secreted proteins (Valdivia, 2008), which also facilitate invasion, internalization, and replication, while countering host cell defenses (Saka et al., 2011; Bastidas et al., 2013). At the end of the developmental cycle, RBs asynchronously transition back into EBs (∼20–44 h) and, through either extrusion or host cell lysis (∼48–70 h), are released to repeat the cycle (Elwell et al., 2016).

Chlamydial transcriptomes have been examined over the developmental cycle, in EBs and RBs, in different chlamydial species (Belland et al., 2003; Albrecht et al., 2010, 2011; Abdelrahman et al., 2011). Epithelial cell transcriptomes responding to plasmid-bearing/plasmid-less *C. trachomatis* has been characterized by microarray (Porcella et al., 2015). Dual RNA-Seq (Humphrys et al., 2013; Marsh et al., 2018) has allowed the transcriptomes of both *C. trachomatis* and infected epithelial cells to be profiled simultaneously,

identifying previously unrecognized early chlamydial gene expression and complex host cell responses (Humphrys et al., 2013). However, in these studies to date, the derived transcriptional profiles represent averaged gene expression over the population of cells sampled (Hebenstreit, 2012). Subsets of cells with dominant gene expression profiles can skew the analysis (Łabaj et al., 2011), possibly obscuring other potentially important cell subsets and their transcriptional profiles (Saliba et al., 2014; Liu and Trapnell, 2016). By examining the expression profiles of individual cells, single cell RNA sequencing (scRNA-Seq) can minimize these biases, enabling a deeper understanding of population heterogeneity, cell states and interactions, and gene regulation (Kolodziejczyk et al., 2015; Regev et al., 2017). scRNA-Seq and other single cell methods have been instrumental in discovering new cell types (Regev et al., 2017) and advancing the understanding of many disease states (Sandberg, 2013), particularly tumor heterogeneity (Patel et al., 2014; Tirosh et al., 2016), hematopoiesis (Kowalczyk et al., 2015), and embryonic development (Yan et al., 2013). Applications of scRNA-Seq to pathogen-infected cells are more limited so far, but are exemplified by studies that show the heterogeneity of macrophage responses to *Salmonella enterica* serovar Typhimurium infection (Saliba et al., 2016), the high degree of cell-cell transcriptional variation induced by influenza virus infection (Russell et al., 2018), and the characterization of lymph node-derived innate responses to bacterial, helminth, and fungal pathogens (Blecher-Gonen et al., 2019).

Here we explore the application of single cell analysis methodologies to *Chlamydia*-infected cells, with the goals of identifying host cell developmental-stage biomarkers, and to assess the utility of these methodologies for deciphering chlamydial biology in cells and tissues. We generated a pilot scRNA-Seq dataset of time-matched infected and mock-infected HEp-2 epithelial cells *in vitro* encompassing the early chlamydial developmental cycle (3, 6, and 12 h). We show that infection responsive changes to the early host cell transcriptome are readily discernible by scRNA-Seq, supporting the potential for host derived infection biomarkers.

RESULTS

Single Cell Capture, Library Construction, Quality Assessment, and Filtering

Chlamydia-infected (*C. trachomatis* serovar E, MOI∼1) and time-matched mock-infected cells spanning three times post-infection were captured using the Fluidigm C1 microfluidic instrument and workflows (**Figure 1A**). We obtained 80 single cells at 3 h (48 *Chlamydia*-infected, 32 mock-infected), 96 cells at 6 h (48 *Chlamydia*-infected, 48 mock-infected), and 88 cells at 12 h (40 *Chlamydia*-infected, 48 mock-infected) for an initial total of 264 cells (**Figure 1B**). Following Illumina library construction and sequencing, the raw sequencing reads from these 264 cells were demultiplexed using DeML (Renaud et al., 2015), yielding 1.03 billion sequence reads (**Supplemental File 1**). Single cell datasets were removed

---

Abbreviations: DE, Differentially expressed; EBs, Elementary bodies; FDR, False discovery rate; IFC, Integrated fluidic circuit; IF, Intermediate filament; MT, Metallothionein; ncRNAs, Non-coding RNAs; PID, Pelvic inflammatory disease; ROS, Reactive oxygen species; RLE, Relative log expression; RBs, Reticulate bodies; STI, Sexually transmitted infections; scRNA-Seq, Single cell RNA sequencing; TFI, Tubal factor infertility.
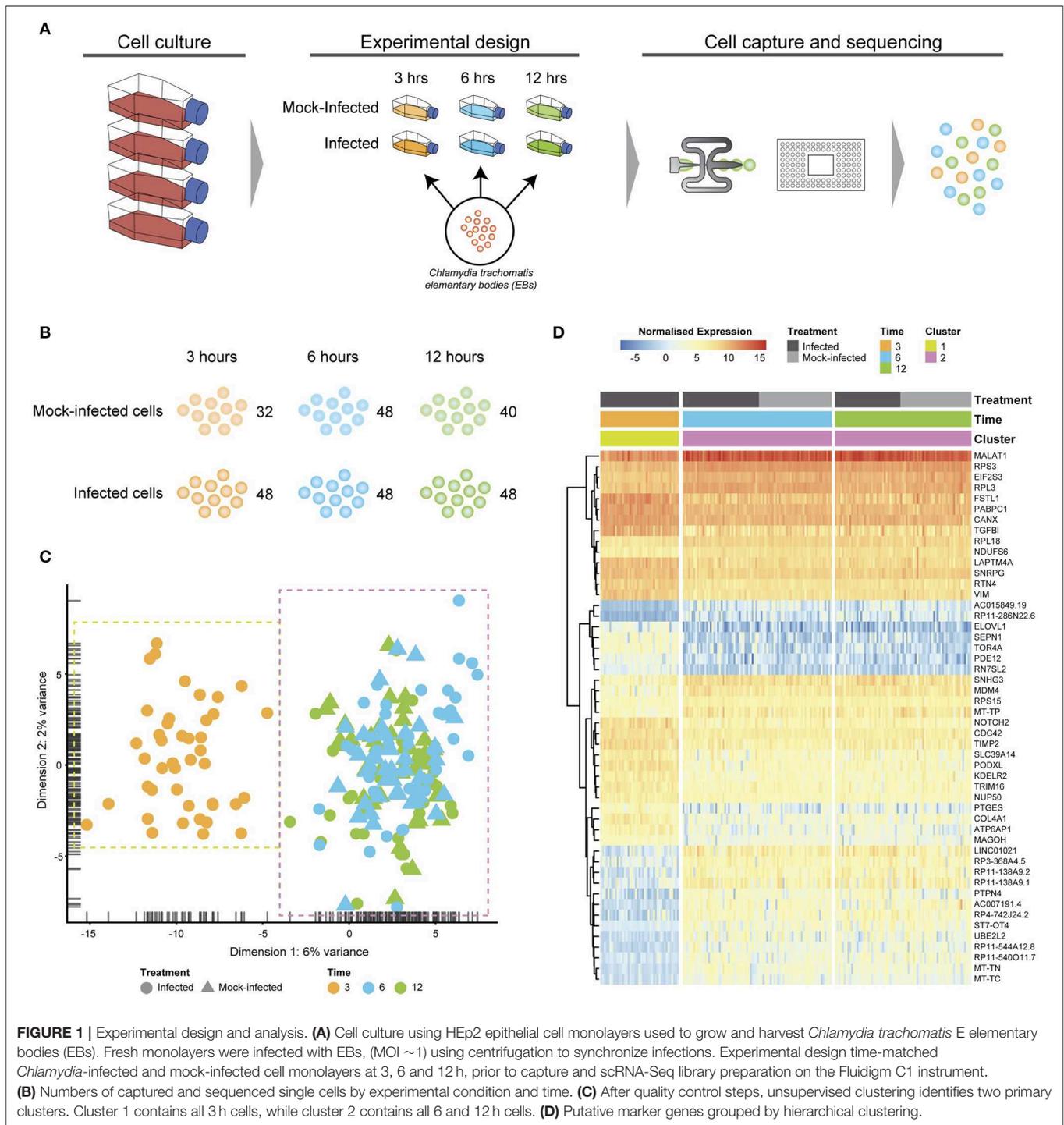
FIGURE 1 | Experimental design and analysis. **(A)** Cell culture using HEp2 epithelial cell monolayers used to grow and harvest *Chlamydia trachomatis* E elementary bodies (EBs). Fresh monolayers were infected with EBs, (MOI ~1) using centrifugation to synchronize infections. Experimental design time-matched *Chlamydia*-infected and mock-infected cell monolayers at 3, 6 and 12 h, prior to capture and scRNA-Seq library preparation on the Fluidigm C1 instrument. **(B)** Numbers of captured and sequenced single cells by experimental condition and time. **(C)** After quality control steps, unsupervised clustering identifies two primary clusters. Cluster 1 contains all 3 h cells, while cluster 2 contains all 6 and 12 h cells. **(D)** Putative marker genes grouped by hierarchical clustering.

from subsequent analyses if they contained <1 million reads after trimming and alignment, and <5,000 counted features (genes). Further quality assessment measures ensured that sequence reads mapped across all chromosomes and that the majority of reads mapped to protein-coding genes (**Supplemental File 2**). Single cell datasets were pooled and

subjected to additional quality assessment steps, including examining rRNA as a measure of depletion success and mitochondrial gene expression as an indicator of cell stress (Zhao et al., 2002), as both are potential sources of bias (**Supplemental File 3**). During quality control, the mock-infected cells at 3 h failed to pass cut-offs, and were excluded

from further downstream analysis (**Supplemental File 3**). After all quality measures, datasets from 200 high quality single cells remained across the three times: 43 *Chlamydia*-infected cells at 3 h; 82.6 h cells (42 *Chlamydia*-infected, 40 mock-infected), and 75.12 h cells (36 *Chlamydia*-infected, 39 mock-infected).

Removal of Confounding Effects

To normalize by library size, Scran's single-cell specific method was used to deconvolute library size factors from cell clusters (Lun et al., 2016). We applied RUVSeq (Risso et al., 2014) to identify and remove further confounding effects, including differences between batches of sequenced cells. Reduction of variation was confirmed in relative log expression (RLE) plots (**Supplemental File 4A**). Density curve plots further show the effect of removing variability from the raw counts, after library size normalization, and after removing further confounding effects (**Supplemental File 4B**). The PCA bi-plot (**Supplemental File 4C**) shows the structure of the data and grouping of the cells based on their transcriptional profiles following these steps. By examining the underlying variables driving PC1 variation, we found that total read counts and time post-infection account for 99% of the total variation (**Supplemental File 4C**), confirming that most variation is not from experimental factors. In addition, doublets (where at least two cells are captured into the same well) can skew the resulting expression profiles, adding a further confounding factor. Although the C1 platform uses integrated fluidic circuits (IFCs) to isolate single cells, it has been associated with a doublet rate as high as 25% (Wang et al., 2019). Due to this high reported rate, we ran different tools to identify doublets, confirming that our data had minimal detected doublets (**Supplemental File 5**).

Cell Cycle Classification

Due to the constraints imposed by chlamydial infection within *in vitro* tissue culture and, given the potential for cell-cell variability despite infection synchronization, we expected to observe a range of cell cycle stages in our data (**Figure 2**). Two of the three stages (G1 and G2/M) show more than double the number of cells from 3 to 6 h, while DNA synthesis (S) is the only cell cycle stage with a decrease in the number of cells from 3 to 12 h. However, despite these trends, no distinct cell cycle clusters are apparent (**Figure 2B**). In addition, there is no clustering between cell cycle state and time post-infection, or infection condition (infected vs. mock-infected). Although we identify cell cycle stage as a likely confounding effect (Barron and Li, 2016) that was removed from our subsequent analyses, it may be relevant to the infection and growth strategies of *Chlamydia*. For example, while infected cells can still grow and divide, the burden of infection causes these cells to proliferate more slowly than uninfected cells, resulting in dividing cells which may be more or less susceptible to infection (Balsara et al., 2006). Additionally, chlamydial infectivity has been related to distinct cell cycle phases, where infection can modulate cell cycle parameters (Johnson et al., 2009).
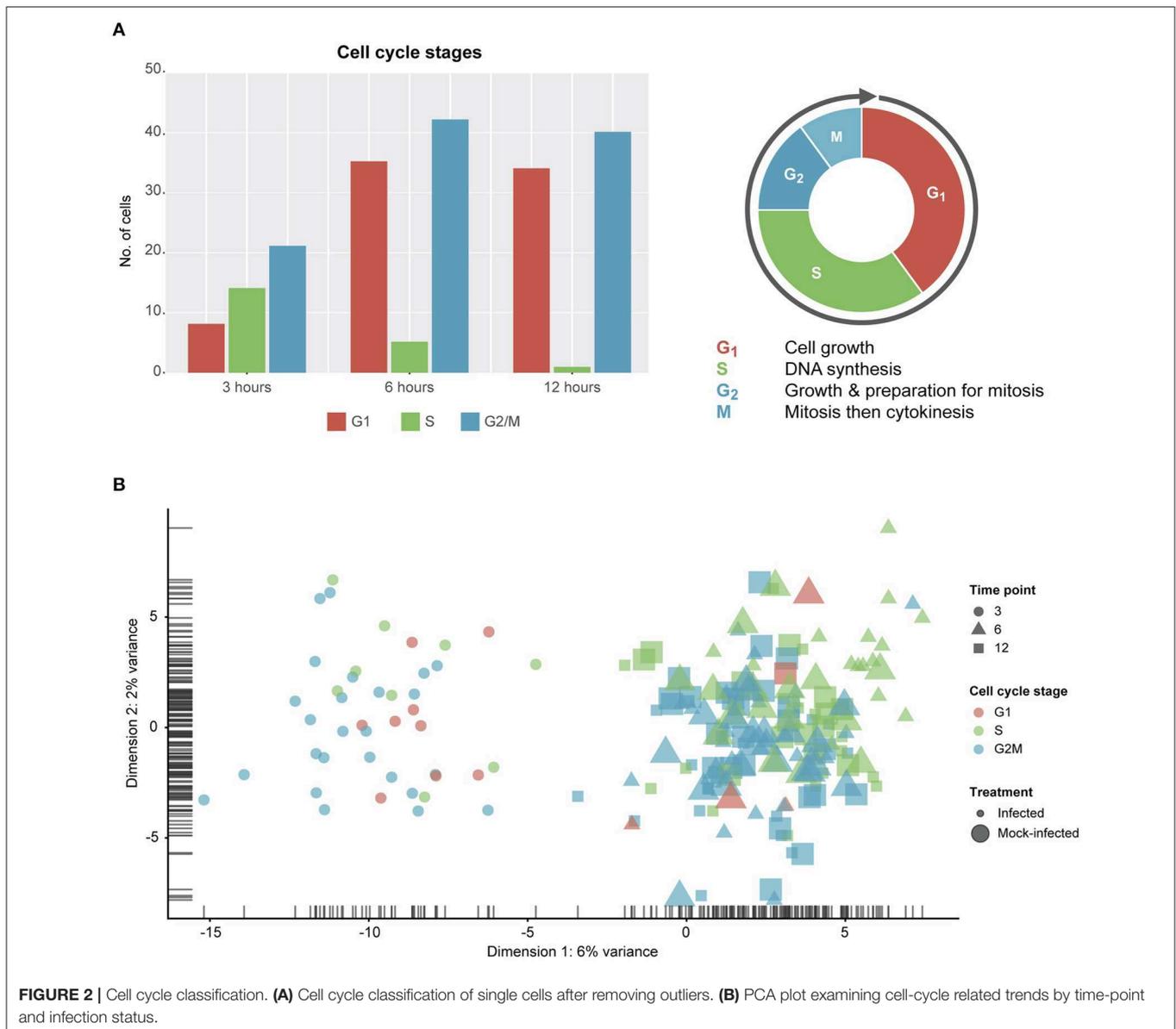
Clustering Demonstrates Transcriptional Heterogeneity of Infected Epithelial Cells Over the Early Chlamydial Developmental Cycle

Unsupervised clustering identified two distinct clusters across the three time points (**Figures 1C**, **3**). Cluster 1 contains only 3-h infected cells, while cluster 2 contains a mixture of cells from 6 and 12 h, with no clear separation between infected and uninfected conditions. We used k-nearest neighbor smoothing (kNN-smoothing) to further reduce scRNA-Seq-specific noise within the expression matrix (Wagner et al., 2018), which is a common occurrence from effects such as dropouts (Gong et al., 2018). The resulting PCA plot recapitulated the clusters identified above, indicating that the previous clustering result was not influenced by noise-related factors. Additional clustering analyses were performed to identify any sub-populations within each cluster on the basis of experimental factors such as time or infection status (**Figure 4**). *Chlamydia*-infected cells again clustered into two main groups, closely matching the overall clustering that separated the 3 h cells from 6 and 12 h, with no further sub-clustering evident.

Pseudotime Analysis Over the Early Chlamydial Developmental Cycle

Unsupervised clustering demonstrates that both infected and uninfected cells have minimal cluster separation at 6 and 12 h (**Figure 1C**). We applied pseudotime analysis to further deconvolute cellular trajectories that may follow a time course or biological mechanisms such as differentiation or infection (Ji and Ji, 2016; Lönnberg et al., 2017). Pseudotime analysis of *Chlamydia*-infected cells alone predicted 3 distinct cell states (**Figure 5A**). Cell state 1 contained 3 h cells, state 2 contained a mixture of 3, 6 and 12 h cells, and state 3 contains a mixture of 6 and 12 h cells (**Figure 5B**). The line connecting the 3 cell states (minimum spanning tree) does not provide a realistic linear trajectory of the infection course from 3 to 12 h. Manually increasing the number of states does provide a more realistic trajectory (**Figure 5C**); however, the similarity of 6 and 12 h cells (both mock-infected and infected states) will require more cells to accurately capture any putative sub-stages of infection.
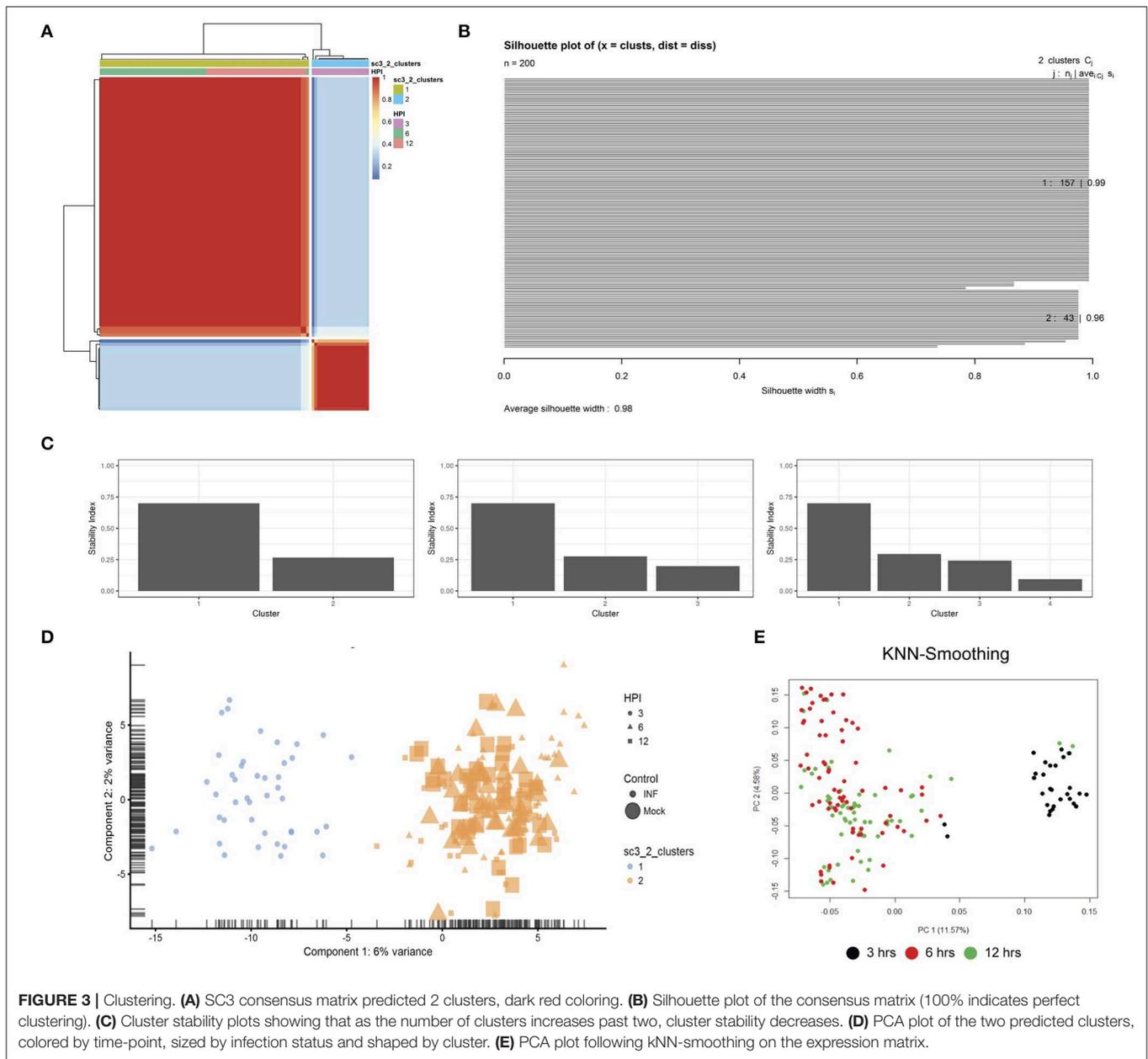
We overlaid the predicted cell cycle states from pseudotime analysis for each cell to identify any shared characteristics of infected cells with mock-infected cells, which could classify cells that were either not infected or had unproductive infections. This analysis identified only two cell states (**Figure 5D**), recapitulating the initial clustering results. By manually increasing the number of cell states to 7, smaller sub-clusters within the two main clusters became evident. However, we still observe a mixture of infected and mock-infected cells within each sub-cluster (albeit with small numbers of cells), further highlighting the transcriptional similarity between infected cells at 6 and 12 h in this dataset.

FIGURE 2 | Cell cycle classification. **(A)** Cell cycle classification of single cells after removing outliers. **(B)** PCA plot examining cell-cycle related trends by time-point and infection status.

Differentially Expressed Genes in *Chlamydia*-Infected and Mock-Infected Cells Highlight Infection Mechanisms

Subsets of genes with significant expression differences between the two primary clusters were examined in order to identify any putative host cell marker genes that distinguish different times post-infection (**Figure 1D**). At 6 and 12 h, genes associated with processes governing RNA and protein metabolism (*RPS3*, *RPL3*, *RPS15*, *RPL18*, *PABPC1*, *MAGOH,* and *EIF2S3*) predominate, with most showing increased expression. Increased expression of vimentin (*VIM*), a type III intermediate filament (IF) present in the cytoskeleton and involved in maintaining cell shape and integrity (Mak and Brüggemann, 2016), distinguishes the 3 h cluster.

We further examined differentially expressed (DE) genes firstly by comparing infected and mock-infected cells at 6 and 12 h, respectively (cluster 2), and secondly by comparing the 3 h infected cells (cluster 1) against cluster 2, as the 3 h mock-infected cells were removed after initial quality control steps. At 6 h, 44 DE genes were identified (13 up-regulated and 21 down-regulated; **Figure 6A**), including three up-regulated metallothionein (MT) genes (*MT1E*, *MT2A,* and *MT1X*). MT up-regulation occurs in response to intracellular zinc concentration increases, reactive oxygen species (ROS) and proinflammatory cytokines (Rice et al., 2016). Intracellular zinc concentrations are an integral component of immunity and inflammation, and zinc deficiency results in an increased susceptibility to infection (Subramanian Vignesh and Deepe, 2017). MTs may also have a

FIGURE 3 | Clustering. **(A)** SC3 consensus matrix predicted 2 clusters, dark red coloring. **(B)** Silhouette plot of the consensus matrix (100% indicates perfect clustering). **(C)** Cluster stability plots showing that as the number of clusters increases past two, cluster stability decreases. **(D)** PCA plot of the two predicted clusters, colored by time-point, sized by infection status and shaped by cluster. **(E)** PCA plot following kNN-smoothing on the expression matrix.

role in protecting against DNA damage and in apoptosis, as well as regulating gene expression during the cell cycle (Cherian and Apostolova, 2000), which are likely to be relevant at 6 h post-infection. Down-regulated pathways at 6 h were dominated by three genes *HSP90AA1* (Heat Shock Protein 90 Alpha Family Class A Member 1), *TUBB* (Tubulin Beta Class I), and *TUBA4A* (Tubulin Alpha 4a), which are linked to the cell cycle, specifically centrosome maturation and microtubule assembly mediating mitosis (**Figure 6B**).

At 12 h, there is an increase in DE genes (245) with 98 up-regulated and 147 down-regulated (**Figure 6A**). We continue to see up-regulated genes that are likely part of a continued immune response to infection, including two MTs (*MT1M* and

MT1E), *TRIM25* (Tripartite Motif Containing 25), *ISG15* (*ISG15* Ubiquitin Like Modifier), *HLA-A* (Major Histocompatibility Complex, Class II, DR Beta 1), *IFIT3* (Interferon Induced Protein With Tetratricopeptide Repeats 3), *OASL* ($2'$-$5'$-Oligoadenylate Synthetase Like), *IL6* (Interleukin 6), and genes associated with cholesterol and fatty acid synthesis (**Figure 6B**). The exploitation of a variety of host lipids by *Chlamydia* to subvert intracellular signaling, survival and growth is well established (Kumar et al., 2006; Cocchiaro et al., 2008; Elwell and Engel, 2012). All down-regulated pathways at 12 h indicate that *Chlamydia*-infected cells are exhibiting stress responses. DNA damage as part of the cell cycle, and repair pathways are enriched, possibly representing a continuation of infection stresses at 6 h, and likely
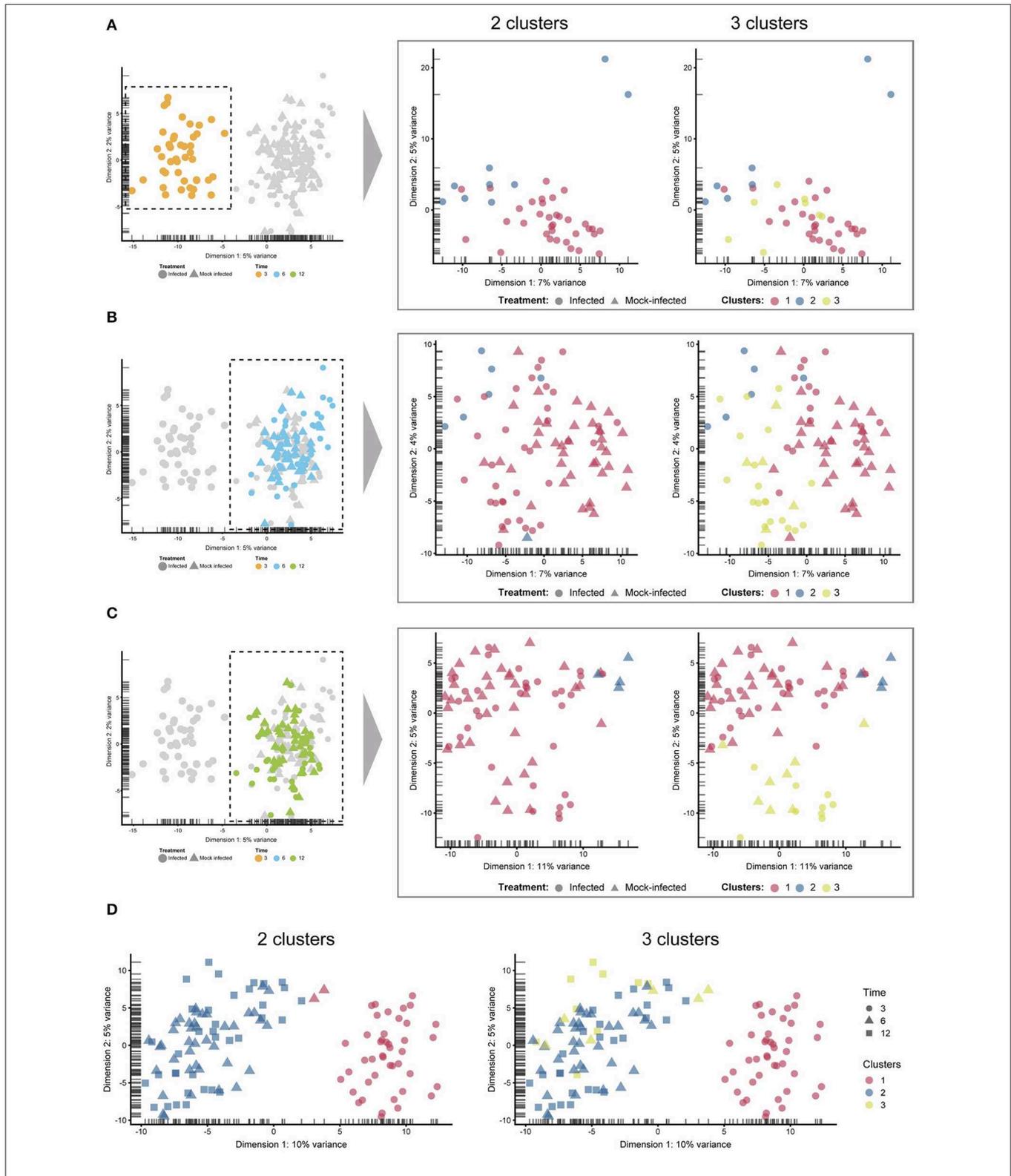
FIGURE 4 | Sub-clustering. The four comparisons shown here were created by manually selecting two and three clusters to examine any sub-clustering events not automatically detected. **(A)** 3 h cells—no sub-clustering evident. **(B)** 6 h cells—no apparent sub-clustering with two clusters; three clusters do display more of a separation (between blue and green), while infected and mock-infected cells cannot be distinguished. **(C)** 12 h cells—some separation evident with 3 clusters, but infection state is not distinguishable. **(D)** Extracting only infected cells show a clear separation of 3 h cells, but not 6 and 12 h cells.
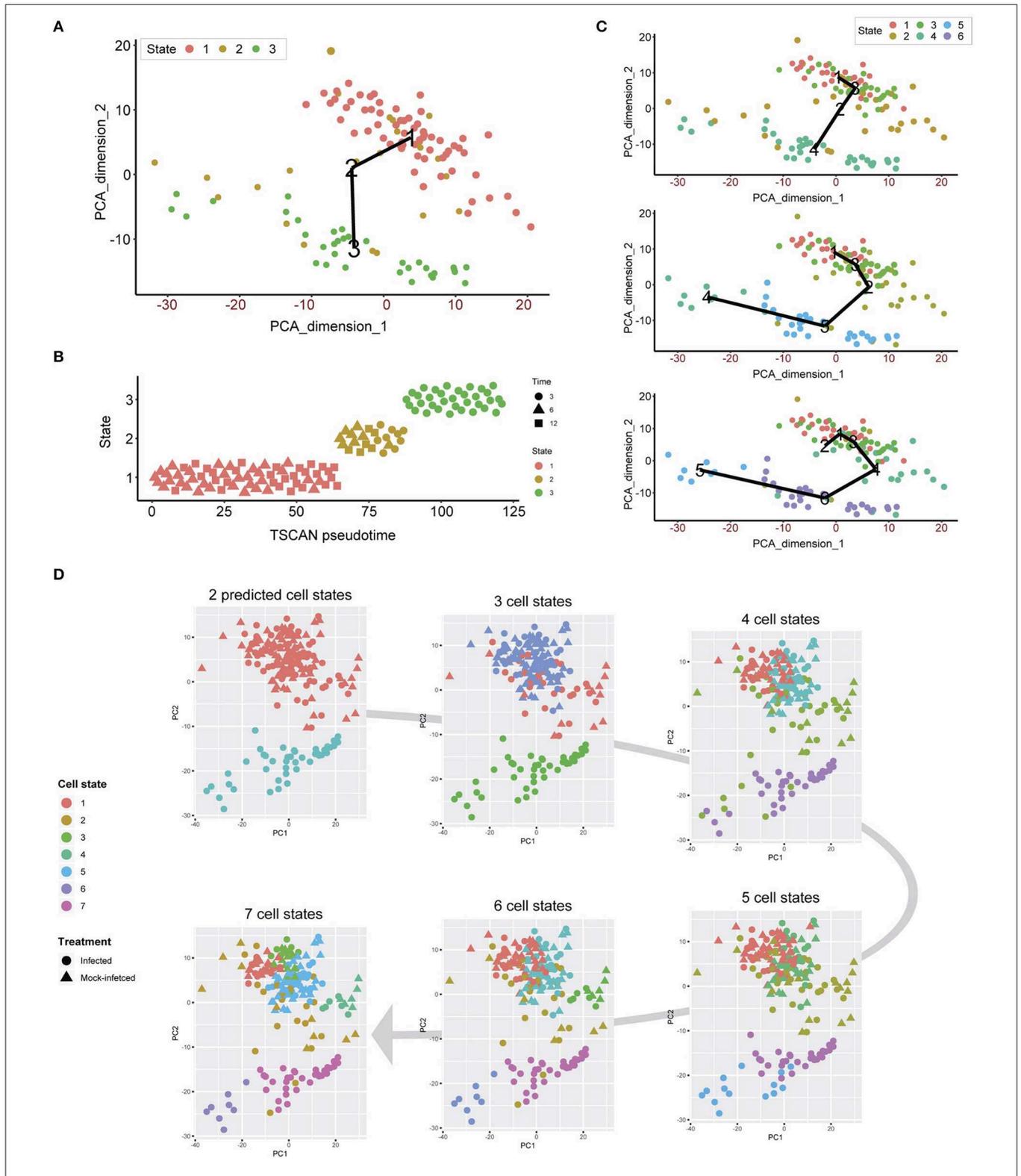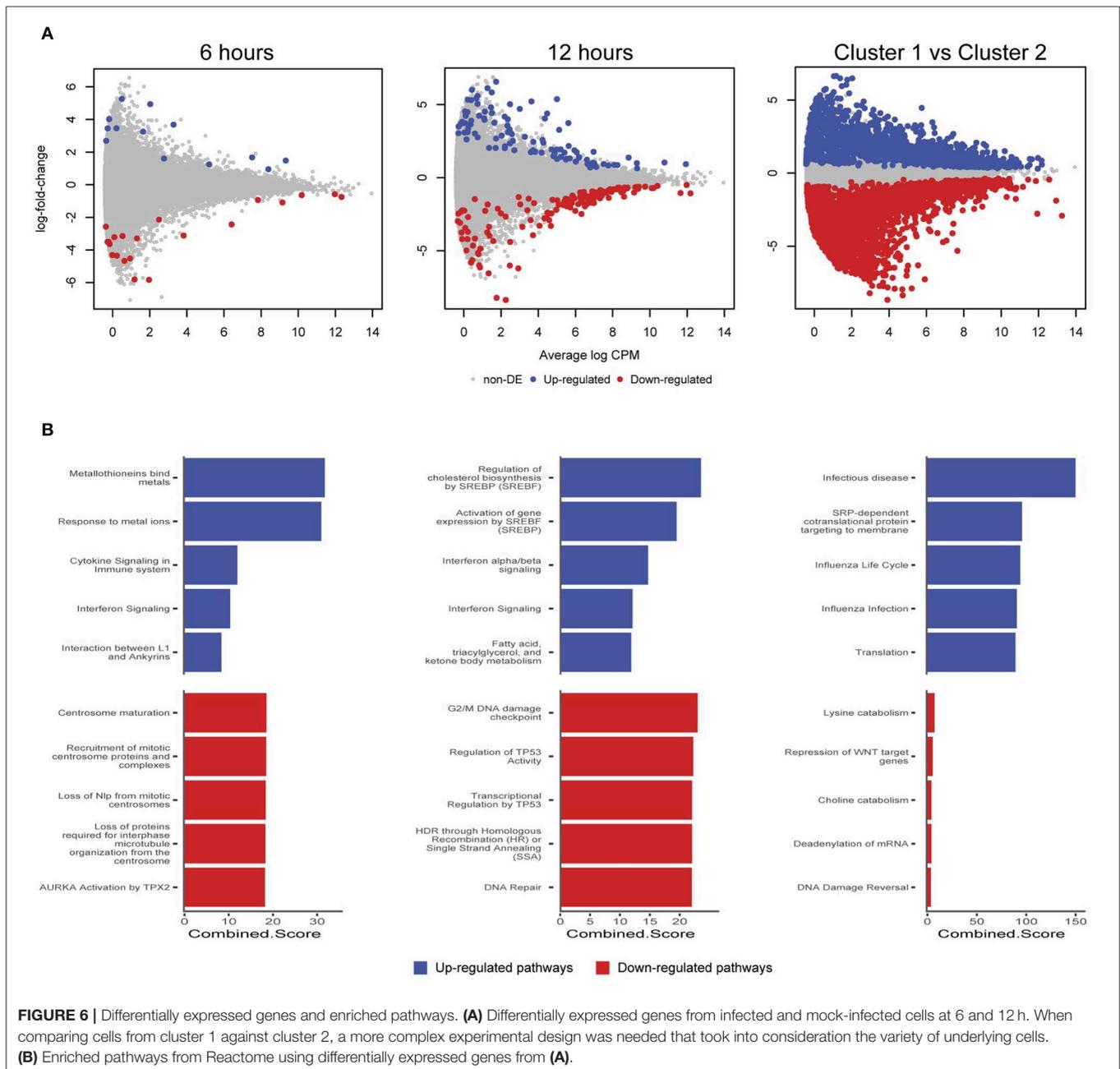
FIGURE 5 | Pseudotime analysis. (A) Pseudotime analysis of infected cells predicts three cell states. The minimum spanning tree (black line) is uninformative and not a true indication of an expected infection trajectory encompassing all cells from 3 to 12 h. (B) Each cell ordered throughout the predicted pseudotime and separated by cell state. (C) Manually increasing the number of cell states to six appears to show a more realistic infection trajectory with a wider number of cells, in addition to showing start and end points correlating to 3 and 12 h cells. (D) When all cells are used, two cell states are predicted that support the initial clustering outcomes. When the number of cell states is manually increased, smaller subsets appear, providing a finer resolution.

indicative of further *Chlamydia*-induced interruption of the cell cycle. Notably, two p53 associated pathways were enriched from associated genes. p53 expression tightly controls the cell cycle and is modulated in response to activities including cell stress, DNA damage, as well as bacterial infection (Zaika et al., 2015). *Chlamydia*-induced down-regulation of p53 may help to protect infected cells against death-inducing host responses, thus allowing chlamydial survival (Siegl et al., 2014). Only four DE genes from 6 and 12 h overlap. The two up-regulated genes were *DUSP5* (Dual Specificity Phosphatase 5) and *MT1E* (metallothionein 1E), while the two down-regulated genes were *TUBA4A* and *HSP90AA1*.

Comparing cluster 1 (3-h infected cells) against cluster 2 (DE genes from 6 and 12 h) demonstrates a substantial number of DE genes (2,291 up and 3,487 down-regulated) (**Figure 6A**). Although the model attempted to account for the loss of the 3 h mock-infected cells, we note a proportion of these DE genes may not be related to a productive chlamydial infection as a result. The down-regulated pathways have low combined scores (a combination of *p*-values and z-scores) compared to the up-regulated pathways, which may be explained by the large number of down-regulated non-coding RNAs (ncRNAs), which are typically not incorporated into the underlying enrichment analyses. Three of the up-regulated pathways are associated with

FIGURE 6 | Differentially expressed genes and enriched pathways. **(A)** Differentially expressed genes from infected and mock-infected cells at 6 and 12 h. When comparing cells from cluster 1 against cluster 2, a more complex experimental design was needed that took into consideration the variety of underlying cells. **(B)** Enriched pathways from Reactome using differentially expressed genes from **(A)**.

infection (*Infectious disease, Influenza life cycle,* and *Influenza infection*; **Figure 6B**), demonstrating that general infection mechanisms are the key differences between these temporally defined clusters.

DISCUSSION

To better understand bacterial pathogenesis and resulting disease outcomes, it is critical to understand functional changes to specific cell populations of infected and neighboring cells, and recruited immune cells in the infected tissue context. This is especially relevant for *Chlamydia* which, due to its obligate intracellular niche and distinct morphologies, has long been refractory to research. As a result, many infection and disease processes at the cellular and tissue level remain largely unknown or poorly characterized *in vivo*. Gene expression profiling of *Chlamydia*-infected cells by microarray (Porcella et al., 2015), dual RNA-Seq (Humphrys et al., 2013; Marsh et al., 2018), or other genome-scale analyses (Hayward et al., 2019) are powerful techniques to help deconvolute these interactions and processes. However, these and similar genome-scale analyses of infected cells have typically been performed on bulk cell populations, i.e., infected cell monolayers *in vitro*, or selectively sorted/purified subsets of cell populations. Such bulk cell approaches can potentially miss cell-cell variability, or cells that contribute to overlapping phenotypic characteristics, potentially masking critical biological heterogeneity as irrelevant signals from non-participating cells that can skew the average. This may influence the understanding of multifactorial and dynamic processes, such as inflammation and fibrosis during ascending chlamydial infection. Single cell approaches can potentially alleviate some of these concerns, but also provide new challenges.

We describe the first application of scRNA-Seq to *Chlamydia*-infected cells. This pilot dataset, comprising 264 single infected and mock-infected cells encompassing three early times of the *in vitro* chlamydial developmental cycle, was designed to examine the feasibility and pitfalls of single cell approaches to investigate chlamydial biology and to ultimately identify host-derived transcriptional biomarkers of chlamydial infection. After quality assessment and filtering measures, we retained 200 high quality, *C. trachomatis*-infected and mock-infected cells.

We note that the experimental design used here will not distinguish *Chlamydia*-mediated effects from infection-specific or non-specific epithelial cell responses. In addition, the *in vitro* infections used as the source of the single cells are centrifugation-synchronized in order to minimize the degree of heterogeneity at each infection time to enable more accurate examination of temporal effects. Despite this, lag time of differentiating EBs to RBs between distinct cells may still influence host responses mediated by temporally expressed/secreted chlamydial factors. Given that the minimal chlamydial generation time during exponential growth has been estimated as 2.6–4.6 h (Wilson et al., 2004; Lambden et al., 2006), it is plausible that cells at each time may cluster with an earlier or later time.

Clustering, pseudotime and cell state prediction analyses demonstrated that *Chlamydia*-treated cells at 3 h are readily distinguishable from *Chlamydia*-treated and mock-infected cells at 6 and 12 h. Curiously, cells at 6 and 12 h clustered together and could not be further deconvoluted from each other, possibly showing that host cell transcription at these times is broadly similar. A recent FAIRE-Seq analysis of *Chlamydia*-infected epithelial cells, examining patterns of host cell chromatin accessibility over the developmental cycle (Hayward et al., 2019), found that 12 h post-infection was relatively quiescent in terms of host cell transcriptional activity. This finding is reflected by our scRNA-Seq analyses here and may be extended to 6 h post-infection. In addition, both *Chlamydia*-infected and mock-infected cells at 6 and 12 h clustered together. One interpretation of this phenomenon is that these early infection times represent a period where the ongoing establishment of the inclusion and chlamydial division after initial entry and infection events is largely cryptic to the host cell as manifested by transcriptional processes.

However, limitations inherent to the experimental design and the technology used here may also influence our results, and should inform future single cell experiments. We used an MOI~1, based upon our previous work with bulk dual RNA-Seq (Humphrys et al., 2013), which typically results in highly infected HEp-2 monolayers (95%+) when using *C. trachomatis* serovar E. When combined with the closed nature of the integrated fluidic circuits used to capture individual cells, the early infection times, and the destructive nature of single cell RNA-Seq, using a lower MOI may have led to populations of both infected and uninfected cells to be sampled. In bulk transcriptomic experiments, any distinct signal from a small number of uninfected cells will be largely overwhelmed. In contrast, uninfected cells in single cell experiments may have an outsized effect, particularly if the total number of cells sampled is insufficient. In addition, the population of infecting EBs may include differentially viable EBs, leading to a divergence of transcriptional profiles between cells with productive infections, compared to cells that are initially infected with non-viable EBs that will not proceed to a productive infection. Similarly, any putative "neighbor" effect of uninfected cells next to infected cells may lead to distinct transcriptional profiles. Given the relatively low number of single cells sampled and the early infection times examined, these factors are a potential source of bias in these pilot experiments. With these limitations in mind, it may be more accurate to describe the *Chlamydia*-infected cells in these experiments as "*Chlamydia*-exposed."

Design of *in vitro* single cell experiments in the *Chlamydia*-infected cell context will benefit firstly from a higher MOI to ensure maximal productive infection. Secondly, collection of much higher numbers of single cells for scRNA-Seq and other single cell genome-scale measurements are now possible, minimizing the potential for introduced sampling biases that arise from low infected cell numbers. Thirdly, the inclusion of additional controls, such as UV-inactivated EBs or opsonized latex beads, will allow host cell transcriptional responses to productive vs. non-productive infection events to be examined, as well as separating host cell infection-specific processes from

non-specific phagocytic responses. Finally, moving away from poly(A) capture library construction to random hexamers instead will allow pathogen transcripts to simultaneously be interrogated in the single cell mode, as recently applied to *Salmonella typhimurium*-infected cells (Avital et al., 2017).

A range of cell cycle states were observed in our data. We attempted to remove these effects as potential confounders through bioinformatic means. *Chlamydia*-infected cells are still able to undergo mitosis, however mitosis-related defects do occur during chlamydial infection. These include an increase in supernumerary centrosomes, abnormal spindle poles, and chromosomal segregation defects, and result in a heavily burdened cell that proliferates more slowly (Grieshaber et al., 2006; Knowlton et al., 2011). Cells that recover from infection are still likely to contain chromosome instabilities, which can then be passed down to uninfected daughter cells (Grieshaber et al., 2006). This may be manifested in our data as we see a number of pathways related to the cell cycle that are down-regulated. While this could be an off-target effect of infection that does not benefit *Chlamydia*, interference with the cell cycle may constitute an infection strategy, as *in vivo* cells will be at different cell cycle stages and thus some may be more or less susceptible to infection. Nevertheless, future *in vitro* investigations of chlamydial infection should attempt to explore and/or mitigate these effects through cell cycle arrest strategies (Johnson et al., 2009) prior to infection and/or single cell separation.

Differential expression comparing cells between clusters identified both conserved and temporally specific gene expression over the times examined. Comparison of these differentially expressed genes with other published *Chlamydia*-infected cell transcriptomic datasets (Humphrys et al., 2013; Porcella et al., 2015) showed little overlap (data not shown), most likely as a consequence of an accumulation of technical differences that make direct comparisons difficult, including the relatively small numbers of single cells sampled here, different times post-infection, different MOIs, and different chlamydial serovars and *in vitro* cell lines. Single cell RNA-Seq approaches may also benefit from parallel bulk RNA-Seq approaches from the same input material to cross-check, compare and validate. Nevertheless, many of the identified pathways and genes are directly relevant to known chlamydial infection processes, including metallothioneins, innate immune processes, cytoskeletal components, lipid biosynthesis, and cellular stress. These analyses demonstrate that, despite the limitations of this pilot dataset, distinct host cell transcriptional responses to infection are readily discernible by single cell approaches, even at the early stages of the chlamydial developmental cycle, yielding robust data and confirming that host cell-derived transcriptional biomarkers of chlamydial infection are identifiable. Thus, single cell genome-scale approaches applied to *Chlamydia*-infected and neighboring cells, recruited immune cells from inflammatory processes, and structural cells obtained from clinical swabs or *ex vivo* tissues, are likely to lend significant insight to the complex processes that underpin chlamydial infection and the associated inflammatory disease outcomes.

METHODS

Cell Culture and Infection

HEp-2 cells (American Type Culture Collection, ATCC No. CCL-23) were grown as monolayers until 90% confluent. Monolayers were infected with *C. trachomatis* serovar E in SPG as previously described (Tan et al., 2009). Additional monolayers were mock-infected with SPG only. The infection was allowed to proceed 48 h prior to EB harvest, as previously described (Tan et al., 2009). *C. trachomatis* EBs and mock-infected cell lysates were subsequently used to infect fresh HEp-2 monolayers. Fresh monolayers were infected with *C. trachomatis* serovar E in 3.5 mL SPG buffer for an MOI $\sim$ 1 as previously described (Tan et al., 2009), using centrifugation to synchronize infections. Infections and subsequent culture were performed in the absence of cycloheximide or DEAE dextran. A matching number of HEp-2 monolayers were also mock-infected and synchronized as above using uninfected cell lysates. Each treatment was incubated at 25°C for 2 h and subsequently washed twice with SPG to remove dead or non-viable EBs. Ten milliliter fresh medium (DMEM + 10% FBS, 25 μg/ml gentamycin, 1.25 μg/ml Fungizone) was added and cell monolayers incubated at 37°C with 5% CO_2. Three biological replicates of infected and mock-infected dishes per time were harvested post-infection into single cell populations by trypsin in sterile PBS prior to immediate single cell capture and library preparation.

Library Preparation and Sequencing

A Fluidigm C1 instrument was used for cell capture. This instrument uses microfluidics on IFCs to capture single cells, lyse and prepare cDNA, using 96 well-plates as input. Only polyadenylated fragments are captured from each cell, typically restricting analysis to eukaryotic mRNA. Cell lysis, reverse transcription, and cDNA amplification were performed using the C1 Single-Cell Auto Prep IFC. The SMARTer Ultra Low RNA Kit (Clontech) was used for cDNA synthesis and Illumina NGS libraries were constructed using the Nextera XT DNA Sample Prep kit. The resulting 264 single cell libraries were sequenced using the Illumina HiSeq 4000 platform (150 bp paired-end reads) across three batches. Each plate was designed with a balanced distribution of time points and conditions.

Pre-processing and Quality Control

Raw sequencing reads were demultiplexed using DeML (Renaud et al., 2015) with default settings. Trim Galore (v.0.4.3) (Krueger, 2012) was used to trim adaptors and low quality reads. Confirmation of the removal of adaptors, low quality reads and other quality control measurements was performed with FastQC (v.0.11.5) (Andrews, 2010). Reads were subsequently aligned to the human genome version (GRCh 38.87) with STAR (v.2.5.1a) (Dobin et al., 2013) retaining paired and unpaired mapped reads that were merged in to a single BAM file.

FastQ-Screen (v.0.11.1) (Wingett, 2011) was used to screen for sources of contamination across all cells. This output and low mapping rates confirmed the removal of all 3-h uninfected cells, due to extremely low mapping rates to the Human genome and high mapping rates to other organisms. Features of the

remaining cells were counted with FeatureCounts (v.1.5.0-p3) (Liao et al., 2014). MultiQC (v.1.0) (Ewels et al., 2016) was used throughout the previous steps, combining output from each piece of software to easily make comparisons between batches and across time points.

Identifying Outlier Cells Based on Filtering

Counted features were imported into Scater (v.1.5.11) (McCarthy et al., 2017), where subsequent quality control reduced the total number of cells from 264 to 200. The filter settings were comprised of four steps: (1) total mapped reads should be >1,000,000; (2) total features >5,000; (3) expression from mitochondrial genes <20% of total expression; and (4) expression from rRNA genes comprise <10%.

Removing Confounding Effects

Cell cycle classification was performed using Cyclone (Lun et al., 2016) prior to filtering out low abundance genes, as recommended. To account for the differences in library sizes between cells, the deconvolution method from Scran (v.1.6.0) (Lun et al., 2016) was used. Further confounding effects such as cell cycle and sequencing batch effects were removed using the RUVs method of RUVSeq (v.1.12.0) (Risso et al., 2014) using $k = 4$. Doublet detection was performed using Scrublet (v.0.1) (Wolock et al., 2019) and DoubletDetection (v.2.4) (Gayoso and Shor, 2019).

Clustering

Unsupervised clustering was performed using the Single-Cell Consensus Clustering (SC3) package (v.1.10.1) (Kiselev et al., 2017). Two clusters (k) were chosen based on automatic prediction by SC3 after iterating through a range of k (2:10). Higher values of K were examined; however, two clusters remained the best fit to this data as assessed by various internal plots including consensus matrices, silhouette plots and cluster stability plots. To further confirm the two clusters, the KNN-Smoothing (Wagner et al., 2018) (K-nearest neighbor smoothing) function (v.1) was also applied to different transformations of the library normalized data.

Pseudotime Analysis

TSCAN (v.1.16.0) (Ji and Ji, 2016) was used to perform pseudotime analysis. When all cells were analyzed, the "pseudocount" and "minexpr_value" flags were set to 0.5 in pre-processing to allow more features to be selected, resulting in an increase of cells with assigned cell states, especially when the number of states was manually increased. The default pre-processing settings were used to examine infected cells alone.

Differential Expression

Most scRNA-Seq differential expression software only allows for direct comparisons, such as cluster comparisons. As our experimental design examined both infected and mock-infected cells and, due to the loss of the 3 h mock cells following QC measures, we used edgeR (v.3.24.3) (Robinson et al., 2010), as it provides better functionality for more complex comparisons than most single-cell specific tools. Initial comparisons were between infected and mock infected cells at 6 and 12 h. The

use of edgeR allowed the comparison between 3 h infected cells (cluster 1) and the remaining cells (cluster 2), taking into consideration the differences between infected and mock infected cells from 6 and 12 h. In addition to including the RUVSeq factors of unwanted variation to the model matrix, the dispersion trend was estimated using "locfit" and "mixed.df" flags set to true. Resulting *p*-values were adjusted using a false discovery rate (FDR) < 0.05 and separated based on their respective fold-changes. Significant genes were examined using enrichR (Chen et al., 2013), with enriched pathways from Reactome sorted by their combined scores (a calculation of *p*-values and z-scores) (Croft et al., 2011).

DATA AVAILABILITY STATEMENT

The data set supporting the results of this article is available in the GEO repository, GSE132525.

AUTHOR CONTRIBUTIONS

RH analyzed, interpreted, and co-wrote the manuscript. JM assisted with analysis and interpretation. MH performed the chlamydial infections and scRNA-Seq laboratory methods. WH assisted with interpretation of the data and contributed to the manuscript. GM conceived the experiments, obtained the funding, oversaw the sequencing, data analysis and interpretation, and co-wrote the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2019.00392/full#supplementary-material

Supplemental File 1 | Demuxing results comparing Illumina and DeML. Demuxing comparison between the standard Illumina software and DeML. 1.8–5.5% additional reads were recovered over each cell batch **(A–C)**.

Supplemental File 2 | Biotype distribution. **(A)** Distribution of different gene biotypes by time. **(B)** Total gene expression by biotype. **(C)** Total gene expression by chromosome location.

Supplemental File 3 | Contamination of 3-h mock-infected cells. Examples of two 3-h mock-infected cells with unusual mapping to 12 different genomes,

indicating cross-contamination. All 3 h mock-infected cells showed similar profiles and were removed from further downstream analysis.

Supplemental File 4 | Identifying and removing confounding effects. **(A)** Relative log expression (RLE) plot of gene expression from all cells after removal of confounding effects using RUVSeq. **(B)** Density curve distribution showing variability associated to key variables from raw counts, after library normalization, and after using RUVSeq respectively. **(C)** PCA plot demonstrating

two-dimensional structure of cell expression profiles after removing confounding effects. Two variables (Total features and Time-point) account for 99% of the variability at component 1 (PC1).

Supplemental File 5 | Doublet detection. **(A)** Scrublet identifies two cells as possible doublets, red bars. **(B)** DoubletDetector found no cells exhibited doublet characteristics.

REFERENCES

Abdelrahman, Y. M., Rose, L. A., and Belland, R. J. (2011). Developmental expression of non-coding RNAs in *Chlamydia trachomatis* during normal and persistent growth. *Nucleic Acids Res.* 39, 1843–1854. doi: 10.1093/nar/gkq1065

Albrecht, M., Sharma, C. M., Dittrich, M. T., Muller, T., Reinhardt, R., Vogel, J., et al. (2011). The transcriptional landscape of *Chlamydia pneumoniae*. *Genome Biol.* 12:R98. doi: 10.1186/gb-2011-12-10-r98

Albrecht, M., Sharma, C. M., Reinhardt, R., Vogel, J., and Rudel, T. (2010). Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Res.* 38, 868–877. doi: 10.1093/nar/gkp1032

Ali, H., Cameron, E., Drovandi, C. C., McCaw, J. M., Guy, R. J., Middleton, M., et al. (2015). A new approach to estimating trends in chlamydia incidence. *Sex. Transm. Infect.* 91, 513–519. doi: 10.1136/sextrans-2014-051631

Andrews, S. (2010). *FastQC A Quality Control tool for High Throughput Sequence Data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed September 30, 2019).

Avital, G., Avraham, R., Fan, A., Hashimshony, T., Hung, D. T., and Yanai, I. (2017). scDual-Seq: mapping the gene regulatory program of Salmonella infection by host and pathogen single-cell RNA-sequencing. *Genome Biol.* 18:200. doi: 10.1186/s13059-017-1340-x

Balsara, Z. R., Misaghi, S., Lafave, J. N., and Starnbach, M. N. (2006). *Chlamydia trachomatis* infection induces cleavage of the mitotic cyclin B1. *Infect. Immun.* 74, 5602–5608. doi: 10.1128/IAI.00266-06

Barron, M., and Li, J. (2016). Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. *Sci. Rep.* 6:33892. doi: 10.1038/srep33892

Bastidas, R. J., Elwell, C. A., Engel, J. N., and Valdivia, R. H. (2013). Chlamydial intracellular survival strategies. *Cold Spring Harb. Perspect. Med.* 3:a010256. doi: 10.1101/cshperspect.a010256

Belland, R. J., Zhong, G., Crane, D. D., Hogan, D., Sturdevant, D., Sharma, J., et al. (2003). Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8478–8483. doi: 10.1073/pnas.1331135100

Blecher-Gonen, R., Bost, P., Hilligan, K. L., David, E., Salame, T. M., Roussel, E., et al. (2019). Single-cell analysis of diverse pathogen responses defines a molecular roadmap for generating antigen-specific immunity. *Cell Syst.* 8, 109–121.e6. doi: 10.1016/j.cels.2019.01.001

Burton, M. J., and Mabey, D. C. W. (2009). The global burden of trachoma: a review. *PLoS Negl. Trop. Dis.* 3:e460. doi: 10.1371/journal.pntd.0000460

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* 14:128. doi: 10.1186/1471-2105-14-128

Cherian, M. G., and Apostolova, M. D. (2000). Nuclear localization of metallothionein during cell proliferation and differentiation. *Cell. Mol. Biol.* 46, 347–356.

Cocchiaro, J. L., Kumar, Y., Fischer, E. R., Hackstadt, T., and Valdivia, R. H. (2008). Cytoplasmic lipid droplets are translocated into the lumen of the *Chlamydia trachomatis* parasitophorous vacuole. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9379 LP−9384. doi: 10.1073/pnas.0712241105

Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., et al. (2011). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39, D691–D697. doi: 10.1093/nar/gkq1018

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Elwell, C., Mirrashidi, K., and Engel, J. (2016). Chlamydia cell biology and pathogenesis. *Nat. Rev. Microbiol.* 14, 385–400. doi: 10.1038/nrmicro2016.30

Elwell, C. A., and Engel, J. N. (2012). Lipid acquisition by intracellular *Chlamydiae*. *Cell. Microbiol.* 14, 1010–1018. doi: 10.1111/j.1462-5822.2012.01794.x

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Gayoso, A., and Shor, J. (2019). *GitHub: DoubletDetection*. Available online at: https://github.com/JonathanShor/DoubletDetection (accessed October 20, 2019).

Gong, W., Kwak, I.-Y., Pota, P., Koyano-Nakagawa, N., and Garry, D. J. (2018). DrImpute: imputing dropout events in single cell RNA sequencing data. *BMC Bioinform.* 19:220. doi: 10.1186/s12859-018-2226-y

Grieshaber, S. S., Grieshaber, N. A., Miller, N., and Hackstadt, T. (2006). *Chlamydia trachomatis* causes centrosomal defects resulting in chromosomal segregation abnormalities. *Traffic* 7, 940–949. doi: 10.1111/j.1600-0854.2006.00439.x

Hafner, L. M., Wilson, D. P., and Timms, P. (2014). Development status and future prospects for a vaccine against *Chlamydia trachomatis* infection. *Vaccine* 32, 1563–1571. doi: 10.1016/j.vaccine.2013.08.020

Hayward, R. J., Marsh, J. W., Humphrys, M. S., Huston, W. M., and Myers, G. S. A. (2019). Chromatin accessibility dynamics of *Chlamydia*-infected epithelial cells. *bioRxiv:681999 [Preprint]*. doi: 10.1101/681999

Hebenstreit, D. (2012). Methods, challenges and potentials of single cell RNA-seq. *Biology* 1, 658–667. doi: 10.3390/biology1030658

Humphrys, M. S., Creasy, T., Sun, Y., Shetty, A. C., Chibucos, M. C., Drabek, E. F., et al. (2013). Simultaneous transcriptional profiling of bacteria and their host cells. *PLoS ONE* 8:e80597. doi: 10.1371/journal.pone.0080597

Ji, Z., and Ji, H. (2016). TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. *Nucleic Acids Res.* 44:e117. doi: 10.1093/nar/gkw430

Johnson, K. A., Tan, M., and Sütterlin, C. (2009). Centrosome abnormalities during a *Chlamydia trachomatis* infection are caused by dysregulation of the normal duplication pathway. *Cell Microbiol.* 11, 1064–1073. doi: 10.1111/j.1462-5822.2009.01307.x

Kiselev, V. Y., Kirschner, K., Schaub, M. T., Andrews, T., Yiu, A., Chandra, T., et al. (2017). SC3: consensus clustering of single-cell RNA-seq data. *Nat. Methods* 14:483. doi: 10.1038/nmeth.4236

Knowlton, A. E., Brown, H. M., Richards, T. S., Andreolas, L. A., Patel, R. K., and Grieshaber, S. S. (2011). *Chlamydia trachomatis* infection causes mitotic spindle pole defects independently from its effects on centrosome amplification. *Traffic* 12, 854–866. doi: 10.1111/j.1600-0854.2011.01204.x

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., and Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Mol. Cell* 58, 610–620. doi: 10.1016/j.molcel.2015.04.005

Kowalczyk, M. S., Tirosh, I., Heckl, D., Rao, T. N., Dixit, A., Haas, B. J., et al. (2015). Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* 25, 1860–1872. doi: 10.1101/gr.192237.115

Krueger, F. (2012). *Trim Galore*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/; http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ (accessed September 20, 2019).

Kumar, Y., Cocchiaro, J., and Valdivia, R. H. (2006). The obligate intracellular pathogen *Chlamydia trachomatis* targets host lipid droplets. *Curr. Biol.* 16, 1646–1651. doi: 10.1016/j.cub.2006.06.060

Łabaj, P. P., Leparc, G. G., Linggi, B. E., Markillie, L. M., Wiley, H. S., and Kreil, D. P. (2011). Characterization and improvement of RNA-Seq precision in quantitative transcript expression profiling. *Bioinformatics* 27, i383–i391. doi: 10.1093/bioinformatics/btr247

Lambden, P. R., Pickett, M. A., and Clarke, I. N. (2006). The effect of penicillin on *Chlamydia trachomatis* DNA replication. *Microbiology* 152, 2573–2578. doi: 10.1099/mic.0.29032-0

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Liu, S., and Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Res.* 5:F1000 Faculty Rev-182. doi: 10.12688/f1000research.7223.1

Lönnberg, T., Svensson, V., James, K. R., Fernandez-Ruiz, D., Sebina, I., Montandon, R., et al. (2017). Single-cell RNA-seq and computational analysis using temporal mixture modeling resolves Th1/Tfh fate bifurcation in Malaria. *Sci. Immunol.* 2:eaal2192. doi: 10.1126/sciimmunol.aal2192

Lun, A. T. L., McCarthy, D. J., and Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with bioconductor. *F1000Res.* 5:2122. doi: 10.12688/f1000research.9501.2

Mak, T. N., and Brüggemann, H. (2016). Vimentin in bacterial infections. *Cells* 5:E18. doi: 10.3390/cells5020018

Marsh, J. W., Hayward, R. J., Shetty, A. C., Mahurkar, A., Humphrys, M. S., and Myers, G. S. A. (2018). Bioinformatic analysis of bacteria and host cell dual RNA-sequencing experiments. *Brief. Bioinform.* 19, 1115–1129. doi: 10.1093/bib/bbx043

McCarthy, D. J., Campbell, K. R., Lun, A. T. L., and Wills, Q. F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33, 1179–1186. doi: 10.1093/bioinformatics/btw777

Menon, S., Timms, P., Allan, J. A., Alexander, K., Rombauts, L., Horner, P., et al. (2015). Human and pathogen factors associated with *Chlamydia trachomatis*-related infertility in women. *Clin. Microbiol. Rev.* 28, 969–985. doi: 10.1128/CMR.00035-15

Patel, A. P., Tirosh, I., Trombetta, J. J., Shalek, A. K., Gillespie, S. M., Wakimoto, H., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344, 1396–1401. doi: 10.1126/science.1254257

Porcella, S. F., Carlson, J. H., Sturdevant, D. E., Sturdevant, G. L., Kanakabandi, K., Virtaneva, K., et al. (2015). Transcriptional profiling of human epithelial cells infected with plasmid-bearing and plasmid-deficient *Chlamydia trachomatis*. *Infect. Immun.* 83, 534–543. doi: 10.1128/IAI.02764-14

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., et al. (2017). The human cell atlas. *Elife* 6:e27041. doi: 10.7554/eLife.27041

Renaud, G., Stenzel, U., Maricic, T., Wiebe, V., and Kelso, J. (2015). deML: robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* 31, 770–772. doi: 10.1093/bioinformatics/btu719

Reyburn, H. (2016). WHO guidelines for the treatment of *Chlamydia trachomatis*. *WHO* 340:c2637. doi: 10.1136/bmj.c2637

Rice, J. M., Zweifach, A., and Lynes, M. A. (2016). Metallothionein regulates intracellular zinc signaling during CD4(+) T cell activation. *BMC Immunol.* 17:13. doi: 10.1186/s12865-016-0151-2

Risso, D., Ngai, J., Speed, T. P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* 32, 896–902. doi: 10.1038/nbt.2931

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Russell, A. B., Trapnell, C., and Bloom, J. D. (2018). Extreme heterogeneity of influenza virus infection in single cells. *Elife* 7:e32303. doi: 10.7554/eLife.32303

Saka, H. A., Thompson, J. W., Chen, Y.-S., Kumar, Y., Dubois, L. G., Moseley, M. A., et al. (2011). Quantitative proteomics reveals metabolic and pathogenic properties of *Chlamydia trachomatis* developmental forms. *Mol. Microbiol.* 82, 1185–1203. doi: 10.1111/j.1365-2958.2011.07877.x

Saliba, A.-E., Westermann, A. J., Gorski, S. A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42, 8845–8860. doi: 10.1093/nar/gku555

Saliba, A. E., Li, L., Westermann, A. J., Appenzeller, S., Stapels, D. A. C., Schulte, L. N., et al. (2016). Single-cell RNA-seq ties macrophagepolarization

to growth rate of intracellular Salmonella. *Nat. Microbiol.* 2, 1–8. doi: 10.1038/nmicrobiol.2016.206

Sandberg, R. (2013). Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11, 22–24. doi: 10.1038/nmeth.2764

Scidmore, M. A., Rockey, D. D., Fischer, E. R., Heinzen, R. A., and Hackstadt, T. (1996). Vesicular interactions of the *Chlamydia trachomatis* inclusion are determined by chlamydial early protein synthesis rather than route of entry. *Infect. Immun.* 64, 5366–5372.

Siegl, C., Prusty, B. K., Karunakaran, K., Wischhusen, J., and Rudel, T. (2014). Tumor suppressor p53 alters host cell metabolism to limit *Chlamydia trachomatis* infection. *Cell Rep.* 9, 918–929. doi: 10.1016/j.celrep.2014.10.004

Subramanian Vignesh, K., and Deepe, G. S. J. (2017). Metallothioneins: emerging modulators in immunity and infection. *Int. J. Mol. Sci.* 18:e2197. doi: 10.3390/ijms18102197

Tan, C., Hsia, R., Shou, H., Haggerty, C. L., Ness, R. B., Gaydos, C. A., et al. (2009). *Chlamydia trachomatis*-infected patients display variable antibody profiles against the nine-member polymorphic membrane protein family. *Infect. Immun.* 77, 3218 LP–3226 LP. doi: 10.1128/IAI.01566-08

Taylor, H. R., Burton, M. J., Haddad, D., West, S., and Wright, H. (2014). Trachoma. *Lancet* 384, 2142–2152. doi: 10.1016/S0140-6736(13)62182-0

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H. II., Treacy, D., Trombetta, J. J., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189–196. doi: 10.1126/science.aad0501

Valdivia, R. H. (2008). Chlamydia effector proteins and new insights into chlamydial cellular microbiology. *Curr. Opin. Microbiol.* 11, 53–59. doi: 10.1016/j.mib.2008.01.003

Wagner, F., Yan, Y., and Yanai, I. (2018). K-nearest neighbor smoothing for high-throughput single-cell RNA-Seq data. *bioRxiv:217737*. doi: 10.1101/217737

Wang, Y. J., Schug, J., Lin, J., Wang, Z., Kossenkov, A., and Kaestner, K. H. (2019). Comparative analysis of commercially available single-cell RNA sequencing platforms for their performance in complex human tissues. *bioRxiv:541433*. doi: 10.1101/541433

Wilson, D. P., Mathews, S., Wan, C., Pettitt, A. N., and McElwain, D. L. S. (2004). Use of a quantitative gene expression assay based on micro-array techniques and a mathematical model for the investigation of chlamydial generation time. *Bull. Math. Biol.* 66, 523–537. doi: 10.1016/j.bulm.2003.09.001

Wingett, S. W. (2011). *FastQ Screen: NGS Reads Quality Control*. Available online at: https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/ (accessed September 13, 2019).

Wolock, S. L., Lopez, R., and Klein, A. M. (2019). Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* 8, 281–291.e9. doi: 10.1016/j.cels.2018.11.005

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1131–1139. doi: 10.1038/nsmb.2660

Zaika, A. I., Wei, J., Noto, J. M., and Peek, R. M. (2015). Microbial regulation of p53 tumor suppressor. *PLoS Pathog.* 11:e1005099. doi: 10.1371/journal.ppat.1005099

Zhao, Q., Wang, J., Levichkin, I. V., Stasinopoulos, S., Ryan, M. T., and Hoogenraad, N. J. (2002). A mitochondrial specific stress response in mammalian cells. *EMBO J.* 21, 4411–4419. doi: 10.1093/emboj/cdf445