

**Proteomic and transcriptomic analysis of *Microviridae*  $\phi$ X174 infection reveals broad up-regulation of host membrane damage and heat shock responses.**

Bradley W Wright<sup>a</sup>, Dominic Y Logel<sup>a</sup>, Mehdi Mirzai<sup>a,b</sup>, Dana Pascovici<sup>b</sup>, Mark P Molloy<sup>c\*</sup>, Paul R Jaschke<sup>a\*#</sup>

<sup>a</sup> Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

<sup>b</sup> Australian Proteome Analysis Facility (APAF), Macquarie University, Sydney, NSW, Australia

<sup>c</sup> Kolling Institute, Northern Clinical School, The University of Sydney, Sydney, NSW, Australia

\* Co-senior.

Running Title: Host response to  $\phi$ X174 infection

# Address correspondence to Paul R. Jaschke, [paul.jaschke@mq.edu.au](mailto:paul.jaschke@mq.edu.au)

## ABSTRACT

Measuring host-bacteriophage dynamics is an important approach to understanding bacterial survival functions and responses to infection. The model *Microviridae* bacteriophage  $\phi$ X174 is endemic to the human gut and has been studied for over seventy years but the host response to infection has never been investigated in detail. To address this gap in our understanding of this important interaction within our microbiome we have measured host *Escherichia coli* C proteomic and transcriptomic response to  $\phi$ X174 infection. We used mass spectrometry and RNA-seq to identify and quantify all 11  $\phi$ X174 proteins and over 1,700 *E. coli* proteins, enabling us to comprehensively map host pathways involved in  $\phi$ X174 infection. Most notably, we see significant host responses centered on membrane damage and remodeling, cellular chaperone and translocon activity, and lipoprotein processing, which we speculate is due to the peptidoglycan-disruptive effects of the  $\phi$ X174 lysis protein E on MraY activity. We also observe the massive upregulation of small heat-shock proteins IbpA/B, along with other heat shock pathway chaperones, and speculate on how the specific characteristics of holdase protein activity may be beneficial for viral infections. Together, this study enables us to begin to understand the proteomic and transcriptomic host responses of *E. coli* to *Microviridae* infections and contributes insights to the activities of this important model phage.

## IMPORTANCE

A major part of the healthy human gut microbiome are the *Microviridae* bacteriophage, exemplified by the model  $\phi$ X174 phage. Although much has been learned from studying  $\phi$ X174 over the last half century, until this work, the *E. coli* host response to infection has never been investigated in detail. We reveal the proteomic and transcriptomic pathways differentially regulated during the  $\phi$ X174 infection cycle, and uncover the details of a coordinated cellular response to membrane damage that results in increased lipoprotein processing and membrane trafficking, likely due to the phage antibiotic-like lysis protein. We also reveal that small heat shock proteins IbpA/B are massively upregulated during infection and that these holdase chaperones are highly conserved across the domains of life, indicating that reliance on them is likely widespread across viruses.

## KEYWORDS

Bacteriophage, chaperone, *E. coli*, gut microbiome, lipoprotein, mass spectrometry, RNA-seq, sHSP

## INTRODUCTION

Successful phage reproduction relies on suppression and evasion of host defense responses (1), and to varying extents, requisitioning host machinery (2). Temporal infection studies employing the use of transcriptomics (3, 4) and proteomics

(5-11) have revealed insights into the establishment of infection and utilization of host machinery. In particular, the use of modern mass spectrometry work-flows and instruments to identify and quantify thousands of proteins, is an invaluable tool for the study of complex biological systems especially when paired with complimentary RNA-sequencing (RNA-seq) methods. Currently, most phage-host interactions, even of model systems, have not yet been characterized using these methods.

Bacteriophage  $\phi$ X174 is a member of the family *Microviridae*, which make up a group of small icosahedral viruses encoded by a single-stranded DNA genome. The small 5,386 nucleotide genome of  $\phi$ X174 containing 11-protein encoding genes makes it a tractable target for proteomic and transcriptomic studies (12, 13). *Microviridae* have been used as model systems for DNA sequencing, genome engineering (14, 15), evolution (16), icosahedral virus packaging (17, 18), and novel virus infection mechanisms (19, 20). Furthermore,  $\phi$ X174 and other *Microviridae* are universal members of the human gut microbiome (21, 22), although their precise role is not well understood.

In this investigation we applied modern proteomics and complementary transcriptomics to map for the first time the temporal response of *Escherichia coli* NCTC122 (C122) to phage  $\phi$ X174 infection. Our detailed measurements of protein and RNA revealed the heat shock chaperone network is selectively up-regulated during infection, along with many proteins involved in membrane repair and remodeling, and membrane transporters. Most notably, we observed massive up-regulation of small heat-shock proteins (sHSPs) IbpA and IbpB at both the protein and RNA level. This

work lays the foundation for further studies of *Microviridae* infections and their interplay within a range of more natural host strains.

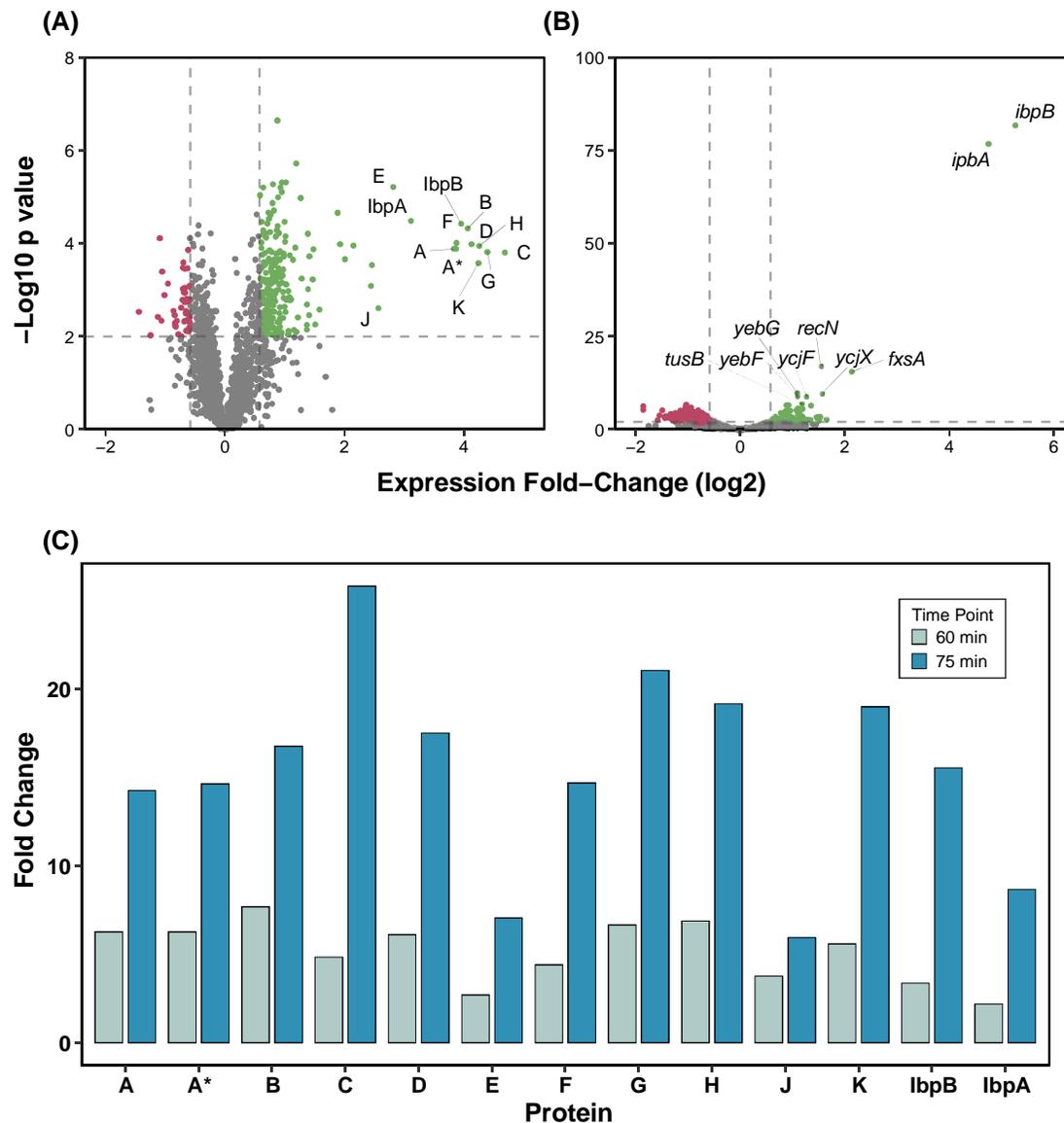
## RESULTS

### *φX174 infection results in moderate changes to the host proteome and transcriptome just prior to lysis*

To measure the host response to  $\phi$ X174 infection, we first used a mass spectrometry proteomic work-flow to compare  $\phi$ X174 infected *E. coli* C122 and mock-infected cells at five time-points (0, 15, 30, 60, and 75-minutes post-infection). We found no differences between the host proteome expression and mock-infections until 60 and 75-minutes, corresponding to the point at which we observed lysis under our conditions (Figure S1A). Mass spectrometry analysis identified 2,184 host proteins (Supplementary File S1), representing ~57 % of the *E. coli* C122 proteome database. Of the identified proteins, we confidently measured the expression of 1,752 host proteins and all 11  $\phi$ X174 proteins. To further validate our proteomics work we also measured the transcriptional response of the host during late lysis period and mapped reads to 85% of *E. coli* protein-encoding genes and 100% of  $\phi$ X174 genes. Furthermore, during infection, phage-specific reads accounted for 18 % of the total mapped reads.

During late-infection we measured at the protein-level 255 up-regulated and 54 down-regulated proteins (Figure 1A and Supplemental File S1), while at the RNA-level 203 genes were found to be up-regulated and 325 genes were down-regulated (Figure 1B and Supplementary Figure S2). The *ibpA* and *ibpB* small heat shock proteins, which

had fold-changes (FCs) of +8.7 / +27.1, and +15.6 / +38.6 at the protein/RNA-levels respectively, were similar in magnitude and significance to the  $\phi$ X174 proteins (Figure 1C). Notably, we did not observe any significant changes to small RNAs within *E. coli* C122 due to  $\phi$ X174 infection (Supplemental File S2).



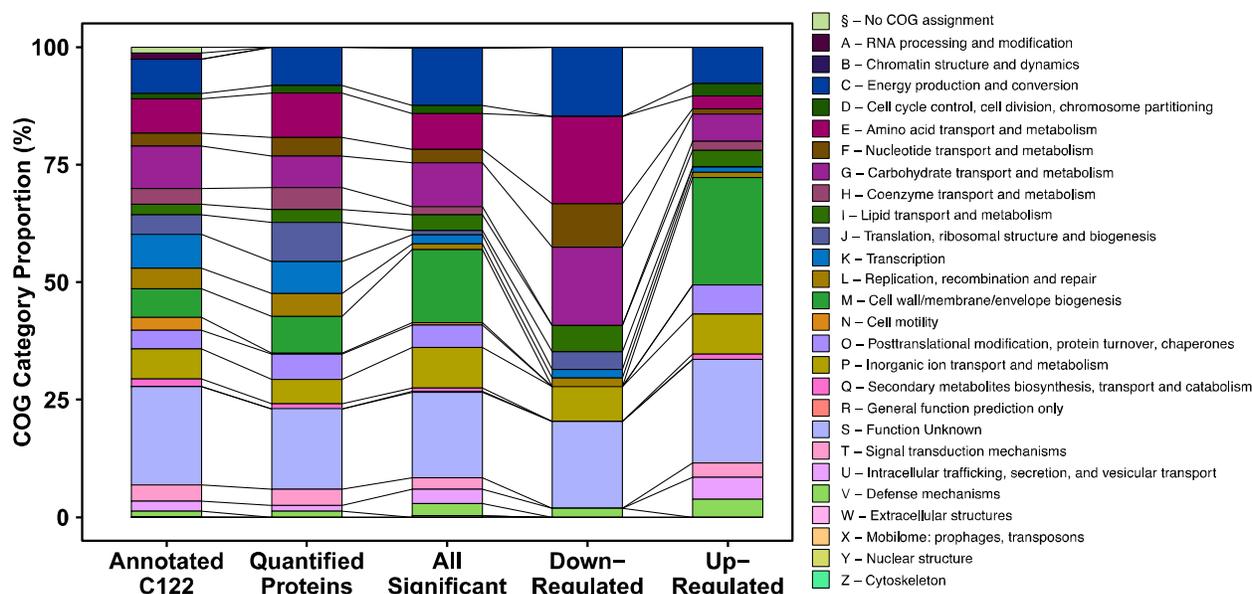
**Figure 1: A moderate number of *E. coli* C122 genes are differentially regulated during late  $\phi$ X174 infection, with *ibpAB* up-regulation equal to  $\phi$ X174 phage genes. (A) Quantified proteins. FC  $\pm$  1.5 and p-value <0.025 significance threshold. (B) Quantified RNA. FC  $\pm$  1.5 and p-value cut-off <0.05 significance threshold. (C) Fold-changes for phage proteins and host proteins IbpA and IbpB at 60-minute and 75-minutes post-infection.**

*Host engages in significant membrane protein up-regulation and metabolic protein down-regulation during infection*

To define the broader biological impacts of  $\phi$ X174 infection, we assigned all 1,752 quantified host proteins, 11  $\phi$ X174 proteins, and 528 differentially expressed RNA transcripts to their respective clusters of orthologous groupings (COG) using EggNOG-mapper (23). The results of this analysis showed no sample preparation enrichment bias as the “Annotated C122” and “Quantified Proteins” annotated biological function category COG distributions were comparable (Figure 2).

Within the group of 309 significantly differentially regulated proteins at 75-minutes post-infection (“All Significant”), we found that all proteins annotated with COG categories M (cell wall/membrane/envelope biogenesis) and O (Post-translational modification, protein turnover, chaperones) were up-regulated during infection (Figure 2). These category enrichments indicate there may be a host response mediating significant membrane restructuring or repair, or poor translocation of membrane bound proteins leading to their accumulation in the cytoplasm. Within the group of 54 down-regulated proteins we found the majority of these proteins (~59%) were in COG categories representative of energy and metabolism (C, E, F, and G) (Figure 2).

These findings were further reinforced with a PANTHER (24) over-representation test, whereby 167 of the 255 up-regulated proteins were mapped according to their GO cellular component annotations. We found the majority (60%) were mapped to the membrane component ( $p$ -value =  $1.54E-57$ ), and interestingly, of the 54 proteins significantly down-regulated as a result of phage infection, a significant proportion (62%) were mapped to the cytoplasm ( $p$ -value =  $2.10E-04$ ).



**Figure 2: Clusters of orthologous groups (COG) distributions of measured *E. coli* C122 proteins.** Annotated C122: 3,666 annotated proteins; Quantified Proteins: 1,763 proteins quantified across all conditions; All Significant: Non-redundant list of 399 significantly differentially regulated proteins across all time-points and samples; Down-regulated: 54 significantly down-regulated proteins from 75-minutes post-infection; Up-regulated: 255 significantly up-regulated proteins from 75-minutes post-infection. COGs were assigned to C122 genes using the EggNOG-mapper (23).

Corresponding with the proteomic findings, the transcriptomics showed COG category enrichment for inorganic ion transport and metabolism (P), amino acid transport and metabolism (E), and energy production and conversion (C) (Supplementary Figure S3). Additionally, the transcriptomics showed enrichment for the COG categories cell motility (N), and intracellular trafficking, secretion, and vesicular transport (U). PANTHER over-representation testing for GO cellular component terms showed significant enrichment within up-regulated genes for the pilus (6.7 fold) and the integral component of the membrane (1.6 fold). The down-regulated genes were enriched for the iron-sulfur cluster assembly complex (16.2 fold), the proton-transporting ATP synthase complex, catalytic core F(1) (13.5 fold), and the plasma membrane

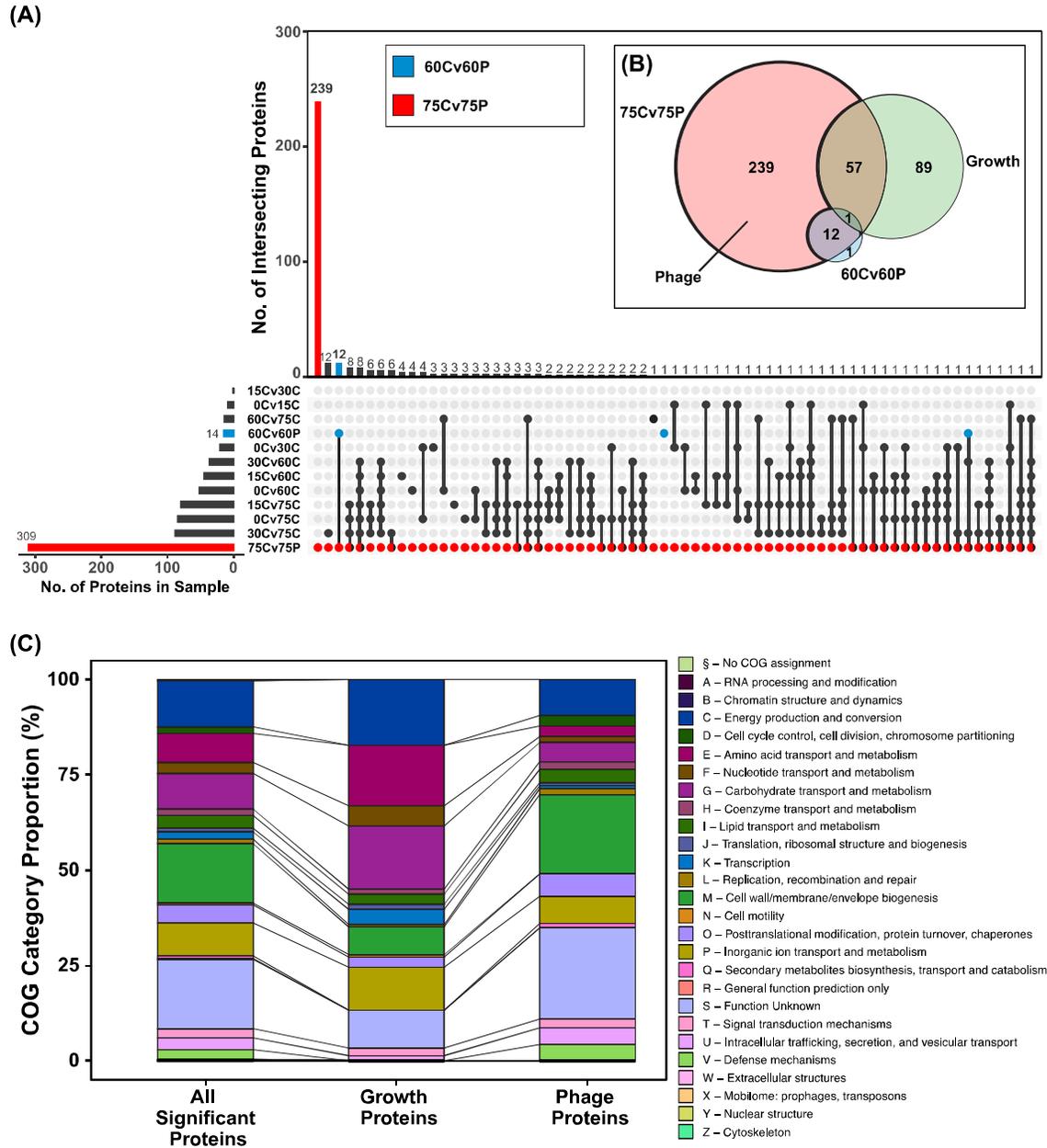
respiratory chain complex I (11.2 fold) (Supplementary Figure S4). All of these enriched cellular components are localized within the bacterial membrane. Membrane component enrichment was reflected in the differentially expressed genes more broadly as well, as gene products destined for the membrane accounted for 42 % and 24 % of the up- and down-regulated datasets, respectively.

To understand the interaction of  $\phi$ X174 protein expression and host responses to infection, we performed comparative analyses using an UpSet plot, which is a tool to visualize and group shared elements within intersecting datasets (25). We used the UpSet visualization to identify the conditions and samples in which significantly differentially regulated proteins are shared. This analysis revealed that the 60-minute post-infection samples (60C and 60P) shared 12 differentially expressed proteins with the 75-minute post-infection samples (75C and 75P): two host proteins (lbpA, and lbpB), and all  $\phi$ X174 proteins except C protein (Figure 3A and Supplementary File S1).

In contrast to the 60-minute samples, the 75-minute post-infection samples showed large-scale changes to the host proteome, with 309 differentially regulated proteins including all 11 phage proteins and the same two host proteins, lbpA and lbpB, seen at 60-minutes (Figure 3A). Within this group of 309 proteins we wanted to identify proteins that were strictly impacted due to the phage infection and not cellular growth, which was also occurring during the phage infection. To identify proteins that changed in expression due to cellular growth alone we compared all permutations of mock-infected time-points and identified 147 proteins that were differentially regulated due to cell growth. We called this group Growth (Figure 3B). Comparing the Growth proteins to the 309 proteins differentially regulated at 75-minutes post-infection, revealed a set of

58 proteins differentially regulated due to both Growth as well as  $\phi$ X174 infection (Figure 3B). Removing these 58 proteins along with the 12 proteins shared with the 60-minutes post-infection sample gave a set of 239 proteins differentially regulated due solely to phage infection. We called this group Phage (Figure 3B).

Examining the biological functions enriched within the Growth protein set revealed proteins involved in metabolic activity such as energy production and conversion (C), and amino acid (E) and carbohydrate (G) transport and metabolism (Figure 3C). In contrast, Phage proteins were enriched for a distinct set of categories: cell wall/membrane/envelope biogenesis (M), intracellular trafficking, secretion, and vesicular transport (U), and defense mechanisms (V) (Figure 3C).



**Figure 3: Sets of significantly differentially regulated proteins due to cellular growth and  $\phi$ X174 infection and their functional groupings. (A)** UpSet visualization of sample groups analyzed by proteomics. Each set represents the group of significantly regulated proteins ( $FC \pm 1.5$ ,  $p$ -value  $< 0.025$ ) extracted from comparative analyses. For example, 75Cv75P (red) is the group of significant proteins passing our stated cut-off thresholds when the 75-minute mock-infected (75C) group is compared to the 75-minute  $\phi$ X174-infected (75P) group. This plot allows visualization of redundant proteins found differentially regulated across multiple groups, and more importantly, facilitates the partitioning out of significant proteins relevant only to  $\phi$ X174 infection. Non-redundant sample size = 399 proteins. **(B)** Euler plot showing relationships between Phage, Growth, 60Cv60P, and 75Cv75P protein sets. **(C)** Clusters of orthologous

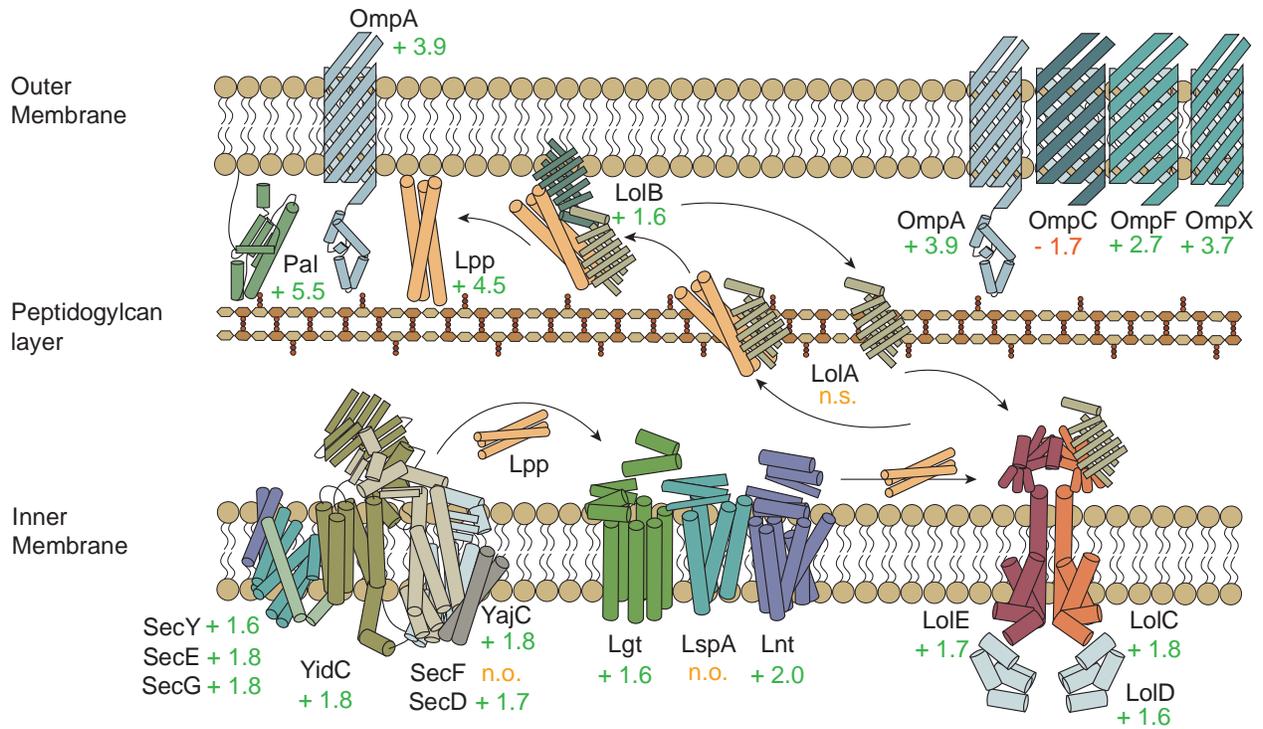
groups (COG) categories of the Growth and Phage protein sets as compared to 399 significant proteins (“All Significant Proteins”). See Supplemental File S1. COGs were assigned to C122 genes using the EggNOG-mapper (23).

*Disruption to peptidoglycan synthesis and up-regulation of membrane tethering lipoproteins associated with  $\phi$ X174 protein E expression*

Looking deeper into the COG category M (cell wall/membrane/envelope biogenesis) we found proteins from the outer membrane lipoprotein maturation and transportation pathway were significantly up-regulated. Typically, a protein is translocated across the inner membrane from the cytoplasm by the Sec pathway, or via the twin-arginine translocation (TAT) pathway. We did not observe differential regulation of the TAT pathway, but did see the up-regulation of all components of the cytoplasmic membrane complex of the Sec pathway (SecYEG, FCs +1.6, +1.8, and +1.8, respectively) and the accessory protein YidC (FC +1.8) (26) (Figure 4). PANTHER over-representation analysis of proteins identified as up-regulated from the Phage proteins set (Supplemental File S1) highlighted that the cell envelope Sec protein transport complex is significantly over-represented ( $p$ -value < 0.0002) during infection. This complex facilitates the translocation of nascent unfolded proteins into the periplasm in an ATP-dependent fashion (27). We did not identify differential expression of SecA, an ATPase that interacts with SecYEG and SecB (28), nor the significant differential expression of SecB itself. SecF was not observed but its binding partner, SecD, was significantly up-regulated (FC +1.7). YajC also interacts with SecDF and we saw it significantly up-regulated (FC +1.8) to a similar magnitude during  $\phi$ X174 infection (Figure 4).

After nascent lipoproteins are translocated through the inner membrane via the Sec pathway machinery, a diacylglycerol group is added to the protein by the inner membrane diacylglycerol transferase protein Lgt (29). We observed the significant up-regulation of Lgt in the phage infected samples (FC +1.6) as well as the Lnt protein that subsequently performs N-terminal acylation (FC +2.0) (29), but we did not observe signal peptide cleaving protein LspA (Figure 4). The ATP-driven LolCDE inner membrane complex (FCs +1.8, +1.6, and +1.7, respectively) releases lipoproteins to the periplasmic chaperone protein LolA (expression not significantly changed) (30). From there, LolA facilitates the lipoprotein's transfer to the outer membrane receptor LolB (FC +1.6) which mediates their integration into the outer membrane (31) (Figure 4).

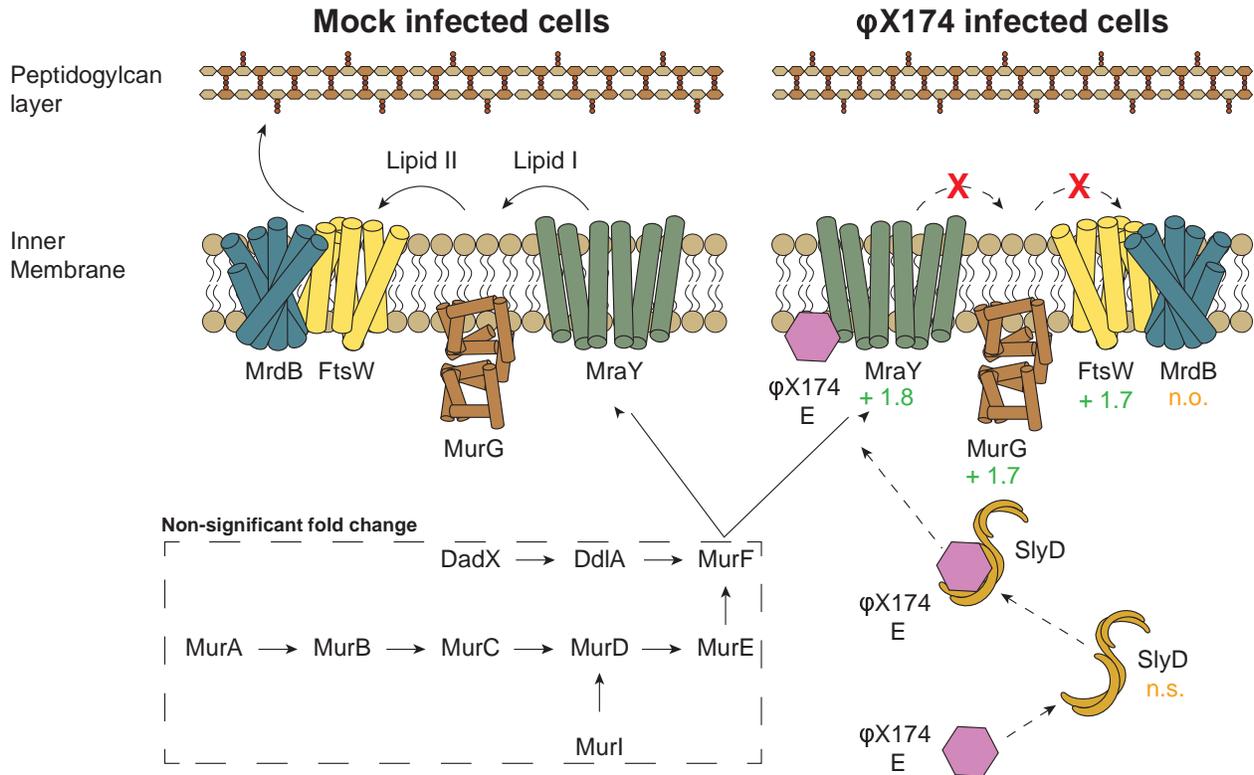
The up-regulation of the lipoprotein processing and trafficking machinery was in concordance with the up-regulation of the major outer membrane lipoprotein (Lpp) in the Phage protein set (FC +4.5) (Figure 4). Lpp tethers the outer membrane to the peptidoglycan layer through a covalent linkage, facilitating structural integrity and maintenance of cell shape (32). Lpp attachment to the outer membrane is dependent on LolB (FC +1.6), which if depleted, is detrimental to the cell due to Lpp mislocalization (33). Similar to Lpp, we saw up-regulation of OmpA and Pal (FCs +3.9 and +5.5), which also function in maintenance of cell membrane integrity, and act through non-covalent interactions connecting the outer membrane and peptidoglycan layers (34, 35) (Figure 4). This result suggests the cell, during  $\phi$ X174 infection, was experiencing membrane integrity problems, possibly of the peptidoglycan layer, resulting in the up-regulation of these membrane anchoring proteins.



**Figure 4: *E. coli* C122 upregulating membrane stabilizing and anchoring proteins Lpp, OmpA, and Pal, as a result of protein E membrane disruption.** Up-regulation of the membrane lipoprotein maturation and transportation pathway. Depicted are nascent lipoproteins translocating across the inner membrane through the Sec pathway, followed by processing, and Lol pathway-mediated export to the outer membrane. Outer membrane proteins are depicted - including OmpA, which has its peptidoglycan anchoring role highlighted along with Lpp and Pal. Values indicate protein fold-change during infection (75P/75C protein samples), n.o.: protein not observed, n.s.: protein observed but not significantly altered in abundance between conditions.

Previous work has demonstrated that phage lysis protein E acts to directly disrupt the activity of integral inner membrane protein MraY (36-38). MraY catalyzes the first membrane associated step of peptidoglycan biosynthesis (39), and subsequent interruption of MraY function by protein E is believed to result in the loss of membrane integrity through synthesis inhibition (37). Interestingly, we saw no up-regulation of the cytoplasmic component of the peptidoglycan biosynthetic pathway that lies immediately upstream of the MraY protein, but did see up-regulation of MraY (FC +1.8) and all

observed downstream inner membrane proteins in this pathway: MurG (FC +1.7), and FtsW (FC +1.7) (Figure 5).



**Figure 5: Selective up-regulation of the membrane bound steps of peptidoglycan synthesis observed during infection.** The upregulated MraY, MurG, and FtsW are putatively a result of protein E inhibited MraY and subsequent peptidoglycan destabilization due to reduced synthesis. Values indicate protein fold-change during infection (75P/75C protein samples); n.o: protein not observed; n.s: protein observed but not significantly altered in abundance between conditions.

In contrast to the extensive proteome-level response, only 16 differentially expressed RNA transcripts were attributed to COG category M. Within this grouping, only two genes, *murI* and *yjdB*, are involved in lipopolysaccharide biosynthesis. *murI*, which is responsible for lipopolysaccharide core region biosynthesis by catalyzing the

racemization of L-glutamate to D-glutamate (40) was down-regulated (FC -1.5) while *eptA* which catalyzes the addition of phosphoethanolamine to the lipid A glucosamine disaccharide (41) was up-regulated (FC +1.6). These results highlight divergence between the captured proteomic and transcriptomic responses.

*Transcriptomics identifies significant up-regulation of the phage-shock protein stress response operon revealing additional cellular response to membrane damage*

A major transcriptional response of *E. coli* C122 to  $\phi$ X174 infection was the up-regulation of gene products responsible for membrane repair functions and the maintenance of the proton motive force. This was primarily seen through the up-regulation of multiple components of the phage shock protein (PSP) response whereby *pspB* (FC +1.7), *pspC* (FC +1.7), and *pspD* (FC +1.8) in the *pspABCDE* operon, and *pspG* (FC +1.8), were up-regulated. Collectively, these genes encode protein products responsible for the binding of PspA to damaged membranes for repair functions. At the proteome-level, significant up-regulation was only observed for PspB (FC +2.1), while PspE was down-regulated, and no change was observed for PspA, PspC, and PspF. Two other genes, *ycjF* and *ycjX*, associated with the PSP response in some species (42), were also upregulated at both the transcriptional and protein levels (RNA FCs +2.4 and +3.0 (Figure S2B); Protein FCs +2.6 and +1.7).

One of the most up-regulated genes during infection was *fxsA* (FC +4.4), which was not identified in the proteomic analysis. Its only ascribed function is to competitively sequester PifA in the cytoplasmic membrane, preventing PifA from performing its host

suicide mechanism, known as F exclusion. F exclusion disrupts the bacteriophage T7 lifecycle by arresting all viral and cellular process upon PifA interaction with T7 products gp1.2 and gp10A (43, 44). No annotated version of *pifA* is in the available C122 genomes and it is normally found on the F plasmid, which was not present in the cells in this work.

### *φX174 infection significantly alters E. coli metabolic gene expression*

To understand the effects of the phage infection on the 58 proteins detected in both Growth and Phage sets (Figure 6D) we classified them based on their three distinct patterns of differential regulation (Figure 6A-C). The first set showed φX174 infection dampening protein expression across 17 genes (Figure 6A). PANTHER GO biological process analysis of this list found enrichment in proteins involved in the tricarboxylic acid cycle (4 proteins, p-value = 5.79E-06), indicative of a disruption to overall metabolic activity during phage infection. The second group of genes showed increased expression in the Phage protein set (Figure 6B). PANTHER GO molecular function analysis on these 14 proteins found enrichment for transporter activity (9 proteins, p-value = 3.59E-06), or more specifically, carboxylic acid transmembrane transporter activity (5 proteins, p-value = 8.50E-06). This is reflective of COG distributions of the entire 58 proteins (Figure 6E), where we see the majority of the highlighted categories being trafficking or transporter groups. The third case showed inverse protein expression pattern between Growth and Phage sets (Figure 6C), which is especially intriguing as it would indicate these proteins are particularly important to phage infection. A PANTHER GO biological process or molecular function analysis of

these 28 genes revealed no significant enrichment. However, 15 of the 28 proteins are located in the cell envelope ( $p$ -value =  $2.82E-11$ ) indicating significant and likely diverse transporter functions from this group of proteins. Of particular interest are aforementioned proteins Lpp and OmpA, as well as OmpA structural homolog, OmpX (FC +3.7) (Figure 4) (45), all of which were found to have large fold-change inductions during phage infection, but were down-regulated during mock-infection (Figure 6C).

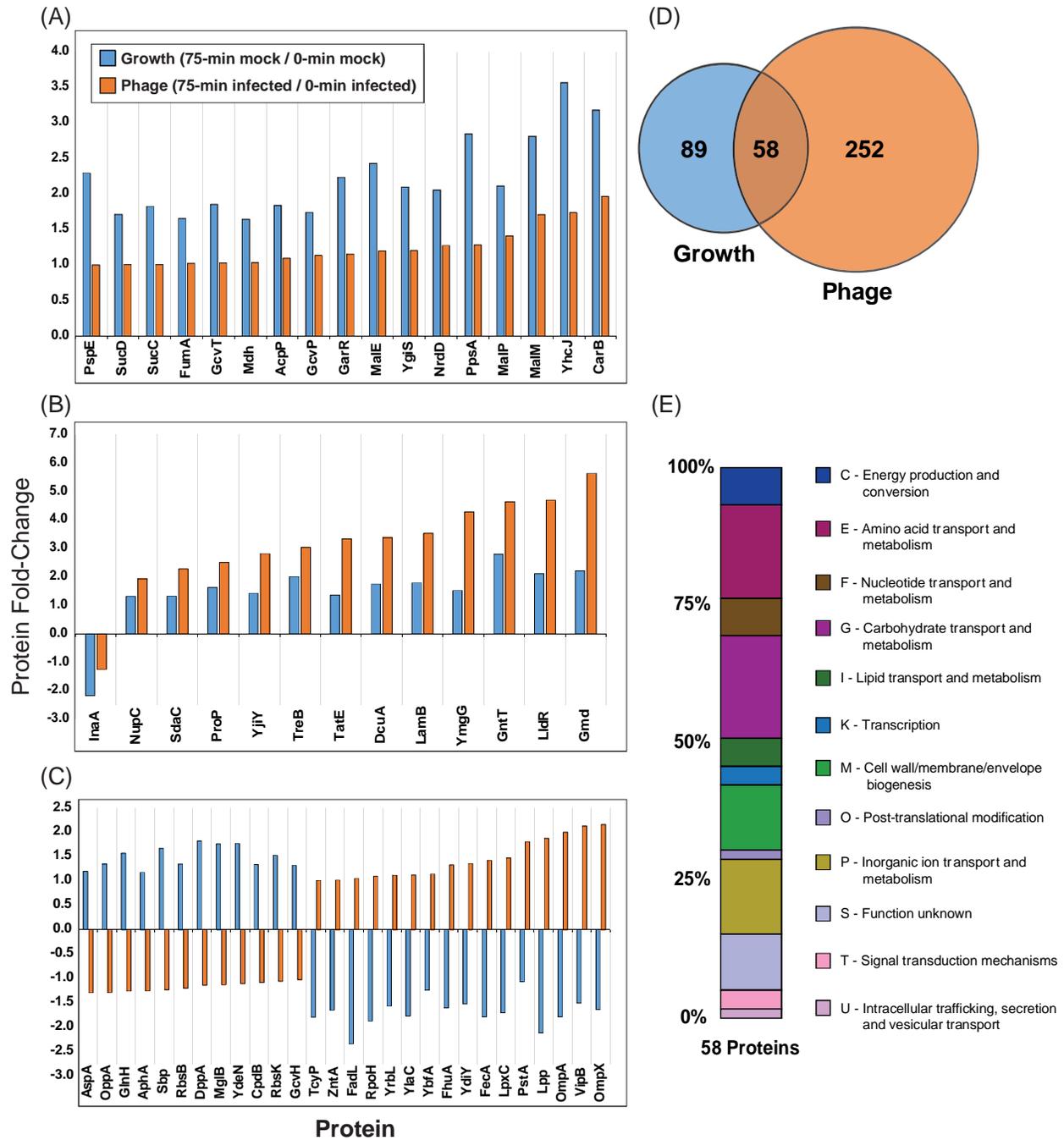
Interestingly, OmpA was recently identified as a competitive binder of the lipoprotein RcsF (45), which we also identified as significantly up-regulated during phage infection (FC +1.6). RcsF is a positive regulator of the Rcs stress response (46), of which we identified the significant up-regulation of Rcs components RcsD (FC +1.7) during infection.

Biological process GO term enrichments of the transcriptomics data showed broad representation of two different host functions which affected host metabolic activity when mapped to COG categories. The first of these host functions corresponded to the COG categories nucleoside transport and metabolism (F), amino acid transport and metabolism (E), and carbohydrate transport and metabolism (G). The COG categories were represented with the enriched biological functions of plasma membrane glucose import (17.9 fold, 4 genes,  $p$ -value =  $4.50E-04$ ), glycine catabolic process (14.3 fold, 4 genes,  $p$ -value =  $7.77E-04$ ), nucleoside transmembrane transport (11.2 fold, 5 genes,  $p$ -value =  $3.59E-04$ ), and nucleobase-containing small molecule interconversion (5.5 fold, 7 genes,  $p$ -value =  $7.54E-04$ ). The second host function mapped to COG categories corresponded to energy production and conversion (C) with the enriched biological functions ATP synthesis coupled proton transport (11.2 fold, 5

genes, p-value = 3.59E-04), gluconeogenesis (7.7 fold, 6 genes, p-value = 4.24E-04), and glycolytic process (6.0 fold, 7 genes, p-value = 4.81E-04) (Figure S4).

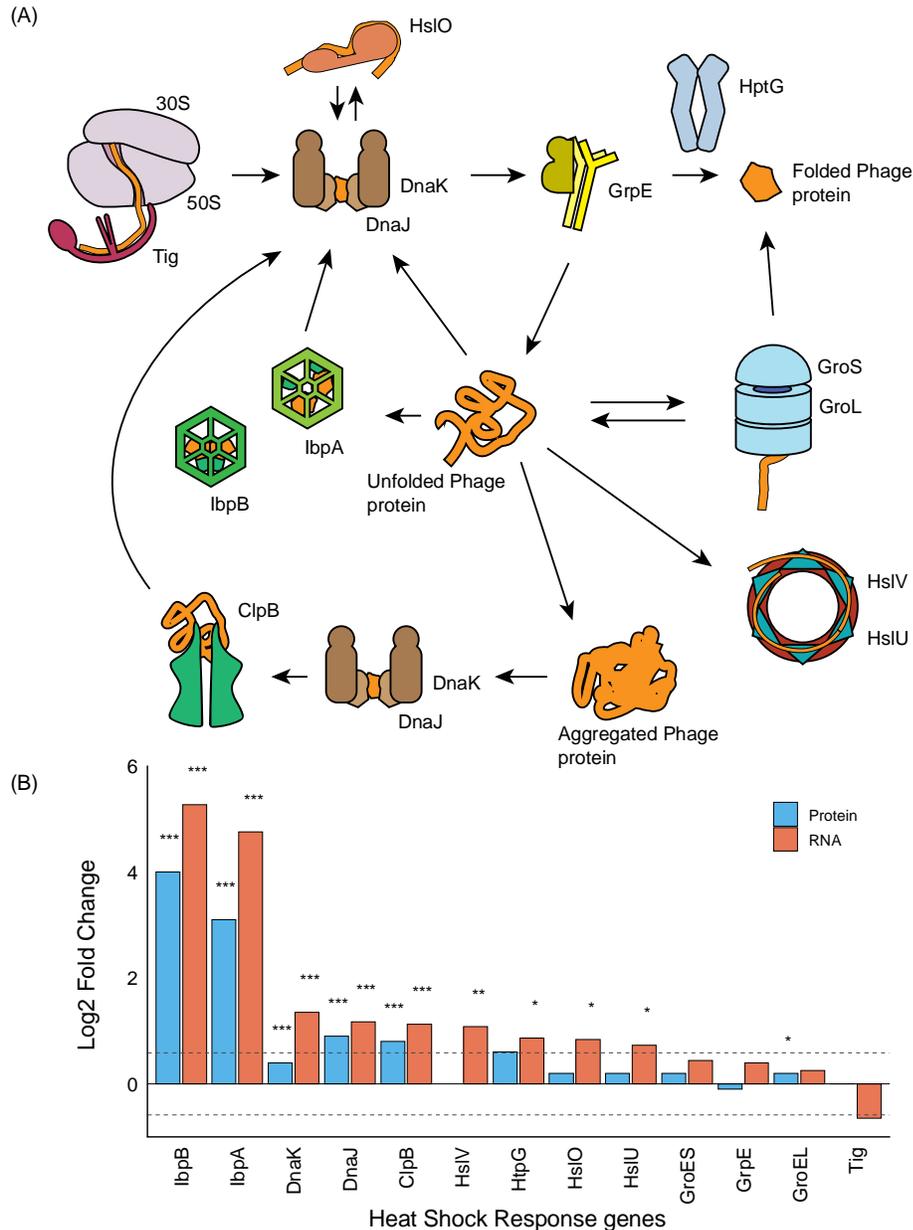
### *φX174 infection strongly activates heat shock chaperone response*

The strongest host response we observed at both the protein and RNA level was the heat shock response, led by the small heat shock proteins (sHSPs) IbpA and IbpB (Figure 1). Our proteomics and transcriptomic data showed that they had the largest increase in expression of any host gene products (Figure 1C). Both *ibpA* and *ibpB* are controlled by the  $\sigma^{32}$  regulon (47) and their large up-regulation during  $\phi$ X174 infection is consistent with the significant increase in RpoH transcription factor  $\sigma^{32}$  (FC +2.0), along with chaperones ClpB (FC +1.7) and DnaJ (FC +1.9) (Figure 7). Notably, only a subset of the  $\sigma^{32}$  heat-shock stress response regulon was activated during  $\phi$ X174 infection, as we observed only 11 proteins and 17 transcripts belonging to this regulon differentially regulated (Figure 7, and Supplementary File S3). Similarly, the general stress response master regulator,  $\sigma^{38}$  (RpoS) was not differentially regulated and only a small proportion of proteins (5%) from the genes under its control were differentially regulated during phage infection (Supplementary File S3).



**Figure 6: Differential expression patterns of the 58 *E. coli* C122 proteins shared by Growth and Phage sets.** Subset of the 58 host proteins where  $\phi$ X174 infection: (A) reduces, (B) increases, (C) inverts expression. (D) Euler plot of all significantly differentially expressed proteins within the Growth and Phage sets. (E) Clusters of orthologous groups (COG) functional categories used to represent major biological functions and their percentage (%) distribution across different groupings for the 58 shared proteins. COGs were assigned to C122 genes using the EggNOG-mapper (23).

Together, these data indicate that  $\phi$ X174 infection was not inducing a general stress response via the  $\sigma^{38}$  or  $\sigma^{32}$  regulon, and that there was selective up-regulation of members of the heat shock chaperone network (Figure 7).



**Figure 7: Large changes to gene expression across the heat shock chaperone and protein folding pathway in *E. coli* C122 during  $\phi$ X174 infection.** Proteins depicted in this pathway are transcriptionally regulated by transcription factor RpoH ( $\sigma^{32}$ ), which we saw up-regulated during infection. (A) The chaperone system revealing

the process controlling protein folding from synthesis to native protein state. The chaperone trigger factor (Tig) binds to the translating ribosome and the growing peptide chain. The DnaK/J chaperone proteins aid in folding by binding to hydrophobic regions of nascent proteins preventing their misfolding and aggregation. DnaK/J may bind to HtpG to reactivate inactive protein substrates on the path to native protein fold (48). The holdase HslO binds to unfolded protein which is released to DnaK for folding (49). GrpE may interact with DnaK thereby dissociating ADP and releasing the bound protein (50). Small heat shock chaperones IbpA/B bind nascent proteins preventing their irreversible aggregation into inclusion bodies, and await the availability of folding chaperones DnaK/J and GroEL/ES. The main folding machine of the system is the GroEL/ES complex. This ATP-dependent system facilitates folding of nascent proteins within the complex's inner surface releasing proteins in their native state. In the case of protein aggregates produced from improperly or partially folded proteins, the chaperone ClpB mediates their disaggregation for their re-folding. Unfolded protein may also shuffle to the HslUV protease/ATPase complex for proteolysis (51). (B) Differential expression of proteins and transcripts measured at lysis. Dashed lines indicate the Log<sub>2</sub> fold change  $\pm$  0.58 values used as significance criteria. P-values are shown as \*\*\* for  $\leq$  0.001, \*\* for  $\leq$  0.01, and \* as  $\leq$  0.05.

## DISCUSSION

In this work, we have comprehensively measured the *E. coli* C122 host response to  $\phi$ X174 infection at the proteomic and transcriptomic level, representing the first characterization of the dynamics of *Microviridae* infection at this level of detail. We identified 2,184 proteins and quantified all 11 phage proteins and 1,752 host proteins across seven time-points spanning the entire infection cycle. Transcriptomics identified significant changes during infection including differential expression of 528 genes, constituting 11 % of the total transcriptome, similar to other studies (52).

We found no evidence for substantial host proteome remodeling during the early stages of infection prior to 60-minutes (Figure 3A), which aligns with previous work showing  $\phi$ X174 has a minimal impact on host protein production until just prior to lysis (53). It is well-established that phage infection can cause drastic and broad metabolic

remodeling of infected cells (11, 54, 55), which is in agreement with what we saw for both our proteomic and transcriptomic datasets.

We showed that membrane damage responses dominated the host response to  $\phi$ X174 infection (Figures 4 and 5). We also found membrane lipoprotein maturation and transportation pathway enrichment (Figure 4) which suggests an up-regulation of this pathway in response to the host's need to traffic certain lipoproteins to the outer membrane, such as major lipoprotein (Lpp). Lpp is one of the most abundant (lipo)proteins of *E. coli* (56) and is reliant on the Lol pathway for outer membrane insertion (33). We saw Lpp down-regulated during growth (FC -2.1), and up-regulated during  $\phi$ X174 infection (FC +1.9), thus emphasizing an infection mediated response. Lpp has recently been shown to regulate cell shape and mechanical rigidity by tethering the outer membrane to the peptidoglycan layer (32), which leads us to suggest that during  $\phi$ X174 infection, Lpp is up-regulated to maintain the cell's membrane integrity.

The most likely model that fits our observations would have lysis protein E from  $\phi$ X174 binding to and inhibiting MraY (36, 37) and causing a decrease in peptidoglycan synthesis which is sensed by the cell and results in upregulation of MraY and inner membrane pathway partners, MurG and FtsW (Figure 5). The cytoplasmic components of the peptidoglycan pathway and chaperone protein, SlyD, which stabilizes the  $\phi$ X174 lysis protein E and is thought to promote its interaction with the inner membrane (57) are not upregulated (Figure 5).

Secondarily, as the  $\phi$ X174-infected *E. coli* cell continues to grow, but is unable to synthesize enough peptidoglycan, the stability of the envelope becomes compromised. The cell senses this and attempts to compensate by upregulating the Lpp, OmpA, and

Pal proteins (Figure 4). OmpA also seems to have a role in maintaining envelope integrity by linking the outer-membrane to the peptidoglycan layer through the non-covalent linkage of its C-terminal domain to peptidoglycan (35). In addition to this activity, OmpA has a direct interaction with peptidoglycan stress sensory protein RcsF (45), which we also saw up-regulated during infection (FC +1.6). Pal protein is also a mediator of outer membrane and peptidoglycan stability by forming a structural linkage between both layers (34).

Host membrane responses to phage infection are common and can involve responses such as induction of polysaccharide capsule biosynthesis (58) and phage shock response (55). We saw the significant upregulation of components *pspB*, *pspC*, *pspG*, and the potentially related genes *ycjX* and *ycjF* (Figure S2), presumably in response to the loss of envelope structural integrity during the lead up to lysis. A similar induction of the phage shock response was observed with phage PRD1 infection of *E. coli* (59) and antibiotics that inhibit peptidoglycan synthesis (60).

Both proteomics and transcriptomics showed a massive up-regulation of small heat shock proteins IbpA (FC +8.7) and IbpB (FC +15.6) and their transcripts (FC +27.1 and +38.6, respectively) during infection. The IbpA/B proteins are well-described and the sHSP family is conserved across all domains of life (61). They are typically described as “holdase” proteins, as they function as chaperones binding mis-folded or un-folded (nascent) proteins to prevent their irreversible aggregation within the cytosol, while they wait for available ATP-mediated folding machinery (62, 63) (Figure 7). The large mismatch between *ibpA/B* RNA and protein fold-changes in our measurements

may be due to the RNA thermometer function encoded within their transcripts that ensures they are only maximally translated at elevated temperatures (64).

We speculate that the up-regulation of IbpA/B proteins is a conserved non-specific bacterial response to viral infection based upon their up-regulation during both phage infection (59, 65) as well as during heterologous protein overexpression (66, 67). Therefore, as a result of their currently defined function (62, 63), IbpA/B up-regulation in *E. coli* C122 during  $\phi$ X174 infection is a host response to higher than normal levels of protein production (in this case, phage protein production), and in the absence of available protein folding machinery, irreversible protein aggregation or mis-folding is prevented through the up-regulation of these holdase proteins.

Phage may also benefit from the specific features of the IbpA/B function by having a queue of proteins held in a near-folded state that can be quickly folded into their final forms by HSP70 (DnaK) and HSP100 (ClpB) activity. Proteins bound by IbpA/B are not spontaneously released and can only be released by DnaK and ClpB (64), introducing a buffering system and control point for virion assembly.

The *ibpA/B* genes have never been identified from genetic screens of host-factors necessary for capsid morphogenesis, but sHSPs are encoded in the genomes of cyanophage infecting *Synechococcus* and *Prochlorococcus* groups (68, 69). Furthermore, a bacteriophage-encoded J-domain protein (Rki) has been identified in T4-related enterobacteriophage RB43 that interacts with *E. coli* J-domain interacting chaperone protein DnaK (HSP70) and stabilizes the  $\sigma^{32}$  heat-shock response (70). MS2 phage relies on DnaJ (HSP40) (71), while the HSP60 chaperones GroEL/ES are required for phages  $\lambda$ , PRD1, HK97, and T7 morphogenesis (72-75). GroEL-like

proteins have also been found encoded within newly annotated bacteriophage genomes (76). In *groEL/groES* mutants phage T4 head morphogenesis is disrupted and results in a random aggregation of phage head proteins attached to the inner membranes (77), pointing to the critical disaggregating and/or folding role of the HSP-chaperones in phage capsid assembly. Young, et al (1989) found that lysis sensitivity increased in *E. coli* when plasmid encoded  $\phi$ X174 lysis protein E and heat-shock genes *dnaK*, *dnaJ*, *groEL*, and *grpE* are present (78). Recently,  $\phi$ X174 has been suggested to use heat shock promoters to co-opt host responses to infection to drive its lifecycle (12, 79). Therefore, together with our data, we propose that the host heat-shock response, and in particular IbpA/B, are important components for  $\phi$ X174 replication.

## MATERIALS AND METHODS

### *$\phi$ X174 infection*

Methods are composed in accordance with (80). Media and buffer components from Sigma Aldrich, unless stated otherwise. *E. coli* C122 (Public Health England NCTC122) was grown overnight at 37°C/250 RPM in phage-LB (81), then inoculated 1/100 in fresh phage-LB and grown to mid-log phase before addition of phage (37 °C/250 RPM). Wild-type  $\phi$ X174 (14) was used for all infections of *E. coli* C122 at an MOI = 5. Phage infection synchronized by pelleting mid-log cultures (4,000 RCF/8-min) and resuspending the pellet to 1/10<sup>th</sup> of the original growth volume with cold HFB buffer (82). Starved cells were incubated 16 °C/30-min to facilitate phage attachment and to prevent DNA ejection (83). After 30-min, infection was initiated with the addition of

phage-LB (37 °C) to the original volume. Controls (mock-infected) had the same volume of phage-LB and 10 % (v/v) glycerol added at mid-log as the phage-infected samples.

### *Protein purification and tandem mass tagging (TMT) labelling*

*E. coli* C122 was grown to mid-log in 200 mL culture volume (37 °C/250 RPM) then split into 2x100 mL volumes: one to be infected with  $\phi$ X174 at MOI = 5, and the other mock-infected with equal volume phage-LB with 10 % (v/v) glycerol (storage solution of  $\phi$ X174). Culture growth was synchronized with 10 mL of ice-cold HFB, and infection was initiated with 90 mL phage-LB (37 °C). At 0, 15, 30, 60, and 75-min post-infection, 12 mL was removed from each culture (without replacement) and placed on ice. At the conclusion of the time-course, intact cells were pelleted 3,220 RCF/8-min. The samples were then suspended in cold 1X phosphate-buffered saline (PBS) and washed an additional two times, discarding the supernatant. Cell pellets were stored overnight at -20 °C. Cell pellets were suspended in 500  $\mu$ L lysis buffer (100 mM Tris-HCl, 1 % (v/v) sodium dodecyl sulfate, 8 M urea, 1X protease inhibitor cocktail solution (Roche, Switzerland). Reconstituted cells were lysed with an ultrasonic probe with 15-pulses at 30 % amplitude. Samples were then pelleted 6,000 RCF/10-min and supernatant moved to new tube. Proteins were reduced 37 °C/1-hr in 10 mM dithiothreitol, followed by alkylation with 30 mM iodoacetamide for 1-hr in dark. Residual iodoacetamide quenched with an equal molarity of dithiothreitol.

Proteins were precipitated by addition of 2 mL methanol, 0.5 mL of chloroform, and 2 mL of water (all ice cold). Samples were pelleted 5,000 RCF/10-min and protein

pellet removed, washed with additional methanol, and dried. Protein pellets were solubilized with 100 mM Tris-HCl/8 M urea. Solutions were pelleted 12,000 RCF/10-min and supernatant collected in new tube. Samples were diluted to urea concentration of 1.6 M through addition of 100 mM Tris-HCl, followed by protein quantitation (Pierce™ BCA Protein Assay Kit, ThermoFisher Scientific). 150 µg of protein from each sample digested overnight/37 °C with 1.5 µg trypsin (Promega), followed by addition of further 1.5 µg of trypsin and digestion for an additional 4-hrs. Peptides were acidified with formic acid at 1 % (v/v) and then C18 stage-tip purified (84). Peptide samples were dried down using vacuum centrifugation, followed by resuspension in 200 mM HEPES buffer (pH 8.8, adjusted with NaOH). Peptides were quantified (Micro BCA™ Protein Assay Kit, ThermoFisher Scientific), and 35 µg peptide from each sample brought up to 140 µL using HEPES buffer. Normalization controls were produced through pooling mock-infected samples (3 µg of peptide per mock-infected sample) – designated N.C, and infected samples (3 µg of peptide per infected sample) – designated N.P.

Chemical labelling of peptide samples was performed using TMT10Plex™ Isobaric Label Reagent Set (ThermoFisher Scientific, USA). 0.2 mg of TMT tag was used to label each sample (Table S1). Note: 128C tag was reserved for N.C. 11.25 µg of N.C was labelled per channel. Similarly, 131 tag was reserved for N.P. 5.375 µg of N.P was labelled per channel. After labelling, 8 µL of 5 % (v/v) hydroxylamine was added to each sample to quench residual tags. Lastly, samples were combined across channels, dried down using vacuum centrifugation, and then purified as described (84).

### *High pH peptide pre-fractionation*

Dried peptide samples (TMT channels 1-4) were suspended in high-pH buffer (5 mM NH<sub>4</sub>OH, pH 10.5), and loaded onto Agilent ZORBAX Extend-C18 column (3.5 μM bead size, 300 Å pore size, 2.1 mM x 150 mM), washed with buffer A (0.1 % (v/v) formic acid, 2 % (v/v) acetonitrile) for 10-min, followed by elution with increasing gradient of buffer B (5 mM NH<sub>4</sub>OH, 90 % (v/v) acetonitrile). The gradient of buffer B was 3 % to 30 % for 55-min then to 70 % for 10-min, and finally to 90 % for 5-min at a flow-rate of 300 μL/minute. Samples were collected every minute and pooled into 13 fractions, vacuum centrifuged, and then suspended in buffer A to concentration of 0.1 μg/μL.

### *Mass spectrometry*

Samples were analyzed on a Q-Exactive mass spectrometer (ThermoFisher Scientific) coupled to EASY-nLC1000 system (ThermoFisher Scientific, USA). Peptide samples were injected onto the LC system using buffer A and were bound on a 75 μM x 100 mM C18 HALO column (2.7 μM bead size, 160 Å pore size). A flow rate of 300 nL/minute using an increasing linear gradient of buffer B (0.1 % (v/v) formic acid, 99.9 % (v/v) acetonitrile) was run from 1 % to 30 % for 110-min followed by 85 % buffer B for 10-min. The mass spectrometer was operated in top-10 mode, with a full scan set at a resolution of 70,000 (at 400 *m/z*) across the *m/z* range of 350-1850 (isolation window of 0.7 *m/z*), and an automatic gains control (AGC) target of 1e6 (or maximum fill time of 250 ms). Selected pre-cursor ions were transferred from the C-trap to the higher energy collision dissociation (HCD) cell for fragmentation at a normalized collision energy of 35

% with precursor dynamic exclusion of 90-sec. MS/MS spectra was collected at a resolution of 70,000 (at 200  $m/z$ ) with an AGC of  $1e5$ , maximum injection time of 250 ms and a fixed first mass of 115.0  $m/z$ .

Raw files exported to Proteome Discoverer v2.1 (ThermoFisher Scientific) for processing. Precursors selected for fragmentation that had greater than 30 % interference were excluded from analysis. Methionine oxidation, N-terminal carbamylation, asparagine and glutamine deamidation, N-terminal acetylation, N-terminal glutamic acid to N-pyro-glutamine, N-terminal glutamine to N-pyro-glutamine, and TMT10plex labelling of primary amines were selected as dynamic modifications. Cysteine carbamidomethylation was designated as a fixed modification. Minimum peptide length was set at 5. Spectra were searched against a custom *E. coli* C122 database that included the phage proteins, totaling 3,806 entries (Supplemental File 4). False discovery rates were fixed 1 % at the peptide and protein level. Runs were normalized by dividing by all channels (e.g. 126, 128N) by the pooled control (128C),  $\log_2$  transformed and further normalized by subtracting by the median of each sample group (Table S1). Sample infected 60<sub>1</sub> was excluded from statistical tests due to phage proteins being designated as outliers. Student's t-test was performed on the  $\log_2$  transformed data. Due to effects of ratio compression and to appropriately control for false-discovery in a (low-power) medium-scale experiment (85), significance was deemed with protein fold-change cut-of  $\geq \pm 1.5$ , p-value  $< 0.025$ .

### *RNA isolation and RNA-seq*

*E. coli* C122 was grown and infected with wild-type  $\phi$ X174 at MOI = 5, 5 mL samples removed from infected and mock replicates at initial infection and at lysis, pelleted (3,500 RCF/5-min/4 °C), and resuspended in 200  $\mu$ L 1X PBS. 400  $\mu$ L of RNAprotect bacterial reagent (Qiagen: 76506) added, vortexed, and incubated 10-min/room temperature. Following protection, RNeasy Mini Kit (Qiagen: 74106) used with optional DNase step. RNA samples were then eluted, RNA quantified (Qubit RNA HS Assay Kit, ThermoFisher Scientific: Q32852), and stored at -80 °C. Sequencing library preparation performed by MacroGen Inc. (South Korea). The rRNA from samples were depleted with a Ribo-Zero Kit (Illumina) and the RNA library generated with a TruSeq Stranded mRNA kit for microbes. Libraries sequenced with Illumina HiSeq 2500 (2x100bp mode). PhiX Control Library not used during the sequencing run to avoid cross-contamination.

Read quality checked with FastQC followed by Rsubread analysis using standard protocol (86). BAM files were processed with Rsubread featureCount function against a manually generated General Feature Format (GTF) file derived from the NZ\_LT906474.1 RefSeq GFF3 file. RNA differential expression was analyzed using DESeq2 (87) and visualized using ggplot2 R package (88). To analyze trends in differential expression, gene names were converted to UniProt identifiers using the Retrieve/ID mapping tool (89) and searched against the PANTHER database Gene list analysis tool (90) using the statistical over-representation test against the internal reference *Escherichia coli* gene list to generate overrepresented GO Terms for both cellular compartments and biological processes. The converted UniProt IDs were used

to extract individual GO Terms as before for all genes in the differentially expressed bins using the EcoCyc database (91).

### *Databases and gene analysis*

*E. coli* C122 gene annotations for proteomics experiment and COG assignments were produced through the EggNOG mapper using default parameters (23). For  $\sigma^{32}$  and  $\sigma^{38}$  transcription factor assignments, we extracted orthologous gene members from the gold-standard *E. coli* K12 gene annotations from the EcoCyc database (91), and assigned matching annotations to C122 genes (Supplemental File S3). The PANTHER (24) database's statistical over-representation tests and functional classifications was used to map protein and transcript enrichments.

Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (92) partner repository with the dataset identifier PXD021681.

## **SUPPLEMENTAL MATERIAL**

Supplementary File S1: TMT proteomics data

Supplementary file S2: Transcriptional dataset of all C122 genes

Supplementary File S3: Sigma32 and Sigma38 analysis

Supplementary File S4: MS protein database

## ACKNOWLEDGEMENTS

We recognize that this research was conducted on the traditional lands of the Wattamattagal clan of the Darug nation. BWW was supported by a Macquarie Research Excellence PhD Scholarship. We thank Dr. Karthik Kamath, Dr. David Cantor, Sasha Tetu, and Varsha Naidu for assistance. Aspects of this research were conducted at the Australian Proteome Analysis Facility. PRJ was supported by the Molecular Sciences Department, Faculty of Science & Engineering, and the Deputy Vice-Chancellor (Research) of Macquarie University. DYL is a recipient of the Macquarie University Research Excellence PhD scholarship (MQRES) and CSIRO PhD Scholarship Program in Synthetic Biology (Synthetic Biology Future Science Platform). BWW was supported by a Macquarie Research Excellence PhD Scholarship.

The contributions of author of this work according to the CRediT contribution taxonomy was: BWW: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – Original draft, Writing – review & editing; DYL: Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing; MM: Investigation, Methodology; DP: Formal analysis, Methodology; MPM: Funding acquisition; Project administration; Resources; Supervision; Writing - review & editing. PRJ: Conceptualization; Formal analysis; Funding acquisition; Methodology; Project administration; Resources; Supervision; Writing - review & editing.

## REFERENCES

1. Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR. 2012. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493:429.
2. De Smet J, Hendrix H, Blasdel BG, Danis-Wlodarczyk K, Lavigne R. 2017. *Pseudomonas* predators: understanding and exploiting phage-host interactions. *Nat Rev Microbiol* 15:517-530.

3. Zhao X, Chen C, Shen W, Huang G, Le S, Lu S, Li M, Zhao Y, Wang J, Rao X, Li G, Shen M, Guo K, Yang Y, Tan Y, Hu F. 2016. Global Transcriptomic Analysis of Interactions between *Pseudomonas aeruginosa* and Bacteriophage PaP3. *Sci Rep* 6:19237.
4. Chevallereau A, Blasdel BG, De Smet J, Monot M, Zimmermann M, Kogadeeva M, Sauer U, Jorth P, Whiteley M, Debarbieux L, Lavigne R. 2016. Next-Generation “-omics” Approaches Reveal a Massive Alteration of Host RNA Metabolism during Bacteriophage Infection of *Pseudomonas aeruginosa*. *PLOS Genetics* 12:e1006134.
5. Wei D, Zhang X. 2010. Proteomic Analysis of Interactions between a Deep-Sea Thermophilic Bacteriophage and Its Host at High Temperature. *Journal of Virology* 84:2365-2373.
6. Young JC, Dill BD, Pan C, Hettich RL, Banfield JF, Shah M, Fremaux C, Horvath P, Barrangou R, VerBerkmoes NC. 2012. Phage-Induced Expression of CRISPR-Associated Proteins Is Revealed by Shotgun Proteomics in *Streptococcus thermophilus*. *PLOS ONE* 7:e38077.
7. Van den Bossche A, Ceysens PJ, De Smet J, Hendrix H, Bellon H, Leimer N, Wagemans J, Delattre AS, Cenens W, Aertsen A, Landuyt B, Minakhin L, Severinov K, Noben JP, Lavigne R. 2014. Systematic identification of hypothetical bacteriophage proteins targeting key protein complexes of *Pseudomonas aeruginosa*. *J Proteome Res* 13:4446-56.
8. Ceysens PJ, Minakhin L, Van den Bossche A, Yakunina M, Klimuk E, Blasdel B, De Smet J, Noben JP, Blasi U, Severinov K, Lavigne R. 2014. Development of giant bacteriophage PhiKZ is independent of the host transcription apparatus. *J Virol* 88:10501-10.
9. Wu D, Yuan Y, Liu P, Wu Y, Gao M. 2014. Cellular responses in *Bacillus thuringiensis* CS33 during bacteriophage BtCS33 infection. *J Proteomics* 101:192-204.
10. Weintraub ST, Mohd Redzuan NH, Barton MK, Md Amin NA, Desmond MI, Adams LE, Ali B, Pardo S, Molleur D, Wu W, Newcomb WW, Osier MV, Black LW, Steven AC, Thomas JA. 2019. Global Proteomic Profiling of *Salmonella* Infection by a Giant Phage. *Journal of Virology* 93:e01833-18.
11. Lemay ML, Otto A, Maass S, Plate K, Becher D, Moineau S. 2019. Investigating *Lactococcus lactis* MG1363 Response to Phage p2 Infection at the Proteome Level. *Mol Cell Proteomics* 18:704-714.
12. Logel DY, Jaschke PR. 2020. A high-resolution map of bacteriophage  $\phi$ X174 transcription. *Virology* 547:47-56.
13. Wright BW, Ruan J, Molloy MP, Jaschke PR. 2020. Genome Modularization Reveals Overlapped Gene Topology Is Necessary for Efficient Viral Reproduction. *ACS Synthetic Biology* doi:10.1021/acssynbio.0c00323.
14. Jaschke PR, Dotson GA, Hung KS, Liu D, Endy D. 2019. Definitive demonstration by synthesis of genome annotation completeness. *Proceedings of the National Academy of Sciences* 116:24206-24213.
15. Jaschke PR, Lieberman EK, Rodriguez J, Sierra A, Endy D. 2012. A fully decompressed synthetic bacteriophage  $\phi$ X174 genome assembled and archived in yeast. *Virology* 434:278-84.
16. Wichman HA, Millstein J, Bull JJ. 2005. Adaptive molecular evolution for 13,000 phage generations: a possible arms race. *Genetics* 170:19-31.
17. Hafenstein S, Fane BA. 2002.  $\phi$ X174 Genome-Capsid Interactions Influence the Biophysical Properties of the Virion: Evidence for a Scaffolding-Like Function for the Genome during the Final Stages of Morphogenesis. *Journal of Virology* 76:5350-5356.
18. Bernal RA, Hafenstein S, Esmeralda R, Fane BA, Rossmann MG. 2004. The  $\phi$ X174 Protein J Mediates DNA Packaging and Viral Attachment to Host Cells. *Journal of Molecular Biology* 337:1109-1122.

19. Sun L, Young LN, Zhang X, Boudko SP, Fokine A, Zbornik E, Roznowski AP, Molineux IJ, Rossmann MG, Fane BA. 2014. Icosahedral bacteriophage PhiX174 forms a tail for DNA transport during infection. *Nature* 505:432-5.
20. Sun Y, Roznowski AP, Tokuda JM, Klose T, Mauney A, Pollack L, Fane BA, Rossmann MG. 2017. Structural changes of tailless bacteriophage PhiX174 during penetration of bacterial cell walls. *Proc Natl Acad Sci U S A* 114:13708-13713.
21. Shkoporov AN, Hill C. 2019. Bacteriophages of the Human Gut: The “Known Unknown” of the Microbiome. *Cell Host & Microbe* 25:195-209.
22. Michel A, Clermont O, Denamur E, Tenaillon O. 2010. Bacteriophage PhiX174's Ecological Niche and the Flexibility of Its *Escherichia coli* Lipopolysaccharide Receptor. *Applied and Environmental Microbiology* 76:7310.
23. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular Biology and Evolution* 34:2115-2122.
24. Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, Thomas PD. 2019. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols* 14:703-721.
25. Lex A, Gehlenborg N, Strobel H, Vuillemot R, Pfister H. 2014. UpSet: Visualization of Intersecting Sets. *IEEE transactions on visualization and computer graphics* 20:1983-1992.
26. Sachelar I, Winter L, Knyazev DG, Zimmermann M, Vogt A, Kuttner R, Ollinger N, Siligan C, Pohl P, Koch H-G. 2017. YidC and SecYEG form a heterotetrameric protein translocation channel. *Scientific reports* 7:101-101.
27. Komar J, Alvira S, Schulze RJ, Martin R, Lycklama ANJA, Lee SC, Dafforn TR, Deckers-Hebestreit G, Berger I, Schaffitzel C, Collinson I. 2016. Membrane protein insertion and assembly by the bacterial holo-translocon SecYEG-SecDF-YajC-YidC. *Biochem J* 473:3341-54.
28. Beckwith J. 2013. The Sec-dependent pathway. *Research in microbiology* 164:497-504.
29. Zuckert WR. 2014. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta* 1843:1509-16.
30. Kaplan E, Greene NP, Crow A, Koronakis V. 2018. Insights into bacterial lipoprotein trafficking from a structure of LolA bound to the LolC periplasmic domain. *Proc Natl Acad Sci U S A* 115:E7389-e7397.
31. Tanaka K, Matsuyama SI, Tokuda H. 2001. Deletion of *lolB*, encoding an outer membrane lipoprotein, is lethal for *Escherichia coli* and causes accumulation of lipoprotein localization intermediates in the periplasm. *J Bacteriol* 183:6538-42.
32. Mathelié-Guinlet M, Asmar AT, Collet J-F, Dufrêne YF. 2020. Lipoprotein Lpp regulates the mechanical properties of the *E. coli* cell envelope. *Nature Communications* 11:1789.
33. Grabowicz M, Silhavy TJ. 2017. Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proceedings of the National Academy of Sciences* 114:4769.
34. Cascales E, Bernadac A, Gavioli M, Lazzaroni J-C, Lloubes R. 2002. Pal Lipoprotein of *Escherichia coli* Plays a Major Role in Outer Membrane Integrity. *Journal of Bacteriology* 184:754.
35. Samsudin F, Ortiz-Suarez Maite L, Piggot Thomas J, Bond Peter J, Khalid S. 2016. OmpA: A Flexible Clamp for Bacterial Cell Wall Attachment. *Structure* 24:2227-2235.
36. Bernhardt TG, Roof WD, Young R. 2000. Genetic evidence that the bacteriophage  $\phi$ X174 lysis protein inhibits cell wall synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 97:4297-4302.
37. Bernhardt TG, Struck DK, Young R. 2001. The lysis protein E of phi X174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. *J Biol Chem* 276:6093-7.

38. Rodolis MT, Mihalyi A, O'Reilly A, Slikas J, Roper DI, Hancock REW, Bugg TDH. 2014. Identification of a Novel Inhibition Site in Translocase MraY Based upon the Site of Interaction with Lysis Protein E from Bacteriophage  $\phi$ X174. *ChemBioChem* 15:1300-1308.
39. Bouhss A, Trunkfield AE, Bugg TDH, Mengin-Lecreux D. 2008. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiology Reviews* 32:208-233.
40. Doublet P, van Heijenoort J, Mengin-Lecreux D. 1992. Identification of the *Escherichia coli* *murl* gene, which is required for the biosynthesis of D-glutamic acid, a specific component of bacterial peptidoglycan. *Journal of Bacteriology* 174:5772.
41. Gibbons HS, Kalb SR, Cotter RJ, Raetz CRH. 2005. Role of Mg<sup>2+</sup> and pH in the modification of *Salmonella* lipid A after endocytosis by macrophage tumour cells. *Molecular Microbiology* 55:425-440.
42. Huvet M, Toni T, Sheng X, Thorne T, Jovanovic G, Engl C, Buck M, Pinney JW, Stumpf MPH. 2011. The Evolution of the Phage Shock Protein Response System: Interplay between Protein Function, Genomic Organization, and System Function. *Molecular Biology and Evolution* 28:1141-1155.
43. García LR, Molineux IJ. 1995. Incomplete entry of bacteriophage T7 DNA into F plasmid-containing *Escherichia coli*. *Journal of Bacteriology* 177:4077.
44. Cheng X, Wang W, Molineux IJ. 2004. F exclusion of bacteriophage T7 occurs at the cell membrane. *Virology* 326:340-52.
45. Dekoninck K, Létoquart J, Laguri C, Demange P, Bevernaegie R, Simorre J-P, Dehu O, Iorga BI, Elias B, Cho S-H, Collet J-F. 2020. Defining the function of OmpA in the Rcs stress response. *eLife* 9:e60861.
46. Wall E, Majdalani N, Gottesman S. 2018. The Complex Rcs Regulatory Cascade. *Annual Review of Microbiology* 72:111-139.
47. Gaubig LC, Waldminghaus T, Narberhaus F. 2011. Multiple layers of control govern expression of the *Escherichia coli* *ibpAB* heat-shock operon. *Microbiology* 157:66-76.
48. Morán Luengo T, Kityk R, Mayer MP, Rüdiger SGD. 2018. Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System. *Molecular Cell* 70:545-552.e9.
49. Moayed F, Bezrukavnikov S, Naqvi MM, Groitl B, Cremers CM, Kramer G, Ghosh K, Jakob U, Tans SJ. 2020. The Anti-Aggregation Holdase Hsp33 Promotes the Formation of Folded Protein Structures. *Biophys J* 118:85-95.
50. Gamer J, Multhaupt G, Tomoyasu T, McCarty JS, Rüdiger S, Schönfeld HJ, Schirra C, Bujard H, Bukau B. 1996. A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor sigma32. *The EMBO journal* 15:607-617.
51. Rohrwild M, Coux O, Huang HC, Moerschell RP, Yoo SJ, Seol JH, Chung CH, Goldberg AL. 1996. HslV-HslU: A novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *Proceedings of the National Academy of Sciences* 93:5808-5813.
52. Lavigne R, Lecoutere E, Wagemans J, Cenens W, Aertsen A, Schoofs L, Landuyt B, Paeshuyse J, Scheer M, Schobert M, Ceysens PJ. 2013. A multifaceted study of *Pseudomonas aeruginosa* shutdown by virulent podovirus LUZ19. *MBio* 4:e00061-13.
53. Mayo RF, Sinsheimer RL. 1970. Process of infection with bacteriophage phiX174. XXXVI. Measurement of virus-specific proteins during a normal cycle of infection. *J Virol* 6:310-9.
54. Mojardín L, Salas M. 2016. Global Transcriptional Analysis of Virus-Host Interactions between Phage  $\phi$ 29 and *Bacillus subtilis*. *Journal of Virology* 90:9293.
55. Fallico V, Ross RP, Fitzgerald GF, McAuliffe O. 2011. Genetic response to bacteriophage infection in *Lactococcus lactis* reveals a four-strand approach involving induction of membrane stress proteins, D-alanylation of the cell wall, maintenance of proton motive force, and energy conservation. *Journal of virology* 85:12032-12042.

56. Li G-W, Burkhardt D, Gross C, Weissman Jonathan S. 2014. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* 157:624-635.
57. Bernhardt TG, Roof WD, Young R. 2002. The *Escherichia coli* FKBP-type PPIase SlyD is required for the stabilization of the E lysis protein of bacteriophage  $\phi$ X174. *Molecular Microbiology* 45:99-108.
58. Porter NT, Hryckowian AJ, Merrill BD, Fuentes JJ, Gardner JO, Glowacki RWP, Singh S, Crawford RD, Snitkin ES, Sonnenburg JL, Martens EC. 2020. Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*. *Nature Microbiology* 5:1170-1181.
59. Poranen MM, Ravantti JJ, Grahn AM, Gupta R, Auvinen P, Bamford DH. 2006. Global changes in cellular gene expression during bacteriophage PRD1 infection. *J Virol* 80:8081-8.
60. Kamenšek S, Žgur-Bertok D. 2013. Global transcriptional responses to the bacteriocin colicin M in *Escherichia coli*. *BMC Microbiology* 13:42.
61. Carra S, Alberti S, Arrigo PA, Benesch JL, Benjamin IJ, Boelens W, Bartelt-Kirbach B, Brundel B, Buchner J, Bukau B, Carver JA, Ecroyd H, Emanuelsson C, Finet S, Golenhofen N, Goloubinoff P, Gusev N, Haslbeck M, Hightower LE, Kampinga HH, Klevit RE, Liberek K, McHaourab HS, McMenimen KA, Poletti A, Quinlan R, Strelkov SV, Toth ME, Vierling E, Tanguay RM. 2017. The growing world of small heat shock proteins: from structure to functions. *Cell Stress Chaperones* 22:601-611.
62. Baneyx F, Mujacic M. 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* 22:1399-408.
63. Mogk A, Bukau B, Kampinga HH. 2018. Cellular Handling of Protein Aggregates by Disaggregation Machines. *Mol Cell* 69:214-226.
64. Mogk A, Ruger-Herreros C, Bukau B. 2019. Cellular Functions and Mechanisms of Action of Small Heat Shock Proteins. *Annual Review of Microbiology* 73:89-110.
65. Ravantti JJ, Ruokoranta TM, Alapuranen AM, Bamford DH. 2008. Global transcriptional responses of *Pseudomonas aeruginosa* to phage PRR1 infection. *J Virol* 82:2324-9.
66. Gill RT, Valdes JJ, Bentley WE. 2000. A Comparative Study of Global Stress Gene Regulation in Response to Overexpression of Recombinant Proteins in *Escherichia coli*. *Metabolic Engineering* 2:178-189.
67. Lethanh H, Neubauer P, Hoffmann F. 2005. The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant *Escherichia coli* and delay degradation of inclusion bodies. *Microb Cell Fact* 4:6.
68. Maaroufi H, Tanguay RM. 2013. Analysis and Phylogeny of Small Heat Shock Proteins from Marine Viruses and Their Cyanobacteria Host. *PLOS ONE* 8:e81207.
69. Chen L-X, Méheust R, Crits-Christoph A, McMahon KD, Nelson TC, Slater GF, Warren LA, Banfield JF. 2020. Large freshwater phages with the potential to augment aerobic methane oxidation. *Nature Microbiology* doi:10.1038/s41564-020-0779-9.
70. Perrody E, Cirinesi A-M, Desplats C, Keppel F, Schwager F, Tranier S, Georgopoulos C, Genevaux P. 2012. A Bacteriophage-Encoded J-Domain Protein Interacts with the DnaK/Hsp70 Chaperone and Stabilizes the Heat-Shock Factor  $\sigma$ 32 of *Escherichia coli*. *PLOS Genetics* 8:e1003037.
71. Chamakura KR, Tran JS, Young R. 2017. MS2 Lysis of *Escherichia coli* Depends on Host Chaperone DnaJ. *Journal of Bacteriology* 199:e00058-17.
72. Zeilstra-Ryalls J, Fayet O, Baird L, Georgopoulos C. 1993. Sequence analysis and phenotypic characterization of groEL mutations that block lambda and T4 bacteriophage growth. *J Bacteriol* 175:1134-43.
73. Hanninen AL, Bamford DH, Bamford JK. 1997. Assembly of membrane-containing bacteriophage PRD1 is dependent on GroEL and GroES. *Virology* 227:207-10.

74. Ding Y, Duda RL, Hendrix RW, Rosenberg JM. 1995. Complexes between chaperonin GroEL and the capsid protein of bacteriophage HK97. *Biochemistry* 34:14918-31.
75. Hanninen A-L, Bamford DH, Bamford JKH. 1997. Assembly of Membrane-Containing Bacteriophage PRD1 Is Dependent on GroEL and GroES. *Virology* 227:207-210.
76. Semenyuk PI, Moiseenko AV, Sokolova OS, Muronetz VI, Kurochkina LP. 2020. Structural and functional diversity of novel and known bacteriophage-encoded chaperonins. *International Journal of Biological Macromolecules* 157:544-552.
77. Takano T, Kakefuda T. 1972. Involvement of a Bacterial Factor in Morphogenesis of Bacteriophage Capsid. *Nature New Biology* 239:34-37.
78. Young KD, Anderson RJ, Hafner RJ. 1989. Lysis of *Escherichia coli* by the bacteriophage phiX174 E protein: inhibition of lysis by heat shock proteins. *Journal of bacteriology* 171:4334-4341.
79. Zhao L, Stancik AD, Brown CJ. 2012. Differential Transcription of Bacteriophage  $\phi$ X174 Genes at 37°C and 42°C. *PLOS ONE* 7:e35909.
80. Hecht A, Filliben J, Munro SA, Salit M. 2018. A minimum information standard for reproducing bench-scale bacterial cell growth and productivity. *Communications Biology* 1:219.
81. Rokyta DR, Abdo Z, Wichman HA. 2009. The genetics of adaptation for eight microvirid bacteriophages. *Journal of Molecular Evolution* 69:229-239.
82. Fane BA, Hayashi M. 1991. Second-site suppressors of a cold-sensitive prohead accessory protein of bacteriophage phi X174. *Genetics* 128:663.
83. Newbold JE, Sinsheimer RL. 1970. The process of infection with bacteriophage  $\phi$ X174: XXXII. Early steps in the infection process: Attachment, eclipse and DNA penetration. *Journal of Molecular Biology* 49:49-66.
84. Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2:1896-906.
85. Pascovici D, Handler DC, Wu JX, Haynes PA. 2016. Multiple testing corrections in quantitative proteomics: A useful but blunt tool. *Proteomics* 16:2448-53.
86. Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* 47:e47-e47.
87. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15:550.
88. Wickham H. 2016. ggplot2: elegant graphics for data analysis doi:10.1007/978-3-319-24277-4. Springer International Publishing.
89. UniProt Consortium T. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* 46:2699-2699.
90. Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A, Lazareva-Ulitsky B. 2006. Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Research* 34:W645-W650.
91. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martinez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti P, Velazquez-Ramirez DA, Weaver D, Collado-Vides J, Paulsen I, Karp PD. 2017. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* 45:D543-d550.
92. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz Ş, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaíno JA. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Research* 47:D442-D450.