A family of viral satellites manipulates invading virus gene expression and affects cholera toxin mobilization

Zachary K Barth[1], Zoe Netter[1], Angus Angermeyer[1], Pooja Bhardwaj[2] Kimberley D Seed[1,3,*]

[1] Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA 94720, USA
[2] Department of Laboratory Medicine, University of California, San Francisco, San Francisco, CA
[3] Chan Zuckerberg Biohub, San Francisco, CA 94158, USA
* To whom correspondence should be addressed Tel: +15106647711, Email: kseed@berkeley.edu

Abstract

Many viruses possess temporally unfolding gene expression patterns aimed at subverting host defenses, commandeering host metabolism, and ultimately producing a large number of progeny virions. High throughput -omics tools, such as RNA-seq, have dramatically enhanced resolution of expression patterns during infection. Less studied have been viral satellites, mobile genomes that parasitize viruses and have far reaching effects on host-cell fitness. By performing RNA-seq on infection time courses, we have obtained the first time-resolved transcriptomes for bacteriophage satellites during lytic infection. Specifically, we have acquired transcriptomes for the lytic *Vibrio cholerae* phage ICP1 and all five known variants of ICP1's parasite, the Phage Inducible Chromosomal Island-Like Elements (PLEs). PLEs rely on ICP1 for both DNA replication and mobilization, and abolish production of ICP1 progeny in infected cells. We investigated PLEs impact on ICP1 gene expression and found that PLEs did not broadly restrict or reduce ICP1 gene expression. A major exception occurred in ICP1's capsid morphogenesis operon, which was downregulated by each of the PLE variants. This transcriptional manipulation, conserved among PLEs, has also evolved independently in at least one other phage satellite, suggesting that viral satellites may be under strong selective pressure to reduce the capsid expression of their larger host viruses. Surprisingly, PLEs were also found to alter the gene expression of CTXϕ, the integrative phage that encodes cholera toxin and is necessary for virulence of toxigenic *V. cholerae*. One PLE, PLE1, upregulated CTXϕ genes involved in replication and integration, and boosted CTXϕ mobility following induction of the SOS response. Our data show that PLEs exhibit conserved manipulation of their host-phage's gene expression, but divergent effects on CTXϕ, revealing that PLEs can influence both their hosts' resistance to phage and the mobility of virulence encoding elements.

Introduction

Viruses are selfish genetic elements that reprogram their host cells for viral reproduction. Turning host cells into viral factories requires viruses to implement both their own tightly regulated gene expression programs and manipulations of host gene expression. Viral genomes can vary from just a couple genes [1] to sizes rivaling those of cellular life [2], so the gene expression strategies of viruses are highly varied. Viral lifecycles exist on a continuum of agency. Some, like the cholera toxin phage (CTXϕ), are relatively passive. CTXϕ exists as an integrated prophage within toxigenic *Vibrio cholerae*. CTXϕ is largely regulated by host stress and virulence regulons, producing cholera toxin during *V. cholerae* infection of mammalian hosts, and replicating during the *V. cholerae* SOS response to DNA damage [3,4]. Aside from coding the two cholera toxin subunits, CTXϕ possesses a minimalist genome with just seven

additional genes, five of which are structural or involved in virion morphogenesis. Upon induction, CTXϕ initiates its replication off the host chromosome. Assembled particles are released through host secretion machinery without killing the cell, allowing horizontal and vertical CTXϕ propagation [4]. In contrast, many lytic phages have deadly mechanisms to shut down host gene expression and maximize expression of their own genes [5]. These mechanisms often unfold in a concerted and controlled manner to give rise to tight, temporal patterns of gene expression during infection; as has long been evidenced through targeted studies, and more recently through global analyses such as RNA-seq [6–12].

Less explored are the transcriptional patterns of viral satellites. These subviral elements parasitize viruses in a similar way to how viruses parasitize their host cells. Like viruses, viral satellites are found in all domains of life and impact their hosts in profound ways. Viral satellites can partially or completely abrogate virion production by the viruses they parasitize [13–15], and can reduce or worsen disease in multicellular organisms [16,17]. Unicellular organisms can be protected against viruses on the population level by endogenous viral satellites, but the efficacy of protection varies depending on the specific virus and satellite genotypes, and infection context [14,18–20]. Given their broad distribution and importance for both their cellular and viral hosts, it is desirable to decipher how the reproductive programs of viral satellites intersect with and differ from the programs of the viruses they parasitize.

A prime model for mechanistic and evolutionary insights into viral satellites are the Phage Inducible Chromosomal Island-Like Elements (PLEs) found in toxigenic *V. cholerae*. PLEs parasitize ICP1 [14], a lytic myophage that is the predominant phage in cholera patient stool samples [21]. Following ICP1 infection, PLEs excise from the host chromosome, and replicate to high copy [14]. Successful PLE parasitism does not abrogate cell lysis, but results in the complete restriction of ICP1 and the release of PLE transducing particles [14] (S1 Fig). The tractable genetics of *V. cholerae* facilitates mechanistic studies of PLE gene products, and insights have been gained regarding chromosomal excision of PLE [22], PLE DNA replication [23,24], and PLE manipulation of lysis kinetics during infection [25]. Notably, ICP1 genome editing is also possible [26], allowing manipulation of both sides of this host-parasite relationship.

To date, five distinct PLEs have been identified within the genomes of *V. cholerae* isolates recovered from cholera patient stool samples dating back to the 1940s [14]. Each individual PLE occurs in isolation; no *V. cholerae* isolate has been found to harbor more than one PLE, and PLEs typically dominate for a time before disappearing and being succeeded by a new PLE genotype [14]. Four of the five known PLEs are integrated into repeats of the superintegron, an array of selfish and mobile elements in the *V. cholerae* small chromosome [14]. PLEs mobility, along with the extensive and growing library of ICP1 isolates, allows PLEs to be compared in shared strain backgrounds during infection by contemporaneous and non-contemporaneous ICP1 isolates [14,22,24]. These experiments have shown that PLE and ICP1 are engaged in a co-evolutionary arms race, with different pairings of PLEs and ICP1 isolates having different infection outcomes. The sole understood method through which ICP1 can overcome PLEs is the ICP1 encoded CRISPR-Cas system, and deletion of that system broadens the PLE and ICP1 interactions that can be studied [14,19,27]. Pairing PLEs against the same host-virus in an isogenic host-cell background allows us to probe for convergence and divergence in how PLEs exploit and restrict ICP1. Thus, the ICP1-PLE system is a powerful model for exploring co-adaptations between a virus and its satellite and is unparalleled for tracking how these adaptations have shaped the evolution of these warring elements.

So far, few insights have been gained into the gene expression programs of ICP1 and PLEs. PLE1 expresses its integrase in uninfected cells [22], expression of PLE1's replication initiator, RepA, is induced following infection of ICP1 [23], and the PLE's lysis modulator, LidI, is detectable by Western blot late during ICP1 infection [25]. The PLE integrase's recombination directionality factor (necessary for directing integrase excision activity) is PexA, an ICP1 protein whose native function is unknown, but whose expression can be detected by 5 minutes post infection [22]. While these limited observations have provided insight into key PLE and ICP1 genes, the gross expression patterns of ICP1 and PLEs remain unknown. Until now, we have not known the degree to which ICP1 alters cellular expression patterns, and whether PLEs alter ICP1's gene expression or reproduce and restrict ICP1 without such alterations. To address these questions, we performed RNA-seq on *V. cholerae* infected by ICP1 over the course of the infection cycle. We sequenced the transcriptome of ICP1 infection in the absence of PLEs, as well as in the presence of each of the five PLEs. This work deciphers ICP1's transcriptional program, the transcriptional program of each PLE, and the *V. cholerae* host transcriptome in each infection context. To our knowledge, this is the first detailed analysis of a viral satellite transcriptome during infection. Following ICP1 infection PLEs exhibit remarkable conservation of temporal transcription patterns and targeted alteration of ICP1 transcription. The patterns described here suggest that like many viruses, viral satellites like PLE have evolved to carefully coordinate gene expression. In contrast, when we compared uninfected PLE strains, we observed disparate interactions between PLEs and other mobile genetic elements in the *V. cholerae* genome. Surprisingly, most PLEs increase expression of the CTXϕ repressor *rstR*; however, the most recently circulating PLE, PLE1, upregulates CTXϕ's replication and integration factors, which we show enhances the mobility of CTXϕ. Collectively our findings show that successive PLEs have conserved interactions with ICP1 and divergent interactions with CTXϕ, providing insights into how satellites manipulate the gene expression of their host viruses, and how they shape the evolution of their host cells.

Results and Discussion

V. cholerae's response to ICP1 infection

The ICP1 infection cycle takes approximately 20 minutes to produce a burst of nearly 90 infectious virions [14]. To capture the temporal range of ICP1's infection cycle, we took samples for RNA sequencing immediately prior to infection and 4, 8, 12, and 16 minutes post infection. Producing a large number of virions in a short period of time would presumably require substantial changes to the host cell transcriptome, and we see such changes occur during ICP1 infection. At 4 minutes post infection there are already dramatic changes to *V. cholerae's* transcriptome, with 17.2% (658/3827) of genes differentially regulated compared to uninfected cells ($Q \leq 0.1$) (Fig 1A and S2 Table). At this 4 minute timepoint, slightly more host transcripts are predicted to be upregulated (345) than downregulated (313). When *V. cholerae* gene expression across infection is normalized to transcripts per kilobase million (TPM) and visualized by heatmap, it appears that the bulk of *V. cholerae* genes are decreasing in transcript abundance following infection, and a small subset is being upregulated (Fig 1B). We interpret the difference between our significant differential expression analysis results and TPM normalized expression profile to result from the differential expression analysis assuming a negative binomial distribution for gene expression changes [28]. While such a model is appropriate for most RNA-seq applications where the majority of genes are not differentially expressed, it may underreport the extreme transcriptional changes that can occur during lytic viral infection. These extreme changes are reflected by the changes in reads abundance over the course of infection. By 4 minutes post infection, ICP1 contributes to more than a quarter of total RNA reads within the culture, and by 16 minutes post infection, ICP1 reads comprise more than 80% of total reads in the culture (Fig 1C). These changes appear even more extreme when

the relative size of the *V. cholerae* and ICP1 genomes are taken into account (Fig 1D). Taken together, we interpret this data to show that ICP1 affects *V. cholera*e transcription by globally reducing *V. cholerae* gene expression while upregulating the expression of a subset of genes.

V. cholerae's most downregulated genes are enriched for tRNA and rRNA processing genes, while several different gene groups are enriched among upregulated transcripts (Fig 1A and S2 Table). The most dramatic differential expression of *V. cholerae* genes occurred in the ArgR regulon responsible for arginine biosynthesis. Arginine biosynthesis genes were highly upregulated upon ICP1 infection, with *argB, argC,* and *argF* expression increasing more than 100-fold at 4 minutes post infection (Fig 1A and S2 Table). Similarly, the arginine transport genes were the most highly upregulated genes encoded in the *V. cholerae* small chromosome. The Na^+/H^+ antiporter *nhaA* was also strongly upregulated. Increases were also seen for genes relating to other amino acid biosynthesis and transport, sulfur compound metabolism, ATP production, flagellar synthesis and motility, and cell division (S2 Table). In contrast to the decrease in tRNA and rRNA processing genes, we see an increase in ribosomal protein coding genes and other genes involved in translation (Fig 1A and S2 Table). Several functional gene classes had many members upregulated and downregulated, for example a large number of genes involved in transport were both up and downregulated, likely reflecting the difference in metabolic needs for lytic virus production versus normal cell growth. We also saw differential expression of cell envelope genes, with a decrease in some lipopolysaccharide biosynthesis genes and an increase in the mannose-sensitive hemagglutinin (MSHA) pilus associated with estuarine growth (Fig 1A and S2 Table) [29,30].

It is hard to predict with certainty which transcriptional changes are a defensive host response to infection and which are due to transcriptional manipulation on the part of ICP1. Many of the differentially expressed genes relate to phenotypes regulated by cyclic di-GMP (di-cGMP) [31]. Recently, another cyclic dinucleotide, cyclic GMP-AMP (cGAMP), which was first discovered in *V. cholerae* [32], has been linked to phage defense [33]. It is interesting to consider that ICP1 infection may be influencing cyclic di-nucleotide signaling by triggering a host defense system, though in this case the defense is not successful. Regardless of the source of these alterations, the expression changes we see are not fully consistent with a shift toward high or low levels of either cyclic dinucleotide. We see several gene changes consistent with high di-cGMP [31] (increased MSHA biosynthesis, increased cold shock, decreased heat shock, and increased type VI secretion expression (S2 Table)) and some changes consistent with low di-cGMP [31] (increased flagellar synthesis and increased expression of the virulence regulator ToxT (S2 Table)). An increase in MSHA is also consistent with low cGAMP, but we also see downregulation of some chemotaxis associated genes, and chemotaxis is repressed by high cGAMP [32]. These changes suggest that the host cell may be receiving competing regulatory inputs that may act above or below the level of cyclic di-nucleotide signaling.

The strongly increased expression of arginine metabolism and transport genes is especially curious and could have positive or negative effects on phage production. ICP1 may be upregulating arginine metabolism to drive the production of polyamines. Arginine often serves as the precursor for polyamine synthesis [34]. Purines may be a limiting resource for phage genome replication, while polyamines are found in the capsids of multiple phages, potentially aiding in DNA condensation, packaging, and ejection [35,36]. We do not see upregulation of genes specific for synthesis of the most common polyamine, putrescine, but the ornithine decarboxylase which is responsible for putrescine synthesis is regulated post-translationally in *Escherichia coli* and mammals to allow rapid adjustment of polyamine pools [34,37]. We also see upregulation of the *V. cholerae* inducible lysine decarboxylase (S2 Table), and in at least one bacterial system, a lysine decarboxylase is able to use ornithine as a substrate [38].

Alternatively, arginine has been found to have a deactivating effect on phage virions under certain conditions [39–41], so the upregulation of arginine may be an attempt by the host cell to curb phage production.

Establishing the ICP1 transcriptional program

Once we examined ICP1's effect on *V. cholerae* transcription, we next sought to document ICP1's transcriptional program. ICP1 has an approximately 126kb genome and more than 200 predicted open reading frames [42]. Less than a quarter of ICP1's putative coding sequences have activities or functions that can be predicted through bioinformatic analysis. As is common for viruses, ICP1's genes fall into distinct temporal groupings based on the timing of peak expression. The early genes, which we define as those that show peak expression at 4 minutes post infection, consist mostly of short genes averaging less than 330bp and encoding hypothetical proteins (Fig 2 and S3 Table). It is difficult to infer the function of these genes, but there are several short immediate early genes in other phage systems that are known to have a role in host cell takeover [43,44]. The next grouping, middle early genes with peak expression at 8 minutes post infection, is mostly comprised of nucleotide metabolism genes including ICP1's DNA polymerase (Fig 2 and S3 Table). The 12 minute time point captures a transitional period in ICP1 transcription. High expression of the middle early genes continues, while expression of late genes has begun but not yet peaked (Fig 2). Few genes hit peak expression at 12 minutes, though an exception are a subset of nucleotide metabolism genes (within the range *gp176–gp211*). Finally, late genes with peak expression at 16 minutes post infection are comprised primarily of putative structural genes and genes known to be involved in lysis (Fig 2 and S3 Table), [25]. The lysis, capsid, and tail genes occur in three separate clusters, all encoded on the (-) strand (Fig 2 and S3 Table). Previously, ICP1 rolling circle replication was predicted to proceed in the (-) direction based on late infection DNA coverage skews [23]. Since ICP1's late genes are transcribed off of the (-) strand, they may be transcribed off of the rolling circle replication template. Such an arrangement would be consistent with the preference for co-directional transcription and DNA leading strand synthesis observed in many bacterial systems [45], and thought to help preserve genome integrity by avoiding replication and transcription conflicts [46]. Overall, the expression patterns we see for ICP1 are consistent with what is known about lytic phage development in general and ICP1's lifecycle in particular.

PLEs exhibit a conserved transcriptional program

Having established the transcriptional program of ICP1, we next sought to examine transcriptional patterns in PLEs. The five PLEs share a similar gene organization. Proximal to the PLE's left attachment site is a gene cluster (denoted here C_{int}) that includes the PLE integrase, *int*. Immediately downstream of this gene cluster is PLE's sole cluster of negative strand encoded genes, here called C_L for 'left' cluster (Fig 3A). This (-) sense cluster flanks an approximately 3kb non-coding region, the right most quarter of which contains the PLE origin of replication [23]. Flanking the PLE origin of replication is a putative *marR*-like gene, and this is followed by two additional rightward facing clusters (C_{R1} and C_{R2}). Exceptions to this arrangement occur with PLE1's standalone *int*, a transposase in PLE2's large non-coding region, and the absence of a *marR*-like gene in PLE3 [14].

Consistent with the PLEs providing interference against this ICP1 isolate [14], we find that all PLEs are transcriptionally activated following infection. Paralleling PLEs' organizational similarities, the transcriptional patterns of PLEs are highly conserved once each PLE has been activated. In uninfected samples, PLEs show some expression of *int*, and, if it is present, the *marR*-like gene, while expression across C_L, C_{R1}, and C_{R2} is variable and often uneven (Fig 3B and S2-S5 Figs). Read counts for PLE genes remain low at 4 minutes post infection (S1 Data), and then at 8 minutes post infection transcriptional patterns start to emerge. For each PLE, C_{R1}

is strongly expressed at 8 minutes post infection, followed by strong expression of C_L at 12 minutes, and high C_{R1} expression at 16 minutes along with the *marR*-like gene (Fig 3 and S2-S5 Figs). It should be noted that all PLE transcripts continue to increase in abundance over the course of infection, though we see differences in the timing of peak gene transcription relative to other genes. This sustained global increase can likely be attributed to increased PLE copy number, as PLE replicates upwards of 1000-fold during ICP1 infection [14].

Early expression of C_{R1} is not surprising, given that this cluster contains PLE's *repA* gene which is necessary for PLE replication [23], and PLE1 replication was previously found to begin before 15 minutes post infection [14]. Interestingly, a highly conserved C_{R1} gene, *orf12.1* in PLE1, has a different expression pattern than the rest of the cluster, peaking at 12 minutes post infection instead of 8. This suggests that *orf12.1* may be under different regulation than the rest of C_{R1}, and may be involved in the transition from early to late PLE gene expression. Overall, the conserved timing of expression of syntenic gene clusters across PLEs suggest that each cluster serves a distinct role in parasitizing ICP1. We suspect that the timing of PLE gene cluster expression has evolved to take advantage of ICP1's own transcriptional program. Coordination between PLE gene expression and ICP1's gene expression would be consistent with PLEs' reliance on ICP1 gene products for key steps of the PLE lifecycle [22–25].

Non-coding RNAs are abundant in ICP1 and PLEs

Non-coding RNAs (ncRNAs) are a prominent feature of both ICP1 and PLE gene transcription. Surprisingly, the most abundant transcripts in the ICP1 and PLE transcriptomes are both predicted to be non-coding. The most abundantly expressed ICP1 transcript is encoded in an approximately 1kb orf-less space between *gp139*, the start of ICP1's lysis cluster [25], and *gp140* (Fig 4). The length of this transcript (~800bp based on the RNA-seq coverage), is comparable to the Giant- Ornate- Lake- and Lactobacialles-Derived (GOLLD) and the Rumen-Originating, Ornate, Large (ROOL) RNAs that have been found in many phage genomes [47]. These RNAs are frequently encoded near tRNAs, although this is not the case for ICP1 since it does not possess any tRNAs. The role of these large ncRNAs is unknown. In one *Lactobacillus brevis* prophage, a GOLLD RNA was found to accumulate during lytic infection, but was dispensable for phage production [47].

In each PLE, the most abundant transcript is located between C_{R1} and C_{R2} (Fig 5 and S6-S9 Figs). This transcript occurs between a set of inverted repeats, and for this reason, we have tentatively named the transcript the inter-inverted repeat (IIR) transcript (S10 Fig). Non-coding RNAs flanked by terminal inverted repeats often occur in miniature inverted-repeat transposable elements (MITEs), and these have been examined in several bacterial species [48]. This similarity suggests that the PLE IIR may have evolved from a selfish mobile ancestor. The PLE IIR transcript also has antisense homology to the leader sequences of several PLE ORFs (S10B Fig). Complementarity between a ncRNA and gene leader sequences is seen in the regulatory RNAs from phages P1, P7, and N15, as well as the phage satellites P4 and ϕR73 [49–51]. With the exception of N15's RNA which promotes the lytic cycle, these regulatory RNAs function to repress the lytic cycle. These roles appear unlikely in PLE, since there is an accumulation of the PLE IIR transcript as infection progresses, but it is not abundant until 12 minutes post infection, which is after PLE early gene expression (Fig 3B, S2-S5 Figs). The PLE IIR transcript's expression pattern appears more consistent with modifying gene expression during infection rather than acting as a global activator or repressor of PLE activity.

Additional transcripts without predicted coding capacity occur in both ICP1 and PLEs. An approximately 300bp region between *gp135* and *gp136* is transcribed in ICP1 during 12 and 16 minutes post infection (Fig 4). Within each PLE, we see abundant transcription approximately

200bp upstream of C_L and 150bp upstream of C_{R1} when these clusters are transcriptionally active (Fig 5 and S6-S9 Figs). Though these 5' UTRs are not conserved on the sequence level, their occurrence in every PLE suggests conservation of function. The 5' UTR of transcripts is a common site for riboswitches [52] suggesting that these untranslated sequences may regulate expression of their downstream genes.

PLE-host interactions in the uninfected cell

By comparing transcriptomes of PLE(+) and PLE(-) strains prior to infection, we were able to determine whether PLEs affected transcription in their *V. cholerae* host prior to phage infection. All PLEs altered the transcription of genes neighboring the PLEs' integration sites (Fig 6A and S11 Fig). Additionally, PLEs altered the expression of several genes within the *V. cholerae* superintegron, including multiple toxin-antitoxin systems (S4 Table). Altered expression of superintegron encoded genes also occurred in the PLE2(+) strain, despite PLE2 being integrated outside of the superintegron [14]. These transcriptional changes may reflect cross talk between PLE-encoded genes or genes flanking the PLE integration site, and other genes encoded within the superintegron. In *V. cholerae*, both mobile genetic elements and genes in the superintegron are known to be repressed by the bacterial chromatin protein H-NS [53–55]. An alternative explanation for why we see changes in superintegron gene expression is that PLE integration alters the nucleoid architecture of this region through its own recruitment of H-NS. Notably, many of the superintegron genes are multi-copy, so the number of differentially regulated genes may be over-reported if the reads mapping cannot differentiate reads from multi-copy genes from different loci. This explanation seems particularly likely for scenarios where genes that are distal to the PLE integration site are predicted to be differentially expressed, and have paralogs located proximal to the integration site. Such a pattern is seen with PLEs 3 and 5 (S4 Table), but can only explain a small number of the differences we see in the superintegron.

Rather surprisingly, we also found that PLEs alter the expression of genes encoded by integrative mobile elements exploiting Xer (IMEXs) (S11 Fig). IMEXs are mobile elements that utilize host Xer recombinases to integrate into the chromosome dimer resolution or *dif* sites of their host-cells, located near the chromosome replication terminus [56]. The strain of *V. cholerae* used in this study has three separate IMEX's: CTXϕ, CTXϕ's satellite RS1ϕ, and the toxin-linked cryptic element (TLC), integrated twice in tandem next to CTXϕ (S12 Fig). RS1ϕ is largely redundant with sequence within CTXϕ. RS1ϕ and CTXϕ have their own copies of the replication initiator *rstA,* a gene required for integration named *rstB*, and the CTXϕ master repressor *rstR* [3,4]. All these IMEXs are integrated on chromosome I, so the PLEs transcriptional effects on these elements must be acting *in trans*. Relative to the PLE(-) strain, most PLEs increased expression of both the RS1ϕ and CTXϕ copies of *rstR*. These same strains also showed upregulation of the TLC gene *tlcR.* An exception to this pattern occurred in PLE1, where rather than *rstR* upregulation we observed upregulation of the CTXϕ replication genes *rstA* and *rstB* (Fig 6A). This observation prompted us to question whether upregulation of *rstA* and *rstB* could prime CTXϕ for mobilization following activation of *V. cholerae's* SOS response. To test this, we used an antibiotic marked copy of CTXϕ and found that following mitomycin C treatment, the presence of PLE1 increased the production of CTXϕ transducing units 10-fold relative to a strain with no PLE or a strain with PLE2. (Fig 6B). These results reveal potentially far reaching effects that mobile elements can have on each other, as well as the hosts they share. Beyond providing the host-cell population with immunity to ICP1 phages, PLEs' integration can enhance the mobility of other mobile genetic elements and, by extension, virulence genes. Thus, PLEs may affect *V. cholerae* fitness in ways that are distinct from their own anti-phage activity and relevant to cholera epidemiology.

PLEs selectively manipulate ICP1 transcription

Having detailed ICP1's and PLEs' transcriptional programs and PLEs' transcriptional effects in uninfected cells, we sought to evaluate whether PLE disrupted ICP1 gene expression. Remarkably, although PLEs abolish ICP1 production [14] we found that PLEs do not broadly restrict or alter ICP1 transcription. At 16 minutes post infection during maximum PLE expression, PLE transcripts comprise roughly 10% of the transcriptome while ICP1's proportion of reads still sits around 80% (Fig 7A). When normalized to the genome size, PLE1 TPM approaches parity with that of ICP1, while the other PLEs achieve a bit less (Fig 7B and S13-S16 Figs) indicating that the relative transcriptional activity of PLEs do not exceed that of ICP1. This is in stark contrast to previously reported DNA levels at 16 minutes post infection, where the amount of PLE1 DNA exceeds that of ICP1, and overall ICP1 DNA replication is substantially reduced by PLE1 [23]. During ICP1 infection of PLE1 positive cells there is a loss of ICP1's rolling circle replication [23]. PLE1's disparate effects on ICP1 replication and transcription can be reconciled by the model that ICP1 late genes are transcribed off the same strand of DNA that serves as the template for rolling circle replication. A block in ICP1 rolling circle replication would not impede ICP1 transcription if the newly synthesized DNA is not expressed, but would permit PLE to interfere with ICP1 packaging as has been hypothesized [23].

In addition to potential mechanistic explanations for how PLE is able to substantially restrict ICP1 replication without broad disruptions of ICP1 transcription, it is interesting to consider why such a discrepancy in PLE activity would be favorable. PLEs are not very transcriptionally active until 8 minutes post infection (Fig 3A, S2-S5 Figs, and S1Data). From 8 minutes post infection onward, ICP1 transcripts are comprised largely of nucleotide metabolism genes, genes involved in DNA replication, and genes encoding virion structural components (Fig 2). Aside from RepA, the replication initiation factor that directly interacts with the PLE origin of replication, PLE does not appear to encode dedicated replication machinery [23]. Further, PLE has been shown to rely on at least some ICP1 gene products for replication [23,24]. PLE also does not encode identifiable structural genes, and requires the same viral receptor as ICP1 for mobilization [14], suggesting that like other phage satellites [57], PLE is packaged within the same virion structural components as its host-phage. The reliance on ICP1's virion production machinery incentivizes PLE to allow ICP1's transcriptional program to progress relatively unperturbed, since ICP1 is already producing the infrastructure for robust virion production. Our data suggests that rather than suppressing the production of ICP1 machinery and replacing it with PLEs' own, PLEs efficiently parasitize that machinery, redirecting it to PLE's own genome and somehow excluding ICP1. Such a strategy would allow PLE to benefit from ICP1's own reproductive adaptations, while restricting ICP1 propagation among the host-cell population.

While PLEs may benefit from permitting ICP1 gene expression, they also might benefit from re-tuning certain aspects of ICP1's transcriptional program to better fit their needs. Consistent with this hypothesis, we observed that one set of ICP1 genes has markedly decreased expression in the presence of all PLEs: *gp122-gp126* (Fig 7C). This group of genes is predicted to be responsible for ICP1 capsid morphogenesis. The genes *gp122* and *gp123* are predicted to encode ICP1's major capsid protein and a capsid decoration protein, respectively, and *gp125* is a predicted protease, likely providing the proteolysis necessary for procapsid maturation as occurs in most tailed phages [58]. Reduced expression of this gene cluster is seen in all PLEs, though the Q values we obtained from the differential analysis for PLE2 and PLE4 were less robust (S13-S16 Figs). Nevertheless, when the production of ICP1's major capsid protein was assessed in the presence of each PLE via Western blot, we found that consistent with the differential transcription observed, all PLEs reduce ICP1 capsid production between 2 and 3-fold

(Fig 7E and S17 Fig). These results show that specific reduction in capsid production is a conserved activity among PLEs.

The PLEs' downregulation of ICP1 capsid genes may be evolutionarily tied to remodeling of the virion capsid. Capsid remodeling, more specifically reduction in capsid dimensions, is a well-studied feature of phage satellite parasitism, having arisen in the phage satellite P4, the Gram-positive Phage Inducible Chromosomal Islands (PICIs), and the Gram-negative PICI's [57,59], though it has yet to be shown for PLEs. Because capsid remodeling has been observed in the three other lineages of tailed phage satellites, capsid remodeling may be a general feature of phage satellite biology, and is likely to be induced by PLEs. Capsid remodeling can restrict host-viruses by assembling virions that are too small for the full-sized genome. Additionally, capsid remodeling likely increases the horizontal mobility of satellites. PLEs are not predicted to encode their own large terminase, suggesting that they rely on ICP1 packaging machinery in addition to structural components. Given that ICP1's genome is about 7 times the length of PLE, and ICP1 is predicted to package virions in a head-full fashion [23], packaging PLEs into ICP1 sized capsids would result in an overabundance of PLE genomes in a single transducing particle. Excess genome packaging would reduce the number of transducing particles that could be assembled for a set level of PLE genome replication.

While reducing capsid size would benefit PLE by reducing the amount of excess PLE DNA packaged per PLE transducing particle, further benefits could be gained by tuning gene expression to reduce production of excess capsid. Prior analyses have shown that protein translation is expected to be the most energy intensive process for viruses within PLE and ICP1's genome size range [60,61], and the major capsid transcript is by far the most abundant mRNA produced by ICP1 (Fig 2). Reducing the amount of capsid produced could help PLEs' recoup the costs of PLE gene expression and additional DNA replication which do not occur during ICP1 infections in the PLE(-) background.

Much as capsid remodeling is a recognized occurrence among tailed phage satellites, our results reveal an emerging pattern of satellite elements tuning down expression of phage late genes. Downregulation of capsid, at least relative to other structural genes, has arisen independently among multiple phage satellites. Satellite P4 encodes a gene named *psu* for polarity suppression unit. Psu acts as a coat decoration protein for the P2 capsid when remodeled by P4, increasing the stability of remodeled capsids [62]. Additionally, Psu is well characterized as a Rho-binding antitermination factor [63–65]. P2 regulates its structural genes through transcriptional attenuation: transcription frequently terminates before reading through the entire operon, so the genes near the front of their operon, including the capsid and scaffold, are expressed to a greater extent than downstream genes [66]. By preventing Rho-dependent termination, P4 retunes structural gene expression so that the ratio of capsid and scaffold relative to other structural genes is reduced, with the ratio of capsid and scaffold to terminase being 5 to 10-fold lower in the presence of P4 [67]. Initially, it was proposed that P4's tuning of structural gene transcription was the mechanism through which P2 capsids were remodeled, but later, remodeling was found to actually be caused by the Sid capsid scaffold, encoded in the same three-gene operon as Psu [68]. Though Psu has undergone extensive biochemical characterization, little has been uncovered about the evolutionary importance of P4 induced antitermination since it was found to be dispensable for capsid remodeling.

SaPIs provide an additional example of phage satellites repressing host-phage genes. In the SaPI host-phage 80α and the related phage 80, late genes are organized into a putative operon starting with the small terminase encoding gene *terS*, followed by additional packaging genes, then head morphogenesis genes, tail morphogenesis genes, and finally lysis genes. Some

SaPIs have been found to repress *terS*, and it was inferred that there was also repression of the late gene operon [69] . It was noted that repression must be incomplete, since SaPIs rely on phage structural, lysis, and large terminase genes for their propagation. While repression of *terS* could benefit SaPIs by preventing the packaging of host-phage genomes, complete repression of the late operon would block SaPI particle production. Notably, only *terS* expression was measured for this operon. If internal promoters exist in this late operon, such that *terS* is silenced by SaPIs, head morphogenesis genes are repressed to an intermediate degree, and tail and lysis genes are unaffected, it would be consistent with the pattern of head morphogenesis repression seen in P4 and PLE, as well as the reproductive needs of SaPIs.

Conclusions

Here, we have provided the first study of phage satellite transcriptomics, obtaining transcriptional programs for ICP1, as well as all five variants of PLE. Aside from broadening our understanding of ICP1 and PLE biology, this work provides surprising insights into the biology of phage satellites. PLE integration alters the gene expression of other mobile elements in *V. cholerae*. Notably, one PLE increases the mobility of CTXϕ, showing that viral satellites can affect the spread of virulence genes that are encoded by unrelated mobile elements. More directly related to the PLE lifecycle, we discovered that PLEs do not induce large scale changes to ICP1's transcriptome, suggesting that PLE has adapted to take advantage of ICP1's lytic program as it occurs under conditions permissive to ICP1 replication. The notable exception is that PLEs downregulate ICP1's capsid morphogenesis operon, and this activity is conserved among PLEs, and convergently evolved in the phage satellite P4. This pattern of evolution suggests a strong selective pressure for viral satellites to tune capsid expression, perhaps as a means to optimize resource use for satellite spread. It will be interesting to see if the patterns established here extend to other viral satellites and what other surprising aspects of viral satellite biology will emerge in the future.

Materials and Methods

Strains and culture conditions

All *V. cholerae* strains, including PLE(+) variants, used in this study are derived from E7946 to ensure comparisons in an otherwise isogenic background. Bacteria were grown on LB agar plates and in LB broth with aeration at 37°C. A detailed list of all strains used throughout this study can be found in Supplementary Table 1. ICP1_2006_E engineered to lack CRISPR-Cas (ΔCRISPR, *Δcas2-3*) [22] was used for all experiments. Phage titers were determined using a soft agar overlay method wherein ICP1 was allowed to adsorb to *V. cholerae* for 10 min at room temperature before the mixture was added to molten LB soft agar (0.3%) and poured onto 100 mm × 15 mm LB agar plates. Plaques were counted after overnight incubation at 37°C.

Generation of mutant strains and constructs

V. cholerae mutants were generated through natural transformation as described previously [70]. For antibiotic marked gene knockouts and overexpression constructs, splicing by overlap extension (SOE) PCR was used.

Sample collection for RNA-seq

Strains were grown to stationary phase in 2mL cultures before being back diluted to an OD_{600} of 0.05 in 6mL LB broth. Strains were then grown to an OD_{600} of 0.47 in 16x 150mm culture tubes with a Biochrom® Ultrospec 10 (equivalent to OD_{600} of 0.3 with a 1cm path length) before initial sample collection and phage infection at a multiplicity of infection (MOI) of 5. Immediately prior to infection, and then at 4, 8, 12, and 16 minutes post infection, 1mL of culture was taken and mixed with 1mL of ice cold methanol, before returning the remaining culture to the incubator. The sample and methanol mixtures were pelleted at 21,694 xg at 4°C for 2 minutes, aspirated,

washed with 1mL ice cold 1X phosphate buffered saline (PBS), and then pelleted and aspirated again. Pellets were snap frozen in liquid N_2 and stored at -80°C until RNA-isolation.

RNA-isolation
RNA was extracted from samples using the Purelink™ RNA Mini Kit (Thermo-Fisher) and DNA was removed from isolated RNA samples using the TURBO DNA-*Free*™ Kit (Thermo-Fisher).

cDNA Library generation and sequencing
RNA samples were submitted to the University of California Berkeley QB3 Core facility for cDNA library generation and sequencing. Ribosomal DNA was removed with an Illumina Ribo-Zero rRNA Removal Kit (Bacteria) prior to cDNA generation. An S220 Focused-Ultrasonicator (Covaris®) was used to fragment the DNA, and library preparation was performed using the KAPA Hyper Prep kit for DNA (KK8504). Truncated universal stub adapters were used for ligation, and indexed primers were used during PCR amplification to complete the adapters and to enrich the libraries for adapter-ligated fragments. Samples were checked for quality on an AATI (now Agilent) Fragment Analyzer. Samples were then transferred to the Vincent J. Coates Genomics Sequencing Laboratory, another QB3-Berkeley Core Research Facility at UC Berkeley, where Illumina sequencing library molarity was measured with quantitative PCR with the Kapa Biosystems Illumina Quant qPCR Kits on a BioRad CFX Connect thermal cycler. Libraries were then pooled evenly by molarity and sequenced on an Illumina HiSeq4000 150PE flowcell. Raw sequencing data was converted into fastq format, sample specific files using the Illumina bcl2fastq2 software on the sequencing centers local linux server system.

RNA-seq analysis
For each sample library, sequencing reads were mapped to separate *V. cholerae*, ICP1_2006E (ΔCRISPR, Δ*cas2-3*), reference files, as well as files for the appropriate PLE genome in CLC Genomics Workbench version 12. Default RNA-seq mapping settings were used, with the exception that multiple mapping of individual reads was disabled. As noted previously [9,12], RNA-seq of lytic infections possess specific challenges because there are multiple genomes (two in most cases, three in the presence of a viral satellite such as PLE), undergoing changes in both their share of total transcripts in culture, as well as the relative expression of their genes. To address this, gene expression was normalized on a per genome basis for differential expression analysis. Differential expression analysis was performed using the DESeq2 [28] R/Bioconductor package with default parameters. For data visualization heat maps of log2 TPM values were plotted using the aheatmap function from the NMF R package. Volcano plots were generated using the EnhancedVolcano package and function. Our reads counts and DESeq2 results are provided within our supporting information (S1 Data). The accession numbers for reference sequences used for mapping can be also be found in our supporting information (S5 Table).

For the generation of reads tracks, RNA Seq reads were mapped to the reference sequences using bowtie2 v2.3.4.1 [71], with the following settings: "--end-to-end --very-sensitive --no-unal --no-mixed --no-discordant". For each sample, read coverage was normalized to sequencing depth and replicates were then averaged. Sequence data for samples used in this work can be found in the Sequence Read Archive under the Bioproject ID: PRJNA609114.

CTX transduction assays
V. cholerae PLE(-), PLE1, and PLE2 CTX(+) donor strains were modified by replacing *ctxAB* with a kanamycin resistance cassette. CTXφ production was induced in these strains by growing up to OD_{600}=0.3 followed by 16 hour incubation at 37°C with aeration in LB supplemented with mitomycin C (20ng/mL) (Sigma). *V. cholerae* CTX(-) recipient strains were

engineered to harbor a cassette inserted in the *lacZ* locus containing a spectinomycin resistance gene and *toxT* under control of P_{tac} and a theophylline-inducible riboswitch. Recipient strains were grown to OD_{600}=0.3, then induced with addition of 1.5mM theophylline and 1mM IPTG at 37°C for 16 hours in LB+10mM $MgCl_2$ without agitation. After mitomycin C treatment, donor strains were centrifuged for 3 minutes at 5,000 x *g* twice to ensure maximum removal of donor cells from CTXφ-containing supernatant. Cleared donor supernatants were mixed with recipient cultures 1:4 and incubated at 37°C for 1 hour without agitation. Transduction mixtures were plated on LB plates supplemented with kanamycin (75µg/mL) and spectinomycin (100µg/mL) and incubated overnight at 37°C to quantify transductants.

Western blots

Isogenic *V. cholerae* strains either lacking PLE or with an integrated PLE (1-5) were grown to an OD_{600}=0.3 and infected with ICP1_2006E_△CRISPR *Δcas2-3* at MOI=1 and returned to the incubator at 37°C with aeration. At 16 minutes post phage addition, 1mL of infected culture was collected and mixed with an equal volume of ice-cold methanol. Samples were centrifuged at 5,000 x *g* for 10 minutes at 4°C to pellet infected cells. Pellets were washed once with ice-cold PBS and resuspended in lysis buffer (50mM Tris, 150mM NaCl, 1mM EDTA, 0.5% Triton X-100, 1x protease inhibitor (Thermo Pierce™ Protease and Phosphatase inhibitor tablet)). Protein concentration was quantified with Pierce BCA Protein Assay Kit (Thermo). 30µg of total protein sample was mixed with Laemmli buffer (Bio-rad) and boiled at 99°C for 10 minutes. Samples were run on Any-kD TGX-SDS-PAGE gels (Bio-rad) and transferred onto nitrocellulose membranes with Transblot Turbo Transfer system (Bio-rad). Custom primary peptide antibody generated in rabbits against ICP1 capsid (Gp122, YP_004251064.1) (GenScript) was diluted 1:1500. Band detection was conducted with a goat α-rabbit-HRP secondary antibody (Bio-rad) at 1:10,000 followed by development with Clarity Western ECL Substrate (Bio-rad) and imaging on a Chemidoc XRS Imaging System (Bio-rad).

Acknowledgements

We thank Andrew Camilli for sharing the CTX minus recipient strain. We thank the University of California Berkeley QB3 Core facility for cDNA library preparation and sequencing. We thank all members of the Seed lab past and present for useful discussions and particularly Kristen LeGault and Caroline Boyd for critical feedback. We also thank Bob Bender, Matt Chapman, and Lyle Simmons for helpful discussions.

References

1. Breitbart M, Delwart E, Rosario K, Segalés J, Varsani A, Ictv Report Consortium. ICTV virus taxonomy profile: circoviridae. J Gen Virol. 2017;98: 1997–1998. doi:10.1099/jgv.0.000871
2. Yutin N, Colson P, Raoult D, Koonin EV. Mimiviridae: clusters of orthologous genes, reconstruction of gene repertoire evolution and proposed expansion of the giant virus family. Virol J. 2013;10: 106. doi:10.1186/1743-422X-10-106
3. McLeod SM, Kimsey HH, Davis BM, Waldor MK. CTXphi and Vibrio cholerae: exploring a newly recognized type of phage-host cell relationship. Mol Microbiol. 2005;57: 347–356. doi:10.1111/j.1365-2958.2005.04676.x
4. Pant A, Das B, Bhadra RK. CTX phage of Vibrio cholerae: Genomics and applications. Vaccine. 2019; doi:10.1016/j.vaccine.2019.06.034
5. Wahl MC, Sen R. Exploiting phage strategies to modulate bacterial transcription. Transcription. 2019;10: 222–230. doi:10.1080/21541264.2019.1684137
6. Ceyssens P-J, Minakhin L, Van den Bossche A, Yakunina M, Klimuk E, Blasdel B, et al. Development of giant bacteriophage □KZ is independent of the host transcription

apparatus. J Virol. 2014;88: 10501–10510. doi:10.1128/JVI.01347-14
7. Leskinen K, Blasdel BG, Lavigne R, Skurnik M. RNA-Sequencing Reveals the Progression of Phage-Host Interactions between φR1-37 and Yersinia enterocolitica. Viruses. 2016;8: 111. doi:10.3390/v8040111
8. Chevallereau A, Blasdel BG, De Smet J, Monot M, Zimmermann M, Kogadeeva M, et al. Next-Generation "-omics" Approaches Reveal a Massive Alteration of Host RNA Metabolism during Bacteriophage Infection of Pseudomonas aeruginosa. PLoS Genet. 2016;12: e1006134. doi:10.1371/journal.pgen.1006134
9. Blasdel BG, Chevallereau A, Monot M, Lavigne R, Debarbieux L. Comparative transcriptomics analyses reveal the conservation of an ancestral infectious strategy in two bacteriophage genera. ISME J. 2017;11: 1988–1996. doi:10.1038/ismej.2017.63
10. Danis-Wlodarczyk K, Blasdel BG, Jang HB, Vandenheuvel D, Noben J-P, Drulis-Kawa Z, et al. Genomic, transcriptomic, and structural analysis of Pseudomonas virus PA5oct highlights the molecular complexity among Jumbo phages. BioRxiv. 2018; doi:10.1101/406421
11. Sacher JC, Flint A, Butcher J, Blasdel B, Reynolds HM, Lavigne R, et al. Transcriptomic Analysis of the Campylobacter jejuni Response to T4-Like Phage NCTC 12673 Infection. Viruses. 2018;10. doi:10.3390/v10060332
12. Blasdel BG, Ceyssens P-J, Chevallereau A, Debarbieux L, Lavigne R. Comparative transcriptomics reveals a conserved Bacterial Adaptive Phage Response (BAPR) to viral predation. BioRxiv. 2018; doi:10.1101/248849
13. Ram G, Chen J, Kumar K, Ross HF, Ubeda C, Damle PK, et al. Staphylococcal pathogenicity island interference with helper phage reproduction is a paradigm of molecular parasitism. Proc Natl Acad Sci USA. 2012;109: 16300–16305. doi:10.1073/pnas.1204615109
14. O'Hara BJ, Barth ZK, McKitterick AC, Seed KD. A highly specific phage defense system is a conserved feature of the Vibrio cholerae mobilome. PLoS Genet. 2017;13: e1006838. doi:10.1371/journal.pgen.1006838
15. La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, et al. The virophage as a unique parasite of the giant mimivirus. Nature. 2008;455: 100–104. doi:10.1038/nature07218
16. Hughes SA, Wedemeyer H, Harrison PM. Hepatitis delta virus. Lancet. 2011;378: 73–85. doi:10.1016/S0140-6736(10)61931-9
17. Simon AE, Roossinck MJ, Havelda Z. Plant virus satellite and defective interfering RNAs: new paradigms for a new century. Annu Rev Phytopathol. 2004;42: 415–437. doi:10.1146/annurev.phyto.42.040803.140402
18. Diana C, Dehò G, Geisselsoder J, Tinelli L, Goldstein R. Viral interference at the level of capsid size determination by satellite phage P4. J Mol Biol. 1978;126: 433–445. doi:10.1016/0022-2836(78)90050-5
19. Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature. 2013;494: 489–491. doi:10.1038/nature11927
20. Gaia M, Benamar S, Boughalmi M, Pagnier I, Croce O, Colson P, et al. Zamilon, a novel virophage with Mimiviridae host specificity. PLoS One. 2014;9: e94923. doi:10.1371/journal.pone.0094923
21. Seed KD, Bodi KL, Kropinski AM, Ackermann H-W, Calderwood SB, Qadri F, et al. Evidence of a dominant lineage of Vibrio cholerae-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. MBio. 2011;2: e00334-10. doi:10.1128/mBio.00334-10
22. McKitterick AC, Seed KD. Anti-phage islands force their target phage to directly mediate island excision and spread. Nat Commun. 2018;9: 2348. doi:10.1038/s41467-018-04786-5

23. Barth ZK, Silvas TV, Angermeyer A, Seed KD. Genome replication dynamics of a bacteriophage and its satellite reveal strategies for parasitism and viral restriction. Nucleic Acids Res. 2019; doi:10.1093/nar/gkz1005
24. McKitterick AC, Hays SG, Johura F-T, Alam M, Seed KD. Viral satellites exploit phage proteins to escape degradation of the bacterial host chromosome. Cell Host Microbe. 2019;26: 504–514.e4. doi:10.1016/j.chom.2019.09.006
25. Hays SG, Seed KD. Dominant *Vibrio cholerae* phage exhibits lysis inhibition sensitive to disruption by a defensive phage satellite. BioRxiv. 2019; doi:10.1101/790493
26. Box AM, McGuffie MJ, O'Hara BJ, Seed KD. Functional Analysis of Bacteriophage Immunity through a Type I-E CRISPR-Cas System in Vibrio cholerae and Its Application in Bacteriophage Genome Engineering. J Bacteriol. 2016;198: 578–590. doi:10.1128/JB.00747-15
27. McKitterick AC, LeGault KN, Angermeyer A, Alam M, Seed KD. Competition between mobile genetic elements drives optimization of a phage-encoded CRISPR-Cas system: insights from a natural arms race. Philos Trans R Soc Lond B, Biol Sci. 2019;374: 20180089. doi:10.1098/rstb.2018.0089
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15: 550. doi:10.1186/s13059-014-0550-8
29. Hsiao A, Liu Z, Joelsson A, Zhu J. Vibrio cholerae virulence regulator-coordinated evasion of host immunity. Proc Natl Acad Sci USA. 2006;103: 14542–14547. doi:10.1073/pnas.0604650103
30. Utada AS, Bennett RR, Fong JCN, Gibiansky ML, Yildiz FH, Golestanian R, et al. Vibrio cholerae use pili and flagella synergistically to effect motility switching and conditional surface attachment. Nat Commun. 2014;5: 4913. doi:10.1038/ncomms5913
31. Conner JG, Zamorano-Sánchez D, Park JH, Sondermann H, Yildiz FH. The ins and outs of cyclic di-GMP signaling in Vibrio cholerae. Curr Opin Microbiol. 2017;36: 20–29. doi:10.1016/j.mib.2017.01.002
32. Davies BW, Bogard RW, Young TS, Mekalanos JJ. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for V. cholerae virulence. Cell. 2012;149: 358–370. doi:10.1016/j.cell.2012.01.053
33. Cohen D, Melamed S, Millman A, Shulman G, Oppenheimer-Shaanan Y, Kacen A, et al. Cyclic GMP-AMP signalling protects bacteria against viral infection. Nature. 2019;574: 691–695. doi:10.1038/s41586-019-1605-5
34. Michael AJ. Biosynthesis of polyamines and polyamine-containing molecules. Biochem J. 2016;473: 2315–2329. doi:10.1042/BCJ20160185
35. Tabor CW, Tabor H. Polyamines in microorganisms. Microbiol Rev. 1985;49: 81–99.
36. Roos WH, Ivanovska IL, Evilevitch A, Wuite GJL. Viral capsids: mechanical characteristics, genome packaging and delivery mechanisms. Cell Mol Life Sci. 2007;64: 1484–1497. doi:10.1007/s00018-007-6451-1
37. Kahana C. Antizyme and antizyme inhibitor, a regulatory tango. Cell Mol Life Sci. 2009;66: 2479–2488. doi:10.1007/s00018-009-0033-3
38. Takatsuka Y, Yamaguchi Y, Ono M, Kamio Y. Gene cloning and molecular characterization of lysine decarboxylase from Selenomonas ruminantium delineate its evolutionary relationship to ornithine decarboxylases from eukaryotes. J Bacteriol. 2000;182: 6732–6741. doi:10.1128/jb.182.23.6732-6741.2000
39. Murata A, Odaka M, Mukuno S. The Bacteriophage-Inactivating Effect of Basic Amino Acids; Arginine, Histidine, and Lysine. Agric Biol Chem. 1974;38: 477–478. doi:10.1271/bbb1961.38.477
40. Watanabe T, Saito H. Inactivation of coliphage T2H by basic amino acids. Microbiol Immunol. 1978;22: 167–172. doi:10.1111/j.1348-0421.1978.tb00360.x
41. Kwiatkowski B, Beilharz H, Stirm S. Disruption of Vi bacteriophage III and localization of

its deacetylase activity. J Gen Virol. 1975;29: 267–280. doi:10.1099/0022-1317-29-3-267
42. Angermeyer A, Das MM, Singh DV, Seed KD. Analysis of 19 Highly Conserved Vibrio cholerae Bacteriophages Isolated from Environmental and Patient Sources Over a Twelve-Year Period. Viruses. 2018;10. doi:10.3390/v10060299
43. Stewart CR, Casjens SR, Cresawn SG, Houtz JM, Smith AL, Ford ME, et al. The genome of Bacillus subtilis bacteriophage SPO1. J Mol Biol. 2009;388: 48–70. doi:10.1016/j.jmb.2009.03.009
44. Kutter E, Bryan D, Ray G, Brewster E, Blasdel B, Guttman B. From Host to Phage Metabolism: Hot Tales of Phage T4's Takeover of E. coli. Viruses. 2018;10. doi:10.3390/v10070387
45. Bentley SD, Parkhill J. Comparative genomic structure of prokaryotes. Annu Rev Genet. 2004;38: 771–792. doi:10.1146/annurev.genet.38.072902.094318
46. Merrikh H, Zhang Y, Grossman AD, Wang JD. Replication-transcription conflicts in bacteria. Nat Rev Microbiol. 2012;10: 449–458. doi:10.1038/nrmicro2800
47. Harris KA, Breaker RR. Large noncoding rnas in bacteria. Microbiol Spectr. 2018;6. doi:10.1128/microbiolspec.RWR-0005-2017
48. Delihas N. Small mobile sequences in bacteria display diverse structure/function motifs. Mol Microbiol. 2008;67: 475–481. doi:10.1111/j.1365-2958.2007.06068.x
49. Wagner EGH, Altuvia S, Romby P. Antisense RNAs in bacteria and their genetic elements. Adv Genet. 2002;46: 361–398.
50. Ravin NV, Svarchevsky AN, Dehò G. The anti-immunity system of phage-plasmid N15: identification of the antirepressor gene and its control by a small processed RNA. Mol Microbiol. 1999;34: 980–994. doi:10.1046/j.1365-2958.1999.01658.x
51. Sabbattini P, Six E, Zangrossi S, Briani F, Ghisotti D, Deho G. Immunity specificity determinants in the P4-like retronphage phi R73. Virology. 1996;216: 389–396. doi:10.1006/viro.1996.0074
52. Sherwood AV, Henkin TM. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. Annu Rev Microbiol. 2016;70: 361–374. doi:10.1146/annurev-micro-091014-104306
53. Ayala JC, Silva AJ, Benitez JA. H-NS: an overarching regulator of the Vibrio cholerae life cycle. Res Microbiol. 2017;168: 16–25. doi:10.1016/j.resmic.2016.07.007
54. Ayala JC, Wang H, Silva AJ, Benitez JA. Repression by H-NS of genes required for the biosynthesis of the Vibrio cholerae biofilm matrix is modulated by the second messenger cyclic diguanylic acid. Mol Microbiol. 2015;97: 630–645. doi:10.1111/mmi.13058
55. Ayala JC, Wang H, Benitez JA, Silva AJ. RNA-Seq analysis and whole genome DNA-binding profile of the Vibrio cholerae histone-like nucleoid structuring protein (H-NS). Genom Data. 2015;5: 147–150. doi:10.1016/j.gdata.2015.05.039
56. Das B, Martínez E, Midonet C, Barre F-X. Integrative mobile elements exploiting Xer recombination. Trends Microbiol. 2013;21: 23–30. doi:10.1016/j.tim.2012.10.003
57. Christie GE, Dokland T. Pirates of the Caudovirales. Virology. 2012;434: 210–221. doi:10.1016/j.virol.2012.10.028
58. Dokland T. Scaffolding proteins and their role in viral assembly. Cell Mol Life Sci. 1999;56: 580–603. doi:10.1007/s000180050455
59. Fillol-Salom A, Martínez-Rubio R, Abdulrahman RF, Chen J, Davies R, Penadés JR. Phage-inducible chromosomal islands are ubiquitous within the bacterial universe. ISME J. 2018;12: 2114–2128. doi:10.1038/s41396-018-0156-3
60. Birch EW, Ruggero NA, Covert MW. Determining host metabolic limitations on viral replication via integrated modeling and experimental perturbation. PLoS Comput Biol. 2012;8: e1002746. doi:10.1371/journal.pcbi.1002746
61. Mahmoudabadi G, Milo R, Phillips R. Energetic cost of building a virus. Proc Natl Acad Sci USA. 2017;114: E4324–E4333. doi:10.1073/pnas.1701670114

62. Lindqvist BH, Dehò G, Calendar R. Mechanisms of genome propagation and helper exploitation by satellite phage P4. Microbiol Rev. 1993;57: 683–702.
63. Linderoth NA, Calendar RL. The Psu protein of bacteriophage P4 is an antitermination factor for rho-dependent transcription termination. J Bacteriol. 1991;173: 6722–6731. doi:10.1128/jb.173.21.6722-6731.1991
64. Ranjan A, Banerjee R, Pani B, Sen U, Sen R. The moonlighting function of bacteriophage P4 capsid protein, Psu, as a transcription antiterminator. Bacteriophage. 2013;3: e25657. doi:10.4161/bact.25657
65. Ghosh G, Reddy J, Sambhare S, Sen R. A bacteriophage capsid protein is an inhibitor of a conserved transcription terminator of various bacterial pathogens. J Bacteriol. 2018;200. doi:10.1128/JB.00380-17
66. Lengyel JA, Calendar R. Control of bacteriophage P2 protein and DNA synthesis. Virology. 1974;57: 305–313. doi:10.1016/0042-6822(74)90170-6
67. Geisselsoder J, Chidambaram M, Goldstein R. Transcriptional control of capsid size in the P2:P4 bacteriophage system. J Mol Biol. 1978;126: 447–456. doi:10.1016/0022-2836(78)90051-7
68. Agarwal M, Arthur M, Arbeit RD, Goldstein R. Regulation of icosahedral virion capsid size by the in vivo activity of a cloned gene product. Proc Natl Acad Sci USA. 1990;87: 2428–2432. doi:10.1073/pnas.87.7.2428
69. Ram G, Chen J, Ross HF, Novick RP. Precisely modulated pathogenicity island interference with late phage gene transcription. Proc Natl Acad Sci USA. 2014;111: 14536–14541. doi:10.1073/pnas.1406749111
70. Dalia AB, McDonough E, Camilli A. Multiplex genome editing by natural transformation. Proc Natl Acad Sci USA. 2014;111: 8937–8942. doi:10.1073/pnas.1406478111
71. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359. doi:10.1038/nmeth.1923

Figure Legends

Fig 1. The *V. cholerae* response to ICP1 infection.
(A) Volcano plot showing gene expression changes between uninfected *V. cholerae* cultures, and the same cultures 4 minutes post infection. Full (bottom) and zoomed (top) views are provided to improve gene feature visibility. Genes are colored according to annotation. A significance cutoff of a Q value less than or equal to 0.1 and a -$\log_2$ fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used. **(B)** Heat map showing changes in *V. cholerae* gene feature $\log_2$ TPM over the course of ICP1 infection. Expression values obtained for three biological replicates are shown at each time point. ICP1 genes were excluded from TPM calculation to highlight relative changes in *V. cholerae* transcript abundance. Colors reflect the Z-score of each gene's $\log_2$ TPM value across replicates and time points. **(C)** Percent reads abundance for both *V. cholerae* chromosomes and ICP1 over the infection time course. **(D)** Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. For all panels, results incorporate values obtained for three biological replicates.

Fig 2. ICP1 transcriptional patterns.
ICP1's genome displaying average reads coverage over the course of infection of PLE (-) *V. cholerae*. Reads are color coded by time point. Gene features are colored based on known or putative gene functions. Results incorporate values obtained for three biological replicates.

Fig 3. PLE transcriptome.
(**A**) PLE1's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color

coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for three biological replicates.

Fig 4. ICP1 noncoding RNA
ICP1's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. ICP1 gene features are colored based on known or putative gene functions as in Fig 2. Regions boxed in orange show transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for three biological replicates.

Fig. 5 PLE1 noncoding RNA
PLE1's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for three biological replicates.

Fig. 6 CTX production is upregulated in PLE1 positive strains
A. Volcano plot showing differential regulation of *V. cholerae* genes in uninfected PLE1(+) cultures relative to PLE (-) cultures. Genes within 2kb of the PLE integration site are colored red, and genes encoded in RS1 and CTXφ are colored blue. Genes that are not significantly differentially regulated are colored light gray. A cutoff of a Q value less than or equal to 0.1 and a -log2 fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used. **B.** The transduction units per mL of CTXφ produced by PLE(-), PLE1(+), and PLE2(+) cultures following induction by mitomycin C.

Fig. 7 PLE downregulates ICP1 capsid expression
(**A**) Percent reads abundance for PLE1, *V. cholerae* chromosomes, and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE1 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE1 (blue) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. (**E**). Representative Western blot against Gp122, ICP1's major capsid protein from infections of PLE(-), PLE1, PLE2, PLE3, PLE4, and PLE5 cultures at 16 minutes post infection. Quantification and replicates are shown in S17 Fig. For panels A-D, results incorporate values obtained for three biological replicates.

Supporting information
S1 Fig. PLE lifecycle.
Model of ICP1 infection in PLE(-) and PLE(+) *V. cholerae*. ICP1 injects its DNA into *V. cholerae*; prior to DNA replication, ICP1 activity leads to PLE activation and excision. ICP1 DNA replication is reduced in the PLE(+) cell where the PLE replicates to high copy. Finally, the cell lyses and releases infectious particles. No ICP1 particles and of PLE transducing particles are released from the PLE(+) cell.

S2 Fig. PLE2 transcriptome
(**A)** PLE2's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color

coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.

S3 Fig. PLE3 Transcriptome
(**A)** PLE3's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.

S4 Fig. PLE4 Transcriptome
(**A)** PLE4's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.

S5 Fig. PLE5 Transcriptome
(**A)** PLE5's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.

S6 Fig. PLE2 noncoding RNA
PLE2's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.

S7 Fig. PLE3 noncoding RNA
PLE3's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.

S8 Fig. PLE4 noncoding RNA

PLE4's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.

S9 Fig. PLE5 noncoding RNA
PLE5's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.

S10 Fig. Conservation of the PLE IIR transcript
(**A**) Clustal alignment of the PLE IIR transcript across all PLEs. Terminal inverted repeats are highlighted in grey. Sequence with antisense homology to PLE gene leader sequences is bolded. (**B**) Alignment of the an 'ideal' leader sequence antisense to the IIR, and the leader sequences of several genes in PLE1. Matching sites are bolded. Mismatches are shown in light gray.

S11 Fig. Some *V. cholerae* genes are differentially expressed in the presence of PLEs
Volcano plots showing differential regulation of *V. cholerae* genes in uninfected PLE(+) cultures relative to PLE(-) cultures. Genes within 2kb of the PLE integration site are colored red, other genes within the superintegron are colored yellow, and genes encoded in RS1,CTXφ, or TLC are colored blue. Genes that are not significantly differentially regulated are colored light gray, and other significantly differentially regulated genes are colored black. A Q value cut off of 0.1 was used to determine significance. PLEs 2 (A), 3 (B), 4 (C), and 5 (D) are shown.

S12 Fig. Integrative mobile elements exploiting Xer in E7946 *V. cholerae*
Diagram showing IMEX elements integrated at E7946's dif site on the large chromosome. Brackets denote mobile elements.

S13 PLE2 specifically downregulates ICP1 capsid production
(**A**) Percent reads abundance for PLE2, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE2 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE2 (orange) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.

S14 PLE3 specifically downregulates ICP1 capsid production
(**A**) Percent reads abundance for PLE3, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE3 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE3 (black) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.

S15 PLE4 specifically downregulates ICP1 capsid production
(**A**) Percent reads abundance for PLE4, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from

each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE4 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE4 (pink) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates

S16 PLE5 specifically downregulates ICP1 capsid production
(**A**) Percent reads abundance for PLE5, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE5 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE5 (blue) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.

S17 PLEs reduce the level of expressed capsid protein
(A) and **(B)** Biological replicates of Western blots against Gp122, ICP1's major capsid protein from infections of PLE(-), PLE1, PLE2, PLE3, PLE4, and PLE5 cultures at 16 minutes post infection. **(C)** Quantification of the Western blot band intensities from (A) (B) and Fig 7E.

S1 Table Strains used in this study.

S2 Table Differentially expressed genes in *V. cholerae* 4 minutes post infection with ICP1. A cutoff of a Q value less than or equal to 0.1 and a $-\log_2$ fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used.

S3 Table ICP1_2006_E gene annotations and classification of temporal gene expression class.

S4 Table Differentially expressed genes in *V. cholerae* PLE (+) compared to PLE (-) in uninfected cells. A cutoff of a Q value less than or equal to 0.1 and a $-\log_2$ fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used.

S5 Table Accession numbers for reference files used for RNA-seq mapping.

S1 Data Tables of read counts and DESeq2 results for samples used in this study.

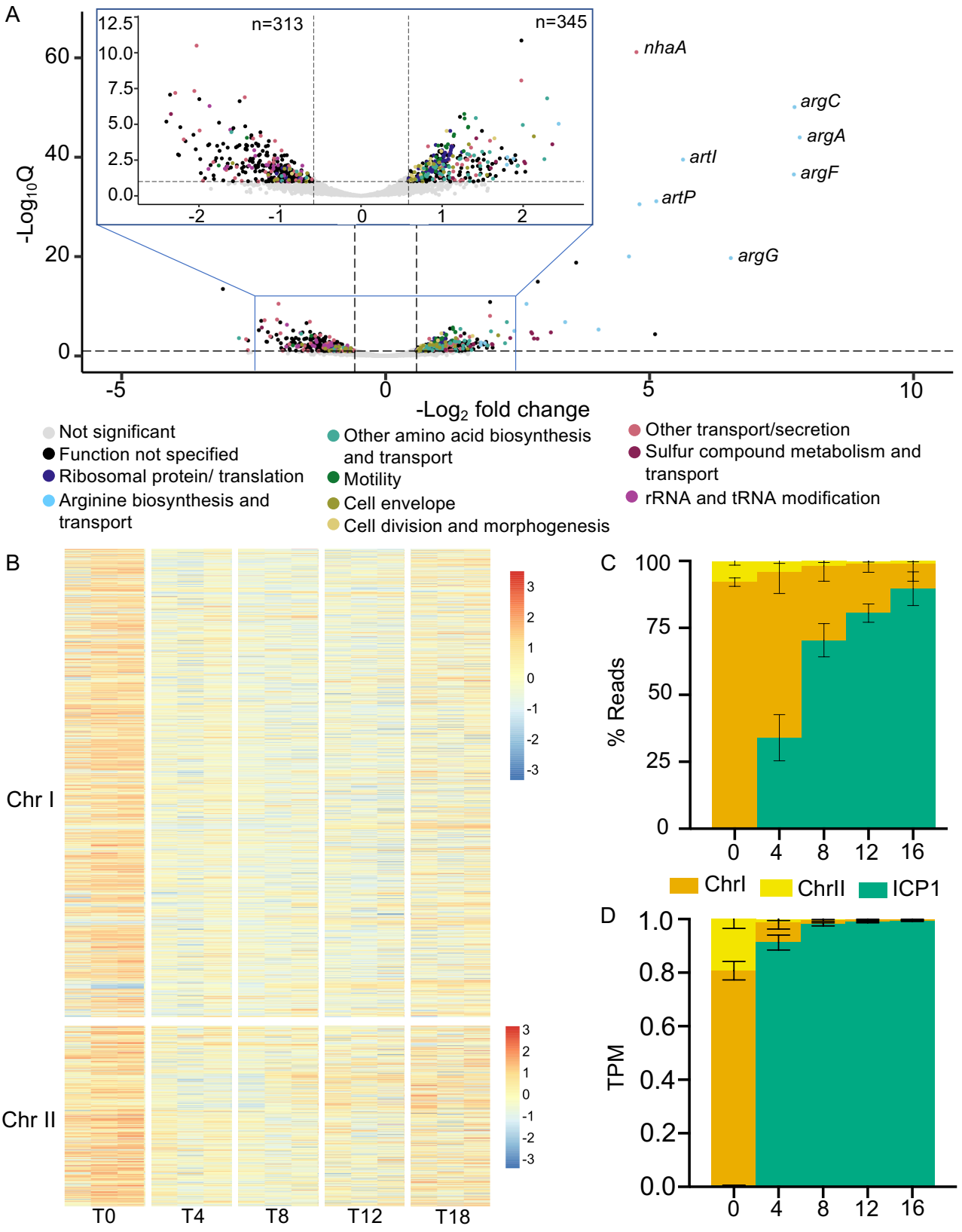


Fig 1. The *V. cholerae* response to ICP1 infection.
(A) Volcano plot showing gene expression changes between uninfected *V. cholerae* cultures, and the same cultures 4 minutes post infection. Full (bottom) and zoomed (top) views are provided to improve gene feature visibility. Genes are colored according to annotation. A significance cutoff of a Q value less than or equal to 0.1 and a -$\log_2$ fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used. **(B)** Heat map showing changes in *V. cholerae* gene feature $\log_2$ TPM over the course of ICP1 infection. Expression values obtained for three biological replicates are shown at each time point. ICP1 genes were excluded from TPM calculation to highlight relative changes in *V. cholerae* transcript abundance. Colors reflect the Z-score of each gene's $\log_2$ TPM value across replicates and time points. **(C)** Percent reads abundance for both *V. cholerae* chromosomes and ICP1 over the infection time course. **(D)** Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. For all panels, results incorporate values obtained for three biological replicates.

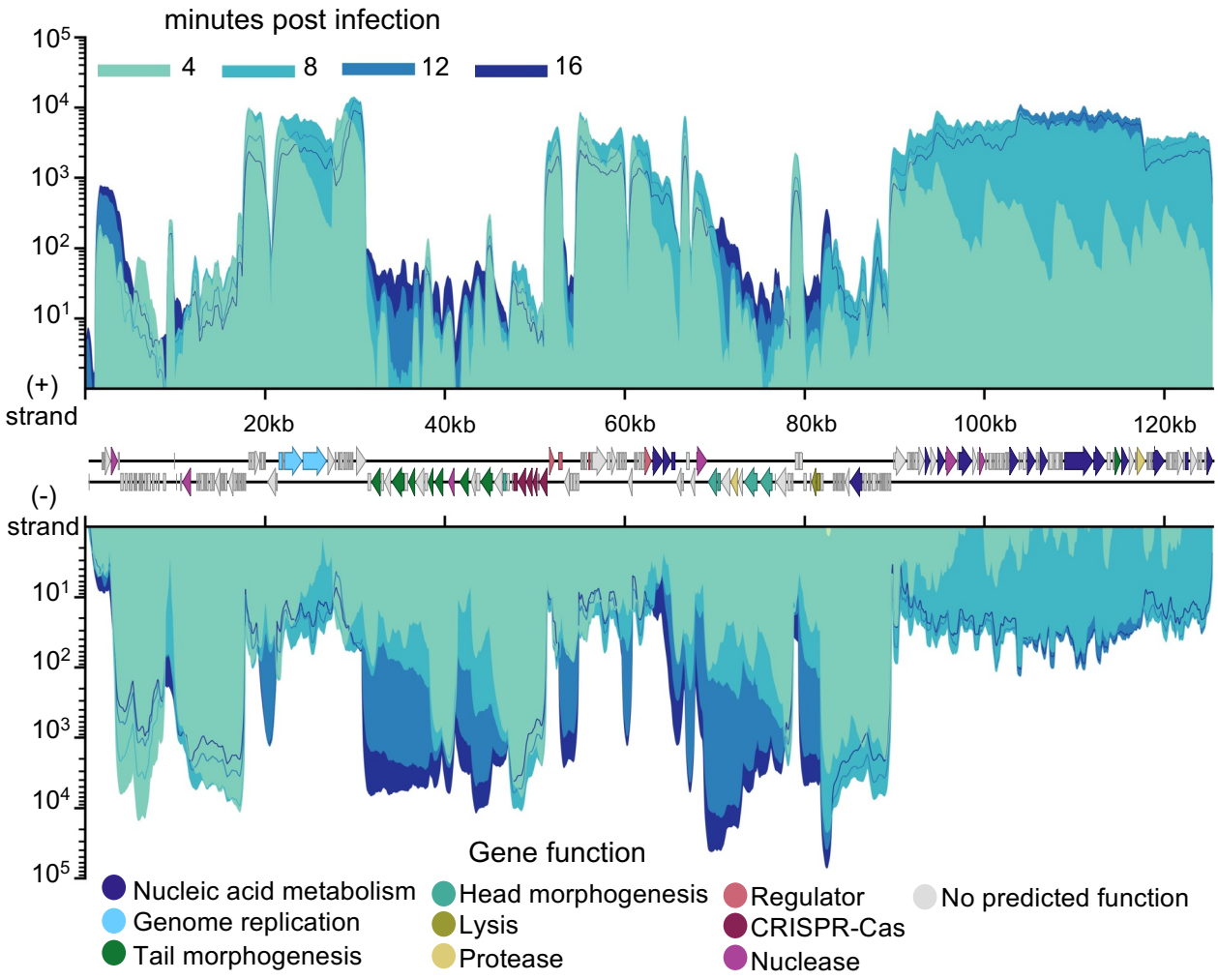


Fig 2. ICP1 transcriptional patterns.
ICP1's genome displaying average reads coverage over the course of infection of PLE (-) *V. cholerae*. Reads are color coded by time point. Gene features are colored based on known or putative gene functions. Results incorporate values obtained for three biological replicates.

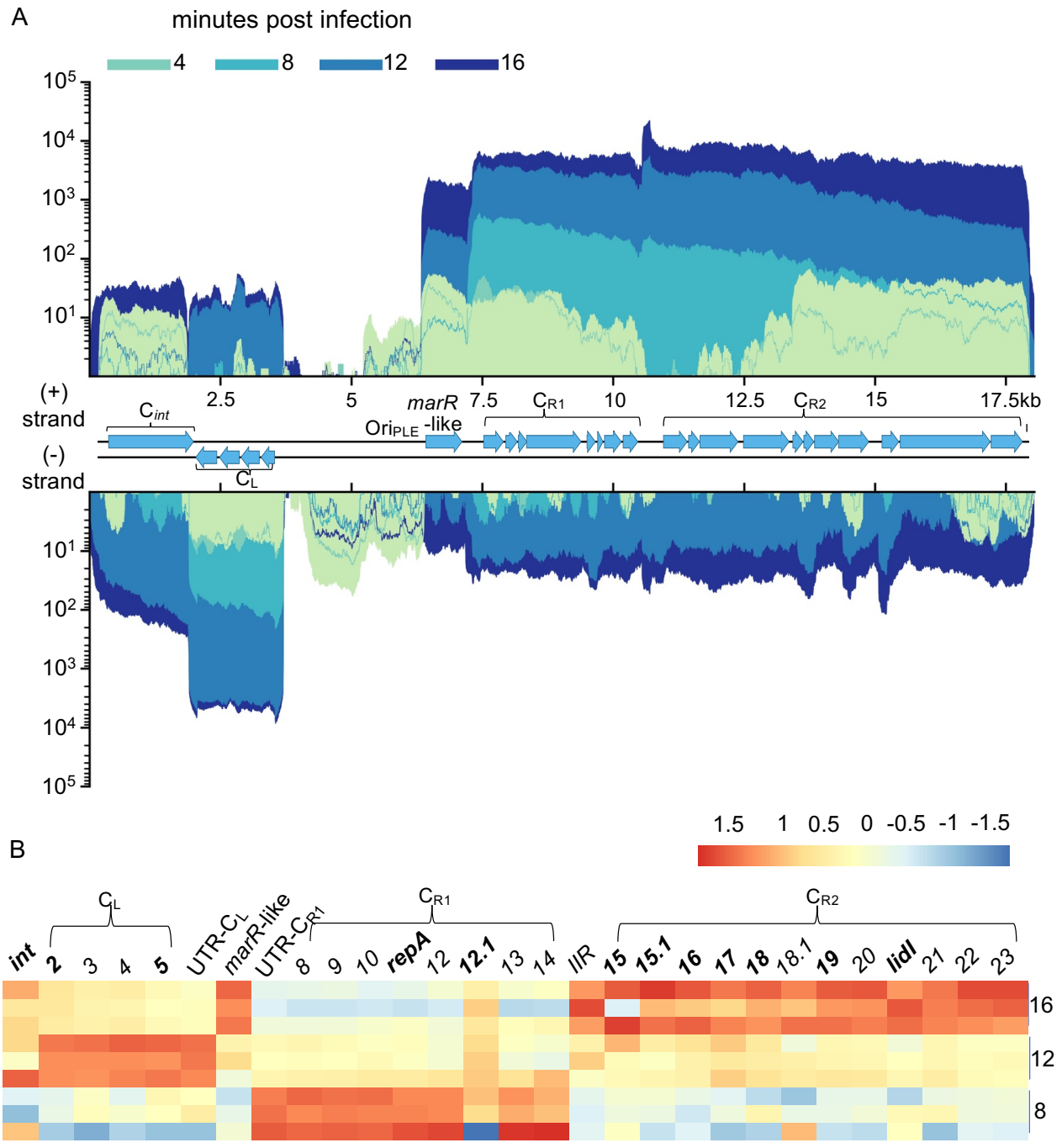


Fig 3. PLE transcriptome.
(**A**) PLE1's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for three biological replicates.

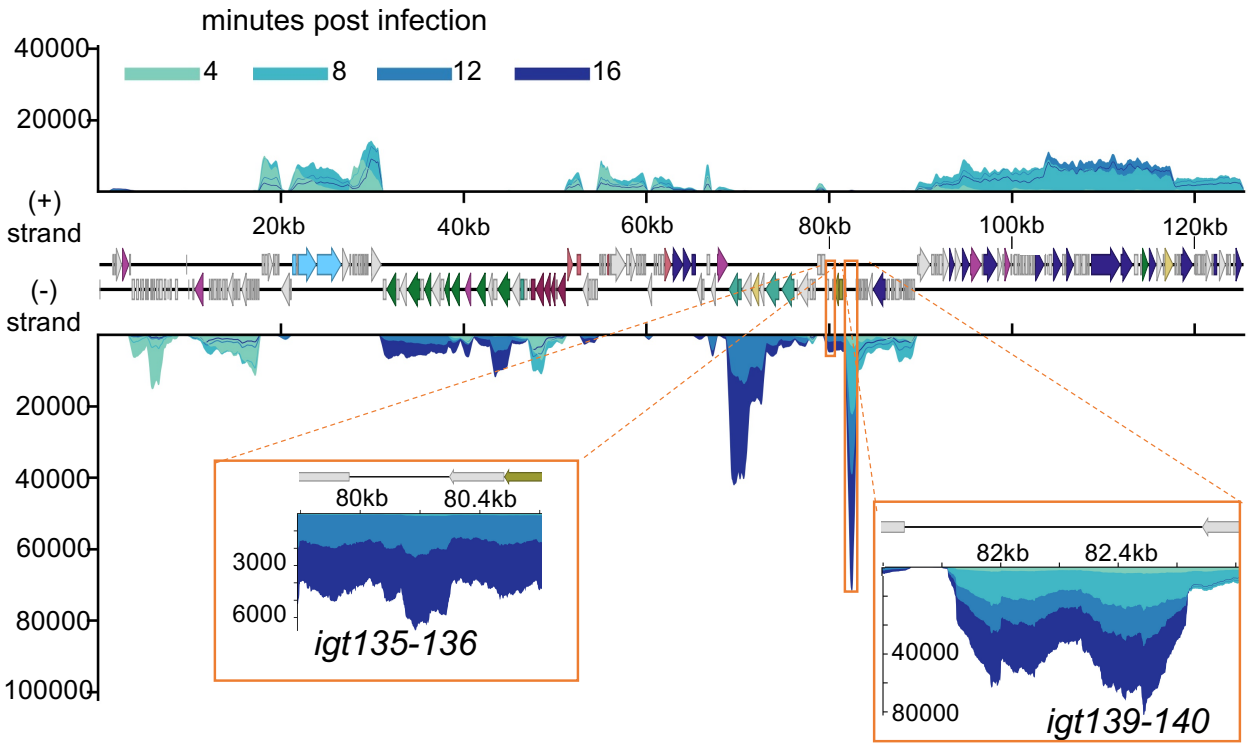


Fig 4. ICP1 noncoding RNA

ICP1's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. ICP1 gene features are colored based on known or putative gene functions as in Fig 2. Regions boxed in orange show transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for three biological replicates.

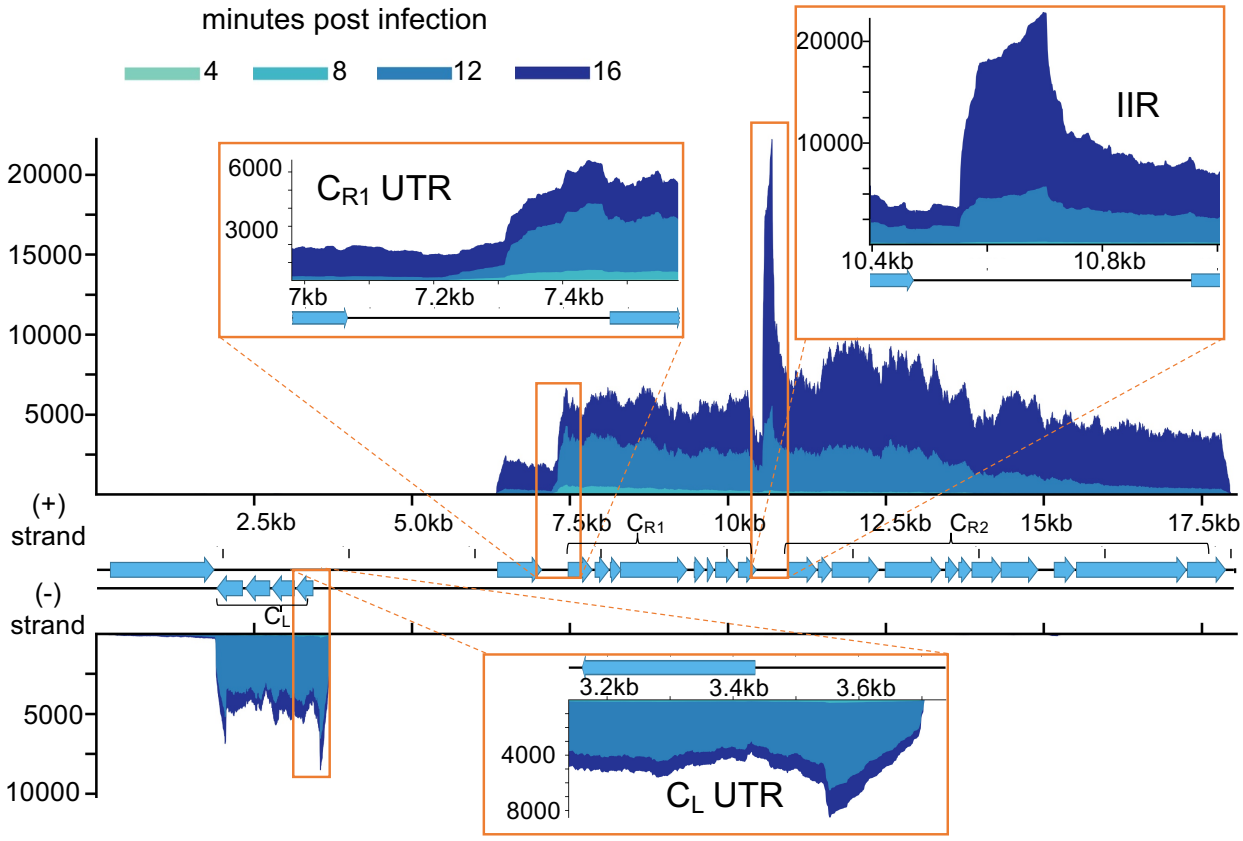


Fig. 5 PLE1 noncoding RNA

PLE1's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for three biological replicates.

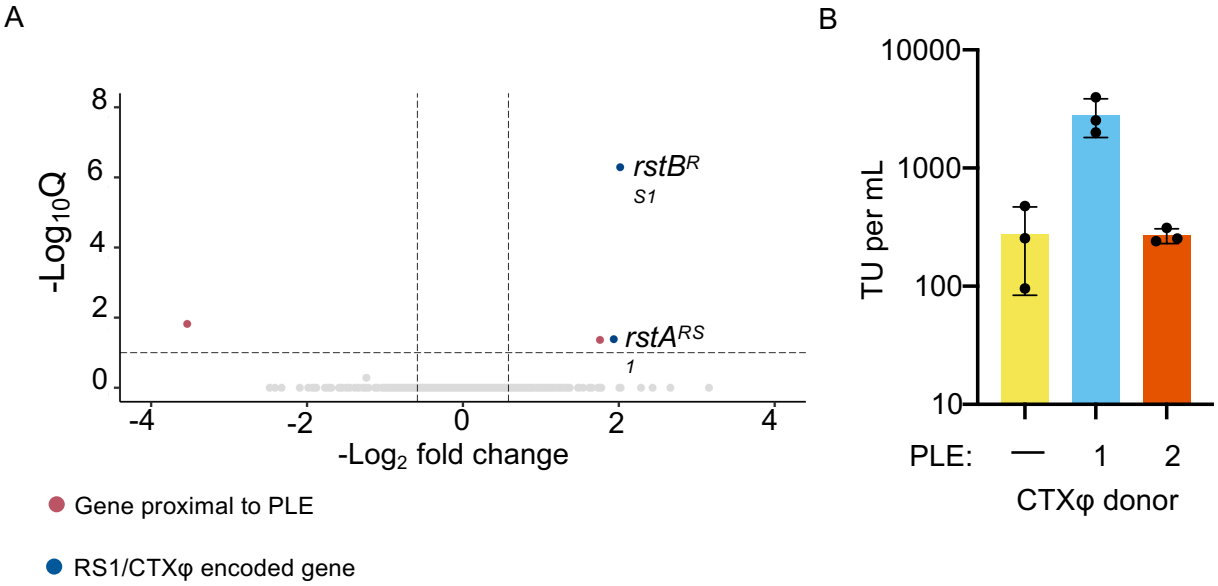


Fig. 6 CTX production is upregulated in PLE1 positive strains
A. Volcano plot showing differential regulation of *V. cholerae* genes in uninfected PLE1(+) cultures relative to PLE (-) cultures. Genes within 2kb of the PLE integration site are colored red, and genes encoded in RS1 and CTXφ are colored blue. Genes that are not significantly differentially regulated are colored light gray. A cutoff of a Q value less than or equal to 0.1 and a -log2 fold change magnitude greater or equal to 0.585 (approximate to 1.5-fold) was used. **B.** The transduction units per mL of CTXφ produced by PLE(-), PLE1(+), and PLE2(+) cultures following induction by mitomycin C.

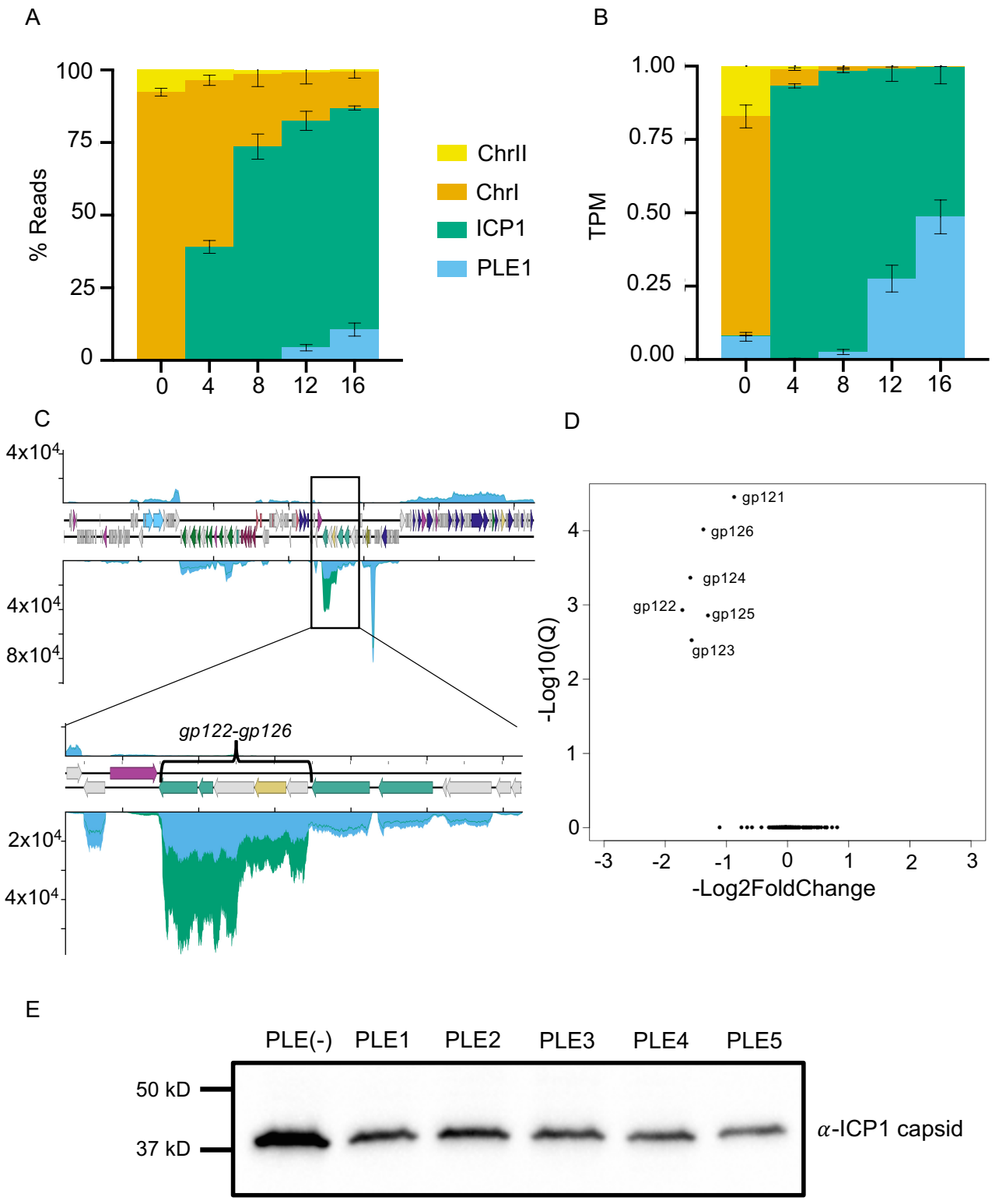
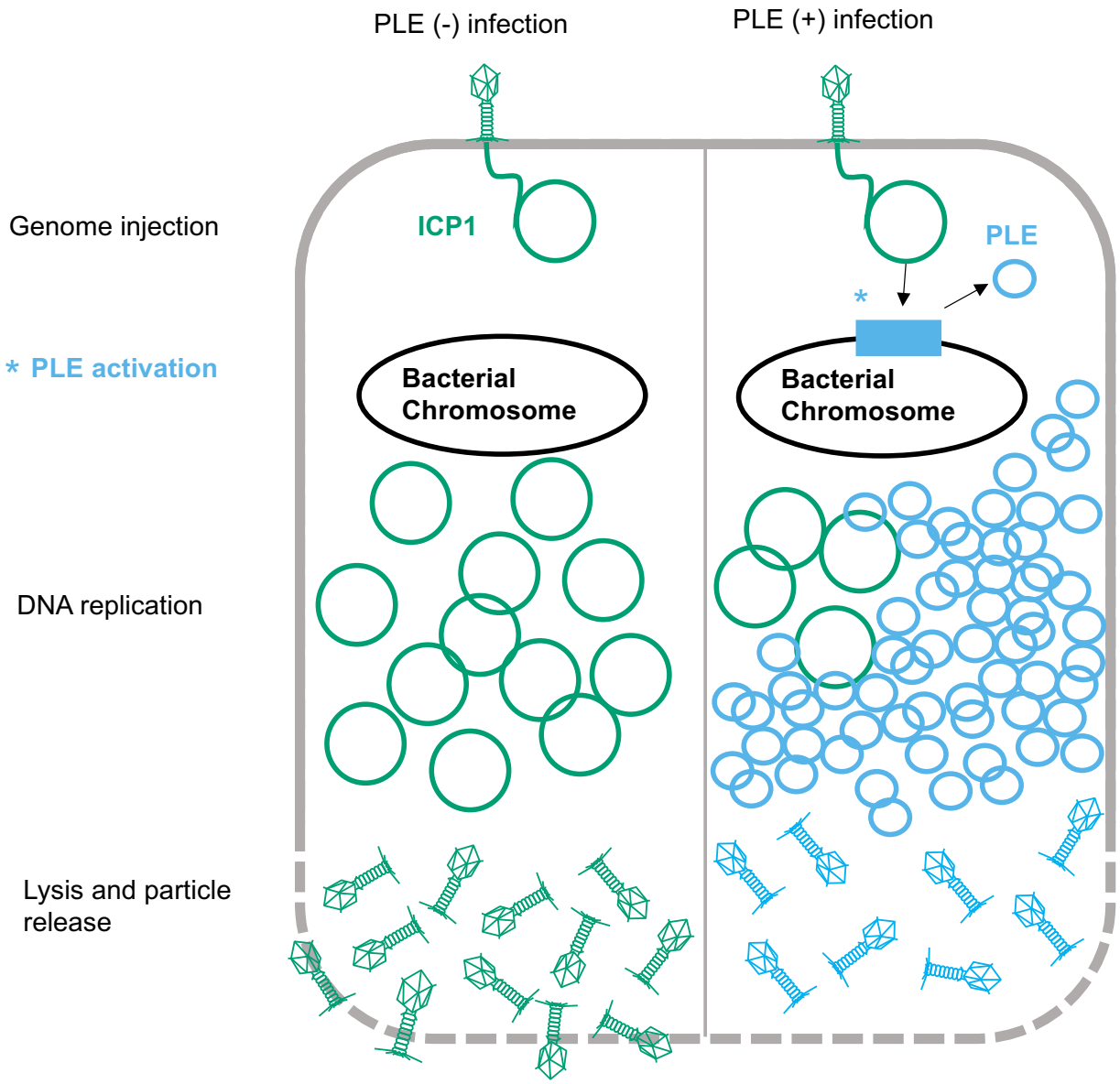
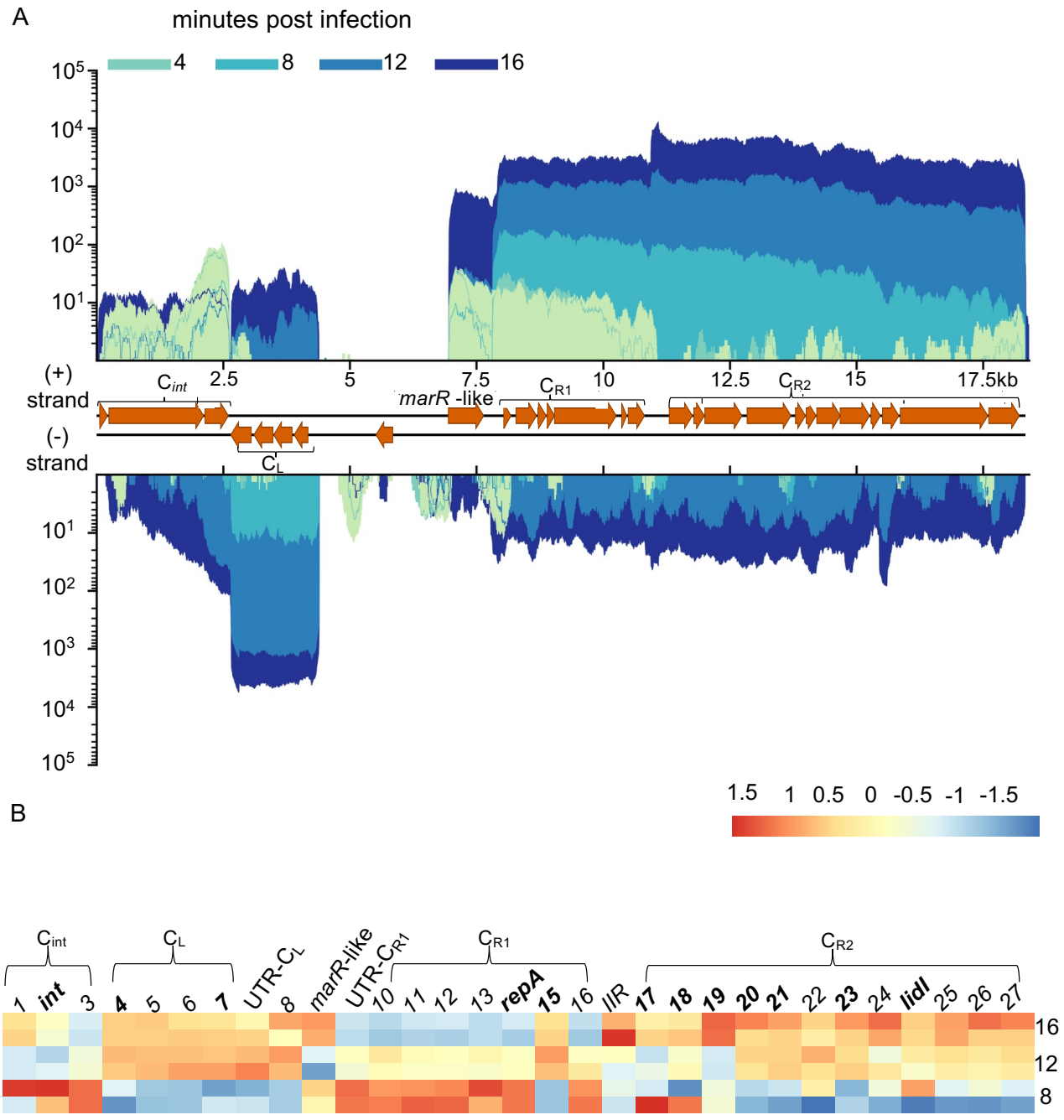


Fig. 7 PLE downregulates ICP1 capsid expression

(**A**) Percent reads abundance for PLE1, *V. cholerae* chromosomes, and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE1 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE1 (blue) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. (**E**). Representative Western blot against Gp122, ICP1's major capsid protein from infections of PLE(-), PLE1, PLE2, PLE3, PLE4, and PLE5 cultures at 16 minutes post infection. Quantification and replicates are shown in S17 Fig. For panels A-D, results incorporate values obtained for three biological replicates.

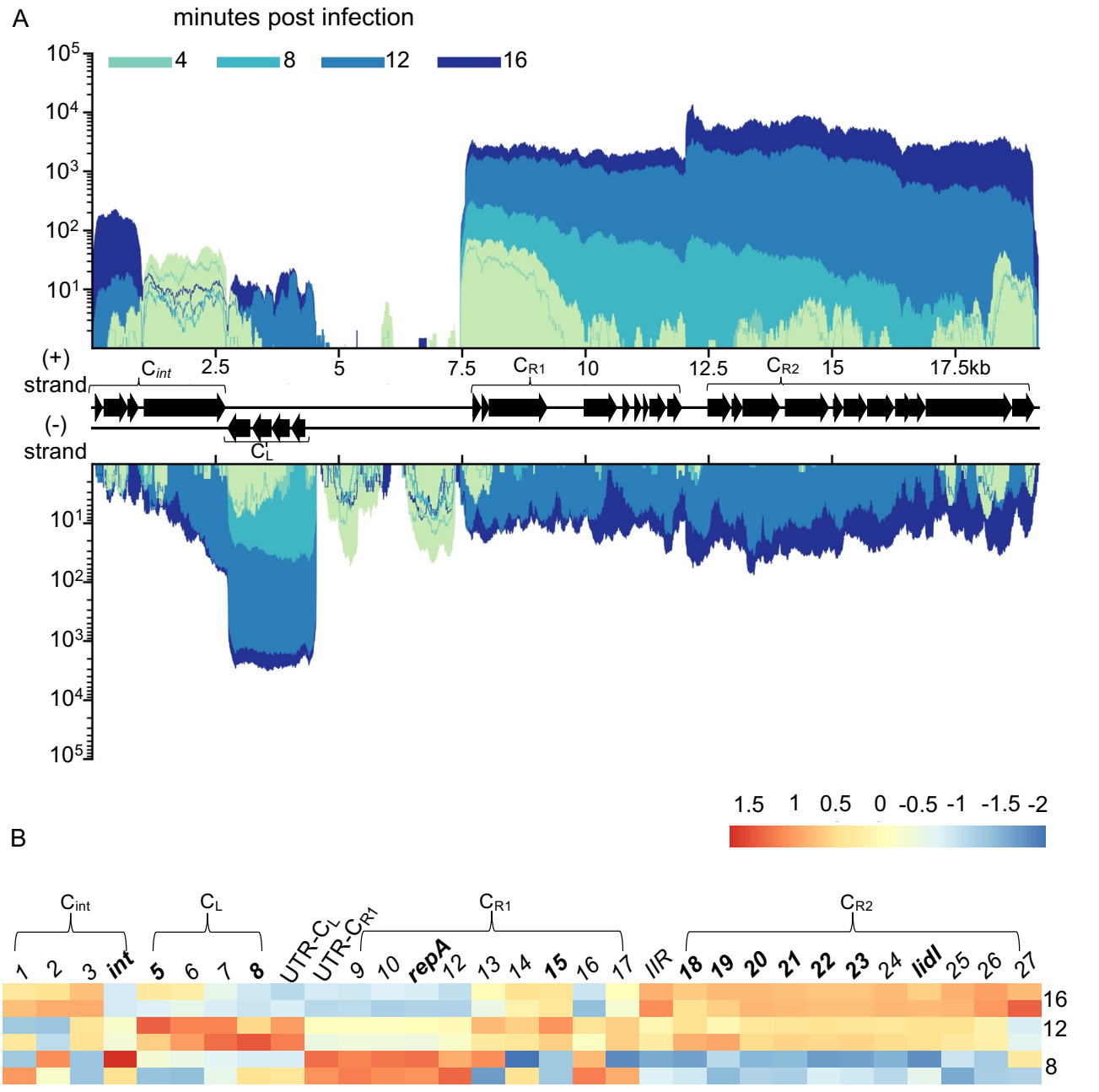


Supporting information
S1 Fig. PLE lifecycle.
Model of ICP1 infection in PLE(-) and PLE(+) *V. cholerae*. ICP1 injects its DNA into *V. cholerae*; prior to DNA replication, ICP1 activity leads to PLE activation and excision. ICP1 DNA replication is reduced in the PLE(+) cell where the PLE replicates to high copy. Finally, the cell lyses and releases infectious particles. No ICP1 particles and of PLE transducing particles are released from the PLE(+) cell.



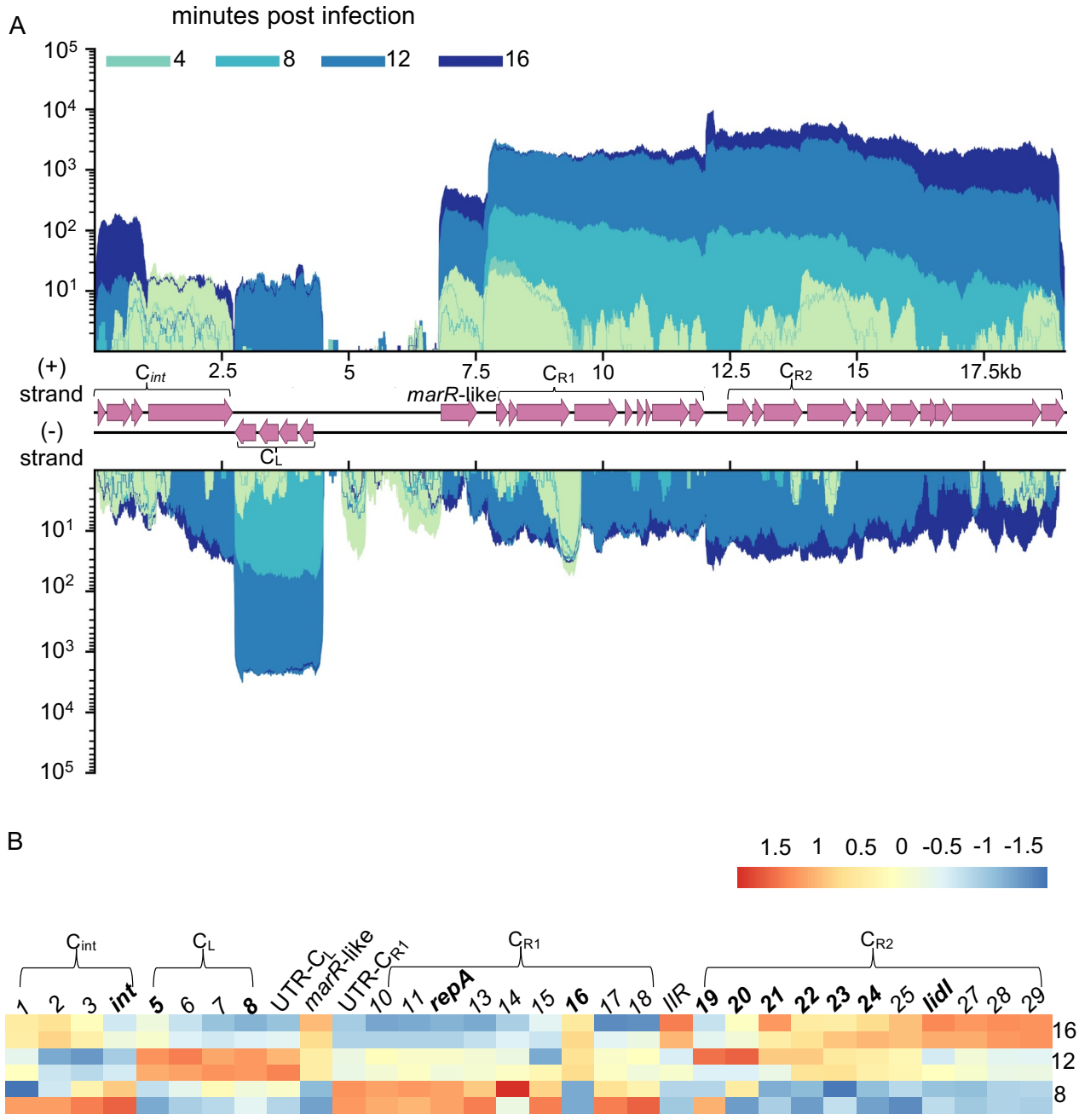
S2 Fig. PLE2 transcriptome

(**A**) PLE2's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.



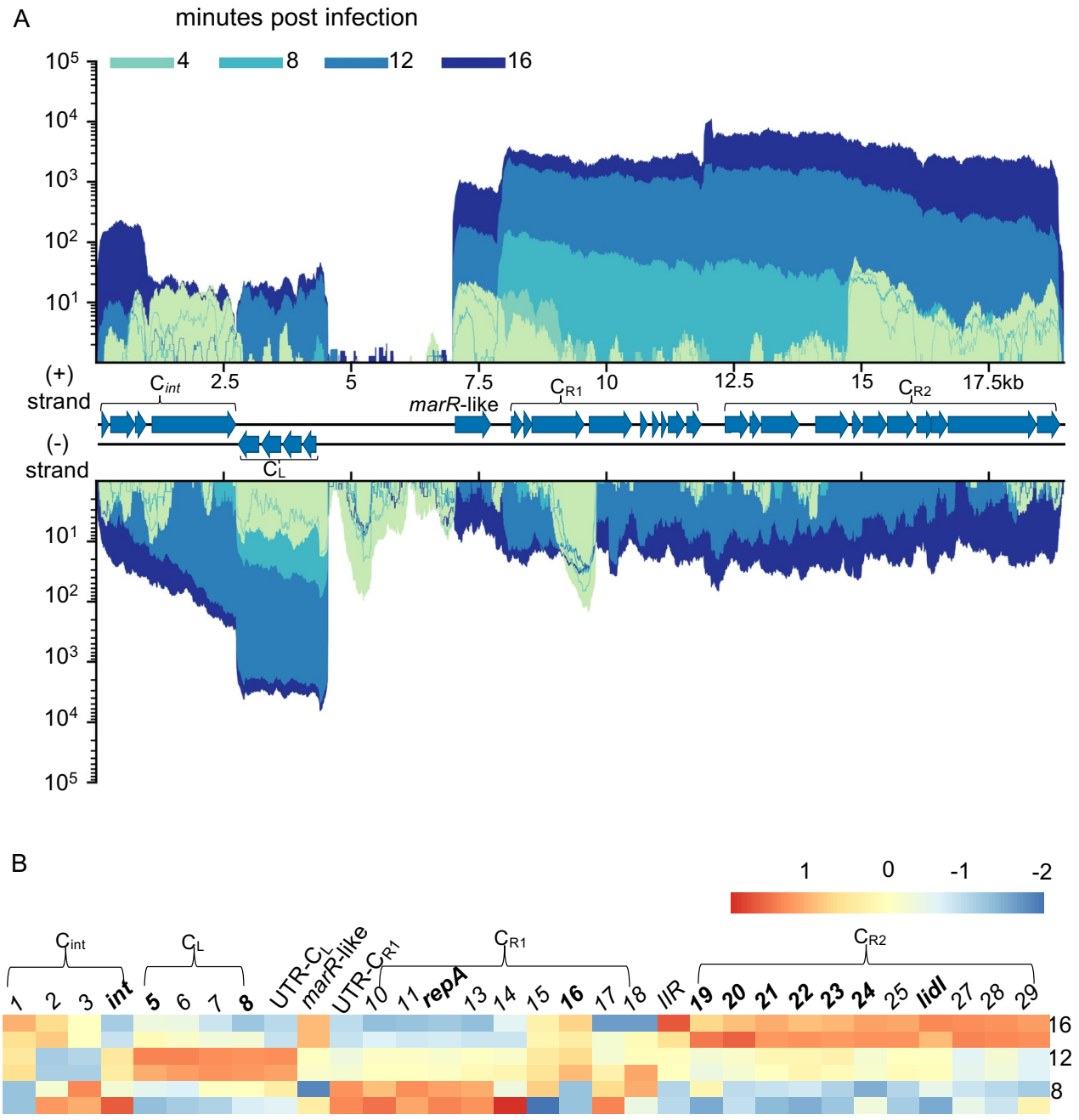
S3 Fig. PLE3 Transcriptome

(**A)** PLE3's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.



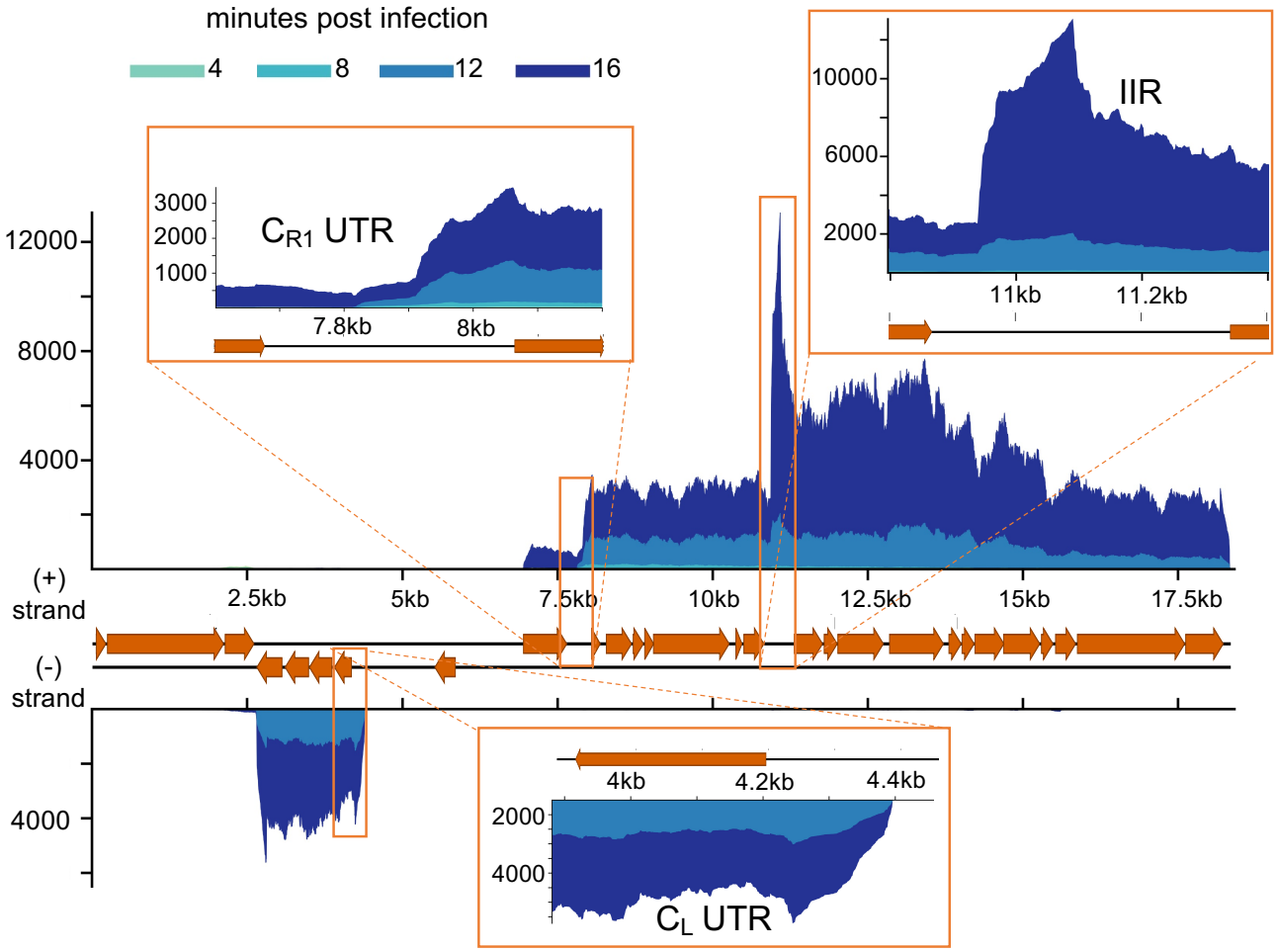
S4 Fig. PLE4 Transcriptome

(**A)** PLE4's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.

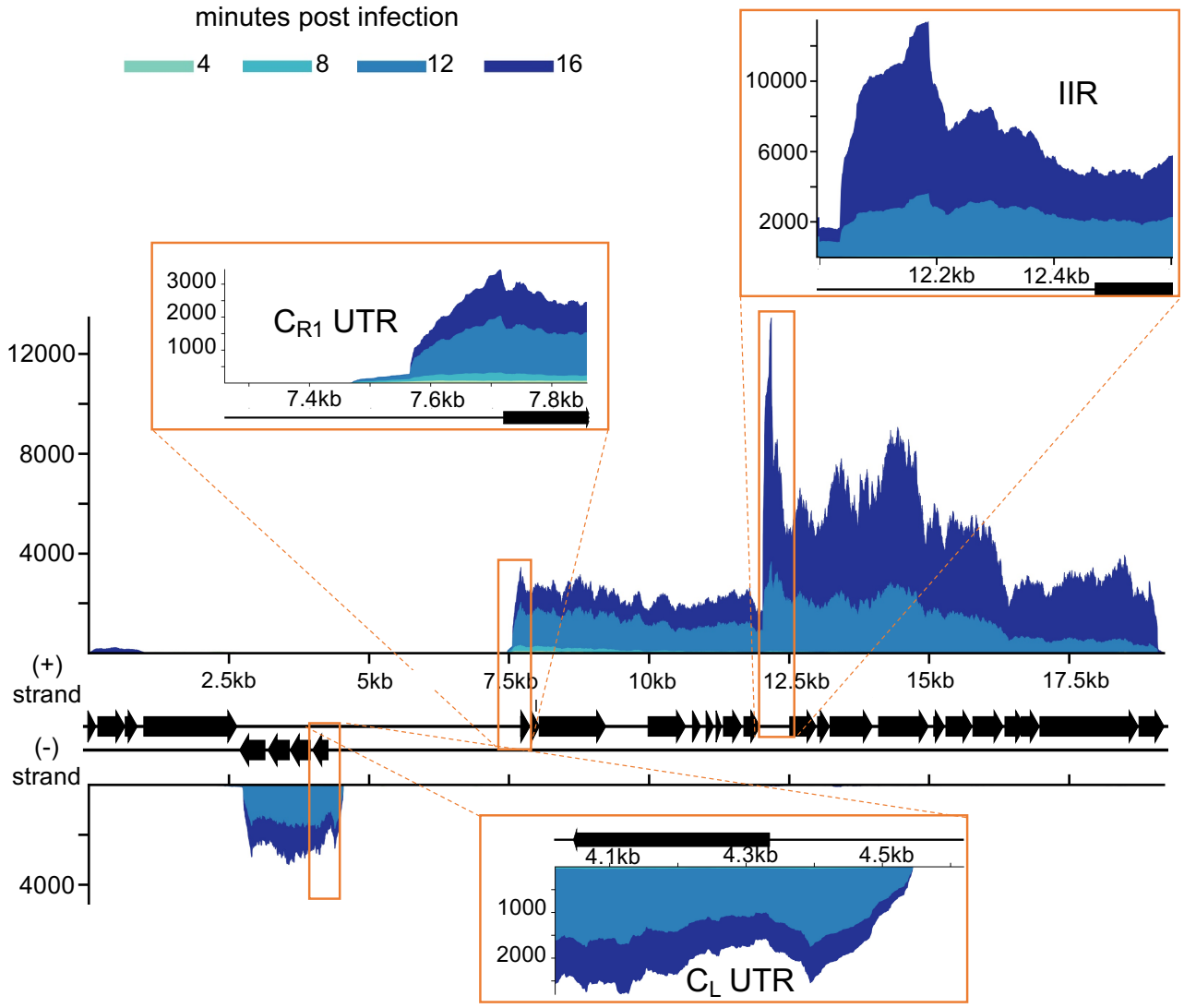


S5 Fig. PLE5 Transcriptome

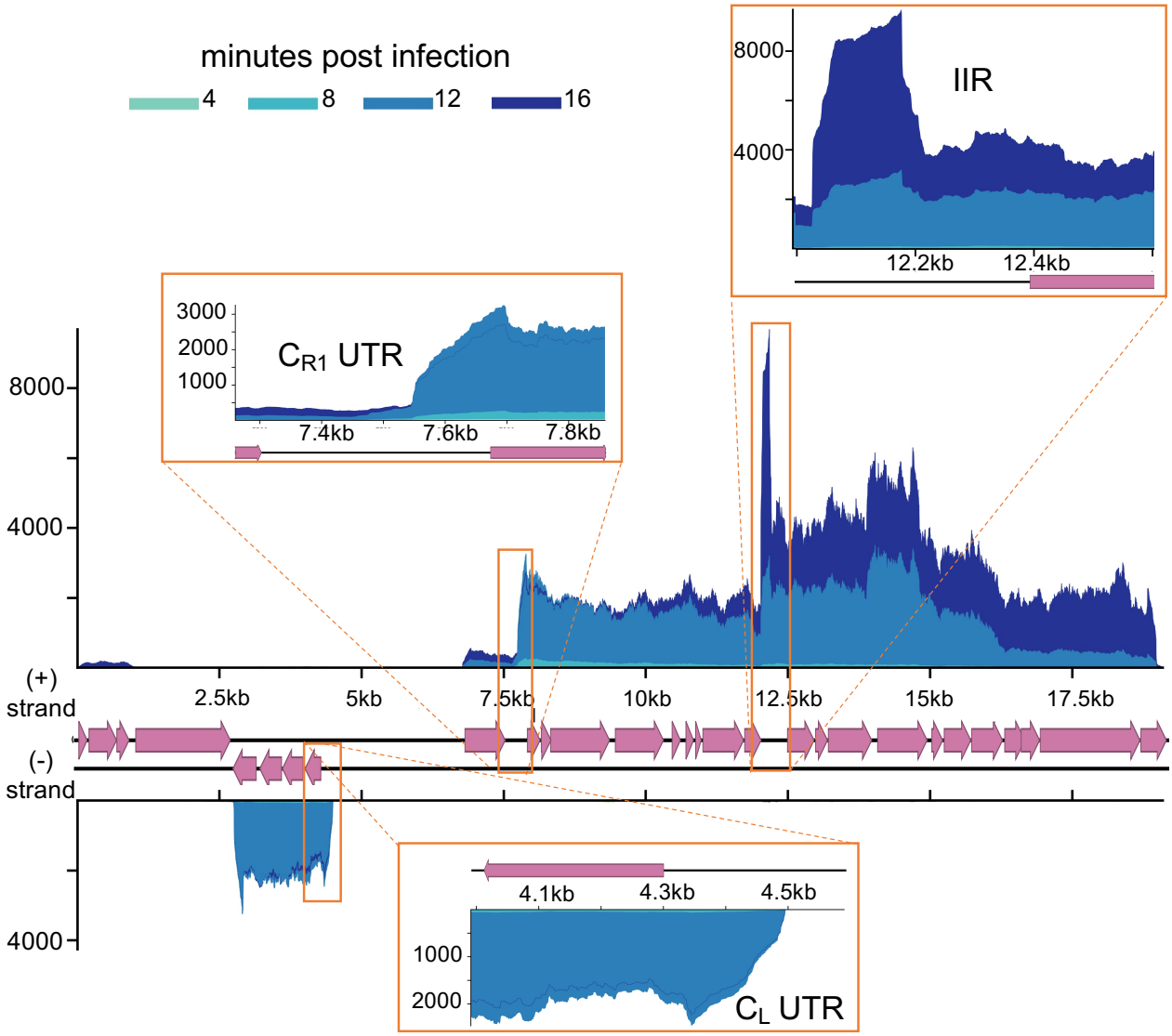
(**A**) PLE5's genome displaying average reads coverage over the course of infection. Reads are depicted on a logarithmic scale, to improve visibility of early expressed genes. Reads are color coded by time point. (**B**) Heat map of PLE1 gene expression over the course of infection. Color reflects the Z-score of each gene's log2 TPM value across replicates and time points. *V. cholerae* and ICP1 genes were excluded from TPM calculation to highlight relative changes in PLE1 transcript abundance. Core genes, protein coding genes with high conservation across PLEs, are bolded. Values for 8, 12, and 16 minutes post infection are shown. Results incorporate gene expression values obtained for two biological replicates.



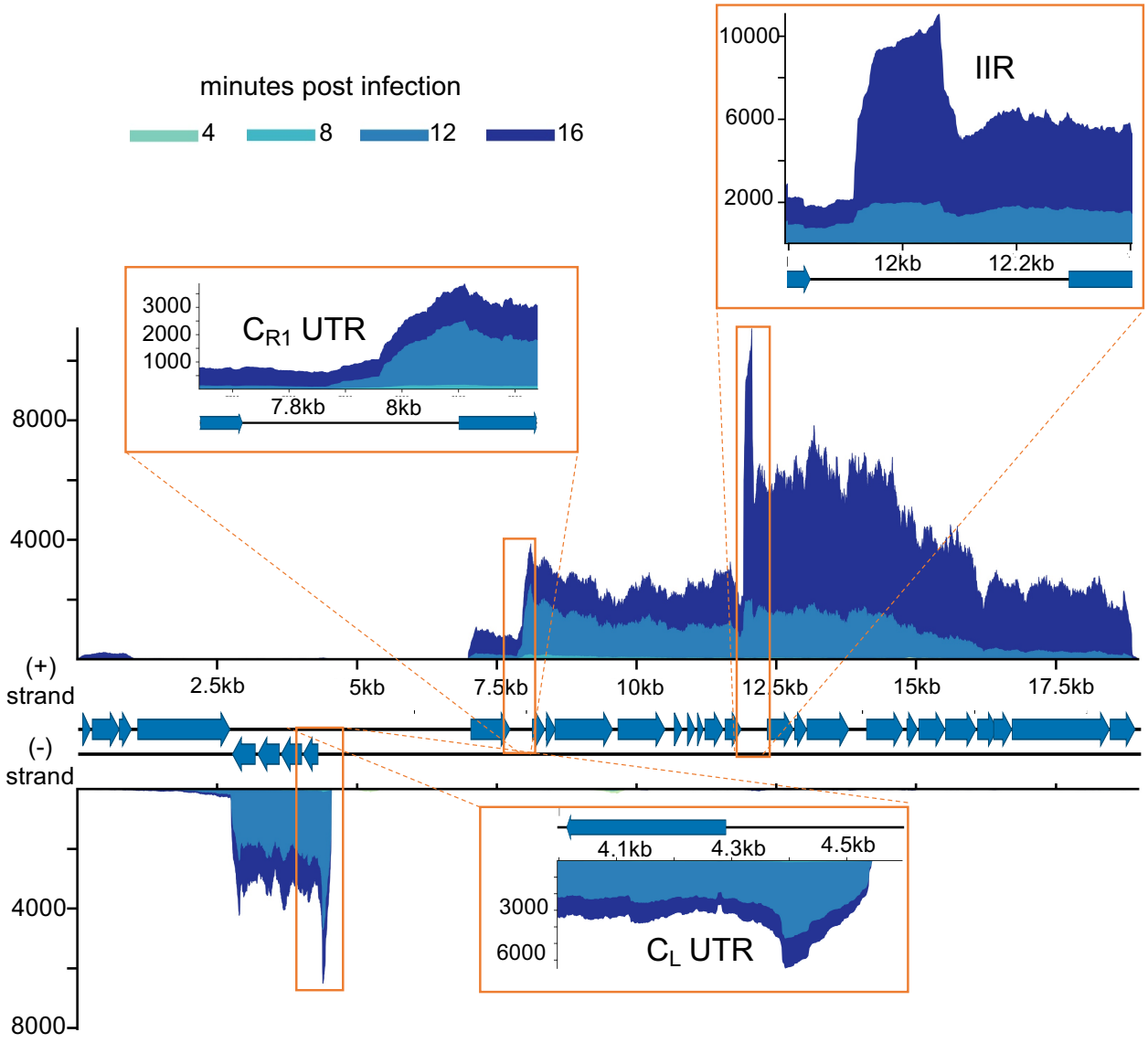
S6 Fig. PLE2 noncoding RNA
PLE2's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.



S7 Fig. PLE3 noncoding RNA
PLE3's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.



S8 Fig. PLE4 noncoding RNA
PLE4's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.



S9 Fig. PLE5 noncoding RNA
PLE5's genome displaying average reads coverage over the course of infection on a linear scale. Reads are color coded by time point. Inserts depict detected transcripts that lack predicted coding sequence. Results incorporate gene expression values obtained for two biological replicates.

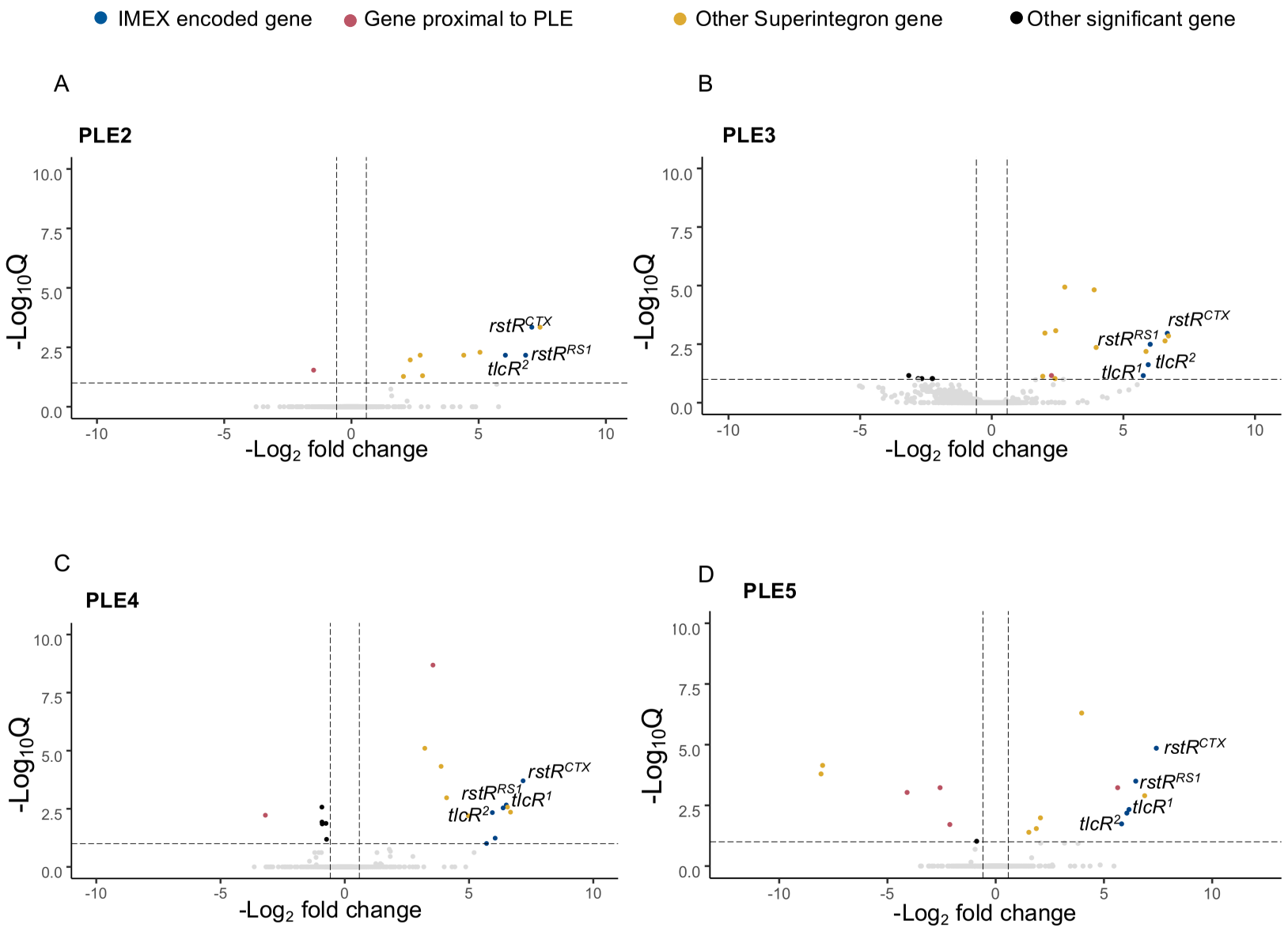
A

B

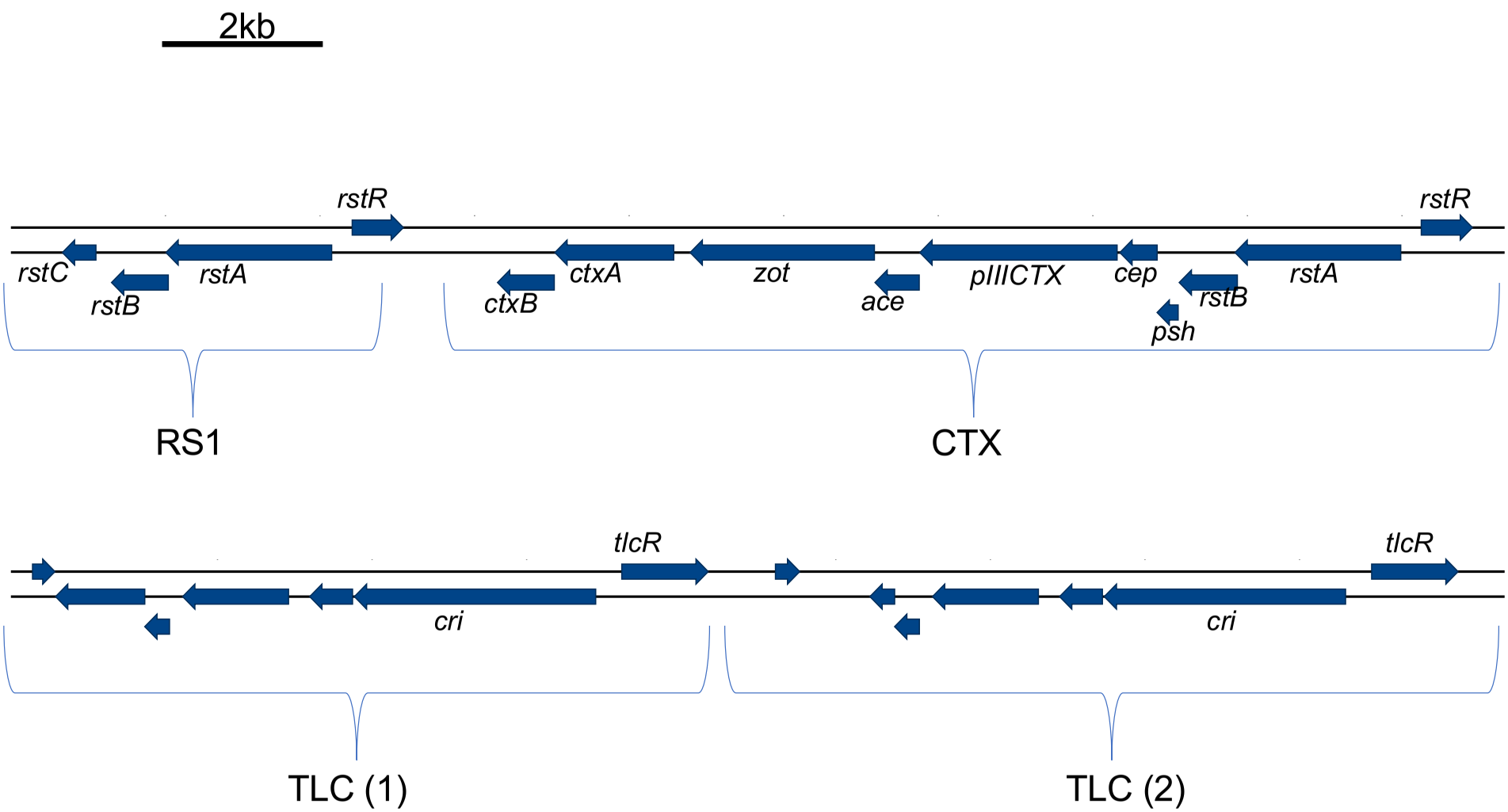
ACATAACAACAAACATAAGGAAA	Ideal PLE1 RNA target
ACATAACAACAACA**ATAAGGAAA**	Upstream 12.1
TA**ACAACAA**A**CAACATAAGGAAA**	Upstream 5
C**CTCAACAACAAA**TT**TAAGGAAA**	Upstream16
AAATAACAACAAA-**AGAAGAAGA**	Upstream 2
AACC**AACAA**T**AA**TA**ATAAGG**AG**A**	Upstream 13
TA**ATAACAACAA**T**CAT**TGA**GGA**G	Upstream 7
ACTATTTT**ACT**A**A**T**ATAAGGAAA**	Upstream 12

S10 Fig. Conservation of the PLE IIR transcript

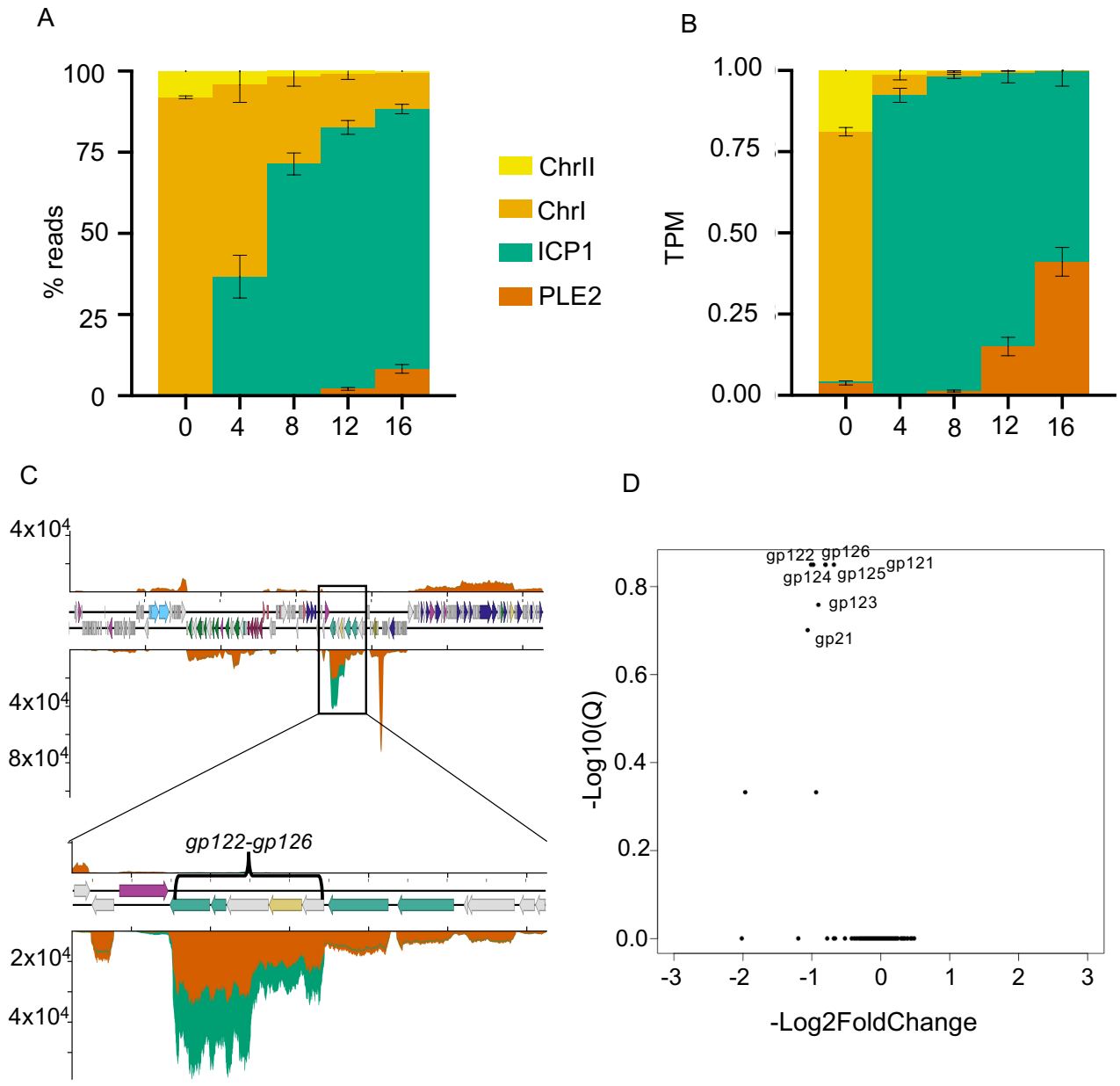
(**A**) Clustal alignment of the PLE IIR transcript across all PLEs. Terminal inverted repeats are highlighted in grey. Sequence with antisense homology to PLE gene leader sequences is bolded. (**B**) Alignment of the an 'ideal' leader sequence antisense to the IIR, and the leader sequences of several genes in PLE1. Matching sites are bolded. Mismatches are shown in light gray.



S11 Fig. Some *V. cholerae* genes are differentially expressed in the presence of PLEs
Volcano plots showing differential regulation of *V. cholerae* genes in uninfected PLE(+) cultures relative to PLE(-) cultures. Genes within 2kb of the PLE integration site are colored red, other genes within the superintegron are colored yellow, and genes encoded in RS1,CTXφ, or TLC are colored blue. Genes that are not significantly differentially regulated are colored light gray, and other significantly differentially regulated genes are colored black. A Q value cut off of 0.1 was used to determine significance. PLEs 2 (A), 3 (B), 4 (C), and 5 (D) are shown.

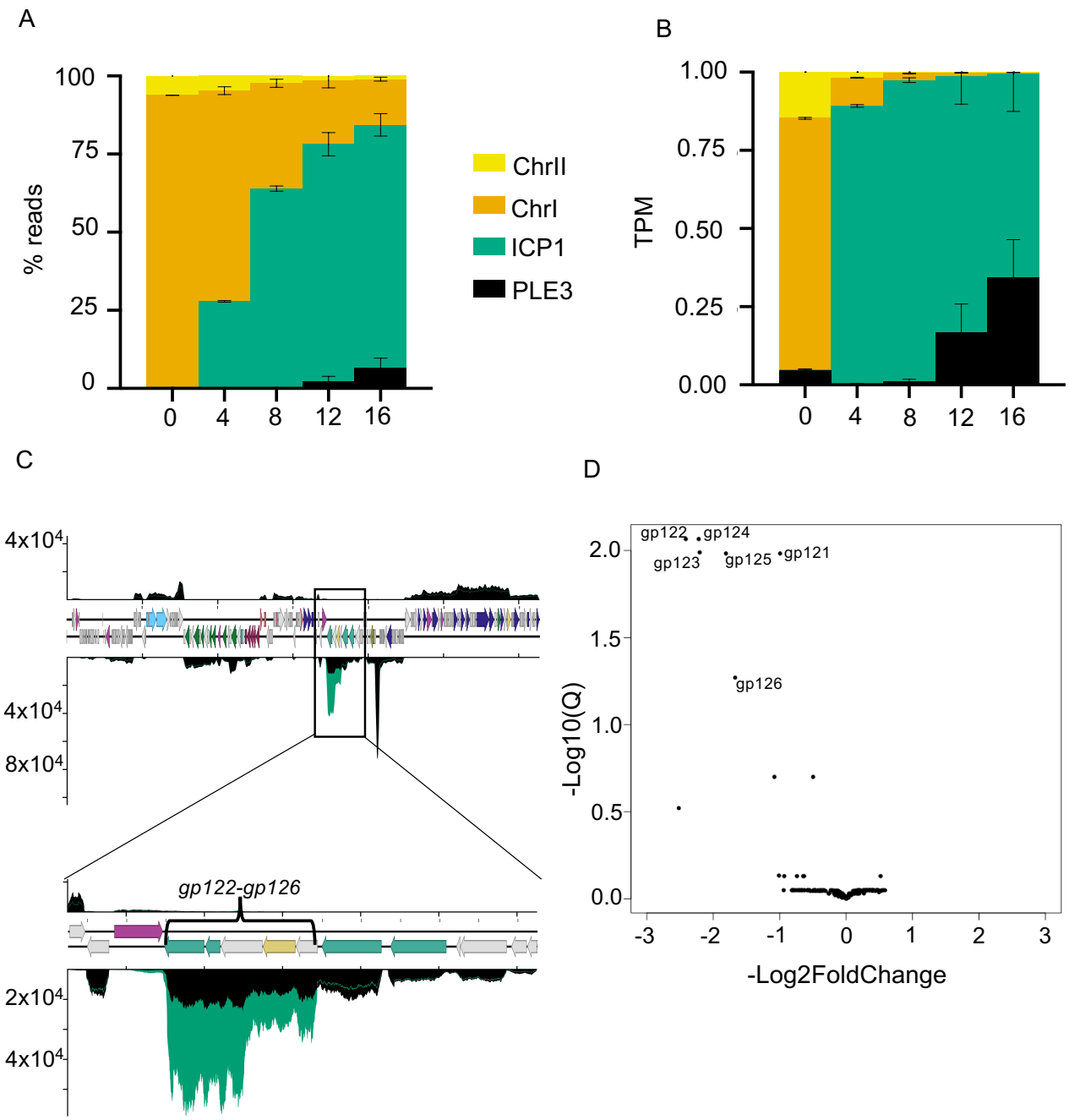


S12 Fig. Integrative mobile elements exploiting Xer in E7946 *V. cholerae*
Diagram showing IMEX elements integrated at E7946's dif site on the large chromosome. Brackets denote mobile elements.

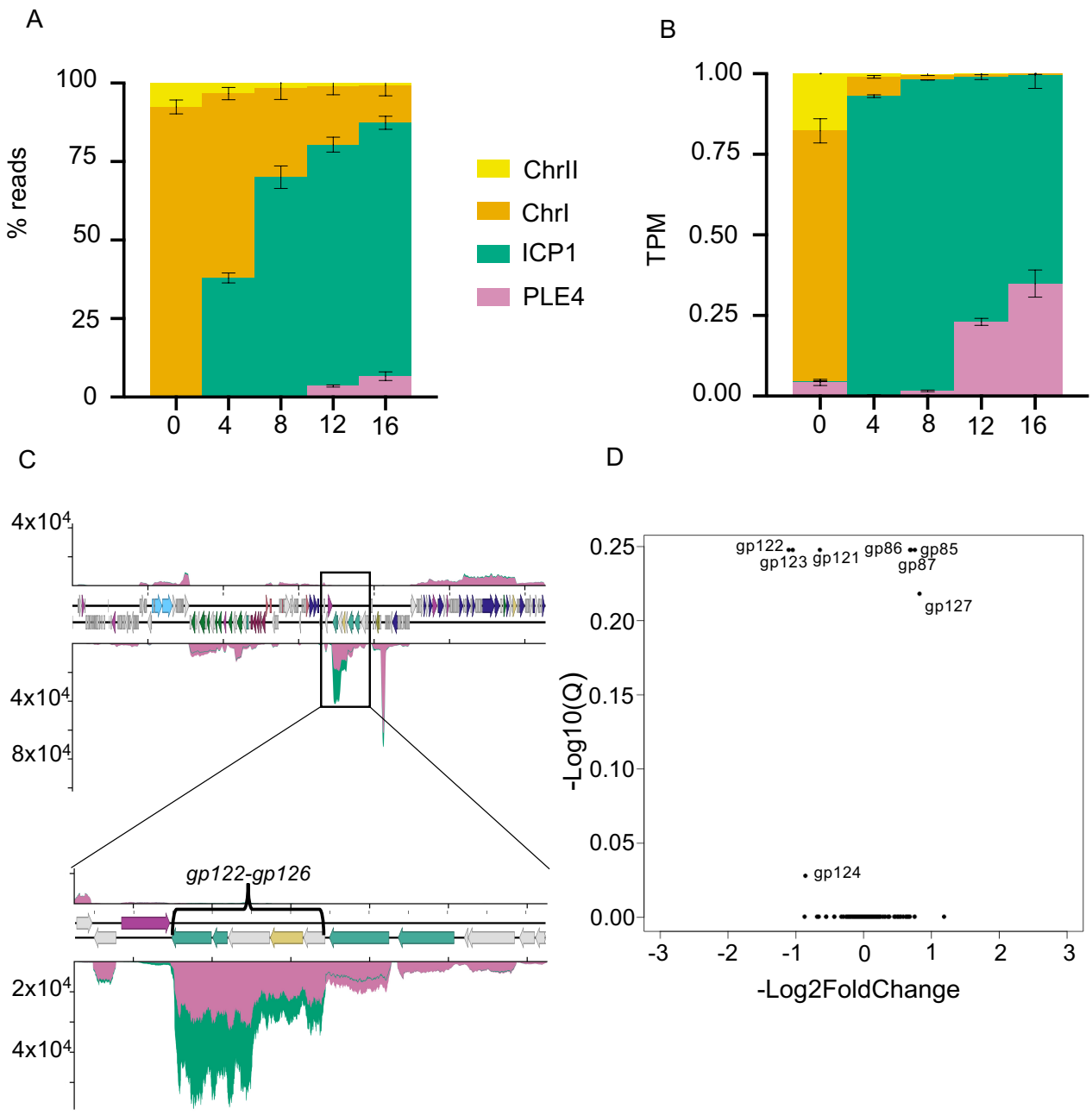


S13 PLE2 specifically downregulates ICP1 capsid production

(**A**) Percent reads abundance for PLE2, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE2 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE2 (orange) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.

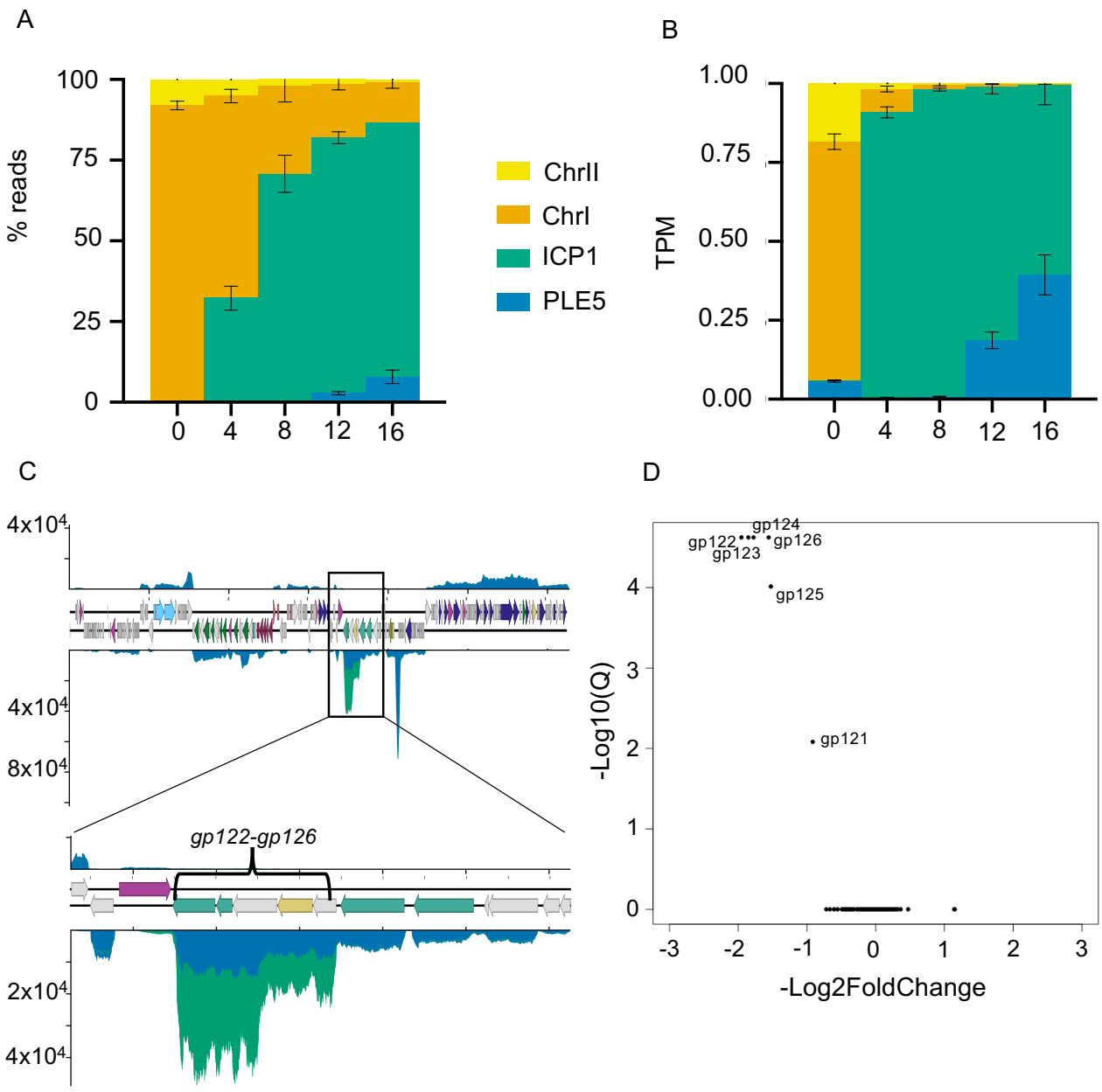


S14 PLE3 specifically downregulates ICP1 capsid production
(**A**) Percent reads abundance for PLE3, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE3 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE3 (black) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.



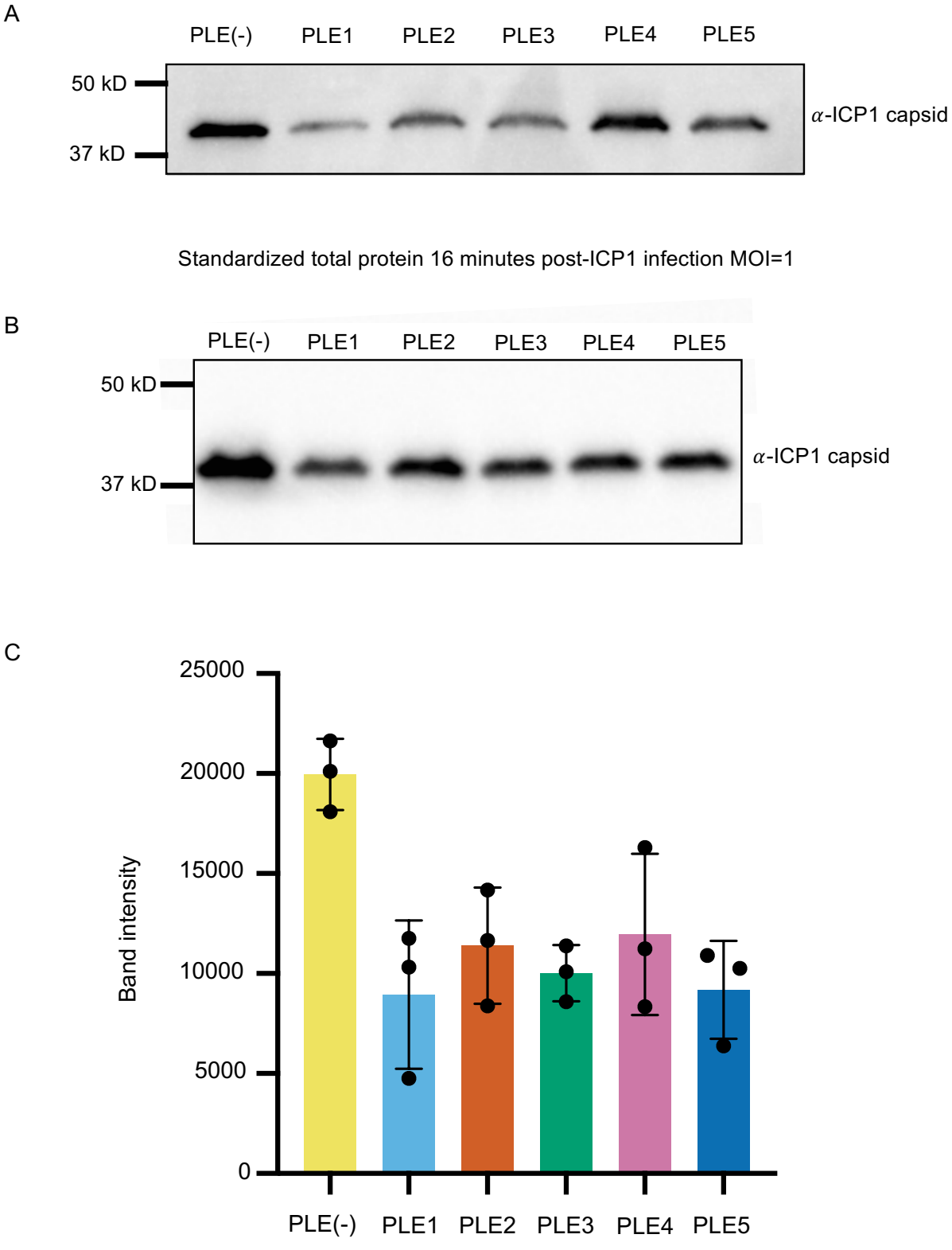
S15 PLE4 specifically downregulates ICP1 capsid production

(**A**) Percent reads abundance for PLE4, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE4 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE4 (pink) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates



S16 PLE5 specifically downregulates ICP1 capsid production

(**A**) Percent reads abundance for PLE5, *V. cholerae* chromosomes and ICP1 over the infection time course. (**B**) Reads normalized to a TPM value based on the total number of reads from each element, and the element's length. (**C**) Volcano plot of ICP1 differential gene expression in the PLE5 culture relative to the PLE(-) culture at 16 minutes post infection. (**D**) Average relative coverage along ICP1's genome in PLE(-) (green) and PLE5 (blue) cultures at 16 minutes post infection. The insert depicts ICP1's head morphogenesis operon. For all panels, results incorporate values obtained for two biological replicates.



S17 PLEs reduce the level of expressed capsid protein
(A) and **(B)** Biological replicates of Western blots against Gp122, ICP1's major capsid protein from infections of PLE(-), PLE1, PLE2, PLE3, PLE4, and PLE5 cultures at 16 minutes post infection. **(C)** Quantification of the Western blot band intensities from (A) (B) and Fig 7E.