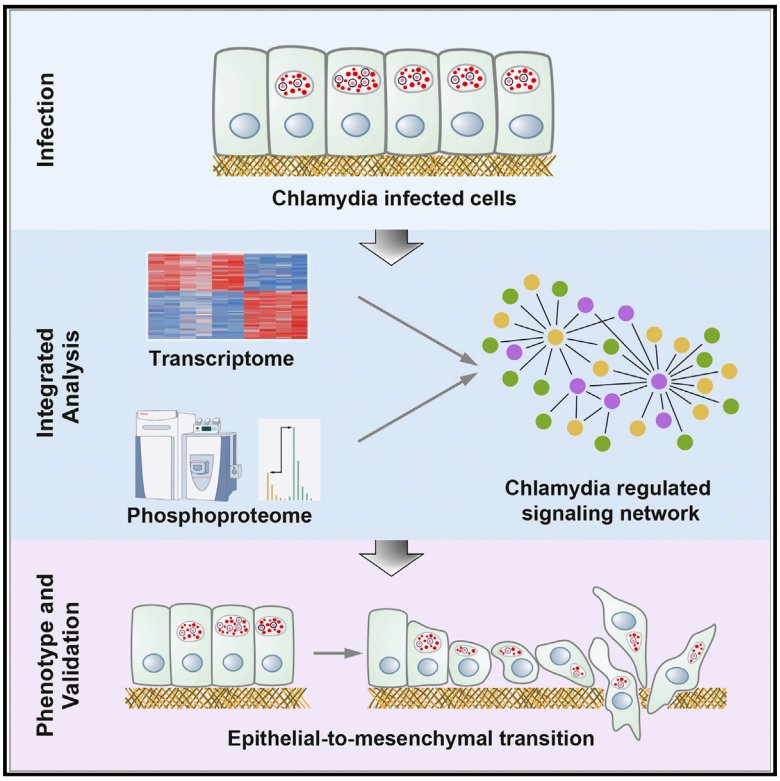
Article

Cell Reports

# Integrated Phosphoproteome and Transcriptome Analysis Reveals *Chlamydia*-Induced Epithelial-to-Mesenchymal Transition in Host Cells

## Graphical Abstract




## Authors

Piotr K. Zadora, Cindrilla Chumduri, Koshi Imami, ..., Matthias Selbach, Thomas F. Meyer, Rajendra Kumar Gurumurthy

## Correspondence

meyer@mpiib-berlin.mpg.de (T.F.M.), rajendra@mpiib-berlin.mpg.de (R.K.G.)

## In Brief

Zadora et al. performed an integrated global phosphoproteomic and transcriptomic analysis, revealing a comprehensive map of *Chlamydia*-induced host cell signaling and identifying transcription factors ETS1 and ERF, which drive epithelial-to-mesenchymal transition. These insights provide mechanistic clues to *Chlamydia* pathogenesis and serve as an important resource for future studies.

## Highlights

- Phosphoproteome and transcriptome analyses reveal *Ctr*-induced host cell signaling
- *Ctr* causes phosphorylation of MAPK/CDK and dephosphorylation of CAMK/PKA/PKC substrates
- *Ctr* induces epithelial-to-mesenchymal transition via ERK-mediated ERF and ETS1 signaling
- *Ctr*-induced EMT conveys host cell invasiveness and disruption of cervical epithelium

Zadora et al., 2019, Cell Reports *26*, 1286–1302
January 29, 2019 
https://doi.org/10.1016/j.celrep.2019.01.006

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

# Integrated Phosphoproteome and Transcriptome Analysis Reveals *Chlamydia*-Induced Epithelial-to-Mesenchymal Transition in Host Cells

Piotr K. Zadora,[1,4] Cindrilla Chumduri,[1,2,4] Koshi Imami,[3] Hilmar Berger,[1] Yang Mi,[1] Matthias Selbach,[3] Thomas F. Meyer,[1,5,*] and Rajendra Kumar Gurumurthy[1,*]
[1]Department of Molecular Biology, Max Planck Institute for Infection Biology, 10117 Berlin, Germany
[2]Department of Hepatology and Gastroenterology, Charité University Medicine, 13353 Berlin, Germany
[3]Proteome Dynamics, Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: meyer@mpiib-berlin.mpg.de (T.F.M.), rajendra@mpiib-berlin.mpg.de (R.K.G.)
https://doi.org/10.1016/j.celrep.2019.01.006

## SUMMARY

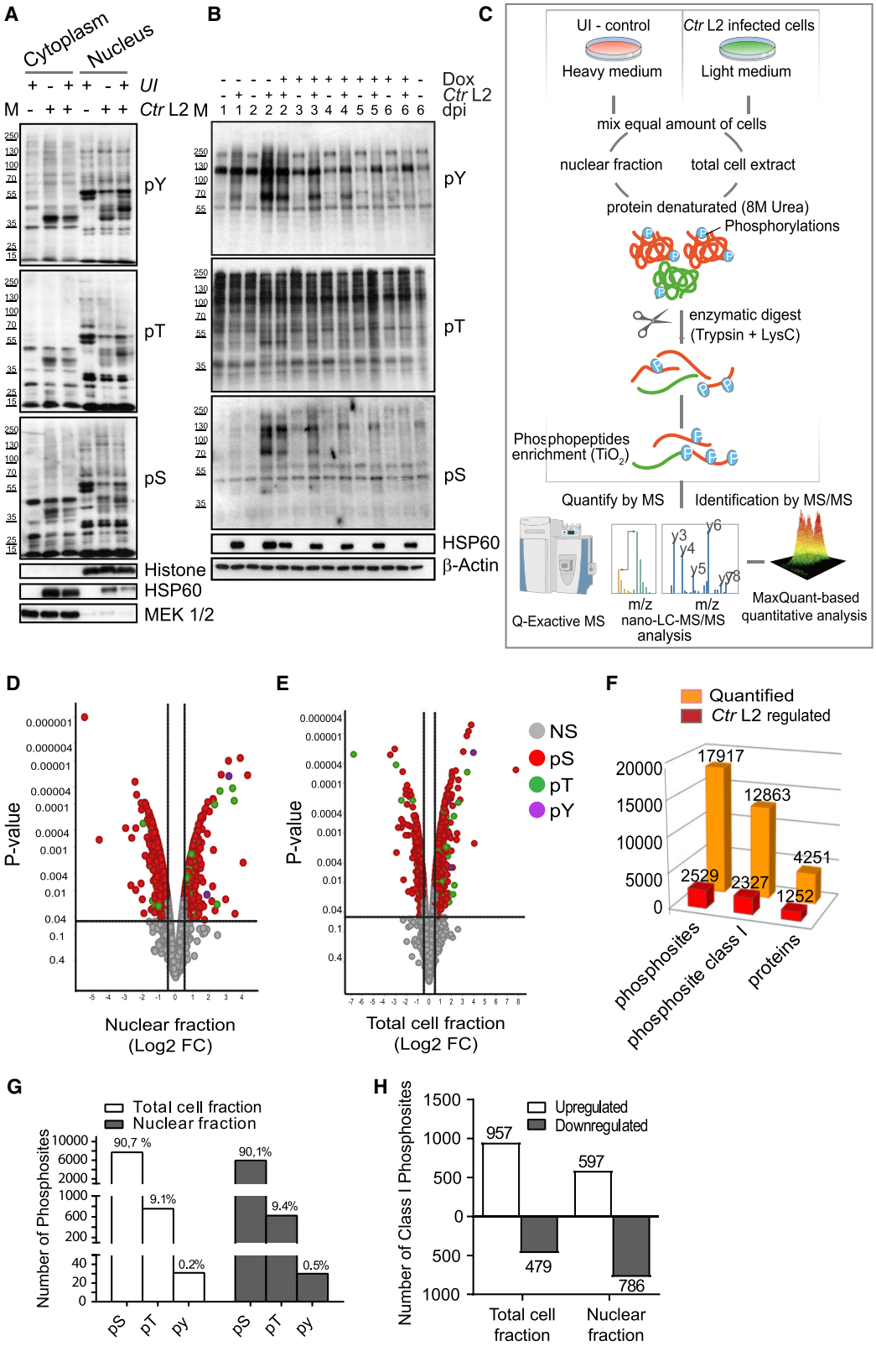
**_Chlamydia trachomatis_ (*Ctr*) causes a range of infectious diseases and is epidemiologically associated with cervical and ovarian cancers. To obtain a panoramic view of *Ctr*-induced signaling, we performed global phosphoproteomic and transcriptomic analyses. We identified numerous *Ctr* phosphoproteins and *Ctr*-regulated host phosphoproteins. Bioinformatics analysis revealed that these proteins were predominantly related to transcription regulation, cellular growth, proliferation, and cytoskeleton organization. *In silico* kinase substrate motif analysis revealed that MAPK and CDK were the most overrepresented upstream kinases for upregulated phosphosites. Several of the regulated host phosphoproteins were transcription factors, including ETS1 and ERF, that are downstream targets of MAPK. Functional analysis of phosphoproteome and transcriptome data confirmed their involvement in epithelial-to-mesenchymal transition (EMT), a phenotype that was validated in infected cells, along with the essential role of ERK1/2, ETS1, and ERF for *Ctr* replication. Our data reveal the extent of *Ctr*-induced signaling and provide insights into its pro-carcinogenic potential.**

## INTRODUCTION

The Gram-negative bacterium *Chlamydia trachomatis* (*Ctr*) infects the epithelium of the genital tract, causing, for example, cervicitis, pelvic inflammatory disease, and scarring, with impact on fertility. Infections frequently remain asymptomatic and become chronic. *Chlamydia* is a strong risk factor for the development of cervical and ovarian cancers, either independently or as a co-factor with human papillomavirus (HPV) infections (Koskela et al., 2000; Shanmughapriya et al., 2012; Zhu et al., 2016). Due to the lack of physiologically relevant infection models, illuminating the underlying mechanisms or even the natural progress of the infection in its different host tissues has remained challenging.

As an obligate intracellular bacterium, *Ctr* has evolved the means for manipulating host cell pathways by altering gene expression and protein stability at the transcriptional, translational, and post-translational levels to ensure that its replicative niche remains alive until the completion of the life cycle (Chumduri et al., 2016; Elwell et al., 2016; Olive et al., 2014). *Ctr* establishes infection by translocating effectors into host cells, thereby triggering cytoskeletal rearrangements and signaling. Upon host cell entry, its effector protein translocated actin-recruiting phosphoprotein (Tarp) is rapidly tyrosine phosphorylated to interact with SH2 domains of human proteins, including the adaptor protein SRC homology 2 domain-containing transforming protein C1 (SHC1), to activate pro-survival extracellular signal-regulated kinase (ERK) signaling (Mehlitz et al., 2010). During the mid- to late stages of infection, the activation of ERK occurs independently of RAS-rapidly accelerated fibrosarcoma (RAF) and plays an essential role in bacterial nutrient acquisition, synthesis of inflammatory cytokines, and expression of anti-apoptotic factors (Gurumurthy et al., 2010; Rajalingam et al., 2008; Su et al., 2004). Furthermore, other mitogenic mitogen-activated protein kinase (MAPK) signaling pathways, involving p38 and JNK, are activated by post-translational modifications, leading to an activation of activator protein-1 (AP1)-dependent transcription, which is essential for *Ctr* development (Buchholz and Stephens, 2007; Chen et al., 2010; Olive et al., 2014). *Ctr* also suppresses the key DNA damage response ataxia-telangiectasia mutated (ATM) protein (Chumduri et al., 2013; González et al., 2014), while degradation of p53 via the Ak strain transforming-mouse double minute 2 homolog (AKT-MDM2) signaling axis induces host metabolism alterations that resemble the Warburg effect seen in cancers (Ojcius et al., 1998; Rother et al., 2018; Siegl et al., 2014). In addition, *Ctr* alters global histone post-translational modifications, which can influence various cellular signals that are essential for the maintenance of genome integrity (Chumduri et al., 2013). Despite this, we know little about the complex multifactorial nature of *Ctr*-induced host cellular signaling. Many signaling events are modulated predominantly by protein phosphorylation. Determining which residues of particular




*(legend on next page)*

proteins are phosphorylated and to what extent would enable us to reveal which kinases are activated following *Ctr* infection.

Here, we used an integrated phosphoproteomics and transcriptomics analysis approach to comprehensively map signaling pathways modulated by *Ctr* and to reveal the complexity of the *Ctr*-induced signaling. Using stable isotope labeling with amino acids in cell culture (SILAC) and phosphopeptide enrichment coupled to tandem mass spectrometry (MS/MS), we identified 2,529 distinct phosphorylation sites that are regulated in response to *Ctr* infection. Most of these were not previously shown to be *Ctr* responsive. Bioinformatics analysis revealed that these proteins were predominantly related to transcription regulation, cellular growth, proliferation, and cytoskeleton organization. *In silico* identification of upstream kinases suggested that MAPK and cyclin-dependent kinase (CDK) were the most overrepresented upstream kinases for upregulated phosphosites, while protein kinase A, G, and C families (AGC) and calmodulin/calcium-regulated kinase (CAMK) were the most overrepresented for the downregulated phosphosites.

Notably, several of the MAPK substrates were found to be transcription factors (TFs), including fos-related antigen 1 (FRA1), ETS2 repressor factor (ERF), and proto-oncogenic transcription factor ETS1, which are implicated in epithelial-to-mesenchymal transition (EMT)-associated gene regulation. In line with this, global analysis of the *Ctr*-regulated transcriptome revealed an enrichment of EMT as one of the top five upregulated hallmark signatures, several of which we identified as targets of FRA1, ERF, and ETS1. These bioinformatics-based predictions were functionally tested to confirm the MAPK-mediated ETS1 and ERF transcriptional regulation and demonstrated their role in EMT. We independently corroborated that *Ctr*-infected cells exhibit other hallmarks of EMT, such as decreased E-cadherin, increased N-cadherin, and expression of SNAIL1. We also observed the disruption of epithelial integrity by *Ctr*, as evidenced by the remodeling of human primary ectocervical cell-derived three-dimensional (3D) raft cultures. The comprehensive picture of *Ctr*-induced host cell signaling emerging from these studies thus provides important clues to the mechanisms underlying its pathogenesis and will serve as an important resource for future studies in this direction.

## RESULTS

### *Ctr*-Responsive Global Host Phosphoproteome

To obtain a global picture of *Ctr*-induced cell signaling, we first investigated the phosphoproteome by western blot analysis of the phosphorylation status of tyrosine (pY), threonine (pT), and serine (pS) amino acid residues on proteins from the cytoplasmic and nuclear fractions of *Ctr*-infected and -uninfected cells, as well as an equal mix of lysates from both. The results indicate extensive phosphorylation changes during acute and persistent infection (Figures 1A and 1B). To quantify the relative fold changes of specific phosphorylation sites, we performed SILAC (Ong et al., 2002) and phosphopeptide enrichment (Rappsilber et al., 2007) coupled to MS/MS-based quantitative phosphoproteomics. To enable comparative analysis, uninfected control End1/E6E7 cells were stable isotope labeled with medium containing $^{13}C_6{}^{15}N_2$ L-lysine/$^{13}C_6{}^{15}N_4$ L-arginine to construct a heavy isotope control phosphoproteome, while cells destined for infection were cultured in light medium to construct a *Ctr*-infected, light phosphoproteome (Figure 1C). Following 32 h of infection with *Ctr*, the light cells were mixed with heavy, uninfected cells at an equal ratio before collecting the enriched phosphopeptides from the total and nuclear fractions to distinguish proteins from heavy and light cells by MS. Peptide scoring, protein identification, and quantification were performed using MaxQuant software (Cox and Mann, 2008). Correlation analysis using the log2 transformed fold change values from two biological replicates revealed a high Pearson's correlation coefficient score, confirming the high quality and reproducibility of the data (Figures S1A and S1B).

From the total and nuclear fractions, we identified 17,917 distinct phosphopeptides that match to 4,564 proteins (Figures

**Figure 1. *Ctr* Infection Leads to Global Alterations in Host Cell Phosphoproteome**

(A) Uninfected or *Ctr*-infected cells for 32 h were collected separately or mixed at a 1:1 ratio, followed by subcellular fractionation and immunoblot analysis. MEK1/2 and histone 4 were used as loading controls for cytoplasmic and nuclear fractions, respectively, and heat shock protein 60 (HSP60) was used as an infection marker. Lysates from each fraction were analyzed for phosphorylation status using antibodies against phosphoserine (pS), phosphothreonine (pT), and phosphotyrosine (pY) antibodies.

(B) Cells were either uninfected or persistently infected with *Ctr*, and protein lysates collected every day until 6 days post-infection (p.i.) were subjected to immunoblot analysis using pS, pT, and pY antibodies. HSP60 and β-actin antibodies were used for the detection of infection particles and loading control.

(A and B) M represents the molecular weight marker represented in kDa. Data are representative of two biological replicates.

(C) Schematic representation of global phosphoproteome analysis upon *Ctr* infection using stable isotope labeling by amino acids in cell culture (SILAC) coupled to liquid chromatography–tandem mass spectrometry (LC-MS/MS) to quantitate differentially regulated phosphoproteomes. 32 h p.i., non-infected (heavy labeled: $^{13}C_6{}^{15}N_2$ L-lysine/$^{13}C_6{}^{15}N_4$ L-arginine) and *Ctr*-infected (light labeled: L-lysine/L-arginine) cells were equally mixed from which either total cell extract or nuclear fraction protein lysates were prepared. Proteins were enzymatically digested into peptides with Lys-C and trypsin followed by LC-MS/MS to measure global proteome or phosphopeptide enrichment using $TiO_2$ columns for global phosphoproteome analysis. The ratio of heavy to light peak area accounts for a relative amount of each peptide analyzed by MaxQuant software.

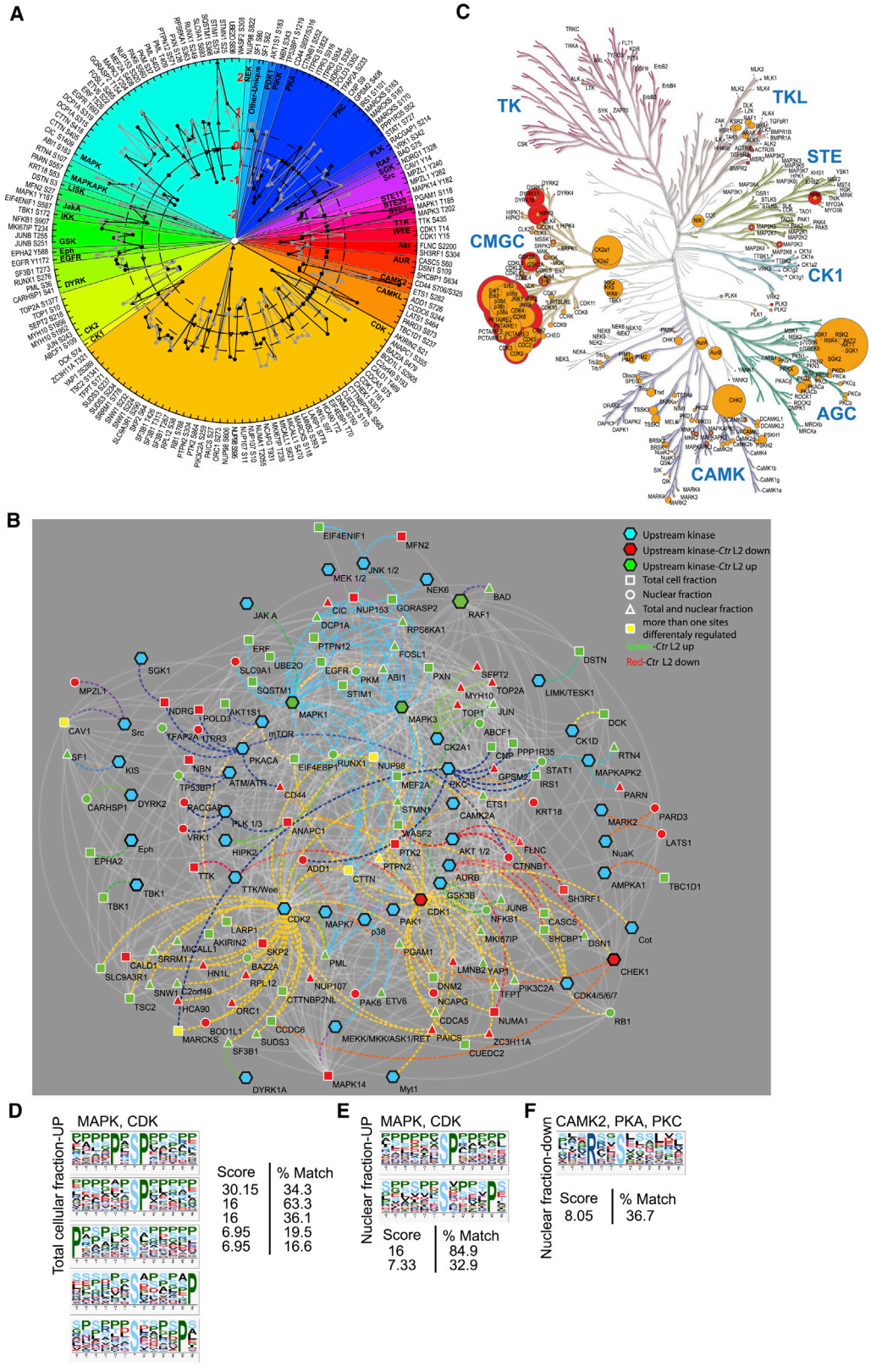
(D and E) Volcano plot representing global phosphoproteome dataset from (D) nuclear fraction and (E) total cell extract according to log2 fold change (FC) (x axis) and p value (y axis). The cutoff of ±0.5 log2 FC (dashed vertical lines) and 0.05 p value (dashed horizontal line) was applied. Phospho hits are highlighted, depending on phosphorylated residues colored as labeled.

(F) A total of 17,917 unique phosphopeptides with 2,529 regulated (±0.5 log2 FC) upon *Ctr* infection were identified from total cell extract and nuclear fraction. Using MaxQuant software, 12,863 class I phosphorylation sites were defined by a localization probability of 0.75, which were considered in further filtering, with 2,327 regulated (±0.5 log2 FC) upon *Ctr* infection.

(G) Phosphosite distribution on S/T/Y phosphorylation residues is represented in graphs for total cell extract and nuclear fraction.

(H) Differentially regulated class I phosphorylation sites with ±0.5 log2 FC in *Ctr*-infected cells compared to uninfected cells.

See also Figure S1.




*(legend on next page)*

1D–1F; Table S1). Based on post-translational modification (PTM) scores obtained from MaxQuant software, we defined 12,863 high confidence class I phosphosites corresponding to 4,251 proteins (with a localization probability of 0.75) (Olsen et al., 2006). Among these, 2,327 class I phosphosites corresponding to 1,252 proteins were significantly regulated upon *Ctr* infection (|log2 FC| > 0.5 and posterior error probability [PEP] < 0.05; Figure 1F; Table S1). The relative frequency of phosphoserine (pS), phosphothreonine (pT), and phosphotyrosine (pY) in the total and nuclear fractions (Figure 1G) is consistent with other studies, with a 90:10:0.05 ratio across S/T/Y sites (Hunter and Sefton, 1980). In the total cell extracts, 1,436 class I phosphorylation sites were regulated at least |log2FC| > 0.5 (PEP < 0.05; FC, fold change) in response to infection, 957 of which were increased, while 479 were decreased. In the nuclear fraction, this corresponded to 1,383 responsive sites, 597 of which were increased and 786 were decreased (Figure 1H; Table S1).

## Characterization of *Ctr*-Induced Kinase Regulation

Pathway overrepresentation analysis of *Ctr*-regulated phosphoproteins revealed that it modulates signaling pathways involved in a wide range of molecular and cellular functions. The top five signaling pathways enriched among the upregulated phosphoproteins from both total cell extract and nuclear fractions of infected cells using Gene Ontology (GO) term analysis for biological processes (GOBP) include regulation of transcription, proliferation, and nucleic acid metabolism. Upregulated phosphoproteins in total cell extracts are involved in the regulation of small GTPase-mediated signal transduction, apoptosis, the stress-activated protein kinase signaling pathway, and JNK/MAPKKK cascades. The downregulated phosphoproteins are mainly involved in pathways associated with cytoskeleton organization, regulation of protein complex disassembly, cell cycle, chromosomal organization, and DNA repair (Figures S2A and S2B; Table S2). Ingenuity pathway analysis (IPA) of *Ctr*-regulated phosphoproteins revealed an overrepresentation of biological processes related to cancer, the reproductive system, gastrointestinal and hepatic diseases, as well as organismal injury and abnormalities (Figure S2B).

We next carried out *in silico* assignment of the upstream kinases to each of the regulated phosphosites using the experimentally annotated site-specific kinase-substrate relation obtained from the PhosphoSitePlus database (Hornbeck et al., 2015). Of 2,327 regulated (log2 > 0.5) phosphosites, only 150 sites on 119 proteins were identified to be experimentally annotated in the database. This analysis revealed 31 kinases with at least one or more substrates, including Akt, CDK, epidermal growth factor receptor (EGFR), glycogen synthase kinase (GSK), MAPK, RAF, and Src, (Figure 2A; Table S3), indicating that an extensive range of kinases is potentially regulated by *Ctr*. We also generated the *Ctr*-regulated kinase interactome for these predicted kinase-substrate relations using the protein-protein interaction network information from the STRING database (Figure 2B; Table S3). This network reveals the protein-protein interactions between predicted kinases and their substrates, which is not evident from analyzing the kinase-substrate relations alone.

However, for the majority of the *Ctr*-regulated phosphosites, the associated upstream kinase is unknown. We therefore assigned the upstream kinase for each of these sites using the group-based prediction system (GPS) with the interaction filter (iGPS) (Song et al., 2012) and motif extractor (motif-x) (Schwartz and Gygi, 2005) bioinformatic tools. The iGPS combines the consensus substrate motif analysis with protein-protein interaction databases to predict the likelihood that a particular kinase or kinase family phosphorylates a given phosphorylation site, whereas motif-x generates potential kinase substrate motifs by measuring the overrepresented patterns of amino acid sequences. Mapping these predicted kinase-substrate relations onto the human kinome tree revealed an overrepresentation in both total cell and nuclear fractions of members of the CMGC kinase group (e.g., MAPK, CDK, GSK3, dual specificity tyrosine-regulated kinase [DYRK], homeodomain-interacting protein kinase [HIPK]), while the nuclear fraction showed enrichment for CAMK, AGS, and tyrosine kinase-like (TKL) kinases (Figure 2C; Table S4). Moreover, motif-x analysis revealed overrepresentation of MAPK and CDK motifs among the upregulated phosphorylation sites in both total and nuclear fractions (Figures 2D and 2E), while CAMK2, protein kinase A (PKA), and PKC motifs were enriched among the downregulated phosphosites in the nuclear fraction (Figure 2F).

*Ctr* expresses a number of effector proteins that are either secreted into the host cytoplasm or linked to the inclusion membrane, with one of their domains exposed to the host cytosol. Moreover, it is widely recognized that the *Ctr* proteins TARP,

**Figure 2. *Chlamydia*-Responsive Kinome Signaling**

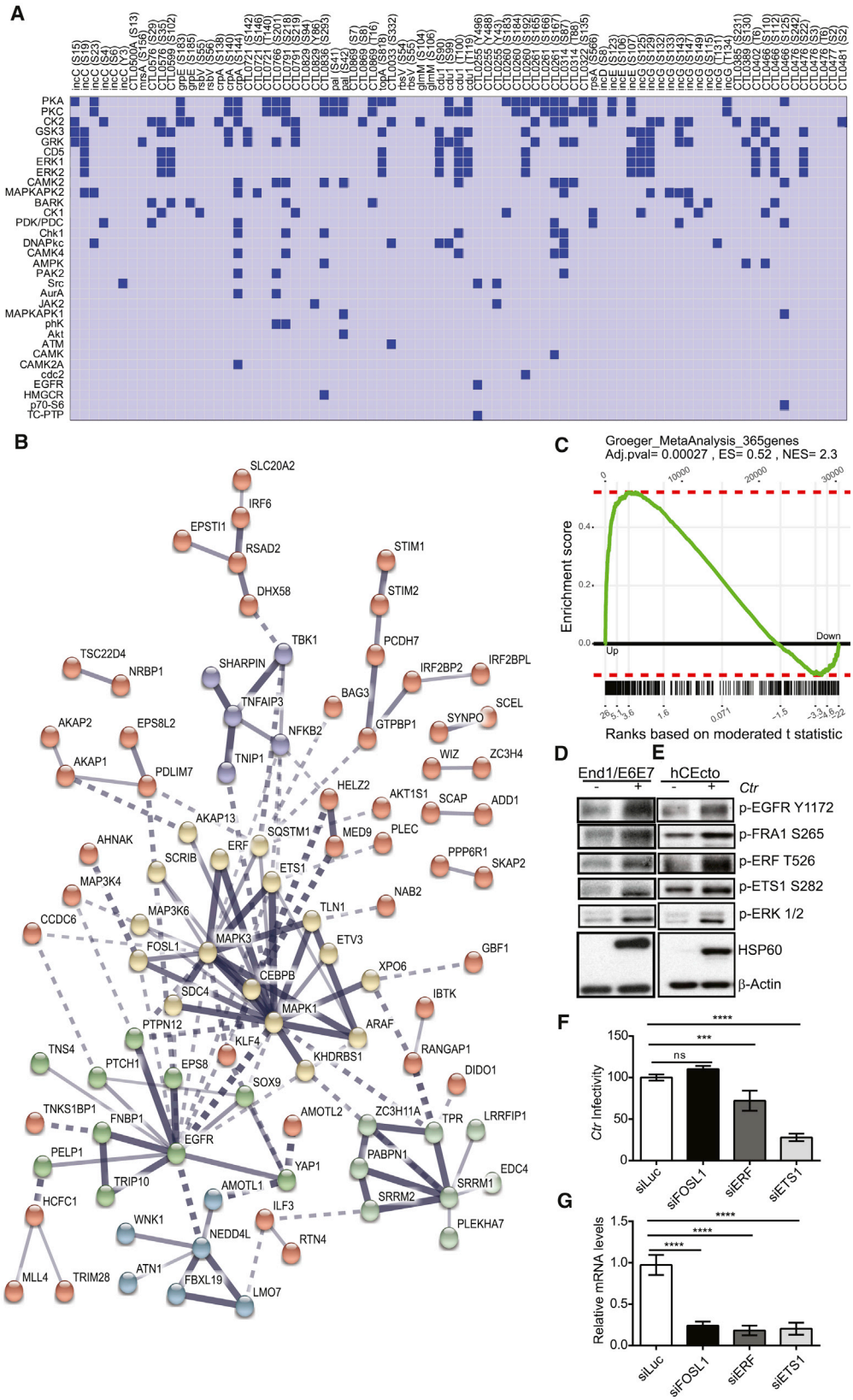
(A) Circular plot representing validated biologically relevant phosphorylation sites retrieved from the PhosphoSitePlus database that were differentially regulated upon *Ctr* infection with at least ±0.5 log2 FC. Different colors correspond to various kinase families that are predicted as upstream regulators of the selected phosphorylation site. A cutoff of ±0.5 log2 FC and localization probability ≥0.75 for total cell extract (gray line) and nuclear fraction (black line) were applied.

(B) A *Ctr*-regulated kinase interactome was generated by integrating kinase-substrate relations retrieved from PhosphoSitePlus and known human protein-protein interactions from STRING of differentially regulated phosphoproteins depicted in (A) generated using Cytoscape (v3.2.1). Predicted upstream kinases are added manually and connected with *Ctr*-regulated phosphorylation sites (dashed lines colored as in A). Up- or downregulated phosphosites are marked green or red, respectively. Proteins with more than one site are shown in yellow. Different shapes correspond to changes in total cell extract (rectangle), nuclear fraction (circle), or both (triangle).

(C) Upstream kinase predictions using iGPS analysis for unannotated phosphosites with ≥2 fold change upon *Ctr* infection from both total cell extract (red circles) and nuclear fraction (orange circles) were mapped to the kinome tree.

(D–F) Motif-x tool was used to identify the overrepresentation of linear signature motifs to predict kinases involved in regulating all of the phosphorylation sites that are upregulated in the total cellular fraction (D), nuclear fraction (E), and downregulated in the nuclear fraction (F) upon *Ctr* infection with $p < 10^{-6}$. Sequences were centered on each phosphorylation site and extended to 15 amino acids (±7 residues).

See also Figure S2.




*(legend on next page)*

translocated early phospho-protein (TepP), inclusion membrane protein G (IncG), and IncA are phosphorylated by host cell kinases (Carpenter et al., 2017; Claywell et al., 2016; Rockey et al., 1997). Therefore, we searched for phosphorylated *Ctr* proteins from the phosphoproteome data of infected cells, revealing 81 *Ctr* proteins to be phosphorylated, which consist predominantly of inclusion membrane proteins. To predict the responsible host kinases for the identified *Ctr* phosphoproteins, we retrieved kinase-substrate relations from the Human Protein Reference Database (HPRD) (Figure 3A; Table S5). This analysis suggests that PKA, PKC, casein kinase 2 (CK2), GSK3, Granta 519 resistant from kidney (GRK), cluster of differentiation 5 (CD5), and ERK1/2, among others, are host kinases that could regulate *Ctr* proteins.

### Integration of *Ctr*-Responsive Transcriptome and Phosphoproteome Identifies EMT Signature

To identify relevant pathways and functions, we mapped the >2-fold regulated phosphoproteins on known protein-protein interactions provided by the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (see Method Details). This analysis revealed five prominent subnetworks. Among these, MAPK1/3 and EGFR, as well as their interaction partners, formed the core networks and were connected directly or indirectly with other modules. The core MAPK1 and MAPK3 interaction network contained five transcription factors—ETS1, FRA1, ERF, ETS variant 3 (ETV3), and CCAAT enhancer binding protein beta (CEBPB)—none of which have been functionally linked to *Ctr* infections thus far (Figure 3B).

Corroborating these results, GO enrichment of all >2 *Ctr*-regulated phosphoproteins showed that transcriptional regulation and MAPKKK cascade were among the top biological processes mediated by *Ctr*-regulated phosphoproteins (Figure S2A). To examine their role in controlling the expression of downstream target genes, transcriptomic analysis of *Ctr*-infected cells was performed. Gene set enrichment analysis (GSEA) of genes differentially expressed during *Ctr* infections using the Molecular Signatures Database (MsigDB) hallmark gene sets revealed EMT as one of the top five upregulated signatures besides inflammation, the tumor necrosis factor-α-nuclear factor κB (TNF-α-NF-κB) axis, interleukin 6 (IL6)-JAK-Stat3, and Kirsten rat sarcoma 2 viral oncogene homolog Phenylalanine (KRAS) signaling (Tables S6 and S7). Consistent with these observations, GSEA of the *Ctr*-induced global transcriptome revealed an enrichment of many genes associated with EMT (Gröger et al., 2012) (Figures 3C and S3A). The transcription factors FRA1, ETS1, and ERF, which are found to have high confidence interactions with ERK1/2 in the STRING analysis (Figure 3B, thick lines), have been implicated in regulating many of the EMT-associated genes (Plotnik et al., 2014; Rajasekaran et al., 2013). Therefore, we decided to perform an in-depth analysis of the possible role of ERK1/2, p-FRA1(S265), p-ETS1(S282), and p-ERF (T526) in EMT modulation during *Ctr* infections.

### FRA1, ETS1, and ERF Transcription Factors and Their Targets Are Regulated during *Ctr* Infection

The regulated phosphosites of the three selected transcription factors from the global analysis were validated by immunoblot with phosphospecific antibodies, using endocervical End1/E6E7 cells (Figures 3D and S3B–S3E). Since this cell line was immortalized with E6/E7 oncogenes of HPV, we further validated these hits using healthy human primary ectocervical cells (hCEctos) derived from HPV$^-$ donors to address and distinguish *Ctr*-specific effects from those induced in the presence of E6/E7 (Figures 3E and S3F). We confirmed that these transcription factors are essential for pathogen development (Figures 3F and 3G). Knockdown of ETS1 and ERF resulted in a significant reduction in *Ctr* infectivity compared to small interfering RNAs (siRNAs) targeting luciferase (siLuci)-treated control cells (Figure 3G), indicating their importance in chlamydial development.

To identify the target genes of FRA1, ETS1, and ERF transcription factors that are specifically regulated during *Ctr* infection, we generated interaction trees of all of the known downstream target genes of FRA1, ETS1, and ERF using IPA. All of the genes differentially regulated ($\geq$1.5 FC and $p \leq 0.05$) (Table S6) upon *Ctr* infection were then overlaid on this network. Shown are the ERF, ETS1, and FRA1 target genes that are regulated by *Ctr* (Figures 4A–4C and S4A). We then annotated the diseases and functions for all of the regulated target genes of FRA1, ETS1, and ERF separately using IPA (Table S8). The results indicate their involvement in the regulation of numerous genes involved in inflammation, angiogenesis, EMT, tumor growth, cell movement, and invasiveness. We validated a subset of these genes by qRT-PCR (Figures 4E, 4F, S4B, and S4C). Genes associated with cellular

**Figure 3. Global Phosphoproteome Validation of Selected Hits and Their Role in *Ctr* Development**

(A) The identified phosphorylation sites on *Ctr* proteins (x axis) and host kinases (y axis) predicted as potential upstream regulators based on the kinase-substrate relations retrieved from the Human Protein Reference Database.

(B) A protein-protein interaction network analysis of 2-fold *Ctr*-regulated phosphoproteins using the STRING database was performed with k-means clustering set to six clusters, while the disconnected nodes were removed, resulting in five prominent interaction networks.

(C) Gene set enrichment analysis of *Ctr*-regulated genes compared with epithelial-to-mesenchymal transition (EMT)-associated gene set reveals significant enrichment of EMT signature genes.

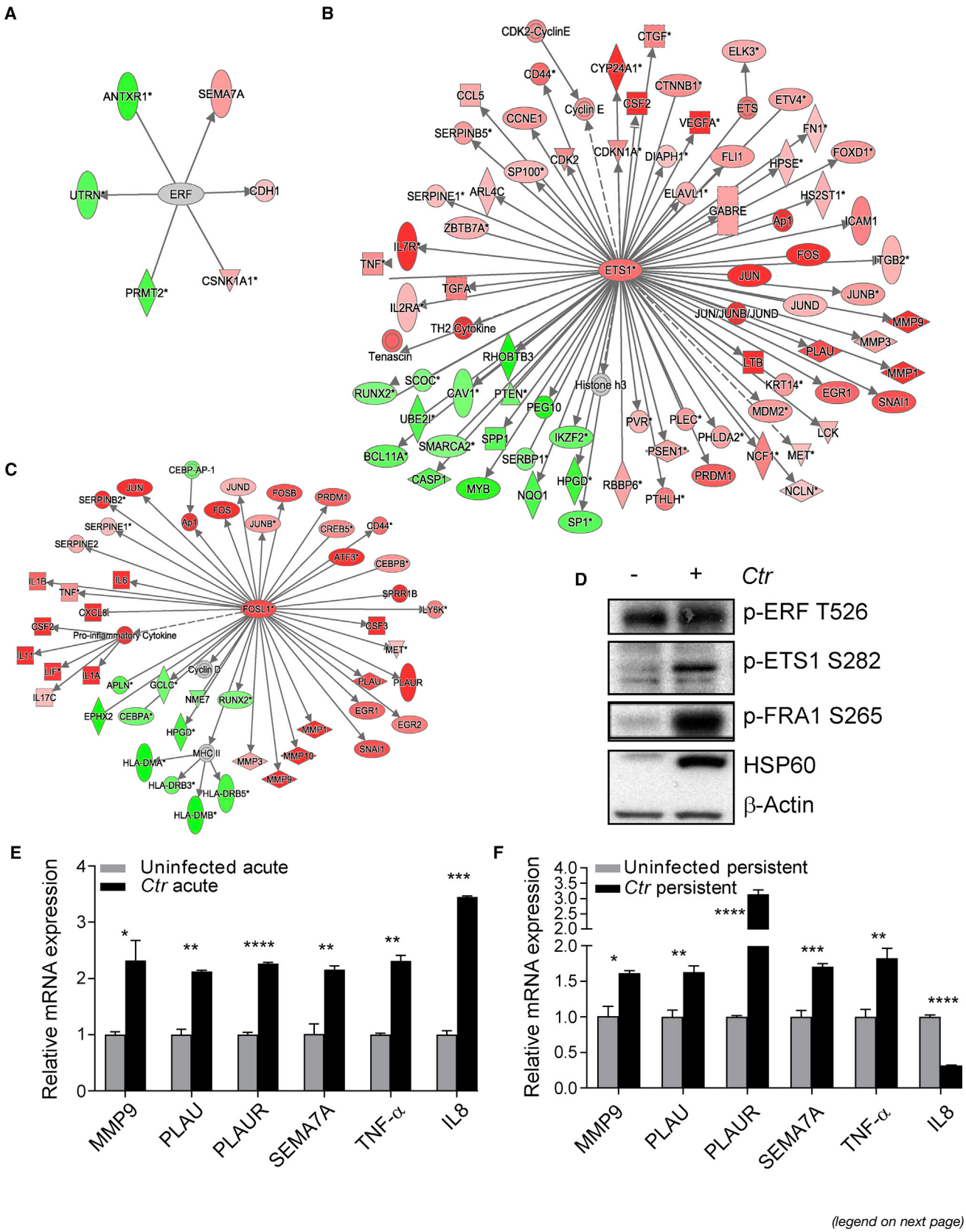
(D and E) Validation of phosphoproteome hits in (D) primary-like HPV E6E7 immortalized human endocervical epithelial cells (End1/E6E7) and (E) human primary ectocervical epithelial cells (hCEctos). Cells uninfected or *Ctr* infected for 32 h were subjected to immunoblot analysis for various proteins, as indicated using phosphospecific antibodies, chlamydial HSP60, and β-actin. Data are representative of three biological replicates.

(F and G) End1/E6E7 cells were transfected with small interfering RNAs (siRNAs) targeting luciferase (siLuci), FRA1/FOSL1, ERF, and ETS1, respectively, for 72 h (F). These cells were subsequently infected with *Ctr* for 48 h. Cell lysate was used to re-infect freshly seeded cells for 24 h to quantify infectivity. Data shown as means ± SDs of three biological replicates (***$p < 0.004$, ****$p < 0.0001$).

(G) Knockdown efficiency was analyzed by qRT-PCR.

(F and G) Data shown as means ± SDs of three biological replicates normalized to siLuci control.

See also Figure S3.




*(legend on next page)*

movement (plasminogen activator, urokinase [PLAU], PLAUR), invasiveness (SEMA7A), inflammation (IL8, TNF-α), tight junctions (E-cadherin), and matrix metalloproteinase (MMP9) were upregulated during both acute and persistent infection in hCEctos (Figures 4E and 4F) and END1/E6E7 cells (Figures S4B and S4C). Accordingly, western blot analysis showed that *Ctr*-induced phosphorylation of FRA1, ETS1, and ERF is maintained during persistent infection (Figures 4D and S4D).

### *Ctr* Induces Loss of Cell Adhesions, Tissue Disruption, and Invasive Phenotype

We thus sought to investigate whether the observed transcriptional upregulation of EMT genes in infected cells induces an EMT phenotype. EMT is a complex process whereby polarized epithelial cells acquire characteristics of an invasive mesenchymal cell phenotype. Epithelial cells undergoing EMT lose polarity and cell adhesion structures, and show enhanced migratory capacity, invasiveness, elevated resistance to apoptosis, and increased production of extracellular matrix (ECM) components (Son and Moon, 2010). Western blot analysis of END1/E6E7 and primary cells showed that persistent *Ctr* infection decreased the levels of the epithelial marker E-cadherin and increased the levels of the mesenchymal marker N-cadherin (Figures 5A and 5B). This was accompanied by reorganization of the actin cytoskeleton from thin cortical bundles to thick, parallel, contractile bundles, which are usually observed in transdifferentiated mesenchymal cells (Figure 5C). To further elucidate this EMT phenotype in a physiological situation, we established an air-liquid interface (ALI) culture model using a defined culture medium that maintains human ectocervical stem cells, which give rise to a differentiated, squamous stratified ectocervical epithelium that recapitulates the tissue architecture. The resulting multi-layered E-cadherin$^+$ epithelium consists of a p63$^+$ basal layer containing stem cells, a parabasal layer, and a terminally differentiated p63-luminal layer (Figure 5D) in which proliferating Ki67$^+$ cells are mainly restricted to the basal compartment (Figure 5E). An infection time course showed that *Ctr* can infect the terminally differentiated luminal epithelium and that the infection proceeds toward the basal stem cell compartment by disrupting epithelial integrity (Figure 5F). To elucidate the migratory capacity and invasiveness of the infected cells, we performed a Matrigel-based assay (Hall and Brooks, 2014), which estimates the capacity of cells to invade through the basement membrane that separates epithelial cells from adjacent connective tissue. Concordant with the other observed EMT traits, persistent *Ctr* infection increased the invasiveness of both End1/E6E7 and hCEcto cells (Figures 5G and 5H).

### ERK Regulates Transcription Factors FRA1, ETS1, and ERF and Initiates EMT during *Ctr* Infection

Next, we sought to confirm our bioinformatics-based kinase-substrate prediction of ERK as the upstream kinase for the regulation of the phosphorylation of p-FRA1(S265), p-ETS1(S282), and p-ERF(T526) in *Ctr* infections (Figures 3B and 3C). Using an ERK1/2 MAPK specific inhibitor, we could completely abrogate the phosphorylation of p-FRA1(S265), p-ETS1(S282), and p-ERF(T526) during *Ctr* infection, thus validating our kinase substrate predictions (Figure 6A). The results were similar in hCEctos (Figure 6B), further confirming that ERK-mediated regulation of transcription factors is *Ctr* specific and does not depend on HPV status.

U0126-mediated ERK inhibition also inhibited the transcriptional activity of FRA1, ETS1, and ERF after *Ctr* infection, as demonstrated by the reduced expression of their downstream target genes MMP9, MMP3, PLAUR, SEMA7A, and IL8 (Figure 6C), and prevented the induction of invasiveness (Figure 6D). These data demonstrate that *Ctr*-induced ERK signaling is crucial for transcriptional and post-translational regulation of a cohort of transcription factors that control EMT.

### *Chlamydia*-Induced Phosphorylation and Nuclear Export of ERF Promote Cellular Invasion

ERF is ubiquitously expressed, exhibits strong transcriptional repressor activity, and is only known to be regulated via ERK-dependent phosphorylation at multiple sites that relieve its transcriptional repressor activity by promoting nuclear export and cytoplasmic accumulation, leading to pro-migratory function (Le Gallic et al., 2004). To investigate whether *Ctr*-induced phosphorylation of the ERF repressor domain at T526 is sufficient to promote nuclear export, we performed subcellular fractionation to obtain nuclear and cytoplasmic proteins from *Ctr*-infected and control cells with or without U0126 treatment. In addition, subcellular fractions of cells treated with EGF served as a positive control. Immunoblot analysis using pERF T526 antibody showed a predominant localization to the cytoplasmic fraction in both *Ctr*-infected and EGF-treated cells, which was abrogated by U0126 (Figure 7A). To further refine the mechanism responsible for the induction of invasiveness by *Ctr*, we generated cells overexpressing ERF mutants using the following plasmids: (1) ERF-m1–7, carrying S/T-to A mutations in seven potential ERK phosphorylation sites that exhibit constitutive nuclear localization; (2) ERF-FSF/FKF (Phenylalanine-Serine-Phenylalanine/Phenylalanine-Lysine-Phenylalanine), carrying a mutation that inhibits the ERF-ERK interaction and thus interrupts signaling to ERF; and (3) ERF T526, carrying a T526-to glutamic acid mutation in

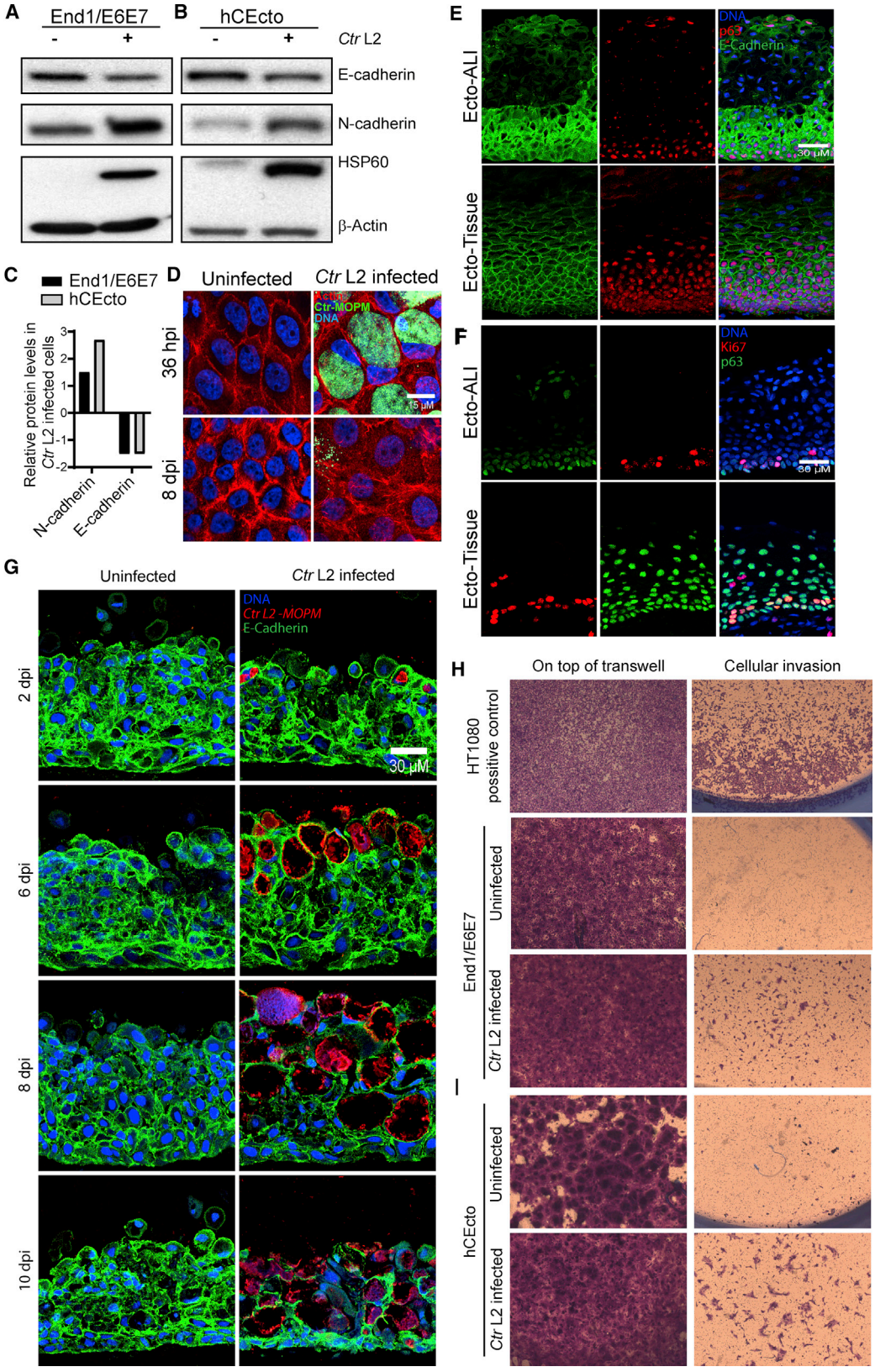
**Figure 4. *Ctr* Regulates Genes, Including Targets of ERF, ETS1, and FRA1, Associated with EMT**

(A–C) Genes transcriptionally regulated by *Ctr* were overlaid onto a network generated using IPA for all known target genes of FRA1, ETS1, and ERF separately. The resulting networks were manually curated to represent only those genes regulated during *Ctr* infection by (A) ERF, (B) ETS1, and (C) FRA1 transcription factors, respectively. Significantly upregulated and downregulated genes are depicted in red and green, respectively.

(D) END1/E6E7 cells uninfected or persistently infected with *Ctr* and cell lysates were subjected to immunoblot analysis for indicated phosphorylations on ERF, ETS1, and FRA1, and chlamydia HSP60 and β-actin as loading control. Data are representative of three biological replicates.

(E and F) hCEcto cells were (E) acutely or (F) persistently infected with *Ctr* for 32 h and 8 days p.i., respectively. Shown is the relative mRNA expression of selected ERF, ETS1, and FRA1 target genes analyzed by qRT-PCR. Data shown are means ± SDs of three biological replicates. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, Student's t test.

See also Figure S4.




*(legend on next page)*

the ERF transcription repressor domain. All of the END1/E6E7 mutants, but also overexpression of wild-type ERF, prevented *Ctr*-induced invasiveness (Figure 7B). This could be due to the excess availability of non-phosphorylated ERF exerting repressive activity. These results indicate that following *Ctr* infection, ERK-mediated phosphorylation of ERF at T526 leads to nuclear export, which relieves repressor activity and promotes invasion.

### ETS1-Dependent Transcription Program Is Crucial for *Ctr*-Induced EMT

The observed *Ctr*-induced phosphorylation of ETS1 at S282 creates binding sites for the COP1, E3 Ubiquitin Ligase (COP1) tumor suppressor protein, which is a ubiquitin ligase component that leads to ETS1 destruction (Lu et al., 2014). Damaged DNA-binding protein 1 (DDB1) and de-etiolated 1 (DET1), components of the COP1 complex, are significantly downregulated in *Ctr*-infected cells (Table S6), indicating that ubiquitin activity is suppressed, which prevents ETS1 degradation. In addition, the transcription factor runt-related transcription factor 1 (Runx1) can cooperatively interact with and effectively activate ETS1 by inducing a phosphorylation-refractory conformation of ETS1 via allosterically enhanced DNA binding stability (Shiina et al., 2015; Shrivastava et al., 2014). Together, the loss of ERF repressor activity and the increased ETS1 protein stability during *Ctr* infection indicate a transcriptional activation of ETS1 that may contribute to the observed EMT phenotype. We therefore created a CRISPR-Cas9-mediated ETS1 knockout cell line as confirmed by immunoblotting against total ETS1 protein (Figure 7C). The loss of ETS1 led to reduced expression of the effector genes IL8, TNF-$\alpha$, MMP3, early growth response 1 (EGR1), and PLAU (Figure 7D) and reduced invasiveness in *Ctr*-infected cells (Figure 7E). Thus, *Ctr* modulates ERK-mediated transcription factor regulation to induce effectors that promote an EMT phenotype with enhanced invasive capacity.

## DISCUSSION

Here, we performed an integrated global phosphoproteomic and transcriptomic analysis, revealing the striking impact of *Ctr* on host cell signaling and cellular behavior. Our comprehensive map of the signaling network of the total and nuclear fraction of host cells was used to generate a *Ctr*-responsive kinome network. Based on this, we identified the phosphorylation status of regulated transcription factors that are ERK/MAPK substrates and demonstrated the role of ETS1 and ERF in the resulting EMT phenotype. The results reveal a substantially greater range of *Ctr*-regulated signaling cascades than previously appreciated and provide a resource for generating deeper insight into their role in pathogenesis and potential host cell transformation.

Post-translational protein modifications have emerged as an additional level of dynamic control over protein function in diverse cell biological contexts. Protein phosphorylation is the most prevalent type of post-translational modification regulated in cellular signaling. By catalyzing the addition of phosphate groups to specific amino acids, usually Ser, Thr and Tyr residues, protein kinases regulate key processes such as cellular proliferation, survival, and migration and can contribute to the various hallmarks of cancer if their activity is deregulated (Fleuren et al., 2016). In line with emerging evidence indicating the ability of *Chlamydia* to interfere with protein function on the level of transcription and post-translational modification to modulate host cellular processes (Chumduri et al., 2016; Elwell et al., 2016; Siegl et al., 2014), we identify here 2,327 class I phosphorylation sites that are significantly affected by *Ctr* infection.

Bioinformatic analysis revealed that upregulated phosphosites during *Ctr* infection are involved in the regulation of transcription, gene expression, proliferation and nucleic acid metabolism, small GTPase-mediated signal transduction, stress-activated protein kinase signaling pathways, and JNK and MAPKKK cascades. Phosphosites downregulated by *Ctr,* however, are involved in cytoskeleton organization, regulation of protein complex disassembly, apoptosis, cell-cycle checkpoints, chromosomal organization, and DNA repair. Our data further highlight the regulation of numerous signaling cascades implicated in cancer, as well as organismal injury and abnormalities.

Host kinases have also been implicated in regulating *Chlamydia* proteins. Tarp and TepP, two effector proteins involved in host cell invasion, are rapidly tyrosine phosphorylated upon host cell entry by unknown host kinases facilitating the interaction with the host adaptor proteins (Chen et al., 2014; Mehlitz et al., 2010). However, we did not find these proteins to be phosphorylated, as they are known to be diminished to undetectable levels during later infection time points (Carpenter et al., 2017; Clifton et al., 2004). We identified 81 *Ctr* proteins, predominantly inclusion membrane proteins, to be phosphorylated. Furthermore, *in silico* kinase-substrate analysis revealed PKA, PKC, CK2, GSK3, GRK, CD5, and ERK1/2, among others, to be potential host kinases that regulate *Ctr* proteins. The functional

**Figure 5. *Chlamydia* Induces Loss of Cell Adhesions, Tissue Damage, and Invasive Phenotype**

(A and B) hCEctos (A) and End1/E6E7 cells (B) persistently infected with *Ctr* for 7 days were subjected to immunoblot analysis for E-cadherin, N-cadherin, chlamydia HSP60, and β-actin.

(C) Shown are the densitometry values of immunoblots shown in (A) and (B).

(D) Uninfected and *Ctr*-infected End1/E6E7 cells were subjected to immunofluorescence analysis 36 h p.i. and 8 days p.i. Representative confocal images for phalloidin (F-actin), *Ctr*-major outer membrane protein (MOMP), and DNA (Hoechst) are shown.

(E and F) Section of paraffin-embedded 3D-ALI cultures of hCEcto cells and human ectocervical tissue were subjected to immunofluorescence analysis. Shown are the representative confocal images of E-cadherin and p63 (E) p63 and Ki67 (F).

(G) Uninfected and *Ctr*-infected 3D-ALI cultures of hCEcto cells were subjected to immunofluorescence analysis for E-cadherin, *Ctr*-MOMP, and DAPI at indicated time points p.i.

(H and I) Uninfected and persistently infected End1/E6E7 (H) and hCEcto (I) cells with *Ctr* were analyzed using a transwell Matrigel-based invasion assay. HT1080 fibrosarcoma cells served as positive control. Images on the left show cells on top of the transwell that did not migrate; images on the right show cells that have migrated through the transwells in response to 10% fetal calf serum as a chemoattractant.

Data shown are representative of three biological replicates.

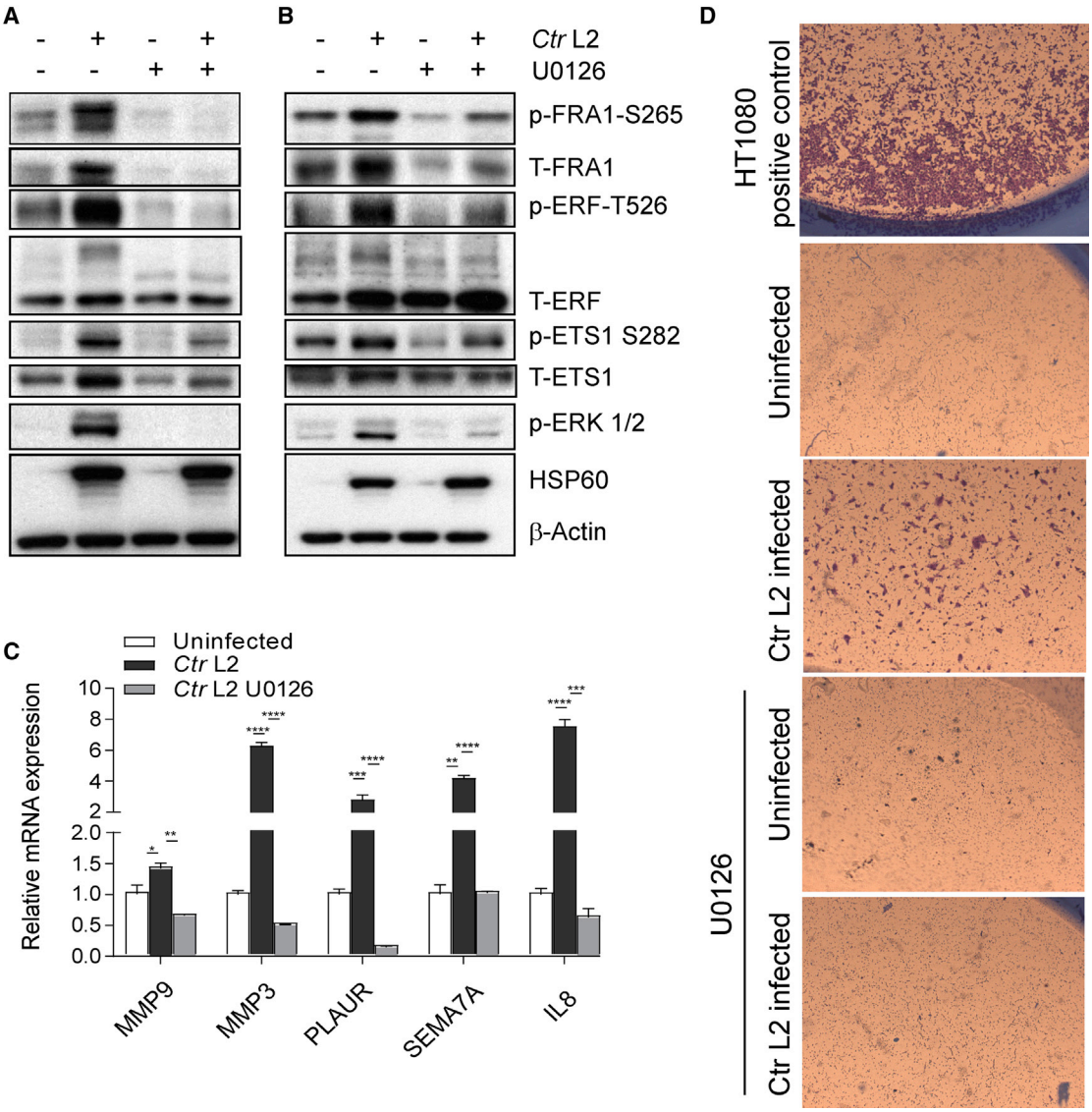



**Figure 6. Chlamydia-Induced ERK Signaling Regulates EMT-Associated Genes and Invasion**

(A and B) End1/E6E7 (A) and hCEcto (B) cells, either uninfected or *Ctr* infected for 32 h with or without additional U0126, were subjected to immunoblot analysis for *Ctr*-regulated phosphoproteins as indicated, chlamydial HSP60, and β-actin. Data are representative of four biological replicates.

(C and D) End1/E6E7 cells either uninfected or persistently *Ctr* infected with or without U0126 treatment.

(C) The expression of ERF, FRA1, and ETS1 target genes analyzed by qRT-PCR. Data are shown as means ± SDs from three technical replicates.

(D) Representative images of transwell Matrigel-based invasion assay. HT1080 fibrosarcoma cells were used as a positive control. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, Student's t test.

implications of the phosphorylation of these *Chlamydia* proteins await further investigation.

Previous studies demonstrated that the deregulation of ERK1/2, AKT, and checkpoint 2 (CHK2) kinases by *Chlamydia* enforces host cell proliferation by interfering with apoptosis and the response to metabolic and oxidative stress and DNA damage (Chumduri et al., 2013; Gurumurthy et al., 2010; Siegl et al., 2014). However, the majority of the regulated phosphoproteins and the predicted upstream kinases identified in the present study have not been previously associated with the response to *Chlamydia*. The *Chlamydia* responsive host kinome network highlights many nodes, including members of the CMGC kinase family such as CDKs and MAPKs, which are predominantly activated in both total cellular and nuclear fractions, while CAMK and AGC members such as CaMKII, CHK2, PKA, and PKC are selectively inactivated in the nuclear fraction. In line with this, cyclic AMP, which is a key regulator of PKA kinase activity, has been found to inhibit *Ctr* maturation (Kaul and Wenman, 1986). Our previous study demonstrated that phosphorylation of the DNA damage response checkpoint protein CHK2 is suppressed,

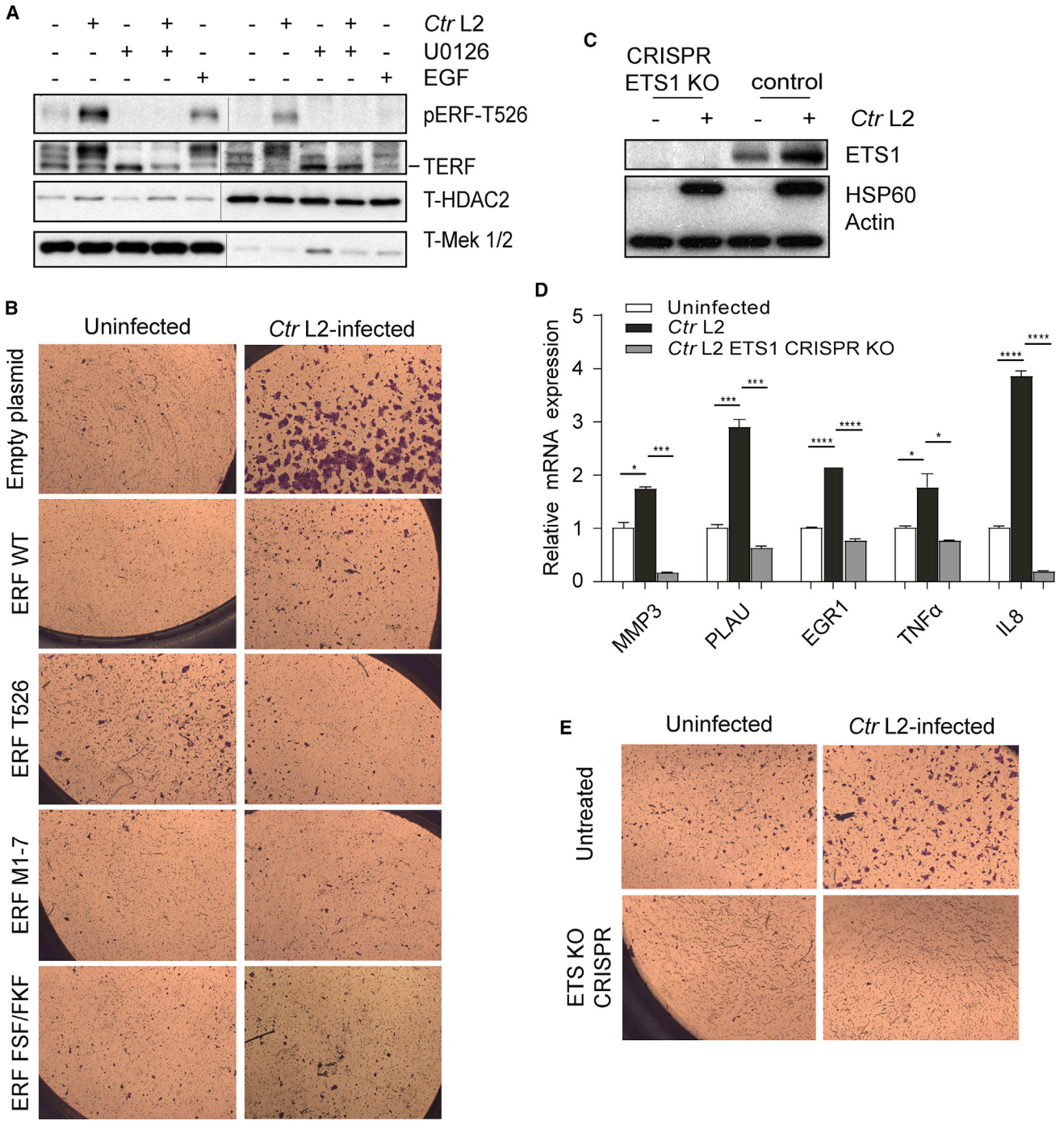



**Figure 7. ERF and ETS1 Are Key Regulators of *Ctr* Epithelial-to-Mesenchymal Transition**

(A) End1/E6E7 cells, either uninfected or *Ctr* infected for 32 h with or without U0126 treatment. Cytoplasmic and nuclear fractions were prepared and subjected to immunoblot analysis for total ERF and pERF T-526. Histone deacetylase 2 (HDAC2) and MEK1/2 were used as loading controls for nuclear and cytoplasmic fractions, respectively. EGF-treated cells were used as positive control. Data are representative of three biological replicates.

(B) End1/E6E7 cell lines overexpressing ERF wild-type (WT), different constructs with loss-of-function mutations in ERF including at ERF T526, all of the phosphosites activated by ERK (EFR M1–M7), and ERK interaction domain (ERF FSF/FKF) or control empty plasmids, respectively, were generated. These cells were either uninfected or persistently infected with *Ctr*, and invasion assay was performed. Representative images of the transwell Matrigel-based invasion assay are shown.

(C–E) CRISPR-Cas9-mediated ETS1 knockout End1/E6E7 cell line was generated.

(C and D) ETS1-CRISPR-Cas9 knockout (KO) and control cells were uninfected or infected with *Ctr* for 32 h.

*(legend continued on next page)*

despite the induction of DNA double-strand breaks upon *Ctr* infection (Chumduri et al., 2013).

Our present analysis revealed ERK1/2 as one of the predominant MAPK kinases activated by *Ctr* and is known to regulate a wide range of targets, thus controlling diverse signaling cascades involved in growth, proliferation, differentiation, survival, and migration. In support of the predicted increased activity of MAPK signaling during *Ctr* infection, other studies have demonstrated rapid activation of pERK, p38, and p-JNK pathways to promote *Chlamydia* growth (Chen et al., 2010; Chumduri et al., 2013; Olive et al., 2014). Our exploration of *Ctr*-induced phosphoproteins identified several transcription factors, including ETS1 and FRA1, that are implicated in EMT as strong interaction partners of MAPK1/3 and have not yet been functionally linked to *Ctr* infections (Figure 3B).

Genes associated with EMT were overrepresented in the *Ctr*-responsive transcriptome and included ETS1, ERF, and FRA1. ETS1 is a proto-oncogenic transcription factor containing a conserved ETS DNA binding domain (EBS) that activates multiple genes involved in senescence, apoptosis, angiogenesis, stem cell development, cell migration, and cancer development (Dejana et al., 2007; Plotnik et al., 2014; Sharrocks, 2001). ERF, another member of the ETS family, is a potent, ubiquitously expressed transcriptional repressor that recognizes promoters with the EBS motif and regulates genes involved in proliferation and Ras-induced tumorigenicity (Allegra et al., 2012; Mavrothalassitis and Ghysdael, 2000; Sgouras et al., 1995). Deregulation of the ETS family of transcription factors has been implicated in the malignant transformation of cells, as they control genes that are important for invasion and metastasis, such as MMPs and PLAUR (Oikawa, 2004). FRA1 is an oncogenic member of the Fos subfamily of basic leucine zipper domain (bZIP) transcription factors. Fos proteins dimerize with Jun proteins to transactivate AP1-dependent genes, including EMT-associated MMPs, urokinase receptors (uPARs), integrins, adhesins and inflammatory genes. FRA1 is involved in cell motility and invasiveness. Elevated FRA1 levels are associated with tumorigenesis and cancer progression (Diesch et al., 2014). We also observed that many AP1-related proteins, such as JunB and D, cFos, Finkel-Biskis-Jinkins murine osteosarcoma viral oncogene homolog B (FOSB), and FRA1, are transcriptionally regulated by *Ctr* along with enhanced ERK1/2-dependent phosphorylation of FRA1 at S265 sites, which is critical for its stability (Basbous et al., 2007). Knockdown of ETS1 or ERF impaired *Chlamydia* development. However, whether FRA1 is of similar importance for the pathogen remains unresolved, as knockdown was lethal for cells, which is consistent with previous observations (Meise et al., 2012).

Our approach also provides insight into key molecular mechanisms through which *Ctr* induces the EMT phenotype. We found that *Ctr*-induced phosphorylation of ERK leads in turn to the phosphorylation of ERF at T526 and ETS1 at S282. ERK phosphorylates ERF at multiple sites to promote nuclear export and cytoplasmic localization, thus relieving its transcriptional repressor function (Le Gallic et al., 1999, 2004). We observed that phosphorylated ERF-T526 accumulates in the cytoplasm and that ERF mutants at ERK target sites, including T626, inhibited the induction of invasiveness by *Ctr*.

Similarly, invasiveness induction was not observed after knockout of ETS1. The phosphorylation of ETS1 at S282 we observed is known to inhibit its DNA binding activity and promote binding of the COP1 ubiquitin complex proteins Cullin 4A (CUL4A), DDB1, and DET1, leading to its degradation (Lu et al., 2014). Despite this, ETS1 protein levels were stabilized after infection, which could be explained by the marked downregulation of DDB1 and DET1 that we observed. It is also known that cooperative transcription factors such as RUNX1 can stabilize its interaction with DNA and thereby override the inhibitory effects of phosphorylation (Shiina et al., 2015; Shrivastava et al., 2014).

EMT is fundamental in development, wound healing, and stem cell behavior and contributes pathologically to fibrosis, tissue scarring, and cancer progression (Lamouille et al., 2014). Wound healing and tumorigenesis share a common phenotype characterized by cells changing from a stationary, differentiated to a migratory, de-differentiated phenotype (Leopold et al., 2012), and malignant tumors frequently arise at sites of chronic tissue injury and excessive wound healing (Schäfer and Werner, 2008). We observed that the enrichment of factors associated with EMT in the *Ctr*-regulated phosphoproteome is accompanied by a gain of invasive capacity of infected cells, together with persistent transcriptional upregulation of genes involved in cellular movement (PLAU), invasiveness (SEMA7A), and extracellular matrix degradation (MMPs). In addition, resistance to apoptosis and senescence, which are also acquired during EMT, are known to be induced by *Ctr* infection (Chumduri et al., 2013; Thiery et al., 2009).

Our phosphoproteomic and transcriptomic data, together with our primary cervical infection model, provide insights into the signaling and mechanisms underlying *Ctr* pathology on several levels. The EMT phenotype of infected cells, in particular the loss of epithelial cell adhesion, is likely to play a role in the epithelial scarring associated with infections (Darville and Hiltke, 2010). In addition, the resulting ability of the bacteria to gain access to the basal stem cells, which are the target cells of the HPV tumor virus, may explain the epidemiological evidence for *Ctr* as a cofactor in cervical cancer (Koskela et al., 2000; Shanmughapriya et al., 2012; Zhu et al., 2016). However, EMT induction in itself may promote epithelial transformation, especially in the context of our previous observations that *Ctr* downregulates the DNA damage response while simultaneously inducing widespread DNA damage (Chumduri et al., 2016). The results of this study will provide a platform to generate new insights into the pathogenesis of *Ctr* infections and their potential synergy with other human genital tract infections.

(C) The blots confirming the loss of ETS1 in ETS1-CRISPR-Cas9 knockout cells compared to control conditions are shown. Chlamydial HSP60 and β-actin served as infection and loading controls, respectively.

(D) Relative mRNA expression of ETS1 target genes analyzed by qRT-PCR. Data are shown as means ± SDs from three biological replicates.

(E) ETS1-CRISPR-Cas9 knockout and control cells were uninfected or persistently infected with *Ctr*, and invasion assay was performed. Representative images of the transwell Matrigel-based invasion assay are shown.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Chlamydia infections
  - Cell lines
  - Human ectocervical (hCEcto) primary cell isolation and propagation
  - Three dimensional air-liquid interface cultures of human derived ectocervix
- METHOD DETAILS
  - Infectivity assays
  - Invasion assay
  - SDS-PAGE and western blotting
  - siRNA transfection and knockdown analysis
  - CRISPR/Cas9 Knockout Cell Line Generation
  - Generation of the ERF overexpression cell lines
  - Cellular fractionation
  - SILAC labeling and labeling efficiency
  - Sample preparation for mass spectrometric (MS) analysis
  - Phosphopeptide enrichment
  - NanoLC-MS/MS analysis
  - Microarray analysis
  - Immunofluorescent histochemistry
  - Automated microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Phosphoproteome/Proteome data analysis
  - Linear signature motif analysis
  - iGPS analysis – prediction of site-specific kinase-substrate relationship
  - GO enrichment analysis
  - STRING protein-protein interaction analysis
  - Ingenuity pathway analysis (IPA)
  - GSEA
  - Statistics
- DATA AND SOFTWARE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and nine tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2019.01.006.

## ACKNOWLEDGMENTS

The authors would like to thank George Mavrothalassitis for kindly providing the ERF constructs; Monika Schmid, Jörg Angermann, Meike Soerensen, and Elke Ziska for technical help; Hans-Joachim Mollenkopf and Ina Wagner for the microarrays; Diane Schad for assistance with graphics; and Rike Zietlow for editing the manuscript. P.K.Z. is a fellow of the German Academic Exchange Service (DAAD). M.Y. is a fellow of the China Scholarship Council.

## AUTHOR CONTRIBUTIONS

R.K.G., C.C., and T.F.M. conceived and designed the study. P.K.Z., R.K.G., C.C., and Y.M. performed the experiments. P.K.Z., C.C., R.K.G., and H.B. performed the data and bioinformatics assessments. K.I. and M.S. performed the MS analysis. C.C., P.K.Z., R.K.G., and T.F.M. wrote the manuscript. T.F.M. supported the project financially. R.K.G. supervised the work.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 2, 2018
Revised: October 5, 2018
Accepted: December 31, 2018
Published: January 29, 2019

## REFERENCES

Allegra, M., Zaragkoulias, A., Vorgia, E., Ioannou, M., Litos, G., Beug, H., and Mavrothalassitis, G. (2012). Semaphorin-7a reverses the ERF-induced inhibition of EMT in Ras-dependent mouse mammary epithelial cells. Mol. Biol. Cell *23*, 3873–3881.

Basbous, J., Chalbos, D., Hipskind, R., Jariel-Encontre, I., and Piechaczyk, M. (2007). Ubiquitin-independent proteasomal degradation of Fra-1 is antagonized by Erk1/2 pathway-mediated phosphorylation of a unique C-terminal destabilizer. Mol. Cell. Biol. *27*, 3936–3950.

Buchholz, K.R., and Stephens, R.S. (2007). The extracellular signal-regulated kinase/mitogen-activated protein kinase pathway induces the inflammatory factor interleukin-8 following Chlamydia trachomatis infection. Infect. Immun. *75*, 5924–5929.

Campeau, E., Ruhl, V.E., Rodier, F., Smith, C.L., Rahmberg, B.L., Fuss, J.O., Campisi, J., Yaswen, P., Cooper, P.K., and Kaufman, P.D. (2009). A versatile viral system for expression and depletion of proteins in mammalian cells. PLoS One *4*, e6529.

Carpenter, V., Chen, Y.S., Dolat, L., and Valdivia, R.H. (2017). The Effector TepP Mediates Recruitment and Activation of Phosphoinositide 3-Kinase on Early *Chlamydia trachomatis* Vacuoles. MSphere *2*, e00207–e00217.

Chen, F., Cheng, W., Zhang, S., Zhong, G., and Yu, P. (2010). [Induction of IL-8 by Chlamydia trachomatis through MAPK pathway rather than NF-kappaB pathway]. Zhong Nan Da Xue Xue Bao Yi Xue Ban *35*, 307–313.

Chen, Y.S., Bastidas, R.J., Saka, H.A., Carpenter, V.K., Richards, K.L., Plano, G.V., and Valdivia, R.H. (2014). The Chlamydia trachomatis type III secretion chaperone Slc1 engages multiple early effectors, including TepP, a tyrosine-phosphorylated protein required for the recruitment of CrkI-II to nascent inclusions and innate immune signaling. PLoS Pathog. *10*, e1003954.

Chumduri, C., Gurumurthy, R.K., Zadora, P.K., Mi, Y., and Meyer, T.F. (2013). Chlamydia infection promotes host DNA damage and proliferation but impairs the DNA damage response. Cell Host Microbe *13*, 746–758.

Chumduri, C., Gurumurthy, R.K., Zietlow, R., and Meyer, T.F. (2016). Subversion of host genome integrity by bacterial pathogens. Nat. Rev. Mol. Cell Biol. *17*, 659–673.

Claywell, J.E., Matschke, L.M., and Fisher, D.J. (2016). The Impact of Protein Phosphorylation on Chlamydial Physiology. Front. Cell. Infect. Microbiol. *6*, 197.

Clifton, D.R., Fields, K.A., Grieshaber, S.S., Dooley, C.A., Fischer, E.R., Mead, D.J., Carabeo, R.A., and Hackstadt, T. (2004). A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. Proc. Natl. Acad. Sci. USA *101*, 10166–10171.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Darville, T., and Hiltke, T.J. (2010). Pathogenesis of genital tract disease due to Chlamydia trachomatis. J. Infect. Dis. *201* (*Suppl 2*), S114–S125.

Dejana, E., Taddei, A., and Randi, A.M. (2007). Foxs and Ets in the transcriptional regulation of endothelial cell differentiation and angiogenesis. Biochim. Biophys. Acta *1775*, 298–312.

Diesch, J., Sanij, E., Gilan, O., Love, C., Tran, H., Fleming, N.I., Ellul, J., Amalia, M., Haviv, I., Pearson, R.B., et al. (2014). Widespread FRA1-dependent control of mesenchymal transdifferentiation programs in colorectal cancer cells. PLoS One *9*, e88950.

Eid, S., Turk, S., Volkamer, A., Rippmann, F., and Fulle, S. (2017). KinMap: a web-based tool for interactive navigation through human kinome data. BMC Bioinformatics *18*, 16.

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

Fleuren, E.D., Zhang, L., Wu, J., and Daly, R.J. (2016). The kinome 'at large' in cancer. Nat. Rev. Cancer *16*, 83–98.

González, E., Rother, M., Kerr, M.C., Al-Zeer, M.A., Abu-Lubad, M., Kessler, M., Brinkmann, V., Loewer, A., and Meyer, T.F. (2014). Chlamydia infection depends on a functional MDM2-p53 axis. Nat. Commun. *5*, 5201.

Gröger, C.J., Grubinger, M., Waldhör, T., Vierlinger, K., and Mikulits, W. (2012). Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. PLoS One *7*, e51136.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812.

Gurumurthy, R.K., Mäurer, A.P., Machuy, N., Hess, S., Pleissner, K.P., Schuchhardt, J., Rudel, T., and Meyer, T.F. (2010). A loss-of-function screen reveals Ras- and Raf-independent MEK-ERK signaling during Chlamydia trachomatis infection. Sci. Signal. *3*, ra21.

Hall, D.M., and Brooks, S.A. (2014). In vitro invasion assay using matrigel™: a reconstituted basement membrane preparation. Methods Mol. Biol. *1070*, 1–11.

Heckl, D., Kowalczyk, M.S., Yudovich, D., Belizaire, R., Puram, R.V., McConkey, M.E., Thielke, A., Aster, J.C., Regev, A., and Ebert, B.L. (2014). Generation of mouse models of myeloid malignancy with a combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nat. Biotechnol *32*, 941–946.

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. *43*, D512–D520.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hunter, T., and Sefton, B.M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proc. Natl. Acad. Sci. USA *77*, 1311–1315.

Kaul, R., and Wenman, W.M. (1986). Cyclic AMP inhibits developmental regulation of Chlamydia trachomatis. J. Bacteriol. *168*, 722–727.

Koskela, P., Anttila, T., Bjørge, T., Brunsvig, A., Dillner, J., Hakama, M., Hakulinen, T., Jellum, E., Lehtinen, M., Lenner, P., et al. (2000). Chlamydia trachomatis infection as a risk factor for invasive cervical cancer. Int. J. Cancer *85*, 35–39.

Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. Nat. Rev. Mol. Cell Biol. *15*, 178–196.

Le Gallic, L., Sgouras, D., Beal, G., Jr., and Mavrothalassitis, G. (1999). Transcriptional repressor ERF is a Ras/mitogen-activated protein kinase target that regulates cellular proliferation. Mol. Cell. Biol. *19*, 4121–4133.

Le Gallic, L., Virgilio, L., Cohen, P., Biteau, B., and Mavrothalassitis, G. (2004). ERF nuclear shuttling, a continuous monitor of Erk activity that links it to cell cycle progression. Mol. Cell. Biol. *24*, 1206–1218.

Leopold, P.L., Vincent, J., and Wang, H. (2012). A comparison of epithelial-to-mesenchymal transition and re-epithelialization. Semin. Cancer Biol. *22*, 471–483.

Lu, G., Zhang, Q., Huang, Y., Song, J., Tomaino, R., Ehrenberger, T., Lim, E., Liu, W., Bronson, R.T., Bowden, M., et al. (2014). Phosphorylation of ETS1 by Src family kinases prevents its recognition by the COP1 tumor suppressor. Cancer Cell *26*, 222–234.

Mavrothalassitis, G., and Ghysdael, J. (2000). Proteins of the ETS family with transcriptional repressor activity. Oncogene *19*, 6524–6532.

Mehlitz, A., Banhart, S., Mäurer, A.P., Kaushansky, A., Gordus, A.G., Zielecki, J., Macbeath, G., and Meyer, T.F. (2010). Tarp regulates early Chlamydia-induced host cell survival through interactions with the human adaptor protein SHC1. J. Cell Biol. *190*, 143–157.

Meise, R., Tomicic, M.T., Kaina, B., and Christmann, M. (2012). The chloroethylating anticancer drug ACNU induces FRA1 that is involved in drug resistance of glioma cells. Biochim. Biophys. Acta *1823*, 1199–1207.

Oikawa, T. (2004). ETS transcription factors: possible targets for cancer therapy. Cancer Sci. *95*, 626–633.

Ojcius, D.M., Degani, H., Mispelter, J., and Dautry-Varsat, A. (1998). Enhancement of ATP levels and glucose metabolism during an infection by Chlamydia. NMR studies of living cells. J. Biol. Chem. *273*, 7052–7058.

Olive, A.J., Haff, M.G., Emanuele, M.J., Sack, L.M., Barker, J.R., Elledge, S.J., and Starnbach, M.N. (2014). Chlamydia trachomatis-induced alterations in the host cell proteome are required for intracellular growth. Cell Host Microbe *15*, 113–124.

Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell *127*, 635–648.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell. Proteomics *1*, 376–386.

Plotnik, J.P., Budka, J.A., Ferris, M.W., and Hollenhorst, P.C. (2014). ETS1 is a genome-wide effector of RAS/ERK signaling in epithelial cells. Nucleic Acids Res. *42*, 11928–11940.

Rajalingam, K., Sharma, M., Lohmann, C., Oswald, M., Thieck, O., Froelich, C.J., and Rudel, T. (2008). Mcl-1 is a key regulator of apoptosis resistance in Chlamydia trachomatis-infected cells. PLoS One *3*, e3102.

Rajasekaran, S., Reddy, N.M., Zhang, W., and Reddy, S.P. (2013). Expression profiling of genes regulated by Fra-1/AP-1 transcription factor during bleomycin-induced pulmonary fibrosis. BMC Genomics *14*, 381.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal. Chem. *75*, 663–670.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.

Rockey, D.D., Grosenbach, D., Hruby, D.E., Peacock, M.G., Heinzen, R.A., and Hackstadt, T. (1997). Chlamydia psittaci IncA is phosphorylated by the host cell and is exposed on the cytoplasmic face of the developing inclusion. Mol. Microbiol. *24*, 217–228.

Rother, M., Gonzalez, E., Teixeira da Costa, A.R., Wask, L., Gravenstein, I., Pardo, M., Pietzke, M., Gurumurthy, R.K., Angermann, J., Laudeley, R., et al. (2018). Combined Human Genome-wide RNAi and Metabolite Analyses Identify IMPDH as a Host-Directed Target against Chlamydia Infection. Cell Host Microbe *23*, 661–671.e8.

Schäfer, M., and Werner, S. (2008). Cancer as an overhealing wound: an old hypothesis revisited. Nat. Rev. Mol. Cell Biol. *9*, 628–638.

Schwartz, D., and Gygi, S.P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat. Biotechnol. *23*, 1391–1398.

Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv. https://doi.org/10.1101/060012.

Sgouras, D.N., Athanasiou, M.A., Beal, G.J., Jr., Fisher, R.J., Blair, D.G., and Mavrothalassitis, G.J. (1995). ERF: an ETS domain protein with strong transcriptional repressor activity, can suppress ets-associated tumorigenesis

and is regulated by phosphorylation during cell cycle and mitogenic stimulation. EMBO J. *14*, 4781–4793.

Shanmughapriya, S., Senthilkumar, G., Vinodhini, K., Das, B.C., Vasanthi, N., and Natarajaseenivasan, K. (2012). Viral and bacterial aetiologies of epithelial ovarian cancer. Eur. J. Clin. Microbiol. Infect. Dis. *31*, 2311–2317.

Sharrocks, A.D. (2001). The ETS-domain transcription factor family. Nat. Rev. Mol. Cell Biol. *2*, 827–837.

Shiina, M., Hamada, K., Inoue-Bungo, T., Shimamura, M., Uchiyama, A., Baba, S., Sato, K., Yamamoto, M., and Ogata, K. (2015). A novel allosteric mechanism on protein-DNA interactions underlying the phosphorylation-dependent regulation of Ets1 target gene expressions. J. Mol. Biol. *427*, 1655–1669.

Shrivastava, T., Mino, K., Babayeva, N.D., Baranovskaya, O.I., Rizzino, A., and Tahirov, T.H. (2014). Structural basis of Ets1 activation by Runx1. Leukemia *28*, 2040–2048.

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

Son, H., and Moon, A. (2010). Epithelial-mesenchymal Transition and Cell Invasion. Toxicol. Res. *26*, 245–252.

Song, C., Ye, M., Liu, Z., Cheng, H., Jiang, X., Han, G., Songyang, Z., Tan, Y., Wang, H., Ren, J., et al. (2012). Systematic analysis of protein phosphorylation networks from phosphoproteomic data. Mol. Cell. Proteomics *11*, 1070–1083.

Su, H., McClarty, G., Dong, F., Hatch, G.M., Pan, Z.K., and Zhong, G. (2004). Activation of Raf/MEK/ERK/cPLA2 signaling pathway is essential for chlamydial acquisition of host glycerophospholipids. J. Biol. Chem. *279*, 9409–9416.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. *43*, D447–D452.

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

Vizcaíno, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. *44* (*D1*), D447–D456.

Zhu, H., Shen, Z., Luo, H., Zhang, W., and Zhu, X. (2016). Chlamydia Trachomatis Infection-Associated Risk of Cervical Cancer: A Meta-Analysis. Medicine (Baltimore) *95*, e3077.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Rabbit-anti-Phospho-Ser | Millipore | Cat# AB1603; RRID:AB_390205 |
| Rabbit-anti- Phospho-Thr | Cell Signaling | Cat# 9381S; RRID:AB_10691696 |
| Mouse-anti- p-Tyr | Santa Cruz Biotechnology | Cat# sc-7020; RRID:AB_628123 |
| Rabbit-anti-Phospho-FRA1(Ser265) | Cell Signaling | Cat# 3880; RRID:AB_2106922 |
| Rabbit-anti-FRA1 | Abcam | Cat# ab124722; RRID:AB_11001005 |
| Rabbit-anti-Phospho-ERF | LSBio | Cat# LS-C342103 |
| Rabbit-anti-ERF | Abcam | Cat# ab153726 |
| Rabbit-anti-Phospho-ETS1(Ser282) | Invitrogen | Cat# 441109G; RRID:AB_2533577 |
| Rabbit-anti-ETS1 | Abcam | Cat# ab124282; RRID:AB_10975199 |
| Mouse-anti-Phospho-ERK1/2 (T185/Y187) | Sigma-Aldrich | Cat# M8159; RRID:AB_477245 |
| Rabbit-anti-Phospho-EGFR (Y1172) | Abcam | Cat# ab47364; RRID:AB_873777 |
| Rabbit-anti-MEK1/2 | Cell Signaling | Cat# 9126; RRID:AB_331778 |
| Mouse-anti-HDAC2 | Cell Signaling | Cat# 5113; RRID:AB_10624871 |
| Mouse-anti-HSP60 | Enzo | Cat# ALX-804-071; RRID:AB_10539940 |
| Mouse monoclonal species-specific KK-12 IgG2a Ctr (anti-MOMP) | D. Grayston, University of Washington, Seattle, WA, USA | N/A |
| Mouse-anti-beta-Actin | Sigma-Aldrich | Cat# A5541 |
| Mouse-anti-E-cadherin | BD Biosciences | Cat# 610181; RRID:AB_397580 |
| Rabbit-anti-N-cadherin | Abcam | Cat# ab18203; RRID:AB_444317 |
| Mouse-anti-Histone H4 | Cell Signaling | Cat# 2935S; RRID:AB_1147658 |
| Rabbit-anti-p63 | Abcam | Cat# ab53039; RRID:AB_881860 |
| Mouse-anti-p63 | Abcam | Cat# ab735; RRID:AB_305870 |
| Rabbit-anti-Ki67 | Abcam | Cat# ab16667; RRID:AB_302459 |
| Cy3-conjugated goat anti-mouse | Jackson Immunoresearch | Cat# 115-165-146; RRID:AB_2491007 |
| Cy3-conjugated goat anti-rabbit | Dianova | Cat# 111-165-144; RRID:AB_2338006 |
| Cy2-conjugated goat anti-mouse | Dianova | Cat# 115-225-062; RRID:AB_2338741 |
| Donkey-anti-mouse IgG-HRP | Santa Cruz Biotechnology | Cat# sc-2314; RRID:AB_641170 |
| Donkey-anti-rabbit IgG-HRP | GE Healthcare | Cat# NA934; RRID:AB_772206 |
| Bacterial and Virus Strains | | |
| *Chlamydia trachomatis* serovar L2 | ATCC | VR-902B |
| Biological Samples | | |
| hCEctos | Department of Gynecology, Charité University Hospital, Berlin, Germany | N/A |
| Chemicals, Peptides, and Recombinant Proteins | | |
| Hoechst 33342 | Sigma | Cat# 23491-52-3 |
| Draq5 | Cell Signaling | Cat# 4084 |
| Collagenase type II | Calbiochem | Cat# 234155 |
| TrypLE | GIBCO | Cat# 12604021 |
| HEPES | Invitrogen | Cat# 15630-056 |
| GlutaMax | Invitrogen | Cat# 35050-038 |
| B-27 | Invitrogen | Cat# 17504-044 |
| N2 | Invitrogen | Cat# 17502048 |
| Hydrocortisone | Sigma | Cat# H0888 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Human EGF | Invitrogen | Cat# PHG0311 |
| Human Noggin | Peprotech | Cat# 120-10C |
| Human FGF-10 | Peprotech | Cat# 100-26-25 |
| N-acetyl-L-cysteine | Sigma | Cat# A9165 |
| Nicotinamide | Sigma | Cat# N0636 |
| TGF-β R kinase inhibitor IV | Calbiochem | Cat# 616454 |
| ROCK inhibitor Y-27632 | Hölzel Diagnostika | Cat# M1817 |
| Forskolin | Sigma | Cat# F6886 |
| Penicillin/streptomycin | GIBCO | Cat# 15140-122 |
| Collagen | Sigma | Cat# C3867 |
| Bovine skin collagen | Sigma | Cat# C4243 |
| Opti-MEM medium | GIBCO | Cat# 31985070 |
| Fugene 6 | Promega | Cat# E2691 |
| HiPerfect | Quiagen | Cat# 301709; Lot. No. 127147353 and Lot. No. 136232611 |
| Esp3I restriction enzyme | ThermoFisher | Cat# FD0454 |
| RMPI 1640 medium | GIBCO | Cat# 52400 |
| DMEM | GIBCO | Cat# 10938 |
| Advanced DMEM/F12 | GIBCO | Cat# 12634 |
| Fetal calf serum | Biochrom AG | Cat# S0155 |
| Glutamine | GIBCO | Cat# 25030081 |
| Sodium pyruvate | GIBCO | Cat# 11360070 |
| PBS | GIBCO | Cat# 14190-094 |
| Matrigel | Corning | Cat# 356231 |
| (H) $^{13}C_6{}^{15}N_2$ L-Lysine | Sigma | Cat# 608041 |
| $^{13}C_6{}^{15}N_4$ L-Arginine | Sigma | Cat# 608033 |
| Critical Commercial Assays | | |
| NE-PER kit | Thermo Fisher | Cat# 78833 |
| Dual-color Quick-Amp Labeling Kit | Agilent Technologies | Cat# 5190-0444 |
| Deposited Data | | |
| Microarray data were deposited in the National Center for Biotechnology Information Gene Expression Omnibus | https://www.ncbi.nlm.nih.gov/geo/ | GSE104166 |
| The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository | http://www.proteomexchange.org/ | PXD011960 |
| Experimental Models: Cell Lines | | |
| HEK293T cells | ATCC | CRL3216; RRID: CVCL_0063 |
| End1 E6/E7 cells | ATCC | CRL-2615 |
| 3T3-J2 cells (gift from Craig Meyers) | Howard Green laboratory, Harvard University | N/A |
| HeLa cells | ATCC | CCL-2; RRID: CVCL_0030 |
| HT1080 cells | ATCC | CCL-121 |
| Oligonucleotides | | |
| siRNA oligonucleotides | See Table S9 | See Table S9 |
| Primers for qPCR | See Table S9 | See Table S9 |
| gRNA oligonucleotides | See Table S9 | See Table S9 |
| ETS1–CRISPR KO Primer | See Table S9 | See Table S9 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Recombinant DNA | | |
| psPax2 | D. Trono, unpublished (http://n2t.net/addgene:12260) | Addgene Cat# 12260 |
| pMD.2G (VSVG) | D. Trono, unpublished (http://n2t.net/addgene:12259) | Addgene Cat# 12259 |
| pL-CRISPR.EFS.GFP plasmid | Heckl et al., 2014 | Addgene Cat# 57818 |
| pLenti-CMV-GFP vector | Campeau et al., 2009 | Addgene Cat# 17448 |
| pSG5-ERF wild-type | Prof. G. Mavrothalasitis (Allegra et al., 2012) | n/a |
| pSG5-ERF-T526 | Prof. G. Mavrothalasitis (Le Gallic et al., 2004) | N/A |
| pSG5-ERF-M1-7 | Prof. G. Mavrothalasitis (Allegra et al., 2012) | N/A |
| pSG5-ERF-FSF/FKF | Prof. G. Mavrothalasitis (Allegra et al., 2012) | N/A |
| Software and Algorithms | | |
| Rosetta Resolver Biosoftware, Build 7.2.2 SP1.31 | Rosetta Biosoftware | No longer available |
| Image Analysis/Feature Extraction software G2567AA v. A.11.5.1.1 | Agilent Technologies | Cat. #G4460 |
| MaxQuant (v1.5.1.2) | Cox and Mann, 2008 | http://www.coxdocs.org/doku.php?id=maxquant:start |
| Motif-X | Schwartz and Gygi, 2005 | http://motif-x.med.harvard.edu/ |
| KinMap | Eid et al., 2017 | http://kinhub.org/kinmap/ |
| DAVID | Huang et al., 2009 | https://david.ncifcrf.gov/ |
| TM4 | http://mev.tm4.org | http://mev.tm4.org |
| STRING | Szklarczyk et al., 2015 | https://string-db.org/ |
| fgsea R package | Sergushichev, 2016 | https://bioconductor.org/packages/release/bioc/html/fgsea.html |
| R (v3.3) | R Core Team | https://cran.r-project.org/ |
| limma R package | Ritchie et al., 2015 | https://bioconductor.org/packages/release/bioc/html/limma.html |
| Circlize R package | Gu et al., 2014 https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu393 | https://cran.r-project.org/web/packages/circlize/ |
| Tibco Spotfire (v7) | TIBCO Software Inc. | https://www.tibco.com/products/tibco-spotfire |
| Ingenuity® Pathway Analysis | https://www.qiagen.com/ingenuity | Cat #: 830102 |
| ScanR Analysis | Olympus Soft Imaging Solutions | https://www.olympus-lifescience.com/en/microscopes/inverted/scanr/#!cms[tab]=%2Fmicroscopes%2Finverted%2Fscanr%2Ffeatures |
| CHOP CHOP tool | http://chopchop.cbu.uib.no | http://chopchop.cbu.uib.no |
| GraphPad Prism | GraphPad Software | https://www.graphpad.com/scientific-software/prism/ |
| iGPS software v1.0.1 | Song et al., 2012 | http://igps.biocuckoo.org/ |
| Other | | |
| Polyethylene terephthalate (PET) track-etched membrane | Corning | Cat# 353097 |
| C18 Stage Tips packed with 10 mg of ReproSil-Pur 120 C18-AQ 5-μm resin | Dr. Maisch GmbH | Cat# r15.aq. |
| monolithic column MonoCap C18 High Resolution 2000 | GL Sciences | Cat# 5020-10015 |
| Cell Culture Inserts, 30 mm, hydrophilic PTFE, 0.4 μm | Merck | Cat# PICM03050 |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Thomas F. Meyer (meyer@mpiib-berlin.mpg.de).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Chlamydia infections

*Ctr* L2 (ATCC VR-902B), stocks were prepared as described earlier (Gurumurthy et al., 2010). Briefly, Ctr was propagated in HeLa cells grown in 150-$cm^2$ cell culture flasks in 24 mL of infection medium [DMEM (GIBCO) supplemented with 5% fetal calf serum (FCS) (Biochrom), 2 mM glutamine, and 1 mM sodium pyruvate]. The cells were detached 48 hours after infection with 3-mm glass beads and centrifuged at 500xg for 10 min at 4°C. Cells were resuspended in sucrose-phosphate-glutamate (SPG) buffer and ruptured by vortexing with glass beads. Cell lysates were then centrifuged as before to sediment nuclei and cell debris. The supernatant was further centrifuged at 20,000xg for 40 min at 4°C, and the resulting bacterial pellet resuspended in 15 mL SPG buffer with a 21- to 22-gauge injection needle. Chlamydia suspensions were stored in aliquots at −75°C until required. Chlamydia infection experiments were performed at a multiplicity of infection (MOI) of 5 unless stated otherwise in infection medium (DMEM supplemented with 5% FCS, 2 mM glutamine, and 1mM sodium pyruvate). The medium was refreshed 2 h p.i., and cells were grown at 35°C in 5% $CO_2$. For persistent *Ctr* infection, cells were infected (MOI 5 unless stated otherwise) for 24 h. 24 h p.i 250 ng/ml doxycycline was added to both uninfected and infected cells and cells allowed to grow for 7 d p.i.

### Cell lines

3T3-J2 (mouse embryo) (kind gift from Craig Meyers), End1/E6E7 (End1) (♀) [American Type Culture Collection (ATCC) CRL-2615], HT1080 (♂) (ATCC, CCL-121) and HeLa (♀) (ATCC, CCL-2) cells were cultured in HEPES-buffered growth medium [DMEM (GIBCO) supplemented with 10% FCS (Biochrome), 2 mM glutamine, and 1 mM sodium pyruvate], at 37°C in a humidified incubator containing 5% $CO_2$.

### Human ectocervical (hCEcto) primary cell isolation and propagation

Human ectocervix samples were provided by the Department of Gynecology, Charité University Hospital, Berlin, Germany. Scientific usage of the samples was approved by the ethics committee of the Charité University Hospital, Berlin (EA1/059/15); informed consent was obtained from all subjects to use their tissue for scientific research. Only anatomically normal cervical tissues were used, within 2-3 h after removal. Human ectocervical biopsy tissue from a 50-year old female patient was washed in 10 cm Petri dish with 1x PBS (GIBCO, # 14190-094) and minced with surgical scissors before incubating in 0.5 mg/ml collagenase type II (Calbiochem, # 234155) for 2.5 h at 37°C in a shaker incubator. Tissue and dissociated cells were pelleted by centrifugation (5 min at 1000 g, 4°C), supernatant was discarded, cells were resuspended in TrypLE express (GIBCO, # 12604021) and incubated for 15 mins at 37°C in a shaker incubator. After dissociation, the cell and tissue pellet was resuspended in ADF (Invitrogen) medium and passed through a 40-μM cell strainer (BD Falc, # 352340) to separate the single dissociated cells from tissue pieces. Cells were pelleted by centrifugation (5 min at 1000xg, 4°C), resuspended in human ectocervical primary cell medium for cell expansion in 75 $cm^2$ flask coated with collagen. At 70%–80% confluence cells were passaged using TrypLE and seeded on lethally irradiated 3T3-J2 mouse fibroblasts in the ectocervical primary cell medium (Consisted of ADF, 12 mM HEPES and 1% GlutaMax, supplemented with 1% B27, 1% N2, 0.5 μg/ml hydrocortisone (Sigma, # H0888-1G), 10 ng/ml human EGF (Invitrogen, # PHG0311), 100 ng/ml human noggin (Peprotech, # 120-10C), 100 ng/ml human FGF-10 (Peprotech, # 100-26-25), 1.25 mM N-acetyl-L-cysteine, 10 mM nicotinamide, 2 μM TGF-β R kinase Inhibitor IV, 10 μM ROCK inhibitor (Y-27632), 10 μM forskolin (Sigma, F6886) and 1% penicillin/streptomycin). For infection experiments, hCEctos were subjected to differential trypsinization to separate fibroblasts from epithelial cells and epithelial cells were seeded on a plastic dish coated with collagen (1:100 in 1x PBS for 1 h at 37°C).

### Three dimensional air-liquid interface cultures of human derived ectocervix

Air-liquid interface (ALI) cultures were established using trans-well organotypic inserts (Merck, # PICM03050). A bovine skin collagen (Sigma # C4243) bed containing 3T3-J2 mouse fibroblasts was plated onto the trans-well insert. Once the collagen solidified, hCEctos were seeded on top. Cells were allowed to grow immersed in ectocervical primary cell culture medium for three days, then the medium on top of the insert was removed to establish an air-liquid interface. The cultures were allowed to grow into 3D multi-layered stratified epithelium for 16 days before infection experiments were initiated.

## METHOD DETAILS

### Infectivity assays

End1/E6E7 cells in six-well plates were infected with *Ctr* for 48 h and then scraped and collected in 15 mL tubes containing sterile glass beads and lysed by vortexing. Dilutions of lysates were transferred to HeLa cells (ATCC, CCL-2) and incubated for 24 h at 35°C and 5% $CO_2$. The cells were fixed in ice-cold methanol overnight at 4°C and immunostained with *Ctr*-major outer membrane protein

(*Ctr* -MOMP) specific antibody and Cy3 labeled secondary antibody. Host cell nuclei were stained with Hoechst. The number and size of chlamydial inclusions and host cells were analyzed with an automated microscope (Olympus Soft Imaging Solutions) as previously described (Gurumurthy et al., 2010). Briefly, images were acquired with DAPI (4′,6-diamidino-2-phenylindole) and Cy3 filter sets (AHF-Analysetechnik) at the same position. Host nuclei positive for Hoechst and inclusions positive for Cy3 were automatically identified and number and size quantified using ScanR Analysis Software (Olympus Soft Imaging Solutions).

### Invasion assay

Uninfected or *Ctr* persistently infected hCEcto, End1/E6E7 and HT1080 cells were detached, counted and added in the upper compartment of a 24-well transwell chamber pre-coated with Matrigel matrix (extracellular matrix) in serum-free medium (SFM). The growth medium supplemented with 10% FCS, pyruvate and glutamine was placed in the bottom compartment of 24 well as a chemoattractant. Cells were allowed to invade for 24 h at 37°C through extracellular matrix into 8 μm polyethylene terephthalate (PET) track-etched membrane (Corning, Cat# 353097). After incubation, cells in the transwell chamber were fixed in 3.7% paraformaldehyde (PFA), followed by cell permeabilization in 100% methanol and stained with 0.2% crystal violet. The microscopic images were taken before and after swab of Matrigel and data was processed using Adobe Illustrator.

### SDS-PAGE and western blotting

Cells grown in six-well plates and treated as per experimental requirement were washed with PBS and lysed with 300 μl of SDS sample buffer (3% 2-mercaptoethanol, 20% glycerine, 0.05% bromophenol blue, 3% SDS). Cell lysates were collected and boiled at 95°C with 1000 rpm shaking for 7 minutes. Samples were stored at −20°C until required. SDS-PAGE and western blotting were performed as described earlier (Gurumurthy et al., 2010). Briefly, proteins from the cell lysates were resolved by SDS–polyacrylamide gel electrophoresis (SDS-PAGE), transferred to polyvinylidene difluoride (PVDF) membranes (PerkinElmer Life Sciences), and blocked with 3% milk powder in Tris-buffered saline (containing 0.5% Tween 20) for 30 min before incubation with the appropriate antibodies. The bound primary antibodies were incubated with the corresponding HRP-conjugated secondary antibodies. Immunoreactive proteins were detected on an X-ray film directly after addition of ECL reagent (Amersham Biosciences).

### siRNA transfection and knockdown analysis

All siRNAs used in this study were purchased from QIAGEN. siRNA transfections were carried out as described previously (Gurumurthy et al., 2010). Briefly, $1 \times 10^5$ cells were seeded into each well of a 12-well plate 24 h before transfection. Cells were then transfected with Hiperfect transfection reagent according to the manufacturer's guidelines. In brief, 1.5 μl of specific siRNA (stock concentration 20 μM) was added to RPMI without serum and incubated with 9 μL of Hiperfect in a total volume of 100 μl. After 10 to 15 min, the liposome-siRNA mixture was added to the cells with 1 mL of cell culture medium, which gave a final concentration of siRNA of 25 nM. After 1 day, cells were trypsinized and seeded into new cell culture plates, depending on the experiments. Three days after transfection, the cells were used for different experiments or to determine knockdown efficiency by RT-qPCR.

### CRISPR/Cas9 Knockout Cell Line Generation

ETS1-targeting guide RNAs (gRNAs) were designed using the CHOP CHOP tool (http://chopchop.cbu.uib.no/) and cloned into the pL-CRISPR.EFS.GFP plasmid (a gift from Benjamin Ebert; Addgene plasmid # 57818 (Heckl et al., 2014)) after digesting the vector with Esp3I restriction enzyme. HEK293T cells (ATCC CRL3216; RRID: CVCL_0063) were transfected with gRNAs containing pL-CRISPR.EFS.GFP plasmids together with packaging vectors in order to produce lentiviruses for transduction of End1 E6/E7. Briefly, HEK293T cells were grown in 10 cm plates until 60%–70% confluent and transfected with lentiviral constructs containing gRNA and lentiviral packaging plasmids (psPax2 and VSVG). The lentiviral vectors were dissolved in Opti-MEM medium together with Fugene 6 transfection reagent and packaging plasmids psPax2 and pMD.2G (VSVG) and incubated for 20-30 min at RT. After incubation, the liposomes formed were added to the cells in growth medium. Next day, the medium was replaced and left for another 24 h at 37 °C in 5% $CO_2$. Two days post-transfection lentiviral particles present in the medium were harvested, filtered (0.45 mm) and used for End1 E6/E7 cell transduction. End1 E6/E7 cells were seeded in 10 cm plates one day before lentiviral particles were ready for use. At 30%–40% confluence lentiviral particles were added onto cells together with 8 μl of polybrene (10 mg/ml), followed by medium exchange after overnight incubation. After four days of lentivirus transduction, GFP-positive cells were FACS-sorted and seeded as single cells into 96-well plates. Single cell clones of transduced cells were expanded, checked for mycoplasma and used for further experiments.

### Generation of the ERF overexpression cell lines

To generate ERF wild-type and mutants overexpression End1/E6E7 cell lines, constructs from plasmids provided by Prof. Mavrothalasitis were recloned into pLenti-CMV-GFP destination vector using gateway recombination system. After plasmid confirmation using enzymatic test digestion and Sanger sequencing obtained plasmids were used for lentiviral particles generation and transduction of End1/E6E7 cell line as described above, followed by FACS sorting of GFP positive cells. Finally, cells were expanded, checked for mycoplasma and used for further experiments.

### Cellular fractionation

Cells were trypsinized and washed twice with 1xPBS prior to cellular fractionation using NE-PER kit from Thermo Fisher according to the manufacturer's protocol.

### SILAC labeling and labeling efficiency

End1/E6E7 cells were labeled by culturing for 8 passages in SILAC DMEM medium (GIBCO) containing either unlabeled (L) or labeled (H) $^{13}C_6{}^{15}N_2$ L-Lysine/ $^{13}C_6{}^{15}N_4$ L-Arginine (Sigma) additionally supplemented with 10% dialyzed FCS (dFCS), 5 nM L-glutamine and 1 mM sodium pyruvate. Labeled and unlabeled cell populations were subjected to labeling efficiency test. For this, a small amount of each cell population was lysed with Laemmli buffer, separated on SDS-PAGE gel, gel bands containing proteins were excised and tryptic digested into peptides. Finally, digested peptides were desalted and applied for MALDI-TOF analysis to estimate the extent of isotope-labeled amino acids incorporation.

### Sample preparation for mass spectrometric (MS) analysis

Proteins were reduced with 10 mM DTT at room temperature for 30 min and alkylated with 50 mM iodoacetamide at room temperature for 30 min in a dark room. Proteins were first digested by lysyl endopeptidase (LysC) at a LysC-to-protein ratio of 100:1 (w/w) for 3 h at room temperature. Then, the sample solution was diluted to a final concentration of 2 M urea with 50 mM ammonium bicarbonate. Trypsin digestion was performed at a trypsin-to-protein ratio of 100:1 (w/w) under constant agitation at room temperature for 16 h. Enzyme activity was quenched by acidification of the samples with trifluoroacetic acid (TFA). The peptides were desalted with C18 Stage Tips (Rappsilber et al., 2003) prior to nanoLC-MS/MS analysis.

### Phosphopeptide enrichment

The tryptic digests corresponding to 300 μg protein were desalted with big C18 Stage Tips packed with 10 mg of ReproSil-Pur 120 C18-AQ 5-μm resin (Dr Maisch GmbH). Peptides were eluted with 300 μL of loading buffer (80% ACN (vol/vol) and 6% TFA (vol/vol) so that the concentration of peptide was 1 μg/μl. Phosphopeptides were enriched using a microcolumn tip packed with 0.5 mg of $TiO_2$. The $TiO_2$ tips were equilibrated with 20 μL of the loading buffer via centrifugation of 100 *g*. 6 × 50 μL of the sample was loaded on a $TiO_2$ tip via centrifugation of 100 *g*. The $TiO_2$ column was washed with 20 μL of the loading buffer, followed by 20 μL of washing buffer (50% ACN (vol/vol) and 0.1% TFA (vol/vol)). The bound phosphopeptides were eluted using successive elution with 30 μL of elution buffer 1 (5% ammonia solution) and 30 μL of elution buffer 2 (5% piperidine) in series. Each fraction was collected into a fresh tube containing 30 μL of 20% formic acid. 3 μL of 100% formic acid was added for further acidification of the samples. The phosphopeptides were desalted with C18 Stage Tips prior to nanoLC-MS/MS analysis.

### NanoLC-MS/MS analysis

Peptides were separated on a 2 m monolithic column MonoCap C18 High Resolution 2000 (GL Sciences), 100 mm i.d. x 2,000 mm at a flow rate of 300 nl/min on an EASY-nLC II system (Thermo Fisher Scientific) by altering the gradient: 5%–6% B in 2 min, 6%–8% B in 28 min, 8%–30% B in 180 min, 30%–45% in 78 min, 45%–60% B in 2 min, 60%–95% B in 1 min. 360-min and 240-min gradient were performed for whole proteome and phosphoproteome analyses, respectively. A Q Exactive plus instrument (Thermo Fisher Scientific) was operated in the data-dependent mode with a full scan in the Orbitrap followed by top 10 MS/MS scans using higher-energy collision dissociation (HCD). For standard proteome analyses, the full scans were performed with a resolution of 70,000, a target value of $3x10^6$ ions and a maximum injection time of 20 ms. The MS/MS scans were performed with a 17,500 resolution, a $1x10^6$ target value and a 20 ms maximum injection time. For phosphoproteome analyses, the full scans were performed with a resolution of 70,000, a target value of $3x10^6$ ions and a maximum injection time of 120 ms. The MS/MS scans were performed with a 35,000 resolution, a $5x10^5$ target value and a 160 ms maximum injection time. Isolation window was set to 2 and normalized collision energy was 26.

### Microarray analysis

Microarray experiments were performed as independent dual-color dye-reversal color-swap hybridizations. Total RNA was isolated with Trizol according to the supplier's protocol using glycogen as co-precipitant. Quality control and quantification of total RNA was assessed using an Agilent 2100 Bioanalyzer (Agilent) and a NanoDrop (Kisker) 1000 UV-Vis spectrophotometer according to the supplier's protocol. RNA labeling was performed with the dual-color Quick-Amp Labeling Kit (Agilent Technologies). In brief, mRNA was reverse transcribed and amplified using an oligo-dTT7 promoter primer, and resulting cRNA was labeled with Cyanine 3-CTP or Cyanine 5-CTP. After precipitation, purification, and quantification, 1.25 μg of each labeled cRNA was fragmented and hybridized to whole genome human 4 × 44k multipack microarrays according to the supplier's protocol (Agilent Technologies). Scanning of microarrays was performed with 5 μm resolution using a G2565CA high-resolution laser microarray scanner (Agilent Technologies) with XDR extended range. Microarray image data were analyzed with the Image Analysis/Feature Extraction software G2567AA v. A.11.5.1.1 (Agilent Technologies) using default settings. The extracted MAGE-ML files were analyzed further with the Rosetta Resolver Biosoftware, Build 7.2.2 SP1.31 (Rosetta Biosoftware). Ratio profiles comprising single hybridizations were combined in an error-weighted fashion to create ratio experiments. A 0.5 log2 fold change expression cut-off for ratio experiments was applied together with anti-correlation of ratio profiles, rendering the microarray analysis highly significant ($p < 0.05$). In addition, microarray data was analyzed using the R package LIMMA (Ritchie et al., 2015). Microarray data have been deposited in the Gene Expression

Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) of the National Center for Biotechnology Information and can be accessed with the GEO accession number GSE104166.

### Immunofluorescent histochemistry

3D-Air Liquid cultures were fixed with 3.7% paraformaldehyde for 1 h at room temperature (RT) followed by washing with PBS twice, embedded orthogonally in Histogel (HG-4000-144) inside a casting mold. Human tissues were extensively washed with PBS and fixed using 3.7% PFA overnight at RT. Samples were subjected to dehydration in an ascending ethanol series followed by isopropanol and xylene (60 mins each) followed by paraffinization using a Leica TP1020 tissue processor. The paraffin blocks were generated inside a casting mold on a Paraffin console (Microm) and 5 μM sections made using a microtome (Microm). For immunostaining, paraffin sections were deparaffinized and rehydrated, followed by antigen retrieval using antigen retrieval solution (Dako, # S1699). Sections were blocked using blocking buffer (1% BSA and 2% FCS in PBS) for 1 h at RT. Primary antibodies were diluted in blocking buffer and incubated for 90 mins at RT followed by five PSB washes before 1 h incubation with secondary antibodies diluted in blocking buffer along with Hoechst or Draq5. Sections were washed with PBS five times and mounted using Mowiol. Images were acquired with a Leica TCS SP8 confocal microscope.

Epithelial cells grown on coverslips were fixed with 3.7% paraformaldehyde for 30 min at RT. Cells were permeabilized and blocked with 0.5% Triton X-100 and 1% BSA in PBS. Primary antibodies were diluted in 1% BSA in PBS and incubated for 1 h at RT followed by three washes in PSBT (0.1% Tween 20 in PBS), followed by 1 h incubation with secondary antibodies and phalloidin were diluted in 1% BSA in PBS along with Hoechst or Draq5. Coverslips were washed three times with PBST and once with PBS and mounted using Mowiol. Images were acquired on a Leica TCS SP8 confocal microscope. Images were processed with Adobe Photoshop.

### Automated microscopy

Images were analyzed by automated microscope from Olympus Biosystems. For each well, six positions were taken and fluorochromes visualized using Cy3 and DAPI filters. The *Ctr*-MOMP image was analyzed for *Ctr* inclusion number and inclusion size. DAPI was used to detect number of nuclei. All data was automatically identified and calculated by Scan R analysis software from Olympus Biosystems, which was further processed in Microsoft Excel 2010.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Phosphoproteome/Proteome data analysis

Raw data were analyzed and processed using MaxQuant (v1.5.1.2). Search parameters included two missed cleavage sites, fixed cysteine carbamidomethyl modification, and variable modifications including L-[$^{13}C_6$,$^{15}N_4$]-arginine, L-[$^{13}C_6$,$^{15}N_2$]-lysine, methionine oxidation, N-terminal protein acetylation, and asparagine/glutamine deamidation. In addition, phosphorylation of serine, threonine, and tyrosine was searched as variable modifications for phosphoproteome analysis. The peptide mass tolerance was 6 ppm for MS scans and 20 ppm for MS/MS scans. The match between runs was enabled. Database search was performed using Andromeda against UniProt human database (October 2014) and Chlamydia database (February 2015) with common contaminants. False discovery rate (FDR) was set to 1% at both peptide and protein level. For protein quantification, a minimum of two ratio counts was set and the 're-quantify' and 'match between runs' functions were enabled. Proteome data are available via ProteomeXchange with identifier PXD011960.

### Linear signature motif analysis

Phosphopeptide sequences with at least $\geq$ 2 fold change were submitted to Motif-X (Schwartz and Gygi, 2005) online tool analysis for the identification of over-represented linear signature motifs to predict upstream kinase regulators. The significance threshold was set to $p < 10^{-6}$.

### iGPS analysis – prediction of site-specific kinase-substrate relationship

To predict kinase-substrate relationships for all the upregulated and downregulated phosphosites based on short linear motifs and protein-protein interactions iGPS software v1.0.1 analysis was performed (Song et al., 2012). For each site in a protein, all predicted kinases get assigned a weight of 1/(number of predictions for this site and protein). Finally, weights are summed up for each kinase included in the predictions and used to define the size of the circle in the kinome tree that was generated using KinMap online tool (Eid et al., 2017).

### GO enrichment analysis

GO enrichment analysis was performed with DAVID (Huang et al., 2009) online tool for 2 fold up and downregulated phosphosites from nuclear fraction and total cell extract. The top five candidates were selected and combined in heatmap using TM4 (http://mev.tm4.org).

### STRING protein-protein interaction analysis

The protein-protein interaction analyses of 2-fold regulated phosphosites were visualized using STRING 10 (Szklarczyk et al., 2015) database with standard settings in confidence view. The interaction network was imported to Adobe Illustrator and modified as shown in the Results section.

### Ingenuity pathway analysis (IPA)

Canonical pathways and biological function of the significantly dysregulated genes and proteins identified in the microarrays and phosphoproteome were investigated using QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, https://www.qiagen.com/ingenuity). Overrepresentation of canonical pathways was obtained by Fisher's exact test and corrected for multiple testing by the Benjamini-Hochberg procedure. The ratio value is calculated based on the number of genes from the dataset that map to the pathway divided by the number of total genes included in the pathway. Moreover, downstream target genes analysis was performed to find genes regulated by selected hits from phosphoproteome analysis using microarray data as a reference. The downstream effects analysis is based on prior knowledge of expected causal effects between genes and biological functions stored in the Ingenuity® Knowledge Base. The analysis examines genes in the user's dataset that are known to affect each biological function and compares their direction of change to what is expected from the literature (https://www.ingenuity.com/).

### GSEA

A published gene set of 365 EMT-associated genes (Gröger et al., 2012) was used to perform GSEA on genes pre-ranked by gene expression-based t-score comparing 48 h infected and non-infected End1 cells, using the fgsea R package (Sergushichev, 2016) with 5,000 permutations.

### Statistics

Results are presented as either mean ± SEM (for normally distributed data) or median with inter-quartile range (for non-normally distributed data). Datasets were compared by unpaired t test, nonparametric Mann–Whitney test or ANOVA. GraphPad Prism was used for statistical tests and plots. P value $\leq 0.05$ was considered statistically significant unless otherwise specificed. Details of tests used can be found in the figure legends. Log2 fold-changes of proteome and phosphoproteome data from two replicates were tested for significance using the R package LIMMA (Ritchie et al., 2015). Full details on statistical analysis for proteome/phosphoproteome and gene expression data performed in R are provided in corresponding scripts (see Data and Software Availability).

## DATA AND SOFTWARE AVAILABILITY

Microarray data presented in this paper have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE104166. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al., 2016) partner repository with the dataset identifier PXD011960. R code used for pre-processing and analysis of data and generation of plots has been deposited under https://github.com/HilmarBerger/Zadora_et_al_Phosphoproteome

**Cell Reports, Volume *26***

## Supplemental Information

# Integrated Phosphoproteome and Transcriptome Analysis Reveals *Chlamydia*-Induced Epithelial-to-Mesenchymal Transition in Host Cells

**Piotr K. Zadora, Cindrilla Chumduri, Koshi Imami, Hilmar Berger, Yang Mi, Matthias Selbach, Thomas F. Meyer, and Rajendra Kumar Gurumurthy**

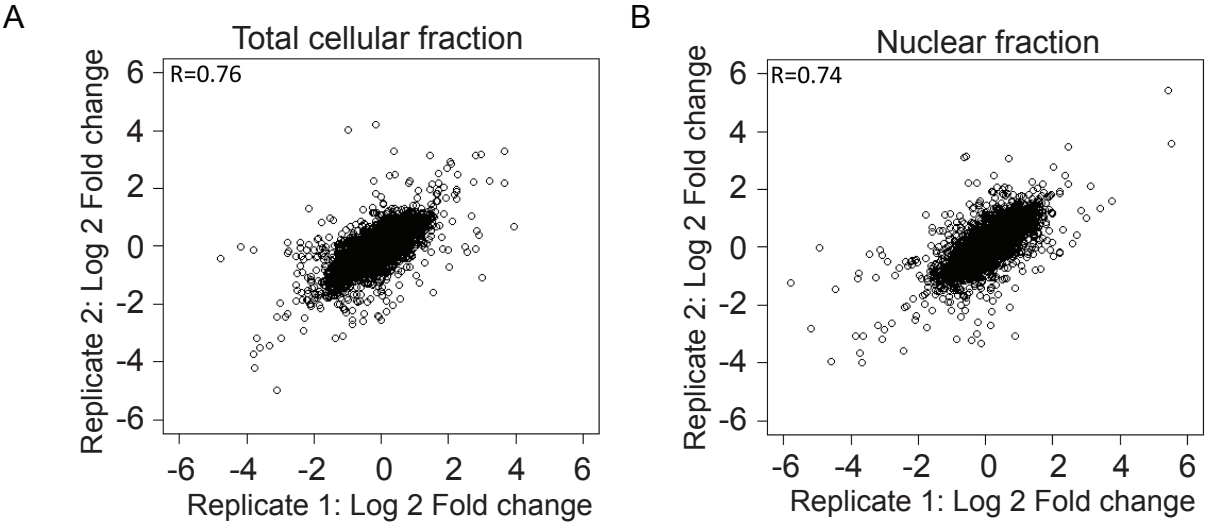
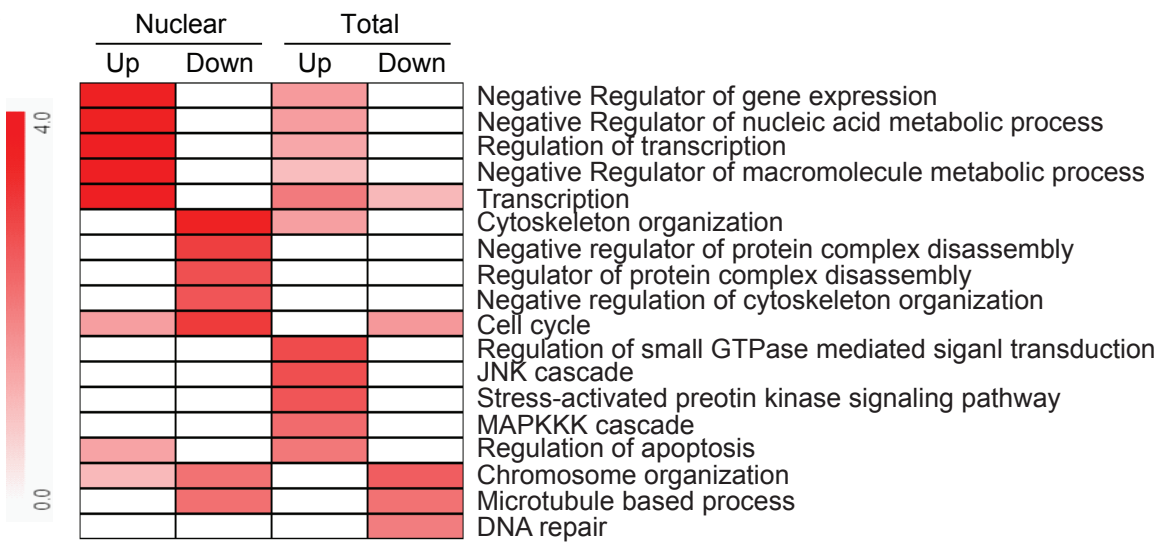



**Figure S1. SILAC labeling efficiency and global phosphoproteome correlation analysis. Related to Figure 1.** (A-B) Pearson's correlation coefficient analysis of log2-transformed phosphopeptide fold changes of two biological replicates from (A) total cellular fraction and (B) nuclear fraction shows high reproducibility.

A




B

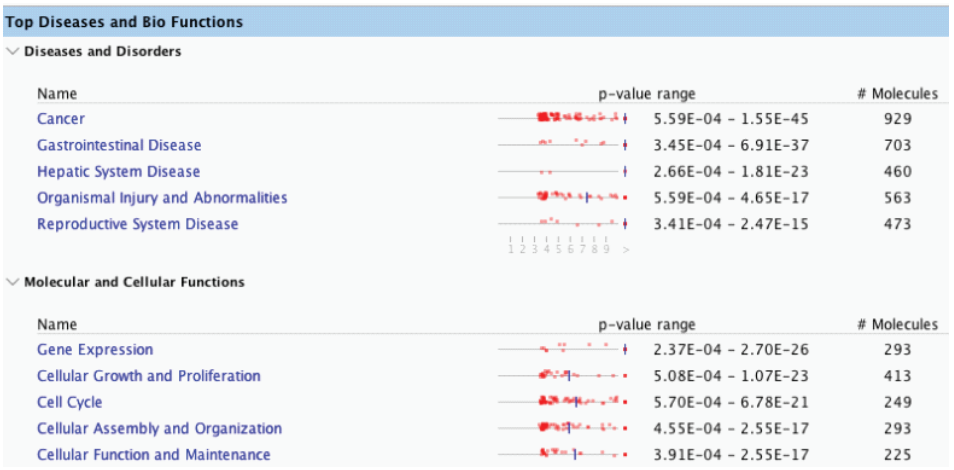



**Figure S2. Functional analysis of *Ctr*-regulated global host phosphoproteome hits. Related to Figure 2.** (A) Shown are the top five biological processes from Gene Ontology (GO) term analysis for up- and downregulated phosphoproteins upon *Ctr* infection with $p < 0.01$. (B) Ingenuity pathway analyses for diseases and disorders as well as molecular and cellular functions enriched among phosphoproteins regulated upon *Ctr* infection. The five top-scoring hits are presented in each category, with significance scores (P-values) and the number of molecules in each class.

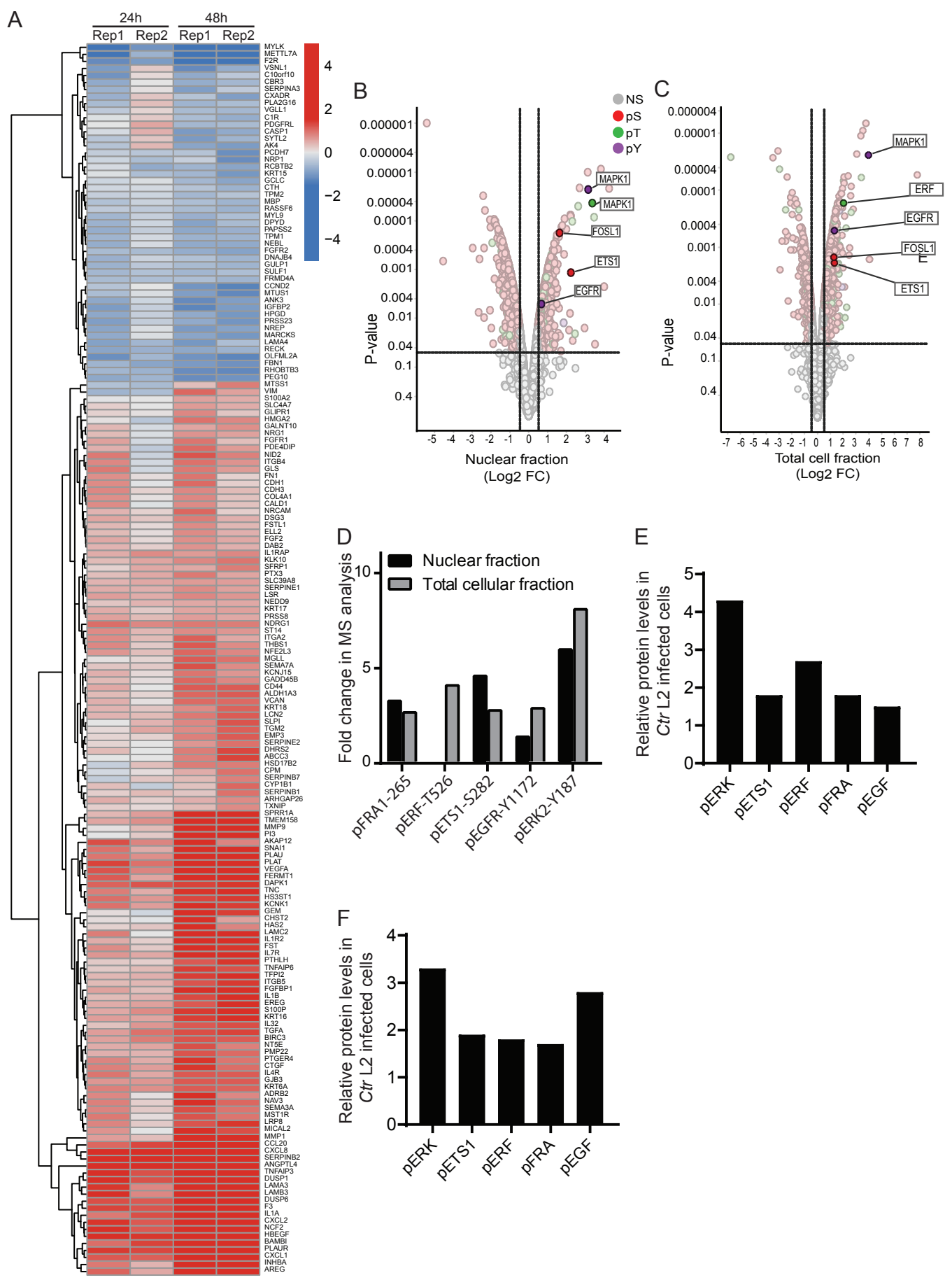



**Figure S3. Related to Figure 3.** (A) Heat map of 365 EMT-associated genes (Gröger et al., 2012) that are differentially regulated in *Ctr*-infected compared to uninfected cells at 24 and 48 h p.i. (B-C) Volcano plot representing global phosphoproteome data set from (B) nuclear fraction and (C) total cell extract according to log2 fold change (FC) (x-axis) and p-value (y-axis). The cut-off of +/- 0.5 log2 FC (dashed vertical lines) and 0.05 p-value (dashed horizontal line) was applied. Phospho hits are highlighted depending on phosphorylated residues colored as labelled. The hits selected for validation are highlighted and labelled. (D) Shown is the fold-change in selected phosphosites after *Ctr* infection from the global phosphoproteome data. (E-F) Densitometric values of immunoblots from End1/E6E7 (E) and hCEcto (F) cells.

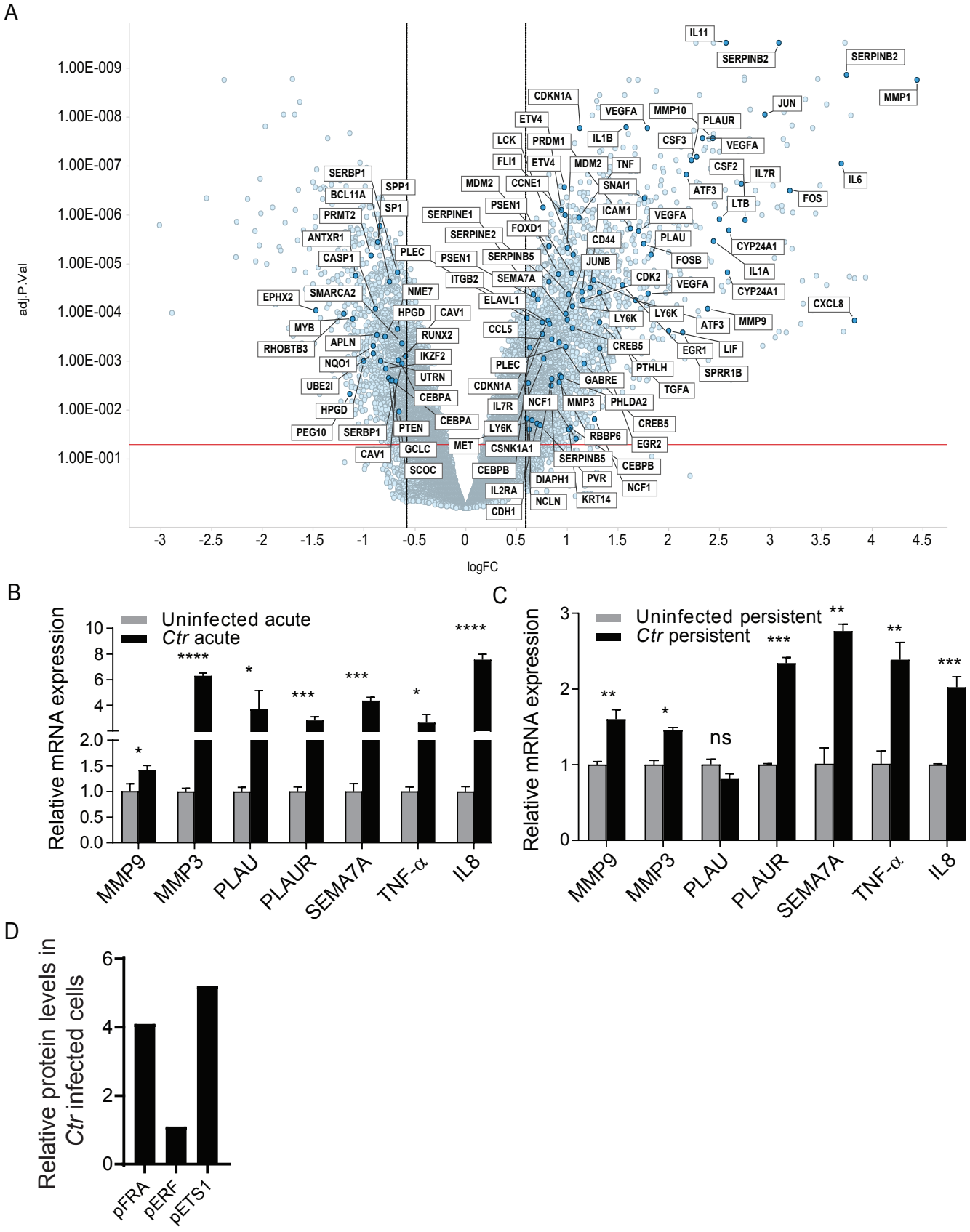



**Figure S4. Transcriptional regulation of epithelial to mesenchymal transition-associated genes upon *Ctr* infection. Related to Figure 4.** (A) Volcano plot highlighting Ctr-regulated genes that are known to be downstream of ERF, ETS1 and FRA1. (B-C) End1/E6E7 cells were (B) acutely or (C) persistently infected with *Ctr* for 32 h and 8 d, respectively. Shown is the relative mRNA expression of selected ERF, ETS1 and FRA1 target genes analyzed by qRT-PCR. Data shown as mean ± SD from three technical replicates. ****p < 0.0001, ***p < 0.001, **p < 0.01 *p < 0.05, p > 0.05, ns - not significant Student's t-test. (D) Densitometric values of immunoblots shown in Figure 4E.