UC Riverside
UC Riverside Previously Published Works

Title

Expression activation of over 70% of Chlamydia trachomatis genes during the first hour of infection.

Permalink

Journal

Authors

Wurihan, Wurihan
Wang, Yuxuan
Yeung, Sydney
et al.

Publication Date

DOI

Expression activation of over 70% of *Chlamydia trachomatis* genes during the first hour of infection

Wurihan Wurihan,[1,2] Yuxuan Wang,[1] Sydney Yeung,[1] Yi Zou,[3] Zhao Lai,[3,4] Joseph D. Fondell,[1] Wei Vivian Li,[5] Guangming Zhong,[6] Huizhou Fan[1]

AUTHOR AFFILIATIONS See affiliation list on p. 13.

ABSTRACT The obligate intracellular bacterium *Chlamydia* has a unique developmental cycle that alternates between two contrasting cell types. With a hardy envelope and highly condensed genome, the small elementary body (EB) maintains limited metabolic activities yet survives in extracellular environments and is infectious. After entering host cells, EBs differentiate into larger and proliferating reticulate bodies (RBs). Progeny EBs are derived from RBs in late developmental stages and eventually exit host cells. How expression of the chlamydial genome consisting of nearly 1,000 genes governs the chlamydial developmental cycle is unclear. A previous microarray study identified only 29 *Chlamydia trachomatis* immediate early genes, defined as genes with increased expression during the first hour postinoculation in cultured cells. In this study, we performed more sensitive RNA sequencing (RNA-Seq) analysis for *C. trachomatis* cultures with high multiplicities of infection. Remarkably, we observed well over 700 *C. trachomatis* genes that underwent 2- to 900-fold activation within 1 hour postinoculation. Quantitative reverse transcription real-time PCR analysis was further used to validate the activated expression of a large subset of the genes identified by RNA-Seq. Importantly, our results demonstrate that the immediate early transcriptome is over 20 times more extensive than previously realized. Gene ontology analysis indicates that the activated expression spans all functional categories. We conclude that over 70% of *C. trachomatis* genes are activated in EBs almost immediately upon entry into host cells, thus implicating their importance in initiating rapid differentiation into RBs and establishing an intracellular niche conducive with chlamydial development and growth.

KEYWORDS *Chlamydia*, transcriptome, gene regulation, transcription, chlamydia developmental cycle

S exually transmitted *Chlamydia trachomatis* infections represent a significant health concern globally. Since its designation as a notifiable infection in 1988, *C. trachomatis* infection in the USA has consistently been the most frequently reported infection to the Centers for Disease Control and Prevention (1). Moreover, ocular *C. trachomatis* infection is a leading infectious cause of blindness in numerous undeveloped regions (2).

As with other *Chlamydia* species, *C. trachomatis* exists in two different cellular forms. The infectious elementary body (EB), approximately 300 nm in diameter and with limited metabolic activities, can temporarily endure in extracellular environments and is responsible for initiating the infection process. Upon binding to the host cell membrane, EBs are endocytosed (3). Over the subsequent hours, bacterium-containing vacuoles (termed inclusions) are transported to perinuclear regions where EBs differentiate into the larger reticulate body (RB). RBs then multiply logarithmically until they asynchronously differentiate back to EBs which exist host cells (4).

The 1 million base pair *C. trachomatis* genome includes over 900 coding sequences (5, 6). To decipher the regulatory patterns of the chlamydial transcriptome, Nicholson

et al. and Belland et al. employed microarray strategies to profile the *C. trachomatis* transcriptome across various developmental cycle stages (7, 8). Nicholson et al.'s earliest data point was at 6 hours postinoculation (hpi), around which time the initial EB-to-RB differentiation typically concludes. Data for earlier time points were not included in this study, possibly due to the limited sensitivity of the microarray in their experimental design which utilized a multiplicity of infection (MOI) of just one inclusion-forming unit (IFU) per host cell (8). In contrast, Belland et al. performed a cDNA microarray analysis as soon as 1 hpi. To bolster transcript detection sensitivity during this immediate early stage, they infected cells with an MOI of 100 (7). At 1 hpi, they identified transcripts from 55 protein-encoding genes. Their subsequent quantitative reverse transcription real-time PCR (qRT-PCR) analysis validated the activated expression of only 29 of these 55 genes, while the other 26 showed declined expression. They defined the 29 activated genes as "immediate early genes" and the other 26 as "late genes" (7).

More recently, Humphrys et al. detected 399 activated gene transcripts using RNA sequencing (RNA-Seq) at 1 hpi (9). Clearly, RNA-Seq offers higher gene expression sensitivity than microarray across all systems (10, 11). However, Humphrys et al. only validated activated expression for 10 of the 399 genes, which were the same 10 genes previously identified by Belland et al. as immediate early genes. Furthermore, they did not investigate which of the remaining 389 mRNAs might also be newly synthesized upon cell entry nor distinguish whether some of the genes may have been simply carried over from the previous developmental cycle (9).

To elucidate the role of immediate early gene expression in the onset of the chlamydial developmental cycle, we revisited the *C. trachomatis* transcriptome during the initial stage of infection. We employed RNA-Seq analysis at MOIs of 50 and 200 to monitor the expression profiles of *C. trachomatis* genes, using the 0 hpi as the baseline from which we could determine differential expression at 1 hpi. We then conducted qRT-PCR analysis with an MOI of 1 to validate RNA-Seq data. Our results demonstrate that over 70% of the *C. trachomatis* genome becomes activated by 1 hpi. Moreover, our findings show broad gene activation in each and every functional ontological category during this presumptively crucial immediate early stage.

RESULTS

RNA-Seq analysis reveals activation of 730 *C. trachomatis* genes from 0 to 1 hpi

Comprehensive identification of differentially expressed genes (DEGs) in *Chlamydia* during the immediate early developmental phase is hindered by two challenges: the inherently low RNA content of EBs (12) and the dominance of host RNA in the samples. To address these challenges here, high infectious doses were used, echoing the approach of Belland et al., who employed an MOI of 100 for microarray at 1 hpi (7).

Our initial RNA-Seq study, using 50 IFUs per cell, resulted in 3.9- and 15.9-fold genome coverage at 0 and 1 hpi, respectively (Table 1). While this captured 785 and 911 genes at the respective time points (Table 1; Table S1), the modest coverage at 0 hpi was a concern. Hence, a subsequent RNA-Seq study was conducted using an increased MOI of 200, which enhanced genome coverage significantly for both time points (Table 1). Comparatively, this higher MOI identified 113 more genes at 0 hpi but just 6 additional at 1 hpi (Table 1; Table S1), suggesting a near saturation point around 15- to 20-fold genome coverage.

The increased genome coverage observed from 0 to 1 hpi suggests substantially increased transcriptional activity during the immediate early developmental phase. To address the unlikely but possible scenario that the increased RNA levels at 1 hpi could be due to a higher number of EBs associated with host cells, we quantified the *C. trachomatis* genome by measuring *ctl0631* gene using quantitative PCR (qPCR). This analysis revealed consistent genome copy numbers in both 0 and 1 hpi cultures (data not shown), lending support to the hypothesis of a broad activation of the *C. trachomatis* transcriptome during the first hour of infection.

TABLE 1 RNA-Seq unique reads mapped to the *C. trachomatis* genome, genome coverage, and genes detected[a]

hpi	50MOI			200MOI		
	Ct reads	Fold coverage	Genes detected	Ct reads	Fold coverage	Genes detected
0	82,377	3.9	785	402,679	19.2	898
1	332,916	15.9	911	1,761,784	105.3	917

[a]L929 cells were inoculated with *C. trachomatis* L2 at an MOI of 50 or 200 and centrifuged at room temperature for 10 min. After three washes, RNA was extracted either immediately (0 hpi) or after a 1-hour incubation at 37°C. RNA-Seq was conducted as outlined in the Materials and Methods section. Average values for the number of reads uniquely mapped to the *C. trachomatis* genome (Ct reads) and fold coverage, derived from biological triplicates, are presented.

We performed DESeq2 analysis (13) to identify DEGs between 0 and 1 hpi. Both the 50MOI and 200MOI studies detected genes undergoing dramatic activation: over 150 and 500 genes with ≥100- and ≥10-fold activation, respectively, and nearly 700 genes with ≥2-fold activation (Table 2; Table S2). The vast majority of the activated genes were consistent between the individual studies; upon analysis of the combined data set, we identified 154, 538, and 730 genes with ≥100-, ≥10-, and ≥2-fold activation, respectively (Table 2; Table S2).

In contrast, only 27 and 29 genes with ≥2-fold reduction in transcripts in the 50MOI study and 200MOI study, respectively (Table 2; Table S2). The high degree of reduction was only eightfold. The downregulated genes are also highly consistent between the studies (Table 2). Analysis of the two data sets combined identified a total of only 30 genes with ≥2-fold downregulation, with 6.9-fold being the largest reduction (Table 2).

Taken together, results from our RNA-Seq studies clearly demonstrate robust activation of the *C. trachomatis* transcriptome during the immediate early developmental phase.

qRT-PCR analysis validates immediate early genes identified from RNA-Seq analysis

To validate our RNA-Seq findings, we conducted qRT-PCR analysis on 48 of the detected activated genes from *C. trachomatis* cultures with an MOI of 1 IFU per cell at 0 and 1 hpi. From the RNA-Seq data, we selected genes for qRT-PCR analysis that had a broad range of FPKM (fragments per kilobase of transcript per million, which normalizes read count based on gene length and the total number of mapped chlamydial reads) from the top 1.0 percentile (i.e., higher expressers) to the bottom 84.5 percentile (i.e., low expressers), as well as a broad range of fold activation (2.9- to 868.4-fold) (Table 3).

Of the 48 genes chosen for qRT-PCR, 44 showed statistically significant increases in transcripts (≥2-fold, $P < 0.05$) from 0 to 1 hpi, consistent with our initial data from the RNA-Seq studies. For the remaining four genes that did not exhibit this pattern, we reevaluated after increasing the MOI to 10. After the adjustment, two genes, *pkn1* and *ldh*, displayed expression activation between from 0 to 1 hpi. However, *sucB* and *murE* remained unchanged (Table 3). The concurrence between our qRT-PCR findings and the RNA-Seq data strongly supports the notion that an overwhelming majority of *C. trachomatis* genes are activated within the first infection hour.

Activation of the rRNA operons during the immediate early phase.

rRNAs, owing to their abundance in nearly all cell types, are typically removed prior to RNA-Seq library construction to enhance detection efficiency of other RNAs. Our method for rRNA removal was notably efficient; we recorded 0 counts for both 16S and 5S rRNA from both 50MOI and 200MOI studies, and a total of only 1211 23S rRNA reads (Table S1). We performed qRT-PCR analysis to quantify all three rRNA with the MOI of 1 to compare their expression levels at 0 and 1 hpi.

The chlamydial rRNA genes are arranged in a 16S, 23S, and 5S rRNA order within two identical rRNA operons (Fig. 1A). Cleavages in the initial transcripts subsequently

TABLE 2 Differentially expressed *C. trachomatis* genes (*P* < 0.05) identified between 0 and 1 hpi in the RNA-Seq studies referenced in the Table 1 legend[a]

	50MOI	200MOI	Common	Combined
≥100-fold increase	203	150	122	154
≥10-fold increase	552	515	475	538
≥2-fold increase	663	681	638	730
2- to 14-fold decrease	27	29	26	30

[a]Numbers of genes with increased or decreased expression in individual studies, common to both studies, or in the combined data set are presented.

produce individual rRNAs (Fig. 1A). We employed three primer pairs, each amplifying sequences present in mature rRNAs and the nascent transcript, and a primer pair targeting only the nascent rRNA (Fig. 1A). We detected a 23.3-fold increase in the nascent rRNA at 1 hpi (Fig. 1B) and also detected 3.5 and 3.3 increases in the 16S and 23S rRNA, respectively (Fig. 1B). A larger increase in the nascent form than the total rRNA is expected since rRNA maturation is relatively efficient, and mature rRNAs are relatively stable. Interestingly, we did not detect any changes in the 5S rRNA despite increased expression of pre-rRNA (Fig. 1B).

We hypothesized that 16S and/or 23S rRNA are limiting. To test that expression increases are sufficient to increase ribosome formation during the immediate early developmental phase, we determined copy ratio of the three *C. trachomatis* rRNA species at both time points. We first performed qPCR analysis on the genomic DNA to determine the amplification efficiency of the primer pairs used for qRT-PCR analysis. The chromosome rRNA gene amplification efficiencies were used to normalize the relative expression levels of the three chlamydial rRNAs. Our analysis revealed an excess of 5S rRNA at both 0 and 1 hpi. In contrast, 23S rRNA appeared limiting, which is most evident at 0 hpi than at 1 hpi (Fig. 1C). These data suggest that increased 16S and 23S rRNAs expression can lead to elevated ribosome numbers; provided expression of genes encoding ribosomal structure and biogenesis proteins are also activated.

Immediate early gene ontology

We next performed gene ontology analysis to better understand the physiological importance of the *C. trachomatis* transcriptome during the immediate early phase. The percentage of up- and down-regulated genes (≥2-fold change, *P* < 0.05) was plotted in each functional category (Fig. 2). DEGs in functional categories and the implications of their expression changes are summarized below.

Activation of genes involved in protein synthesis

Topping of the list of up-regulated gene categories is the translation and ribosomal structure and biogenesis category, with 133 of the 138 (96.4%) genes activated (Fig. 2A; Table S3). Consistent with this finding, 28 of 37 (75.7%) *C. trachomatis* tRNAs are also increased (Fig. 2A; Table S3). Together with aforementioned increases in 16S and 23S rRNAs, these findings strongly suggest an enhanced protein synthesis capacity starting the immediate early developmental phase.

Activation of genes with functions in nutrient acquisition and metabolism and energy production and conversion

Other ontological categories that ranked high in the list of activated genes at 1 hpi are those with functions in nutrient acquisition and metabolism. Specifically, they include inorganic ion transport and metabolism, nucleotide transport and metabolism, lipid transport and metabolism, and amino acid transport and metabolism categories with 20 of 21 (95.2%), 14 of 15 (93.3%), 32 of 35 (91.4%), and 47 of 52 (90.4%) genes activated, respectively (Fig. 2A). In addition, 36 of 42 (85.7%) genes in the coenzyme transport and metabolism category and 29 of 36 (80.6%) genes in the carbohydrate transport and

TABLE 3 Validation of gene activation identified by RNA-Seq using qRT-PCR analysis[a]

Gene	Description	RNA-Seq			qRT-PCR	
		FPKM percentile	Fold increase	P	Fold increase	P
rpsR	30S ribosomal protein S18	1.0	323.5	1.2E-251	111.4	2.2E-02
rpmA	50S ribosomal protein L27	1.7	190.7	3.8E-193	71.3	1.7E-06
incB	Inclusion membrane protein B	1.9	648.8	3.7E-246	215.4	7.0E-07
rpmI	50S ribosomal protein L35	3.0	427.9	1.9E-123	441.3	1.7E-05
rplV	50S ribosomal protein L22	3.6	305.0	8.4E-149	63.2	4.3E-07
pyrH	Uridylate kinase	4.9	235.2	0.0E-00	110.0	1.2E-03
infC	Translation initiation factor IF-3	5.6	322.8	3.0E-242	111.4	1.3E-05
rpmJ	50S ribosomal protein L36	6.0	122.6	4.3E-79	85.0	9.2E-11
gspD	Type II secretion system protein D	8.7	284.4	0.0E-00	58.2	6.7E-05
accD	Acetyl-CoA carboxylase subunit beta	9.5	193.0	2.1E-144	5.4	1.8E-06
fusA	Elongation factor G	10.8	369.4	0.0E-00	207.7	3.5E-04
rpoA	DNA-directed RNA polymerase subunit alpha	12.2	158.3	2.0E-260	94.0	1.0E-05
groEL	Chaperonin GroEL	16.8	5.3	2.3E-05	16.1	1.8E-02
nusB	Transcription termination factor NusB	17.3	334.7	1.3E-101	57.9	4.4E-11
nusG	GTPase ObgE	19.1	194.1	1.4E-79	893.4	7.6E-05
rpoZ	RNA polymerase omega subunit	19.9	138.9	6.4E-48	85.1	2.5E-07
gapA	Glyceraldehyde 3-phosphate dehydrogenase	20.8	144.0	3.3E-179	286.1	1.3E-04
engA	GTPase Der	21.4	317.8	3.5E-172	104.0	3.8E-03
rpoB	DNA-directed RNA polymerase subunit β	21.8	231.7	3.1E-299	88.9	1.3E-05
mdhC	Malate dehydrogenase	22.0	3.7	6.0E-23	4.4	1.9E-03
nusA	Transcription termination factor NusB	22.6	251.6	8.9E-126	788.0	2.9E-05
rpoC	DNA-directed RNA polymerase subunit beta'	23.3	317.9	5.8E-266	726.9	4.2E-07
gltT	Sodium:dicarboxylate symport protein	23.5	55.1	2.5E-131	22.8	2.3E-06
gyrB	DNA topoisomerase IV subunit B	24.6	51.8	1.2E-218	22.9	1.8E-07
def	Peptide deformylase	25.6	40.7	8.9E-46	21.1	1.8E-06
grgA	General activator of genes	25.7	89.6	5.3E-106	50.4	7.6E-04
plsB	Glycerol-3-phosphate acyltransferase	26.0	815.9	3.2E-59	770.2	6.1E-04
secG	Preprotein translocase subunit SecG	27.0	22.0	9.4E-15	21.3	1.4E-03
infB	Translation initiation factor IF-2	27.3	172.2	6.7E-205	248.6	6.7E-05
tal	Transaldolase	31.3	36.5	6.8E-66	32.1	2.0E-12
tyrP	Tyrosine-specific transport protein	34.4	44.2	1.3E-36	21.4	3.4E-05
rho	Transcription termination factor	34.9	23.2	6.8E-73	30.6	4.0E-03
greA	Transcription elongation factor GreA	39.0	42.0	2.3E-88	36.1	5.5E-03
adk	Adenylate kinase	40.7	95.4	1.2E-33	309.1	5.6E-04
rpoD	RNA polymerase sigma factor	44.7	27.3	2.8E-77	33.2	9.6E-09
ihfA	DNA-binding protein HU	46.2	10.1	1.0E-07	5.4	6.4E-06
atpI	V-type ATP synthase subunit I	49.5	16.4	1.6E-43	36.6	9.2E-05

(Continued on next page)

TABLE 3 Validation of gene activation identified by RNA-Seq using qRT-PCR analysis[a] (*Continued*)

Gene	Description	RNA-Seq			qRT-PCR	
		FPKM percentile	Fold increase	P	Fold increase	P
pgi	Glucose-6-phosphate isomerase	52.3	10.5	1.9E-20	27.2	1.4E-04
sucB	2-Oxoglutarate dehydrogenase complex subunit dihydrolipoyllysine-residue succinyltransferase	55.0	18.0	2.9E-42	1.3*	7.5E-01*
pdhC	Pyruvate dehydrogenase complex dihydrolipoyllysine-residue acetyltransferase component	55.0	23.5	2.4E-15	68.4	1.7E-04
pmpI	Outer membrane protein PmpI	59.0	20.8	1.7E-115	7.7	2.4E-03
nrdB	Ribonucleoside-diphosphate reductase subunit beta	66.0	6.0	2.1E-06	9.5	1.5E-03
lpdA	2-Oxoglutarate dehydrogenase complex dihydrolipoyl dehydrogenase	67.1	6.7	1.9E-11	31.2	9.6E-05
pkn1	Serine/threonine-protein kinase	67.7	6.8	4.2E-11	9.8*	6.5E-04*
murE	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate---2,6-diaminopimelate ligase	73.9	8.4	1.9E-04	2.1*	3.7E-1*
pknD	Serine/threonine-protein kinase	75.3	5.1	2.8E-11	5.9	3.8E-04
ldh	Leucine dehydrogenase	79.2	9.7	1.6E-04	6.0*	7.7E-04*
sucD	Succinyl-CoA ligase subunit alpha	84.5	3.8	2.8E-02	1.9	1.6E-02

[a]L929 cells were inoculated with *C. trachomatis* L2 at an MOI of 1. RNA was extracted at 0 and 1 hpi as outline in Table 1 legend. Forty-eight genes were selected, based on their FPKM ranking, for qRT-PCR analysis. Forty-four of the 48 genes demonstrated significant increases in expression. The remaining four genes, which did not show significant increases in the initial analysis, were reevaluated using cultures infected at an MOI of 10. Results from the higher MOI cultures are indicated with asterisks.
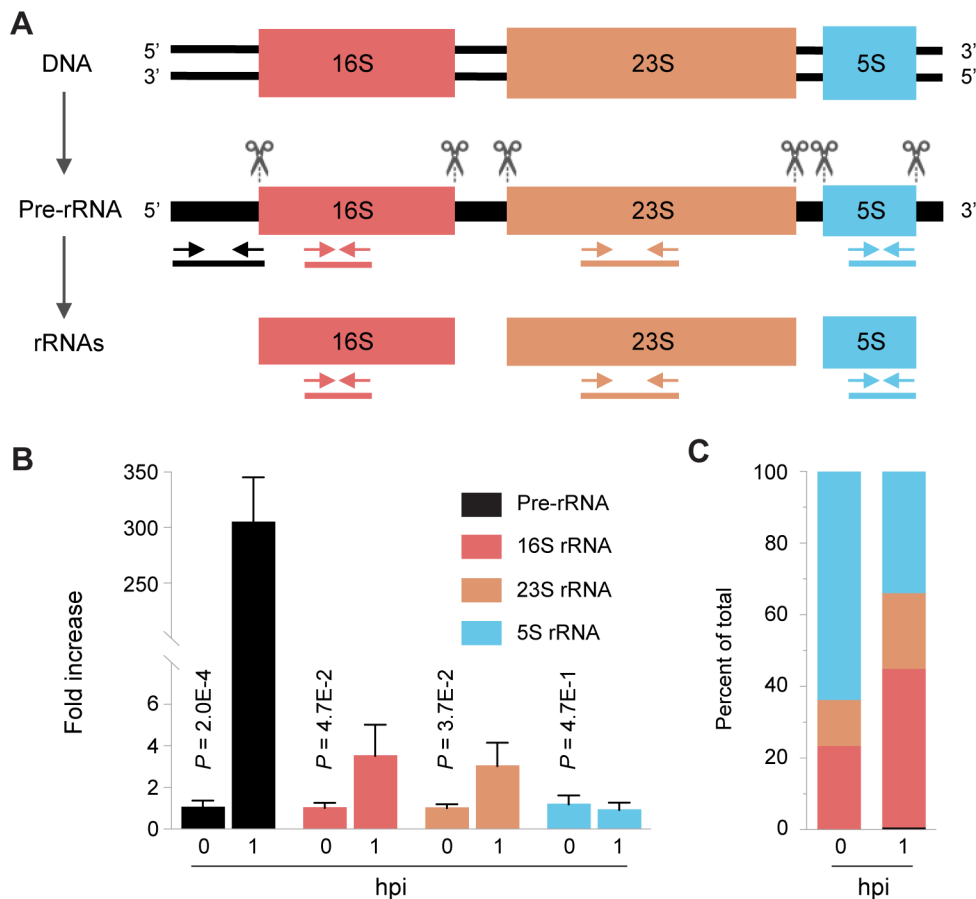
FIG 1 23S rRNA is the rate-limiting rRNA in *C. trachomatis* during the immediate early developmental phase. (A) *C. trachomatis* chromosomal rRNA gene organization and rRNA genesis. The schematic illustrates the rRNA genes on the chromosome and the processing steps of pre-rRNA to mature rRNAs. Scissors symbols indicate RNA cleavage at indicated sites. Arrows represent primers targeting specific sequences for amplifying pre-rRNA and mature rRNA. Lines below the arrows signify reverse transcription PCR products using different respective primer pairs. (B) Temporal changes in rRNA levels (0–1 hpi). Bar graph shows the largest increase in pre-rRNA, moderate increases in 16S and 23S rRNA, and unchanged 5S rRNA from 0 to 1 hpi. (C) Proportional shifts in mature rRNAs. Stacked bar graph depicts the changing proportions of the total rRNA pool, with concurrent 23S and 16S rRNA increases and 5S rRNA decrease. The proportions of pre-rRNA in the graph are too small to be visible at both 0 and 1 hpi. (B, C) pre-rRNA and mature rRNAs (16S, 23S, and 5S) were quantified via qRT-PCR, utilizing primers indicated in (A).

metabolism category are also activated (Fig. 2A; Table S3). This widespread activation suggests that *C. trachomatis* rapidly adapts to meet heightened nutrient requirements, support the increased RNA and protein synthesis, and initiate morphological transformation. Additionally, 29 of 36 (69.4%) genes essential for energy production and conversion were activated, ensuring adequate energy for the increased cellular activities.

Activation of pathogenicity and virulence genes

Incs (inclusion membrane proteins) are required for establishment and maintenance of a chlamydial intracellular growth niche (14). Secretion of many Incs and other virulence determinants depends on the type III secretion system (T3SS) and/or trafficking systems (15–17). The chlamydial plasmid encodes a virulence protein Pgp3, a transcription regulator of chromosomal gene, and proteins essential for the maintenance of the plasmid (18). Notably, 35 of 39 (89.7%) Inc genes and 22 of 48 (45.8%) genes in T3SS, 4 of 5 (80%) genes involved in defense, 14 of 18 (77.8%) genes with functions in intracellular trafficking, secretion, and vesicular transport, and 7 of 8 (87.5%) genes encoded by the
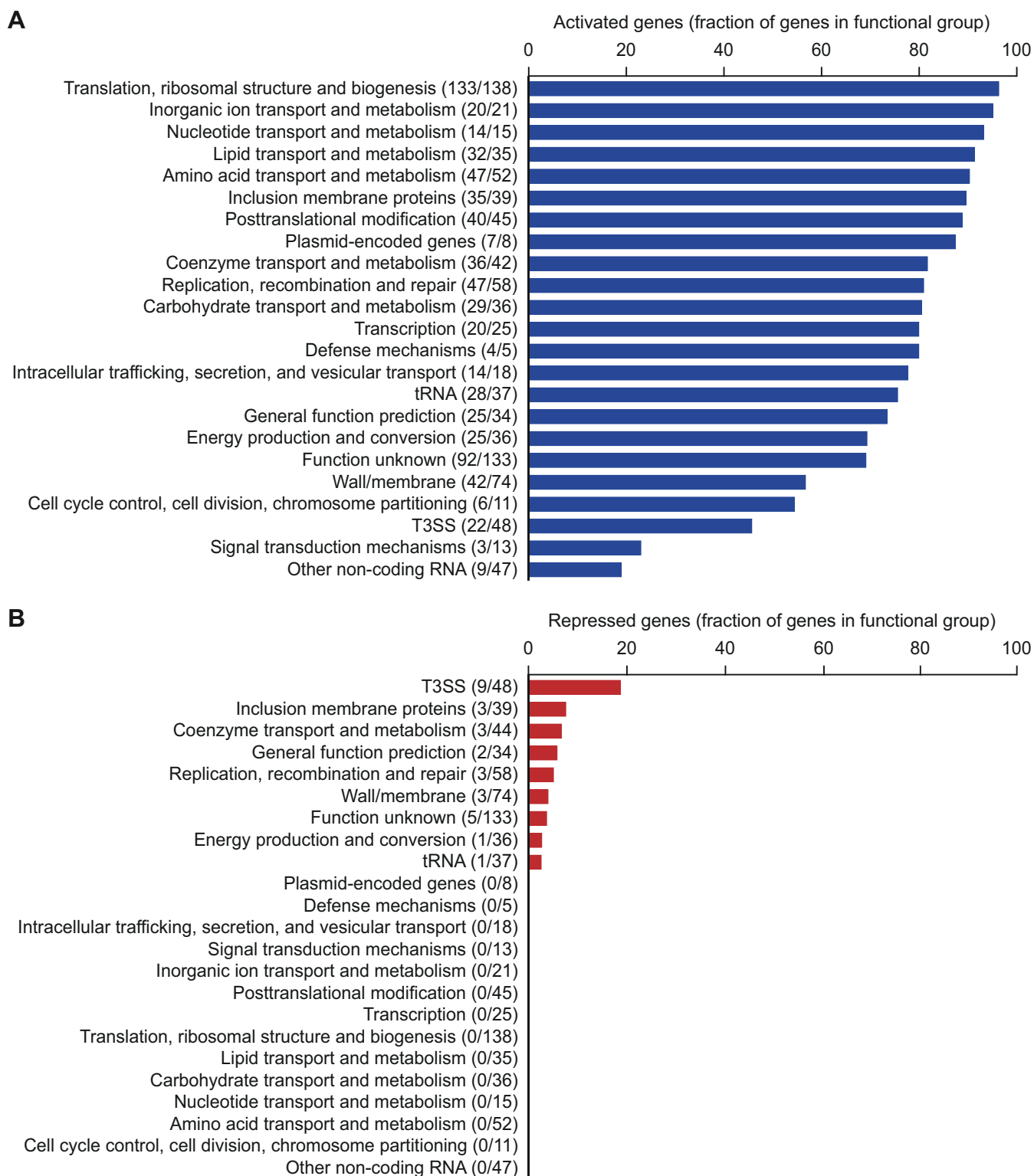
A

Activated genes (fraction of genes in functional group)

B

Repressed genes (fraction of genes in functional group)

FIG 2 Functional categorization of activated and repressed genes. (A) Distribution of activated genes across all 23 functional categories. (B) Distribution of activated genes in only nine functional categories. (A, B) The numbers adjacent to each category indicate the count of repressed genes out of the total in that category, based on the criteria of ≥2-fold change in expression ($P < 0.05$).

virulence plasmid are activated by 1 hpi (Fig. 2A). Collectively, these results underscore the rapid expression of a diverse array of pathogenicity and virulence factors during the

immediate early phase, which presumably positions the pathogen for a successful infection.

Activation of genes with functions in transcription and transcriptional regulation

Twenty of 25 (80%) genes in the transcription category are activated by 1 hpi. These include genes encoding all the four subunits of the RNA polymerase core enzyme composed (i.e., *rpoA*, *rpoB*, *rpoC*, and the newly identified *rpoZ*) (19). *rpoD* encodes σ66, the principal sigma factor that binds the RNA polymerase core enzyme to form a holoenzyme. The σ66 holoenzyme is thought to be necessary for initiating transcription from all early gene promoters, while alternative holoenzymes containing either σ28 or σ54 are required for initiating transcription from certain late gene promoters. Consistent with this notion, *rpoD* demonstrated a 28.1- and 33.2-fold activation in RNA-Seq and qRT-PCR analyses, respectively, from 0 to 1 hpi, while expression levels of *fliA* (σ28) and *rpoN* (σ54) remain unchanged.

Genes encoding eight additional transcriptional factors are activated from 0 to 1 hpi (Table S1). We found that *euo* is strongly activated (Table S3), consistent with numerous previous studies showing that it represses both midcycle and late genes (20–23). In addition, *grgA*, which regulates RB growth and progeny EB formation (24), and *hrcA*, which encodes the heat-inducible transcription repressor HrcA that regulates the expression of protein chaperones (25, 26), are both strongly activated by 1 hpi (Table S3). Furthermore, the transcription elongation factor gene *greA* and all four transcription termination factor-encoding genes (i.e., *rho, nusA, nusB,* and *nusG*) are all activated during the immediate early phase (Table S3).

Taken together, the activation of a broad spectrum of genes encoding components of the RNA polymerase core and holoenzyme, as well as transcription factors regulating initiation, elongation, and termination, suggests that increasing the immediate early capacity of the transcription machinery is critical for the subsequent growth and survival of *C. trachomatis*.

Activation of genes with functions in the category of DNA replication, recombination, and repair (DRRR) and in the category of cell cycle, division, chromosome partitioning (CCDCP)

The primary differentiation of the infectious EB to the replicative RB takes about 6 hours. Interestingly, 48 of 58 (82.8%) DRRR genes are activated from 0 to 1 hpi (Fig. 2A). In addition, 6 of 11 (54.5%) genes related to CCDCP are also activated (Fig. 2A). The reason and significance for the seemly untimely DRRR and CCDCP gene activation are explored below.

Downregulation of late genes

The 30 genes with moderate (2.0- to 6.9-fold) downregulation from 0 to 1 hpi fall into only 9 of 23 (39.1%) functional categories (Fig. 2B). The top ranked decreased category contains mostly putative T3SS effectors (*copB, copD, ctl0034, ctl0064, ctl0255, ctl0338, ctl0338A, ctl0339, ctl0886*) and a T3SS chaperone (*scc2*). As expected (7, 24, 27), three EB-enriched outer membrane protein genes (*omcA, pmpD,* and *ctl0815*) showed decreased expression. Similarly, the percentages of decreased genes in the remaining eight categories are all in single digits (2.7%–7.1%) (Fig. 2B). Significant decreases in the expression of these genes during the immediate early phase suggest that continued expression of these genes is not important for the immediate early (and early) *C. trachomatis* development.

Lack of correlation between genomic organization of DRRR and CCDCP genes and their immediate early activated

We hypothesized that some DRRR and CCDCP genes are activated within the first hour because they are components of operons containing genes with critical roles during the immediate early phase. To test this hypothesis, we surveyed the organization of increased DRRR genes on the chromosome. Our analysis revealed that 30 of the 48 increased DRRR genes are located in operons with non-DRRR genes (Fig. 3A), while only 19 are standalone non-operon genes or are in operons containing multiple DRRR genes
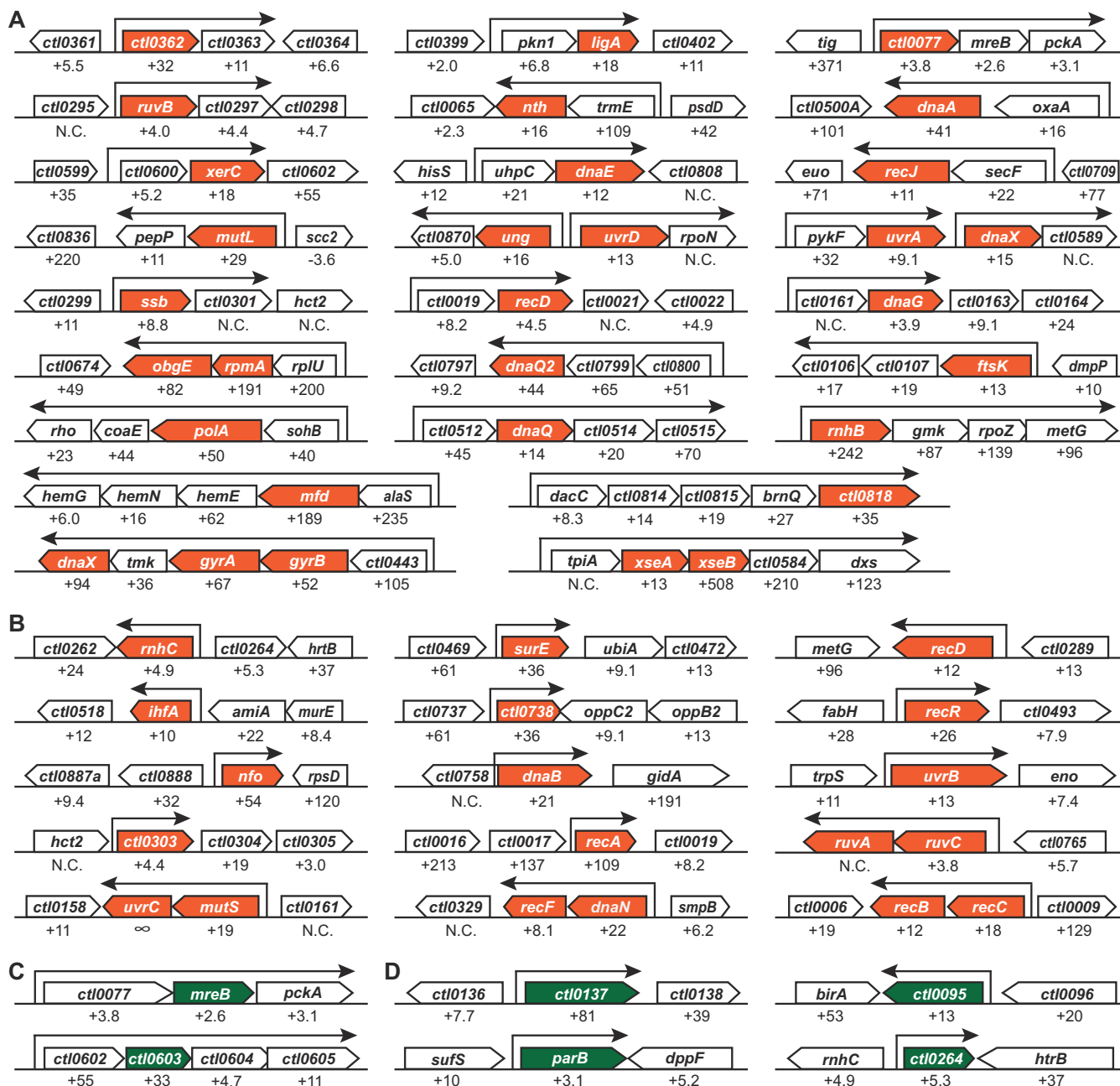
FIG 3 Genomic organization of DRRR and CCDCP genes. (A) Operon-based DRRR genes interspersed with genes from other functional categories. (B) Standalone DRRR genes and operons exclusively composed of DRRR genes. (C) Operon-based CCDCP genes interspersed with genes from other functional categories. (D) Standalone CCDCP genes and operons exclusively composed of CCDCP genes. (A–D) arrows indicate the transcription direction and the extent of each operon. The number below each gene represents fold activation; "N.C." denotes no change.

(Fig. 3B). Among six increased CCDCP genes, only two are in operons containing genes of other functional categories (Fig. 3C), while the remaining four are all non-operon genes (Fig. 3D). Therefore, possible collateral activation does not fully account for increased expression of DRRR and CCDCP genes from 0 to 1 hpi.

DISCUSSION

In this study, we present evidence that a significant majority of *C. trachomatis* genes are activated by 1 hpi (Tables 1 to 3; Tables S1 and S2), thus indicating that the immediate early chlamydial transcriptome is much more robust than previously recognized in the pioneering *Chlamydia* transcriptomic study by Belland et al. (7). We achieved this insight by adopting Belland et al.'s strategy of using high MOIs for 0 and 1 hpi and by employing RNA-Seq, a technique with significantly superior sensitivity and breadth compared to the slide microarray technology employed in their study.

There is also a key difference in our approach to preparing RNA samples from 0 hpi cultures. Belland et al. used RNA from purified EBs, likely raising the sensitivity in their microarray analysis and potentially masking the detection of genes with increased expression at 1 hpi. We prepared RNA samples from infected cells at both 0 and 1 hpi, which allowed for a more accurate identification of immediate early gene activation.

Practical implications

We anticipate that our report will ignite broad interests in exploring the influence of chlamydial gene product, host factors, and potential antichlamydial drugs on the immediate early chlamydial transcriptome. Analysis of our RNA-Seq data sets (Tables 1 and 2; Tables S1 and S2) indicates that with biological triplicates, a 4- to 5-fold genome coverage at 0 hpi and 15- to 20-fold genome coverage at 1 hpi can identify most DEGs, including both activated genes and repressed genes. However, to bolster confidence in the RNA-Seq data, we recommend fivefold higher coverages.

Even with increased MOI, we recommend qRT-PCR analysis to be performed to validate DEGs identified by RNA-Seq analysis. This may be particularly important for genes with relatively low expression levels, which can be identified with FPKM values (ranked around 50 percentile or lower) and for genes with relatively small changes (<10-fold change).

Fundamental implications

Once inside host cells, successful *Chlamydia* infection depends on execution of two interrelated tasks: avoidance of clearance by the host cellular defense system and conversion of EBs into RBs (14). While the previous microarray study recognized the importance of immediate early gene expression in the regulation of host cell signaling (7), results presented in this report show for the first time that immediate early transcription activation extends to all functional gene categories involved in the EB-to-RB conversion (Fig. 1 and 2). We infer that EB-to-RB conversion requires broad activation of the chlamydial transcriptome. For example, the surface area of RBs is about 10 times larger than that of EBs. Expansion of the surface area necessitates synthesis and/or acquisition of large amounts of lipids and robust synthesis of membrane proteins and glycans. The demands for increased protein synthesis can only be accomplished through broad expression increases in the components comprising the protein synthesis machinery, which include ribosomal RNAs and proteins and tRNAs. The increases in transcription and protein synthesis further demands a surge in acquisition and synthesis of metabolic precursors, cofactor, and energy, which in turn depends on new protein synthesis.

The activation of a significant majority of chlamydial genes during the immediate early development phase suggests a physiological importance for the rapid enhancement of multiple biological processes. Nonetheless, the observed upregulation of DRRR and CCDCP genes during this period (Fig. 2; Tables S2 and S3) was somewhat

unexpected, given the prevailing view that DNA replication does not begin until after the completion of the EB-to-RB conversion. Our genomic analysis of the activated DRRR and CCDCP genes suggests that activation of numerous standalone and operon-contained DRRR and CCDCP genes (Fig. 3) might facilitate some level of DNA synthesis activity during the immediate early phase. This invites speculation that chlamydial chromosomal de-condensation, necessary for broad transcription activation, may involve strand breakage and subsequent DNA repair, which would necessitate increased DRRR and CCDCP gene expression. Indeed, this rationale might explain the observance of DRRR and CCDCP genes within operons containing other genes essential for the immediate early phase. A precedent for this scenario is revealed with the genes *gmk* (guanylate kinase), *rpoZ* (RNA polymerase omega subunit), and *metG* (methionine-tRNA ligase), all of which are co-transcribed with *rnhB* (RNase H). RnhB is necessary for DNA synthesis (28), while Gmk and RpoZ play essential roles in RNA synthesis, and metG is crucial for protein synthesis (19).

While this study was conducted in cultured cells, its insights may be relevant to chlamydial pathogenesis *in vivo*. Interferon gamma (IFN-γ), produced in response to chlamydial and other infections, inhibits *C. trachomatis* growth by promoting tryptophan degradation and altering the metabolism of certain other amino acids (29–32). Conversely, *C. trachomatis* has developed the ability to enter a persistent state where RBs halt division and conversion to EBs until a balanced nutritional environment is reestablished (29–32). Notably, *ctl0225*, one of numerous amino acid transporters activated by 1 hpi (Tables S1 to S3), regulates chlamydial persistence (30, 31). The upregulation of *ctl0225* during the immediate early phase could be a strategic adaptation of *C. trachomatis* to establish infection in environments with elevated IFN-γ levels.

In summary, our study has uncovered that a vast majority of chlamydial genes are activated within the first hour following the entry of EBs into host cells. This transcriptional activation spans across all functional gene categories. While this widespread activation of the chlamydial transcriptome facilitates a variety of biological processes essential for the primary differentiation and establishment of an intracellular environment favorable for development and growth, the notable increase in the expression of DRRR and CCDCP genes raises the possibility of some DNA synthesis occurring during this early phase. Additionally, our experimental strategies as well as our recommendations for genome coverage, informed by our data, will be invaluable for studies of early transcriptomes, not only in chlamydiae but also in a broader spectrum of intracellular pathogens.

MATERIALS AND METHODS

Chlamydia and culture

C. trachomatis L2 (strain 434/BU) was purchased from ATCC (33). The bacterium was grown using L929 cells and Dulbecco's modified Eagle medium containing 4.5 g/L glucose and 110 mg/L sodium pyruvate and supplemented with fetal bovine serum (final concentration, 5%), gentamicin (20 μg/mL), and cycloheximide (1 μg/mL). EBs for this study were sequentially purified through ultracentrifugation using 35% MD-76 and 44%/40%/52% MD-76 gradients.

RNA preparation

Total chlamydial and host RNA was prepared as previously described with modifications (25). The overnight culture media in 6-w plates with L929 monolayers were replaced with fresh medium containing EBs to achieve an MOI of 1, 10, 50, or 200. The plates were centrifuged at $900 \times g$ at room temperature for 10 min and then washed three times with 100 μg/mL heparin in Hank's balanced salt solution to remove free EBs. At the end of washes, cells in one plate were lysed with Tri Reagent (Millipore Sigma), and those in the other plate were lysed with Tri Reagent after incubation at 37°C in a CO_2 incubator. RNA

in Tri Reagent was further purified by following manufacturer's instructions. Contaminating DNA was removed through two rounds of DNase I-XT. Complete DNA removal was confirmed by lack of amplification of *ctl0631*. RNA concentration was determined using Qubit RNA assay kit (ThermoFisher). Aliquots of the DNA-free RNA samples were stored at −80°C.

RNA-Sequencing

RNA-Seq was performed as described with minor modifications (20, 25). Briefly, total RNA integrity was determined using Fragment Analyzer (Agilent) prior to RNA-Seq library preparation. Illumina MRZE706 Ribo-Zero Gold Epidemiology rRNA Removal kit was used to remove mouse and chlamydial rRNAs. Oligo(dT) beads were used to remove mouse mRNA. RNA-Seq libraries were prepared using Illumina TruSeq stranded mRNA-seq sample preparation protocol, subjected to quantification process, pooled for cBot amplification, and sequenced using Illumina HiSeq 3000 platform with a 50 bp single-read sequencing module. Short read sequences were first aligned to the CtL2 434/Bu genome including the chromosome (GCF_000068585.1_ASM6858v1) and the pL2 plasmid (AM886278) and the mouse genome (GCF_000001635.27) using TopHat2 aligner and then quantified for gene expression by HTSeq to obtain raw read counts per gene, and then converted to FPKM (Fragment Per Kilobase of gene length per Million reads of the library) (34–36). DESeq2, an R package commonly used for analysis of data from RNA-Seq studies and test for differential expression (13), was used to normalize data and find group-pairwise differential gene expression based on three criteria: $P < 0.05$, average FPKM > 1, and fold change 1.

Quantitative reverse transcription real-time PCR

qRT-PCR was performed using QuantStudio 5 real-time PCR System (ThermoFisher Bioscientific) and Luna Universal one-step qRT-PCR kit (New England BioLabs) as previously described (20). Refer to Table S4 for information for qRT-PCR primers. *t*-Tests were conducted using Microsoft Office Excel to evaluate the qRT-PCR results. *P*-values were adjusted for multiple comparisons by Benjamini-Hochberg procedure to control the false discovery rate.

Gene ontology analysis

Gene ontology analysis was performed based on Clusters of Orthologous Genes functional classification of the *C. trachomatis* proteome (37) as recently reported (24).

AUTHOR AFFILIATIONS

[1]Department of Pharmacology, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

[2]Institute of Translational Medicine, Medical College, Yangzhou University, Yangzhou, China

[3]Greehey Children's Cancer Research Institute, University of Texas Health San Antonio, San Antonio, Texas, USA

⁴Department of Molecular Medicine, University of Texas Health San Antonio, San Antonio, Texas, USA
⁵Department of Statistics, University of California Riverside, Riverside, California, USA
⁶Department of Microbiology and Immunology, University of Texas Health San Antonio, San Antonio, Texas, USA

AUTHOR ORCIDs

Wurihan Wurihan ⓘ http://orcid.org/0000-0001-7853-0130
Yuxuan Wang ⓘ http://orcid.org/0000-0003-0771-1078
Joseph D. Fondell ⓘ http://orcid.org/0000-0002-7046-9854
Wei Vivian Li ⓘ http://orcid.org/0000-0002-2087-2709
Guangming Zhong ⓘ http://orcid.org/0000-0001-7053-5009
Huizhou Fan ⓘ http://orcid.org/0000-0002-4903-926X

DIRECT CONTRIBUTION

This article is a direct contribution from Huizhou Fan, a member of the *Infection and Immunity* Editorial Board, who arranged for and secured reviews by Daniel Rockey, Oregon State University, and David Nelson, Indiana University School of Medicine.

DATA AVAILABILITY

RNA-seq data have been deposited into the NCBI Gene Expression Omnibus under accession number GSE248988.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Table S1 (IAI00539-23-S0001.xlsx). RNA-Seq readCount, FPKM, and baseMean values of the 50MOI and 200MOI RNA-Seq studies.
Table S2 (IAI00539-23-S0002.xlsx). Results of DESeq2 analysis of RNA-Seq data obtained at 0 and 1 hpi.
Table S3 (IAI00539-23-S0003.xlsx). Differentially expressed genes in functional categories.
Table S4 (IAI00539-23-S0004.xlsx). qRT-PCR primers.

REFERENCES

1. CDC. 2023. Sexually Transmitted Disease Surveillance 2021. Atlanta: US Department of Health and Human Services; 2023.
2. Flueckiger RM, Giorgi E, Cano J, Abdala M, Amiel ON, Baayenda G, Bakhtiari A, Batcho W, Bennawi KH, Dejene M, et al. 2019. Understanding the spatial distribution of trichiasis and its association with trachomatous inflammation-follicular. BMC Infect Dis 19:364. https://doi.org/10.1186/s12879-019-3935-1
3. Hybiske K, Stephens RS. 2007. Mechanisms of *Chlamydia trachomatis* entry into nonphagocytic cells. Infect Immun 75:3925–3934. https://doi.org/10.1128/IAI.00106-07
4. Hybiske K, Stephens RS. 2007. Mechanisms of host cell exit by the intracellular bacterium *Chlamydia*. Proc Natl Acad Sci U S A 104:11430–11435. https://doi.org/10.1073/pnas.0703218104
5. Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, Mitchell W, Olinger L, Tatusov RL, Zhao Q, Koonin EV, Davis RW. 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science 282:754–759. https://doi.org/10.1126/science.282.5389.754
6. Thomson NR, Holden MTG, Carder C, Lennard N, Lockey SJ, Marsh P, Skipp P, O'Connor CD, Goodhead I, Norbertzcak H, Harris B, Ormond D, Rance R, Quail MA, Parkhill J, Stephens RS, Clarke IN. 2008. *Chlamydia trachomatis*: genome sequence analysis of iymphogranuloma venereum isolates. Genome Res 18:161–171. https://doi.org/10.1101/gr.7020108
7. Belland RJ, Zhong G, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL, Caldwell HD. 2003. Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. Proc Natl Acad Sci U S A 100:8478–8483. https://doi.org/10.1073/pnas.1331135100
8. Nicholson TL, Olinger L, Chong K, Schoolnik G, Stephens RS. 2003. Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. J Bacteriol 185:3179–3189. https://doi.org/10.1128/JB.185.10.3179-3189.2003
9. Humphrys MS, Creasy T, Sun Y, Shetty AC, Chibucos MC, Drabek EF, Fraser CM, Farooq U, Sengamalay N, Ott S, Shou H, Bavoil PM, Mahurkar A, Myers GSA. 2013. Simultaneous transcriptional profiling of bacteria and their host cells. PLoS One 8:e80597. https://doi.org/10.1371/journal.pone.0080597
10. Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63. https://doi.org/10.1038/nrg2484
11. Cloonan N, Grimmond SM. 2008. Transcriptome content and dynamics at single-nucleotide resolution. Genome Biol 9:234. https://doi.org/10.1186/gb-2008-9-9-234

12. McClarty G. 1994. Chlamydiae and the biochemistry of intracellular Parasitism. Trends Microbiol 2:157–164. https://doi.org/10.1016/0966-842x(94)90665-3

13. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8

14. Elwell C, Mirrashidi K, Engel J. 2016. *Chlamydia* cell biology and pathogenesis. Nat Rev Microbiol 14:385–400. https://doi.org/10.1038/nrmicro.2016.30

15. Rucks EA. 2023. Type III secretion in *Chlamydia*. Microbiol Mol Biol Rev 87. https://doi.org/10.1128/mmbr.00034-23

16. Yang Z, Tang L, Sun X, Chai J, Zhong G. 2015. Characterization of CPAF critical residues and secretion during *Chlamydia trachomatis* infection. Infect Immun 83:2234–2241. https://doi.org/10.1128/IAI.00275-15

17. Bugalhão JN, Mota LJ. 2019. The multiple functions of the numerous *Chlamydia trachomatis* secreted proteins: the tip of the iceberg. Microb Cell 6:414–449. https://doi.org/10.15698/mic2019.09.691

18. Song L, Carlson JH, Whitmire WM, Kari L, Virtaneva K, Sturdevant DE, Watkins H, Zhou B, Sturdevant GL, Porcella SF, McClarty G, Caldwell HD. 2013. *Chlamydia trachomatis* plasmid-encoded Pgp4 is a transcriptional regulator of virulence-associated genes. Infect Immun 81:636–644. https://doi.org/10.1128/IAI.01305-12

19. Cheng A, Wan D, Ghatak A, Wang C, Feng D, Fondell JD, Ebright RH, Fan H, Babitzke P. 2023. Identification and structural modeling of the RNA polymerase omega subunits in chlamydiae and other obligate intracellular bacteria. mBio 14:e0349922. https://doi.org/10.1128/mbio.03499-22

20. Wurihan W, Zou Y, Weber AM, Weldon K, Huang Y, Bao X, Zhu C, Wu X, Wang Y, Lai Z, Fan H. 2021. Identification of a GrgA-Euo-HrcA transcriptional regulatory network in *Chlamydia*. mSystems 6:e0073821. https://doi.org/10.1128/mSystems.00738-21

21. Rosario CJ, Hanson BR, Tan M. 2014. The transcriptional repressor EUO regulates both Subsets of *Chlamydia* late genes. Mol Microbiol 94:888–897. https://doi.org/10.1111/mmi.12804

22. Rosario CJ, Tan M. 2012. The early gene product EUO is a transcriptional repressor that selectively regulates promoters of *Chlamydia* late genes. Mol Microbiol 84:1097–1107. https://doi.org/10.1111/j.1365-2958.2012.08077.x

23. Hakiem OR, Rizvi SMA, Ramirez C, Tan M. 2023. Euo is a developmental regulator that represses late genes and activates midcycle genes in *Chlamydia trachomatis*. mBio 14:e0046523. https://doi.org/10.1128/mbio.00465-23

24. Lu B, Wang Y, Wurihan W, Cheng A, Yeung S, Fondell JD, Lai Z, Wan D, Wu X, Li WV, Fan H. 2024. Requirement of GrgA for *Chlamydia* infectious progeny production, optimal growth, and efficient plasmid maintenance. mBio:e0203623. https://doi.org/10.1128/mbio.02036-23

25. Huang Y, Wurihan W, Lu B, Zou Y, Wang Y, Weldon K, Fondell JD, Lai Z, Wu X, Fan H. 2021. Robust heat shock response in *Chlamydia* lacking a typical heat shock Sigma factor. Front. Microbiol 12:812448. https://doi.org/10.3389/fmicb.2021.812448

26. Wilson AC, Tan M. 2002. Functional analysis of the heat shock regulator HrcA of *Chlamydia trachomatis*. J Bacteriol 184:6566–6571. https://doi.org/10.1128/JB.184.23.6566-6571.2002

27. Soules KR, LaBrie SD, May BH, Hefty PS. 2020. Sigma 54-regulated transcription is associated with membrane reorganization and type III secretion effectors during conversion to infectious forms of *Chlamydia trachomatis*. mBio 11:e01725-20. https://doi.org/10.1128/mBio.01725-20

28. Cerritelli SM, Crouch RJ. 2016. The balancing act of ribonucleotides in DNA. Trends Biochem Sci 41:434–445. https://doi.org/10.1016/j.tibs.2016.02.005

29. Byrne GI, Lehmann LK, Landry GJ. 1986. Induction of tryptophan catabolism is the mechanism for gamma-interferon-mediated inhibition of intracellular *Chlamydia* psittaci replication in T24 cells. Infect Immun 53:347–351. https://doi.org/10.1128/iai.53.2.347-351.1986

30. Banerjee A, Sun Y, Muramatsu MK, Toh E, Nelson DE, Roy CR. 2023. A member of an ancient family of bacterial amino acids transporters contributes to *Chlamydia* nutritional virulence and immune evasion. Infect Immun 91:e0048322. https://doi.org/10.1128/iai.00483-22

31. Muramatsu MK, Brothwell JA, Stein BD, Putman TE, Rockey DD, Nelson DE. 2016. Beyond tryptophan synthase: identification of genes that contribute to *Chlamydia trachomatis* survival during gamma interferon-induced persistence and reactivation. Infect Immun 84:2791–2801. https://doi.org/10.1128/IAI.00356-16

32. Belland RJ, Nelson DE, Virok D, Crane DD, Hogan D, Sturdevant D, Beatty WL, Caldwell HD. 2003. Transcriptome analysis of *Chlamydial* growth during IFN-γ-mediated persistence and reactivation. Proc Natl Acad Sci U S A 100:15971–15976. https://doi.org/10.1073/pnas.2535394100

33. Balakrishnan A, Patel B, Sieber SA, Chen D, Pachikara N, Zhong G, Cravatt BF, Fan H. 2006. Metalloprotease inhibitors Gm6001 and TAPI-0 inhibit the obligate intracellular human pathogen *Chlamydia trachomatis* by targeting peptide deformylase of the bacterium. J Biol Chem 281:16691–16699. https://doi.org/10.1074/jbc.M513648200

34. Anders S, Pyl PT, Huber W. 2015. Htseq--a python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169. https://doi.org/10.1093/bioinformatics/btu638

35. Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome Biol 11:R106–R106. https://doi.org/10.1186/gb-2010-11-10-r106

36. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578. https://doi.org/10.1038/nprot.2012.016

37. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43:D261–9. https://doi.org/10.1093/nar/gku1223