Targeting of RNA Polymerase II by a nuclear *Legionella pneumophila* Dot/Icm effector SnpL

Ralf Schuelein[1^], Hugh Spencer[1], Laura F. Dagley[2], Peng fei Li[1,3], Lin Luo[4], Jennifer L. Stow[4], Gilu Abraham[5], Thomas Naderer[5], Laura Gomez-Valero[6], Carmen Buchrieser[6], Chihiro Sugimoto[7,8], Junya Yamagishi[7,8], Andrew I. Webb[2], Shivani Pasricha[1,3] and Elizabeth L. Hartland[1,3,6,9*]

[1]Department of Microbiology and Immunology, University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Melbourne 3000, Australia

[2]The Walter and Eliza Hall Institute of Medical Research and Department of Medical Biology, University of Melbourne, Parkville 3052, Australia

[3]Centre for Innate Immunity and Infectious Diseases, Hudson Institute of Medical Research, Clayton 3168, Australia

[4]Institute for Molecular Bioscience, The University of Queensland, Brisbane 4072, Australia

[5]Biomedicine Discovery Institute and Department of Biochemistry & Molecular Biology, Monash University, Clayton, 3800, Australia

[6]Institut Pasteur, Biologie des Bactéries Intracellulaires and [2]CNRS UMR 3525, 75724, Paris, France

[7]Global Station for Zoonosis Control, GI-CoRE, Hokkaido University, North 20, West 10 Kita-ku, Sapporo, Hokkaido 001-0020, Japan

[8]Research Center for Zoonosis Control, Hokkaido University, North 20, West 10 Kita-ku, Sapporo, Hokkaido 001-0020, Japan

[9]Department of Molecular and Translational Science, Monash University, Clayton 3168, Australia

Running title: *Legionella* effector SnpL targets host gene transcription

*Correspondence: Prof Elizabeth Hartland, Tel: +61 3 8572 2800, Email: elizabeth.hartland@hudson.org.au

[^]Current address; Department of Microbiology, Monash University, Clayton 3168, Australia

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cmi.12852

SUMMARY

The intracellular pathogen *Legionella pneumophila* influences numerous eukaryotic cellular processes through the Dot/Icm-dependent translocation of more than 300 effector proteins into the host cell. Although many translocated effectors localize to the *Legionella* replicative vacuole, other effectors can affect remote intracellular sites. Following infection, a subset of effector proteins localizes to the nucleus where they subvert host cell transcriptional responses to infection. Here we identified Lpg2519 (Lpp2587/Lpw27461), as a new nuclear-localized effector that we have termed SnpL. Upon ectopic expression or during *L. pneumophila* infection, SnpL showed strong nuclear localization by immunofluorescence microscopy but was excluded from nucleoli. Using immunoprecipitation and mass spectrometry, we determined the host-binding partner of SnpL as the eukaryotic transcription elongation factor, SUPT5H/Spt5. SUPT5H is an evolutionarily conserved component of the DRB sensitivity-inducing factor complex (DSIF complex) that regulates RNA polymerase II (Pol II) dependent mRNA processing and transcription elongation. Protein interaction studies showed that SnpL bound to the central KOW motif region of SUPT5H. Ectopic expression of SnpL led to massive upregulation of host gene expression and macrophage cell death. The activity of SnpL further highlights the ability of *L. pneumophila* to control fundamental eukaryotic processes such as transcription that, in the case of SnpL, leads to global upregulation of host gene expression.

INTRODUCTION

Legionella pneumophila is an environmental pathogen of soil and water amoebae and an opportunistic pathogen of humans. Following inhalation in contaminated aerosols, *L. pneumophila* is internalized by alveolar macrophages in which the bacteria replicate within a specialised intracellular vacuole termed the *Legionella* containing vacuole (LCV). LCV biogenesis is driven by the bacterial type IV Dot/Icm (Defective in organelle trafficking/Intracellular multiplication) secretion system that translocates more than 300 effector proteins into the host cell (Burstein *et al.*, 2009, Huang *et al.*, 2011, Zhu *et al.*, 2011). Many Dot/Icm effectors are localized to the LCV and play various roles in LCV ubiquitination, host vesicle and membrane tethering and fusion and phosphoinositide dynamics (Hubber *et al.*, 2010, Finsel *et al.*, 2015, Qiu *et al.*, 2017). Other Dot/Icm effectors act in remote parts of the cell to influence a range of fundamental cellular processes that also presumably aid LCV stability and bacterial replication. These include Dot/Icm effector proteins that specifically traffic to organelles such as mitochondria and the nucleus (Dolezal *et al.*, 2012, Rolando *et al.*, 2013).

Several Dot/Icm effectors interfere with host gene transcription and protein translation, thereby shaping a unique transcriptional response in infected cells. For example, the glucosyltransferases, Lgt1, Lgt2, and Lgt3 modify a conserved serine residue (Ser^{53}) in elongation factor 1A (eEF1A) and thereby halt protein synthesis (Belyi *et al.* 2013). SidI also targets eEF1A as well as eEF1Bγ (Shen *et al.*, 2009) and SidL inhibits protein translation through an unknown mechanism (Fontana *et al.*, 2011). Together with two other effectors, these five effectors trigger an innate immune signature in infected murine macrophages characterized by the expression of *Il1α, Il1β, Tnf, Il23a, Csf1* and *Csf2* and genes for mitogen activated kinases (MAPK) (Fontana *et al.*, 2011, Fontana *et al.*, 2012, Barry *et al.*,2013,

Barry *et al.*, 2017). However, this response is unproductive for infected cells as only selected cytokines such as IL1α and IL1β escape the inhibition of protein synthesis by the same effectors (Barry *et al.,* 2013, Asrat *et al.*, 2014, Copenhaver *et al.*, 2015, Barry *et al.*, 2017). Other Dot/Icm effector proteins directly manipulate the transcriptional machinery of the host cell. The SET domain effector, RomA/LegAS4, is a histone lysine methyltransferase that modifies the epigenetic landscape of the infected cell with two alternative effects on transcription reported. In one report, RomA/LegAS4 localized to the nucleolus of the cell and activated ribosomal RNA gene transcription (rDNA) through interaction with heterochromatin protein 1 and modification of histone H3 predominantly as H3K4me2 at the rDNA promoter (Li *et al.*, 2013). However, another study reported that RomA/LegAS4 mediated a novel histone mark of H3K14me3 thereby preventing acetylation at that site. Since H3K14 acetylation is an activating mark, RomA/LegAS4 expression resulted in global gene repression, including genes involved in the innate immune response (Rolando *et al.*, 2013, Rolando *et al.* 2014).

In this study, we describe a previously predicted but uncharacterized Dot/Icm effector protein, SnpL (Lpw27461/Lpg2519/Lpp2587) that localizes to the cell nucleus and binds to SUPT5H. SUPT5H is a component of the DRB (5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF) complex that regulates promoter proximal pausing of RNA polymerase II (Pol II) throughout the genome but particularly on highly expressed genes and genes involved in stimulus-response pathways (Jonkers *et al.*, 2015). During the initial stages of mRNA elongation, Pol II can pause and accumulate to high levels at sites 30-60 nucleotides downstream of the transcription start site especially under conditions of nutrient limitation (Adelman *et al.*, 2012, Kwak *et al.*, 2013). The release of paused Pol II is determined by the positive transcription elongation factor-b

(P-TEFb), which phosphorylates the carboxy-terminal domain (CTD) of Pol II as well as negative elongation factor (NELF) and DSIF leading to the ejection of NELF and the conversion of DSIF from a negative to positive transcription elongation factor (Jonkers *et al.*, 2015). A high degree of pausing and regulated release of Pol II is believed to aid rapid and synchronous gene expression (Shao *et al.*, 2017).

The DSIF complex in mammalian cells comprises two proteins, SUPT4H and SUPT5H, which interact with Pol II (Hartzog *et al.*, 2013). SUPT5H is a member of the NusG family of proteins, which is conserved across the three domains of life (Hartzog *et al.*, 2013), and is a large multidomain protein comprising an N-terminal acidic domain, a NusG N-terminal (NGN) domain, multiple Kyprides, Ouzounis, Woese (KOW) domains and a series of short C-terminal repeats (CTR) (Ponting, 2002). SUPT4H on the other hand is a small zinc finger domain protein that is only (Malone *et al.*, 1993) found in archaea and eukaryotes (Ponting, 2002, Hartzog *et al.*, 2013). In yeast, the equivalent proteins Spt4 and Spt5 also form a heterodimeric complex and are highly conserved in structure and function to human SUPT4H and SUPT5H respectively and likely to be broadly conserved across all lower eukaryotes (Hartzog *et al.*, 2013). Here we examined the transcriptional response to SnpL in mammalian cells, which suggested that *L. pneumophila* targets host SUPT5H to drive global activation of gene expression.

RESULTS

The conserved Dot/Icm effector, Lpw27461/SnpL, localizes to the host nucleus

Dot/Icm effector protein-encoding genes are typically found clustered within the *L. pneumophila* genome. Here three putative effector proteins (Lpw27441, Lpw27451, Lpw27461), adjacent to one another in the genome of *L. pneumophila* strain 130b, were

examined for both their conservation among other *L. pneumophila* strains, and their ability to translocate into host cells upon infection. Pair-wise BlastP analysis (NCBI) revealed that Lpw27441, Lpw27451 and Lpw27461 shared more than 70% amino acid sequence identity with homologous proteins from *L. pneumophila* strains Philadelphia 1, Paris, Corby, and Lens. Comparative analysis of proteins from *L. pneumophila* strain 130b with *L. longbeachae* strain NSW150 revealed lower amino acid conservation among homologues, with the exception of Lpw27441 which shared 78% amino acid sequence identity with Llo2410 (Fig. 1A). Interestingly the corresponding genes in *L. longbeachae* are not contiguous (Fig. 1A).

To determine whether Lpw27441, Lpw27451, Lpw27461 were translocated into host cells during *L. pneumophila* infection, a fluorescent based TEM beta-lactamase (BlaM) translocation assay was performed (de Felipe *et al.*, 2008). Genes of interest were cloned into pXDC61 to generate plasmids encoding N-terminal TEM1 BlaM translational fusions (pTEM1-Lpw27441, pTEM1-Lpw27451 and pTEM1-Lpw27461) and transformed into wild type *L. pneumophila* 130b and an isogenic Δ*dotA* mutant. *L. pneumophila* 130b carrying pTEM1-RalF was included as a positive control and Δ*dotA* pTEM1-RalF as negative a control (Table 1). J77A4.1 murine macrophages were infected with the above strains and fluorescence was measured at 450 nm and 520 nm. Background fluorescence was subtracted using a media-only sample and the ratio of blue/green fluorescence was normalized to *L. pneumophila* 130b carrying pXDC61 only. The resulting ratio indicated that Lpw27451 and Lpw27461 were translocated into host cells in a Dot/Icm dependent manner, while Lpw27441 was not (Fig.1B).

To determine the possible localization of Lpw27451 and Lpw27461 within the host cell after translocation, 3xFLAG-tagged Lpw27451 and Lpw27461 were expressed ectopically in

HEK293T cells via transfection. Using immunostaining with anti-FLAG antibodies, we observed that Lpw27451 staining occurred throughout the host cytoplasm, whereas Lpw27461 localized strongly to the host cell nucleus (Fig. 2A). During *L. pneumophila* 130b infection, 4xHA-tagged Lpw27461 expressed by the bacteria also showed nuclear localization following 8 h infection of iBMDMs (Fig. 2B). Interestingly, by either ectopic expression or *L. pneumophila* 130b infection, nuclear Lpw27461 was excluded from host nucleoli (Fig. 2A, B).

Bioinformatic inspection of Lpw27461 identified no clear nuclear localization signal (NLS) motifs (NLStradamus method) or conserved catalytic domains (NCBI domain finder). In order to examine the evolutionary conservation of Lpw27461, homologous sequences were recruited by BlastP searches. Homologues of Lpw27461 were found in all 300 *L. pneumophila* strains used for the search, suggesting a key role of this gene in the *L. pneumophila* species, as well as its presence in their common ancestor. Putative Lpw27461 homologues were found in a further 24 *Legionella* species. Lpw27461 sequences from these 24 *Legionella* species and from 15 selected *L. pneumophila* strains were aligned and the corresponding phylogenetic tree was obtained by likelihood. The main clades obtained correlated well with overall species relatedness (Fig. 2C), indicating that Lpw27461 has probably followed a vertical evolution. However, the degree of conservation of this protein at interspecies level ranged from moderate to low (Fig. 2C) and therefore functional divergence in some *Legionella* species cannot be discounted.

Lpw27461/SnpL associates with the host transcriptional regulator SUPT5H

To identify the host target of Lpw27461, FLAG-tagged constructs of Lpw27461 were expressed in HEK293T cells via transfection. HEK293T cells were also transfected with the

empty p3xFLAG-Myc-CMV™-24 vector for comparison. Following immunoprecipitation with anti-FLAG antibodies, samples were separated by SDS-PAGE and stained with SYPRO Ruby (ThermoFisher Scientific). A band of approximately 120-150 kDa was identified in the immunoprecipitated sample of FLAG- Lpw27461 transfected cells, along with an expected band of 26 kDa corresponding to FLAG-Lpw27461 (Fig. 3A)**.** The approx. 120-150 kDa band was excised from the gel, along with a corresponding gel excision from p3xFlag transfected cells, and analysed via mass spectrometry (LC-MS/MS). Immunoprecipitate from FLAG-Lpw27461 expressing cells returned 80 peptide hits with at least 95% confidence for the human homolog of the Suppressor of Ty5 (SUPT5H) regulator of transcription elongation**.** This was absent in cells transfected with empty p3xFLAG vector, strongly suggesting that SUPT5H was an interaction partner of Lpw27461. Given this, we termed Lpw27461, SnpL, for SUPT5H Interaction partner of *Legionella.* The SUPT5H-SnpL interaction was confirmed by LC-MS/MS analysis of the entire immunoprecipitate elution, which also revealed SUPT4H as enriched in FLAG-SnpL pull down (Fig. 3B, PSM listed in Table S1). In humans, SUPT5H associates with SUPT4H to form a dimeric complex termed the DRB-sensitivity inducing factor (DSIF) (Wada *et al.*, 1998). DSIF is implicated in the pausing of RNA polymerase II at selected promoters through its interaction with Pol II and other transcription factors (Hartzog *et al.*, 2013). Through FLAG-immunoprecipitation, we showed that FLAG-SnpL homologues of *L. pneumophila* strains Paris, Corby and Philadelphia JR32, as well as *L. longbeachae* NSW150 also interacted with SUPT5H, suggesting that this interaction is highly conserved amongst *Legionella* ssp. (Fig. 3C).

Identification of a SnpL binding region in SUPT5H

SUPT5H is a large multidomain protein comprising an N-terminal acidic domain, a NusG N-terminal (NGN) domain, multiple Kyprides, Ouzounis, Woese (KOW) domains and a series

of short C-terminal repeats (CTR) (Fig. 4A). In order to identify the SnpL binding region of SUPT5H, far western immunoblots were performed using truncated fragments of 4HA-tagged SUPT5H expressed in HEK293T cells. Cell lysates were subjected to SDS-PAGE before being transferred to PVDF membrane. GST-SnpL was then applied to the membrane, before detection with anti-GST antibody. Purified GST alone was applied as a negative control. GST-SnpL interacted full length 4HA-SUPT5H (160-170 kDa, amino acids 1-1087) and with a truncation of SUPT5H encompassing amino acids 418-1087 (4HA-SUPT5H$_{418-1087}$) (75 kDa). GST-SnpL did not appear to interact with a truncation encompassing amino acids 751-1087 (4HA-SUPT5H$_{751-1087}$), although its expression following transfection was weaker compared to other constructs. In addition, no interaction was observed with a truncation of SUPT5H encompassing the central KOW domains, 1-754 (4HA-SUPT5H$_{1-754}$). Taken together, these results suggest the central KOW domains may be important but not sufficient for SnpL binding (Fig. 4B). GST alone showed no specific interactions with any 4HA-SUPTH derivatives (Fig. 4B).

Ectopic expression of SnpL results in the upregulation of host gene expression

Given its interaction with a component of the DSIF complex, we predicted that SnpL would affect the host cell transcriptome. To elucidate the effect of SnpL on the transcriptome of mammalian cells, we constructed an iBMDM cell line stably expressing doxycycline inducible FLAG -SnpL. iBMDM_FLAG-SnpL showed strong expression of the 28 kDa FLAG-SnpL fusion protein 6 h after induction with 50 ng/mL doxycycline (Fig. 5A). RNA-Seq analysis was then performed on iBMDM left untreated or treated with 50 ng/mL doxycycline for 7 h or 12 h to induce expression of FLAG-SnpL. iBMDM cells expressing doxycycline-inducible GFP were used as an additional control. Resulting reads were aligned with the *Mus musculus* genome and the expression profiles of iBMDM expressing FLAG-

SnpL versus uninduced samples or iBMDM expressing GFP were curated using DESeq (Anders *et al.*, 2010) (Table S2-S5).

After 7 h of induction of FLAG-SnpL, 28% and 26% of expressed genes showed statistically significant differential expression compared to uninduced FLAG-SnpL iBMDM and induced GFP controls, respectively (5% FDR cutoff and adjusted *p*-value less than 0.05) (Fig. 5B). After 12 h of induction of FLAG-SnpL, almost double the number of genes showed significant differential expression compared to uninduced FLAG-SnpL iBMDM and induced GFP controls (57% and 43%, respectively) (Fig. 5B). Plotting the fold-change of iBMDM expressing FLAG-SnpL versus uninduced iBMDM and GFP expressing controls revealed that the pattern of differential expression was for the most part, consistent at 7 h and 12 h (Fig. 5C). In addition, this suggested that SnpL-induced differential regulation of gene transcription was not an artifact of ectopic gene expression. The few genes that did show inconsistent changes in expression against the two controls were omitted from further analyses.

Genes that showed a fold change in expression > 8 or < 1/8 in FLAG-SnpL induced versus uninduced FLAG-SnpL were selected for further analysis (111 genes at 7 h and 339 genes at 12 h). Inspection of this subset clearly illustrated that ectopic expression of SnpL resulted in the robust upregulation of host genes at 7 h (96.4%) rather than downregulation, with the total number of upregulated genes increased at 12 h (97.6%) (Fig. 6A). To understand the biological impact of SnpL expression, upregulated genes were analyzed using the PANTHER classification system (Mi *et al.*, 2013). The output suggested that the genes fall into broad/global biological process and molecular function groups, showing strong similarity at 7 h and 12 h (Fig. 6B and Fig. 6C). GO term enrichment analysis, using the DAVID Gene Set

Enrichment analysis tool (Huang da *et al.*, 2009) consistently resulted in the enrichment of very broad biological processes such as the immune response, cell adhesion and phosphate metabolic process (data not shown). Taken together these findings suggest that SnpL induction does not result in the targeted regulation of a particular pathway or biological process. Instead, the observed upregulation of host genes is global and genome wide.

Effect of SnpL on macrophage proliferation and bacterial replication

To examine SnpL activity further in a biological context, we used iBMDM expressing inducible FLAG-SnpL or GFP to assess the effect of SnpL expression on iBMDM proliferation, adhesion and cell viability. iBMDM numbers were quantified after 16 h of doxycycline induction and presented as a ratio of induced versus non-induced cells. A significant decrease in cell numbers (35.5%) was observed for iBMDM expressing FLAG-SnpL compared with iBMDM expressing GFP (Fig. 7A and 7B), suggesting that SnpL expression adversely affected iBMDM adhesion. To determine whether SnpL expression results in iBMDM cell death we performed lactate dehydrogenase (LDH) release assays. After 14 h of doxycycline induction, a significant increase in cytotoxicity was observed for iBMDM expressing FLAG-SnpL compared with iBMDM expressing GFP (cytotoxicity presented as a ratio of secreted LDH by induced versus non-induced cells) (Fig. 7C).

Despite the observed impact of SnpL expression on macrophage viability, SnpL was not essential for bacterial replication during iBMDM infection at 37°C, or *Acanthamoeba castellanii* (Neff) infection at 30°C and 37°C (Fig. 8). At 24, 48 and 72 h post infection no significant differences were observed between replication of the Δ*snpL* mutant and wild type *L. pneumophila* 130b. This was not overly surprising as Dot/Icm effectors are often functionally redundant (Ensminger, 2016) and this may indicate that other Dot/Icm effectors target host gene expression pathways.

DISCUSSION

The ability of *L. pneumophila* to replicate intracellularly relies on a functional Dot/Icm system and the translocation of numerous Dot/Icm effector proteins into the infected cell. Many of these localize to the LCV membrane while others traffic to different subcellular compartments. Here, we identified SnpL as a Dot/Icm effector that localizes to the host cell nucleus and induces the global activation of host gene expression. SnpL bound to the DSIF component SUPT5H, which, together with SUPT4H, regulates promoter proximal pausing of Pol II throughout the genome (Jonkers *et al.* 2015).

Consistent with global gene activation, pathway analysis of the transcriptome upon SnpL expression identified a range of cell activities that were upregulated, including fundamental biological processes such as cell division, adhesion and survival. Given this, it was not surprising that SnpL induction led to morphological changes in macrophages such as cell rounding and detachment and ultimately cell death. How this activity assists infection is unclear as a Δ*snpL* mutant replicated as efficiently as wild type *L. pneumophila* in macrophages. The role of SnpL during infection of amoebae is also unclear. Preliminary genome analysis suggests that homologues of *SUPT5H* and members of the *NELF* family are conserved in *A. castellanii* (Neff). SnpL may also drive mRNA expression in amoebae, and result in clearer biological consequences such as holding the amoebae in a cell cycle phase that supports bacterial replication, to the detriment of host cell survival. Indeed, recent reports show that *L. pneumophila* prevents cell division in *A. castellanii* in a Dot/Icm dependent manner (Mengue *et al*. 2016).

Although genetically linked, SnpL did not show the same properties as Lpw27441/Lpg2517 and Lpw27451/Lpg2518 encoded in the same cluster. Like SnpL, the hypothetical protein,

Lpw27451/Lpg2518 was translocated into cells in a Dot/Icm dependent manner, but this effector showed cytoplasmic rather than nuclear localisation. Lpw27441/Lpg2517 on the other hand was not translocated and sequence-wise aligned with the AsnC family of transcriptional regulators. Hence Lpw27441/Lpg2517 may play a role in regulating expression of the cluster, although this remains to be determined.

It would appear that the activity of SnpL antagonizes effectors such as RomA that repress host gene transcription (Rolando *et al.* 2014). Unlike RomA, no putative nuclear localization signal (NLS) sites were identified in SnpL and it is possible that SnpL either enters the nucleus via passive diffusion, due to its small size, or via an atypical NLS sequence. Undoubtedly, other Dot/Icm effectors localize to the nucleus during infection, although these remain to be identified. Until we have a complete knowledge of the nuclear Dot/Icm effectors, the full effect of *L. pneumophila* manipulation of the host cell transcriptome will not be known.

So-called "nucleomodulin" effectors have been described in other human, animal and plant pathogens, including *Shigella*, *Listeria*, *Xanthomonas* and *Agrobacterium*, showing that nuclear modulation is a generalizable virulence strategy (Bierne *et al.* 2012). These effectors variously target chromatin, transcription and cell cycle complexes. For example, the effector AnkA from the multi-host Rickettsial pathogen, *Anaplasma phagocytophilum*, drives global transcriptional changes in host gene expression by direct binding to DNA (Sinclair *et al.*, 2014). AnkA binds AT rich regions of DNA with structural rather than sequence specific qualities, and recruits histone deacetylase-1 to these site (Sinclair *et al.*, 2014). Another nuclear effector that suppresses host defence mechanisms is OspF from *Shigella* species. OspF is a phosphothreonine lyase that irreversibly modifies mitogen activated kinases

(MAPK) in the host cell nucleus to prevent phosphorylation and hence activation (Arbibe *et al.*, 2007, Li *et al.*, 2007). Other effector proteins promote gene transcription by acting as transcription factors in the host cell nucleus. The Transcription Activator-Like (TAL) proteins of *Xanthomonas* species bind consensus sequences in plant cells to activate gene expression (Kay *et al.*, 2007).

In summary, SnpL is the latest nucleomodulin effector to be described in *L. pneumophila*. SnpL appears to mediate its effect by direct binding to the highly conserved eukaryotic protein, SUPT5H, thereby subverting the activity of the DSIF complex. Although the consequences of SnpL activity are not clear, carriage of the effector among all *L. pneumophila* strains and a wide range of *Legionella* species suggests its function assists survival of the bacteria in eukaryotic hosts and thus the persistence of Legionellae in the environment.

EXPERIMENTAL PROCEDURES

Bacterial strains and growth conditions.

Bacterial strains and plasmids used in this study are listed in Table 1. *Legionella* strains were grown in ACES [N-(2-acetamido)-2-aminoethanesulfonic acid]-buffered yeast extract (AYE) broth or on buffered charcoal yeast extract (BCYE) agar (Feeley *et al.*, 1979). Where necessary, chloramphenicol, kanamycin and/or isopropyl-β-D-thiogalactopyranoside (IPTG) were added to media at 6 µg mL, 25 µg mL and 1 mM respectively. *Escherichia coli* cultures were grown in Luria-Bertani (LB) broth or agar supplemented with ampicillin at 100 µg mL and/or chloramphenicol at 12.5 µg mL where required. *L. pneumophila* grown on solid media were incubated aerobically at 37 °C for 72 h, while *E. coli* grown on solid media were

incubated aerobically at 37 °C overnight. Unless stated otherwise, broth cultures of *L. pneumophila* and *E. coli* were grown aerobically at 37 °C overnight with shaking.

Tissue culture of human and murine-derived immortalised cells

Human embryonic kidney 293T cells (HEK293T), HeLa cells, murine J774A.1 and immortalised murine macrophages derived from wild type C57BL/6 mice (iBMDM) were maintained in Dulbecco's Modified Eagle Medium (DMEM) with GlutaMAX-1 (Gibco), supplemented with 10% heat-inactivated Fetal Bovine Serum (Gibco) and grown in T75cm^2 flasks (Corning). Lentiviral stable cell lines were maintained in Roswell Park Memorial Institute 1640 (Gibco) media with 10% FBS and 3 µg /mL puromycin (Gibco). All cell lines were incubated in 5% CO_2 at 37°C and split at a cell-media ratio of 1:10 after 48-72 h growth.

Genetic manipulation

The plasmids and primers used in this study are listed in Table 1 and Table 2, respectively. Unless stated otherwise, PCR products were amplified and ligated into pGEM®-T Easy for sequence verification by Sanger sequencing before being cloned into different vector backbones. Cloning was carried out using primer incorporated restriction sites (underlined sequences in Table 2) and genomic DNA of *L. pneumophila* strain 130b as template if not indicated differently. TEM1 fusion constructs for translocation assays were made through the amplification of full-length *snpL* (*lpw27461*) (primers $SnpL_{(pTEM1)F}$ and $SnpL_{(pTEM1)R}$), *lpw27451* (primers $Lpw27451_{(pTEM1)F}$ and $Lpw27451_{(pTEM1)R}$), and *lpw27441* (primers $Lpw27441_{(pTEM1)F}$ and $Lpw27441_{(pTEM1)R}$) before cloning corresponding products into pXDC61 harbouring TEM-1. Cell transfection constructs expressing FLAG-tagged SnpL and LPW27451 were generated through amplification of *snpL* (primers $SnpL_{(pFLAG)F}$ and

SnpL$_{(pFLAG)R}$) and *lpw27451* (primers Lpw27451$_{(pFLAG)F}$ and Lpw27451$_{(pFLAG)R}$) before cloning corresponding products into p3xFLAG-Myc-CMV™. To create FLAG-tagged versions of different SnpL homologues *lpg2519, lpp2587, lpc1951* and *llo2415* were cloned into p3xFLAG-Myc-CMV™ after PCR amplification using genomic DNA from corresponding *Legionella* strains as template. To create plasmid p4HA-SnpL expressing HA-tagged SnpL during Legionella infections full length *snpL* was amplified using primers SnpL(p4HA)$_F$ and SnpL(p4HA)$_R$ and the corresponding product was cloned into the *Bam*HI and *Xba*I sites of pMMB207c.

Vectors expressing HA-tagged fusions with either full length SUPT5H or parts of it were constructed for immunoprecipitation and far western analysis. Therefore, full-length *SUPT5H* or different parts of it were amplified by using *SUPT5H* specific primers and pLX304:SUPT5H as template before the corresponding products were cloned into the *Kpn*I/*Hpa*I sites of pCMV:4xHA-SnpL excising SnpL. To create pCMV:4HA-SnpL a 715bp *Eco*RI/*Xba*I-fragment of p4HA was cloned into the corresponding sites of pEGFP_N1. GST-tagged SnpLwas purified from pGEX-4T-1:*snpL*. This expression vector was created using *Eco*RI and *Sal*I to cut *snpL* out of p3xFLAG:*snpL* and transferred into corresponding sites of pGEX-4T-1.

For the generation of inducible stable cell lines, vector pF_TRE3G:3xFLAG-*snp*L was created by amplifying FLAG-tagged *snpL* from p3xFLAG:*snpL* using primers SnpL$_{(pF_TRE3G)F}$ and SnpL$_{(pF_TRE3G)R}$ before the corresponding product was cloned into pF_TRE3G_PGK_puro. pF_TRE3G:*GFP* was generated by ligating the BamHI/XbaI *GFP*-fragment from pEGFP_N1 into the *Bam*HI/*Nhe*I-sites of pF_TRE3G_PGK_puro.

The process of creating an unmarked deletion in *L. pneumophila* 130b has been previously described (Riedmaier *et al.* 2014). Briefly, by using primers $SnpL_{\Delta 1\text{-}4}$ a PCR product of the region flanking *snpL* was generated containing an in-frame deletion of *snpL* itself. This *snpL*-mutant cassette was inserted into the *Sal*I site of suicide vector pSR47s (Merriam *et al.*, 1997). The resulting plasmid pSR47s*:*Δ*snpL* was used to replace the native *snpL* region by the mutant copy containing the *snpL* in-frame deletion.

Plasmid constructs were maintained in *E. coli* through transformation of plasmid DNA into chemically competent XL1-Blue via heat-shock, as per manufacturer's protocols. Plasmids were extracted from *E. coli* XL-1 Blue using the AxyPrep™ plasmid DNA purification miniprep kit (Axygen). Transformation of plasmid DNA into *L. pneumophila* was performed via electroporation. Overnight cultures of *L. pneumophila* were grown in AYE broth, pelleted in PBS, washed with ice-cold water and resuspended in water +10% (v/v) glycerol. Plasmid DNA was added and cells pulsed at 2.3kV, before cells were grown in AYE broth and subsequently plated on BCYE with appropriate antibiotics for 48 h.

TEM-1 beta lactamase translocation assay

J774A.1 cells were seeded at 4 x 10^4 cells/well in 96-well tissue culture plates (Corning) and incubated for 24 h. *L. pneumophila* strains were grown in AYE broth with 1mM IPTG and 12.5 µg/mL chloramphenicol. Cells were inoculated in triplicate with *L. pneumophila* strains at a MOI of 1:40 for 1 h. After infection, cells were washed with Hanks Balanced Salt Solution (HBSS) supplemented with 5% (v/v) HEPES. Beta-lactamase loading solutions (Life Technologies) and CCF2-AM substrate (Invitrogen) were added with 2.5 mM probenecid according to manufacturer's protocols. Cells were incubated at room temp for 105 minutes before loading solutions were removed and replaced with HBSS + 5% HEPES + 2.5

mM probenicid. A ClarioSTAR microplate reader (BMG Labtech) was then used to detect emissions at 450 nm and 520 nm after samples were excited with transmission of 410 nm light. The ratio of 450-520 nm light was calculated, background fluorescence subtracted and normalised against negative controls. An unpaired two-tailed Student's t test was applied to determine significance.

Immunoprecipitation of FLAG-tagged fusion proteins

HEK293T cells were seeded in 10 cm tissue culture dishes. After 24 h, cells were transfected with 11 µg of purified p3xFLAG-Myc-CMV™-24, pFLAG-SnpL or the same vector backbone expressing the individual SpnL homologues using using FuGENE® 6 (Promega). After an additional 24 h of incubation cells were washed 3 times with cold PBS before adding 600 µL cold KalB lysis buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 1% vol/vol Triton X-100, 1 mM EDTA) supplemented with 1 mM Na3VO4, 1 mM NaF, 1 mM PMSF and Complete Protease Inhibitor (Roche). Transfected cells were then pelleted and supernatants collected. Anti-FLAG M2 monoclonal magnetic beads (Sigma) were added to the cell lysate supernatant, as per manufacturer's protocol, and incubated on rotating wheel at 4°C for approximately 24 h. Cell lysate was then removed and FLAG beads were washed 3x with lysis buffer. FLAG peptide (Sigma) was added to the beads at a concentration of 100 µg/mL and incubated for 30 mins at 4°C to elute FLAG-tagged proteins which were then separated by SDS-PAGE and analysed by mass spectrometry. Input and IP samples were also stained with Sypro Ruby (Life technologies) following manufacturer's protocols. Samples for subsequent whole cell enrichment analyses were generated as above in triplicate, with samples not subjected to SDS-PAGE. Rather, whole cell lysate and IP eluent samples were stored at -80° for analysis via mass spectrometry, as described below.

Immunoblotting

Samples were prepared for SDS-PAGE gel electrophoresis with the addition of lithium dodecyl sulphate (LDS) sample buffer and dithiothreitol (DTT) reducing agent and heating to 70°C for 10 min. Samples were then loaded onto NuPAGE™ 4-12% bis-tris gels (Life Technologies) and separated via electrophoresis at 165 V for 30 mins in MES buffer (Life Technologies). Proteins were transferred to polyvinylidene difluoride (PVDF) membranes using the iBLOT-2™ transfer system (Life Technologies). After transfer, membranes were placed in a blocking solution of tris-buffered saline (TBS) with 5% (w/v) milk power and 0.1% Tween-20 (Chem Supply) for 1 h. After washing with TBS + 0.1% Tween-20, primary antibodies were applied overnight in TBS +0.1% Tween-20 + 5% BSA at 4°C. Primary antibodies used were: anti-FLAG-HRP conjugate (Sigma-Aldrich) at 1:2000, anti-FLAG (Sigma-Aldrich) at 1:2000, anti-β-actin at 1:4000 (AC-15), anti-HA.11 (Covance) at 1:2000, anti-beta lactamase (QED Bioscience) at 1:1500, anti-GFP (BioRad) and anti-GST at 1:1000 (Cell Signalling Technology). After further washing, secondary antibodies were then applied for 1 h. HRP-conjugated rabbit α-mouse and goat α-rabbit (Biorad) antibodies were applied were appropriate at 1:2000 dilutions in TBS + 0.1% Tween-20 + 5% BSA. Development reagents (GE Healthcare) were applied to membranes in accordance with manufacturer's protocols and chemiluminescence was then visualised using a ChemiBIS imaging system (DNR-Bioimaging).

Far western blotting

pGEX-4T-1-SnpL encoding a GST-SnpL fusion protein, was expressed in *E. coli* B21 and cultured overnight in LB, along with *E. coli* B21 carrying pGEX-4T-1. Cultures were further diluted 1:100 in LB supplemented with ampicillin (100 µg/mL) and incubated at 37°C until reaching an optical density of OD_{600} 0.6, whereupon 1 mM IPTG was added and cultures

were grown for a further 2 h. After centrifugation, bacterial cells were lysed via homogenisation (EmulsiFlex-C3™, Avestin) and lysate centrifuged at 13000g for 30 mins. Supernatant was recovered and GST/GST-SnpL protein purified via glutathione affinity chromatography following manufacturer's protocols (Novagen), before protein was dialysed against TBS. Protein concentration was determined via spectrophotometry (Nanodrop™, Thermo scientific).

HEK293T cells were then seeded in 6-well tissue culture plates and incubated overnight. Cells were then transfected with 2 µg plasmids encoding HA-tagged SUPT5H derivatives using FuGENE 6™. 24 h later, cells were lysed with lysis buffer and the lysate was subjected to SDS-PAGE and transferred to PVDF membrane. Expression of HA-SUPT5H derivatives was confirmed via immunoblot. Purified GST-SnpL (14 µg) or GST (7 µg) was then applied to the membranes and incubated in protein binding buffer (100 mM NaCl, 20 mM Tris [pH 7.6], 5 mM EDTA, 10% glycerol, 0.1% Tween-20, 2% milk powder and 1 mM DTT) overnight at 4°C. Unbound protein was washed with PBS + 0.1% Tween-20 before bound GST-tagged proteins were detected with the application of anti-GST (Cell Signalling) and anti-rabbit-HRP antibodies, as per immunoblot (as described below).

Inducible cell lines expressing GFP or SnpL

To generate recombinant viral vectors, HEK293T cells were seeded in 10 cm tissue culture plates overnight at 2 x 10^5 cells mL-1 and then transfected with 3.5 µg pCMVδR8.2, 2 µg pVSV-G, 7.5 µg of either pF_TRE3G:3xFLAG-*snpL* or pF_TRE3G:GFP using FuGENE® 6 (Promega) according to the manufacturers protocol. Media was exchanged after 24 h and the viral supernatant was collected, filtered through a 0.45 µM filter and stored at -80°C after 48 h.

To create stable and inducible lines, iBMDM were seeded in 6 well plates and grown for 24 h to 70% confluency. Media was then removed and replaced with 2 mL of viral supernatant and 2 µg/mL polybrene (Sigma) before plates were centrifuged at 1500rpm for 1 h. 24 h later, viral supernatant was replaced by fresh medium. Another 24 h later, iBMDM_FLAG-SnpL and iBMDM_GFP cells were selected in Roswell Park Memorial Institute (RPMI) 1640 Medium plus GlutaMAX^{TM}-1 (Gibco) supplemented with 10% heat-inactivated FBS (Gibco) and 3 µg ml-1 puromycin as described previously (Speir *et al.* 2016).

To confirm doxycycline-induced expression of SnpL and GFP, transduced cell lines were seeded in 12 well tissue culture plates. After 24 h incubation, doxycycline (Sigma) was applied to cells at 50 ng/mL for a 24 h period. Post-doxycycline application, growth media was removed on ice and cells washed with cold PBS. Cells were then lysed using lysis buffer and pelleted. Supernatant was collected and prepared for western analysis with the addition of lithium dodecyl sulphate (LDS) sample buffer and dithiothreitol (DTT) reducing agent, and heated to 70°C for 10 mins. Samples were subjected to SDS-PAGE and transferred to PVDF membranes before application of anti-GFP and anti-FLAG M2 antibodies at 1:1000 dilution in tris-buffered saline (TBS) + 0.1% Tween-20 (Chem Supply). Anti-mouse-HRP secondary antibodies (BioRad) were applied in TBS + 0.1% Tween 20 and chemiluminescence detected.

Immunofluorescence microscopy

To determine effector subcellular localisation, HEK293T cells were grown in 24 well tissue culture plates on coverslips and transfected with either p3xFLAG-*Myc*-CMV™-24, pFLAG-SnpL or pFLAG-*lpw27451* using FuGENE® 6 (Promega). For determination by infection, iBMDMs were seeded in 24 well tissue culture plates on coverslips and infected 16 h later with *L. pneumophila* 130b harbouring p4HA-SnpL that had been cultured with or without 1

mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at a multiplicity of infection of 10 for 8 h. Both transfected and infected cells were fixed in 4% paraformaldehyde-PBS for 20 min, then treated with 0.1% TritonX-100-PBS for 20 min and blocked with 3% BSA-PBS for 30 min. Cells were then incubated for 60 min in staining solution containing 0.2% BSA and a 1:800 dilution of anti-FLAG monoclonal antibody (Sigma-Aldrich) for the detection of FLAG-fusions or a 1:50 dilution of anti-HA.11 monoclonal antibody (Covance) for the detection of HA-fusions before the bound primary antibody was detected using a 1:1000 dilution of Alexa Fluor 488-conjugated anti-mouse antibody. For the last 10 minutes HOECHST stain (1:20000) was added to the staining solution. Coverslips were mounted onto glass slides with Dako Fluorescent Mounting Medium (Dako). Immunofluorescence images were acquired using a Zeiss confocal laser scanning microscope with a 100×/EC Epiplan-Apochromat oil immersion objective.

Quantification of macrophage proliferation

5 x 10^4 iBMDM_FLAG-SnpL and iBMDM_GFP cells were seeded in technical and biological triplicate in 24 well tissue culture plates before application of 50 ng/mL doxycycline to induce expression of FLAG-SnpL and GFP. After 16 h incubation, cells were harvested and enumerated using a haemocytometer (Marienfeld).

Cytotoxicity assay

3 x 10^5 iBMDM_FLAG-SnpL and iBMDM_GFP cells were seeded in 24-well tissue culture plates before the application of 50 ng/mL doxycycline to induce expression of FLAG-SnpL or GFP for 14, 16 or 20 h. Cytotoxicity was measured by LDH release using CytoTox 96 nonradioactive cytotoxicity assay (Promega) per the manufacturer's instructions. The level of LDH release was detected using a ClarioSTAR microplate reader (BMG Labtech)

(absorbance 490nm). The maximal LDH release positive control (100% cytotoxicity) was generated by adding the lysis solution (Cytotox One kit) to uninduced iBMDM_FLAG-SnpL. The percent of cytotoxicity of all values of the experimental wells were normalised to the LDH release positive control. Background fluorescence was subtracted using the average value of the culture medium background control. The assays were performed in technical and biological triplicate. An unpaired two-tailed Student's t test was applied to determine significance.

Intracellular replication assays

iBMDM were seeded in duplicate in 24-well tissue culture plates and infected the next day with stationary-phase *L. pneumophila* at a multiplicity of infection of 1. After 2 h in 5% CO_2 at 37°C cells were then treated with 100 µg/ml gentamicin for 1 h to kill extracellular bacteria. At indicated time points cells were lysed with 0.01% digitonin. Serial dilutions of the inoculum and bacteria recovered from lysed cells were plated on BCYE agar (of at least 3 independent experiments).

A. castellanii was seeded in 24-well tissue culture plates (6 x 10^5 cells per mL) in Peptone Yeast Glucose medium (2% peptone, 0.1% sodium citrate, 0.1% yeast extract) supplemented with 0.1 M glucose, 0.4 mM $CaCl_2$, 2.5 mM KH_2PO_4, 4 mM MgSO4, 2.5 mM Na_2HPO_4, 0.05 mM $Fe_4O_{21}P_6$). 24 h later, *A. castellanii* was washed with the infection buffer (0.1% sodium citrate supplemented with 0.4 mM $CaCl_2$, 2.5 mM KH_2PO_4, 4 mM MgSO4, 2.5 mM Na_2HPO_4, 0.05 mM $Fe_4O_{21}P_6$) to remove non-adherent cells. Infection was carried out with stationary phase *L. pneumophila* at a multiplicity of infection of 2 in the infection buffer. The plates were incubated at 37℃ or 30℃ with 5% CO_2 for 2 h. Media was then replaced with infection buffer containing gentamicin (100 µg/ml). After 1 h incubation, cells were washed

three times with the infection buffer. Bacteria were released from amoebae by vortexing at each time point. Serial dilutions of the inoculum and bacteria recovered from lysed cells were plated on BCYE agar. Results were expressed as the mean of at least 3 independent experiments.

Mass spectrometry analysis

Samples were prepared for proteomic analysis via mass spectrometry using the FASP protein digestion kit (Protein discovery). Protocol was modified, with eluent reduced in 5 mM Tris(2-carboxyethyl)phosphine (TCEP), and lysates were digested overnight at 37°C with trypsin Gold (Promega) in 50 mM NH_4HCO_3, before washing with NH_4HCO_3. Purified peptides were re-suspended in Buffer A and separated using Waters nanoAcquity BEH column 75 µm inner diameter, 360 µm outer diameter, 1.7µm C_{18} reverse phase analytical column connected to a nESI tip. Samples were loaded directly onto the column using an ACQUITY UPLC M-Class System (Waters) at 600 nl /min for 20 mins with Buffer A (0.1% FA) and eluted at 300 nl/min using a gradient altering the concentration of Buffer B (99.9% ACN, 0.1% FA) from 0% to 32% B over 90 mins. Separated peptides were measured with a Q-Exactive (Thermo Scientific) mass spectrometer and data acquired using data dependent acquisition. One full precursor scan (resolution 70,000; 350-2,000 m/z, AGC target of 3×10^6) followed by 10 data-dependent HCD MS-MS events (resolution 17.5 k AGC target of 1×10^5 with a maximum injection time of 200 ms, NCE 27 with 10% stepping) were allowed with 60 seconds dynamic exclusion enabled. Output mass spectra data was processed and analysed by MaxQuant (ver. 1.5.0.0) using the Andromeda search engine to identify peptides and employing the UniProtKB/Swiss-Prot reference *Homo sapiens* and *Legionella pneumophila (strain lens)* databases (Cox *et al.*, 2011). The search included variable modifications of

oxidation (methionine), amino-terminal acetylation and a fixed modification of carbamidomethyl (cysteine).

Label-free quantitative proteomics pipeline

Statistically-relevant protein expression changes were identified using a custom in-house designed pipeline as previously described (Delconte *et al.*, 2016) where quantitation was performed at the peptide level. Probability values were corrected for multiple testing using Benjamini–Hochberg method. Cut-off lines with the function $y = -\log_{10}(0.05)+c/(x-x_0)$ (Keilhauer *et al.* 2015) were introduced to identify significantly enriched proteins. c was set to 0.2 while x_0 was set to 1, representing proteins with a twofold (log2 protein ratios of 1 or more) or fourfold (log2 protein ratio of 2) change in protein expression, respectively.

RNA-sequencing

Sample preparation and total RNA extraction: iBMDM_FLAG-SnpL and iBMDM_GFP cells were seeded at a concentration of $3.6x10^6$ (n=3) and incubated in 5% CO_2 at 37°C O/N. 50 ng/mL of doxycycline was added to the medium to induce expression of transduced fusion constructs. 7 h and 12 h post induction, RNA was extracted using TRIsure™ (Bioline) according to the manufacturer's instructions. Samples were treated with DNase (Ambion TURBO DNA-free kit) and the quality of the RNA was assessed using an Agilent 2100 Bioanalyzer with the RNA 6000 Nano Chip (Agilent Technologies, Waldbronn, Germany). RNA quantity was determined using a spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

For transcriptome sequencing and alignment, cDNA library preparation and sequencing was performed as previously described (Tanaka *et al.*, 2013). Briefly, 1 ug of total RNA was used

for poly(A) selection and cDNA libraries were constructed using the Illumina sample preparation kit. The HiSeq 2500 (Illumina) platform was then used to perform 100 bp paired-end sequencing, employing the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) according to manufacturer's instructions. Acquired reads were mapped on mouse genome GRCm38 by Tophat2 v2.1.1 (Kim *et al.*, 2013), with the annotation described in gtf release-85 allowing 2 mismatches. Reads were counted using HTSeq (Anders et al., 2015). Differential expression analysis was performed using the R software package tool, DESeq (Anders *et al.*, 2010). After data normalisation, volcano plot-based methods (using Excel) were used to visualise $\log_2$ fold-change expression data against the false discovery rate (FDR, DESeq-PADJ) and a 5% false discovery rate (FDR) cut-off was set. The $\log_2$ fold-change expression data was further analysed for consistent trends of expression against two independent control conditions. Samples were omitted if they showed inconsistent trends of expression between controls or if the $\log_2$ fold-change was <3 and > -3. The PANTHER Classification System (Mi *et al.*, 2013) was used to visualise the functions of differently expressed genes. The DAVID Gene Set Enrichment analysis tool (Huang da *et al.*, 2009) was used to identify significantly enriched functional gene groups in accordance with GO ontology (Ashburner *et al.*, 2000). Gene clusters with an adjusted *P*-value of < 0.05 were considered significant.

Bioinformatic analysis of SnpL homologues

Homologues of SnpL from *L. pneumophila* strain 130b were searched by BlastP against a local database containing 300 *L. pneumophila* strains representing the species diversity (David *et al.*, 2017). Putative homologues in other species were searched by BlastP against the NCBI non-redundant database. Additional BlastP searches, using the sequence from other *Legionella* species as seed sequence, were performed to complete the data set. All recruited

significant homologues were aligned using Muscle (Edgar, 2004). The resulting alignment was curated with Gblocks (Castresana, 2000) and phylogenetic reconstruction was done by likelihood implemented in the PhyML program (Guindon *et al.* 2009). Branch support was calculated using the approximate *likelihood*-ratio test (aLRT). Searches for nuclear localisation sequences in Lpw27461 (SnpL) was achieved through use of the NLStradamus tool of the University of Toronto (Nguyen Ba *et al.*, 2009).

Acknowledgements

This work was supported by the Australian National Health and Medical Research Council (Project Grants APP1098826 and APP1044061 to ELH); Program Grant 606788 to ELH and JLS and GI-CORE. Work in CB's laboratory is supported by the Institut Pasteur, by the Fondation pour la Recherche Médicale (FRM) grant N° DEQ20120323697, the grant ANR-10-LABX-62-IBEID, and the Infect-ERA project EUGENPATH (ANR-13-IFEC-0003-02). The authors declare no financial interests related to this work. PL is supported by the China Scholarship Council.

References

Adelman, K. and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13,** 720-731.

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome biology* **11,** R106.

Arbibe, L., Kim, D.W., Batsche, E., Pedron, T., Mateescu, B., Muchardt, C., *et al.* (2007). An injected bacterial effector targets chromatin access for transcription factor NF-kappaB to alter transcription of host genes involved in immune responses. *Nat Immunol* **8,** 47-56.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25,** 25-29.

Asrat, S., Dugan, A.S. and Isberg, R.R. (2014). The frustrated host response to *Legionella pneumophila* is bypassed by MyD88-dependent translation of pro-inflammatory cytokines. *PLoS Pathog* **10,** e1004229.

Barry, K.C., Fontana, M.F., Portman, J.L., Dugan, A.S. and Vance, R.E. (2013). IL-1alpha signaling initiates the inflammatory response to virulent *Legionella pneumophila* in vivo. *J Immunol* **190,** 6329-6339.

Barry, K.C., Ingolia, N.T. and Vance, R.E. (2017). Global analysis of gene expression reveals mRNA superinduction is required for the inducible immune response to a bacterial pathogen. *Elife* **6**.

Belyi, Y., Jank, T. and Aktories, K. (2013). Cytotoxic glucosyltransferases of *Legionella pneumophila*. *Curr Top Microbiol Immunol* **376,** 211-226.

Berger, K.H. and Isberg, R.R. (1993). Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*. *Mol Microbiol* **7,** 7-19.

Bierne, H. and Cossart, P. (2012). When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14,** 622-633.

Burstein, D., Zusman, T., Degtyar, E., Viner, R., Segal, G. and Pupko, T. (2009). Genome-scale identification of *Legionella pneumophila* effectors using a machine learning approach. *PLoS Pathog* **5,** e1000508.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17,** 540-552.

Cazalet, C., Gomez-Valero, L., Rusniok, C., Lomma, M., Dervins-Ravault, D., Newton, H.J., *et al.* (2010). Analysis of the *Legionella longbeachae* genome and transcriptome uncovers unique strategies to cause Legionnaires' disease. *PLoS Genet* **6,** e1000851.

Cazalet, C., Rusniok, C., Bruggemann, H., Zidane, N., Magnier, A., Ma, L., *et al.* (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nature Genetics* **36,** 1165-1173.

Copenhaver, A.M., Casson, C.N., Nguyen, H.T., Duda, M.M. and Shin, S. (2015). IL-1R signaling enables bystander cells to overcome bacterial blockade of host protein synthesis. *Proc Natl Acad Sci U S A* **112,** 7557-7562.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V. and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10,** 1794-1805.

David, S., Sanchez-Buso, L., Harris, S.R., Marttinen, P., Rusniok, C., Buchrieser, C., *et al.* (2017). Dynamics and impact of homologous recombination on the evolution of *Legionella pneumophila*. *PLoS Genet* **13,** e1006855.

de Felipe, K.S., Glover, R.T., Charpentier, X., Anderson, O.R., Reyes, M., Pericone, C.D. and Shuman, H.A. (2008). *Legionella* eukaryotic-like type IV substrates interfere with organelle trafficking. *PLoS Pathog* **4,** e1000117.

Delconte, R.B., Kolesnik, T.B., Dagley, L.F., Rautela, J., Shi, W., Putz, E.M., *et al.* (2016). CIS is a potent checkpoint in NK cell-mediated tumor immunity. *Nat Immunol* **17,** 816-824.

Dolezal, P., Aili, M., Tong, J., Jiang, J.H., Marobbio, C.M., Lee, S.F., *et al.* (2012). *Legionella pneumophila* Secretes a Mitochondrial Carrier Protein during Infection. *PLoS Pathogens* **8,** e1002459.

Edelstein, P.H., Nakahama, C., Tobin, J.O., Calarco, K., Beer, K.B., Joly, J.R. and Selander, R.K. (1986). Paleoepidemiologic investigation of Legionnaires disease at Wadsworth Veterans Administration Hospital by using three typing methods for comparison of legionellae from clinical and environmental sources. *J Clin Microbiol* **23,** 1121-1126.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32,** 1792-1797.

Ensminger, A.W. (2016). *Legionella pneumophila*, armed to the hilt: justifying the largest arsenal of effectors in the bacterial world. *Curr Opin Microbiol* **29,** 74-80.

Feeley, J.C., Gibson, R.J., Gorman, G.W., Langford, N.C., Rasheed, J.K., Mackel, D.C. and Baine, W.B. (1979). Charcoal-yeast extract agar: primary isolation medium for *Legionella pneumophila*. *J Clin Microbiol* **10,** 437-441.

Finsel, I. and Hilbi, H. (2015). Formation of a pathogen vacuole according to *Legionella pneumophila*: how to kill one bird with many stones. *Cell Microbiol* **17,** 935-950.

Fontana, M.F., Banga, S., Barry, K.C., Shen, X., Tan, Y., Luo, Z.Q. and Vance, R.E. (2011). Secreted bacterial effectors that inhibit host protein synthesis are critical for induction of the innate immune response to virulent *Legionella pneumophila*. *PLoS Pathogens* **7,** e1001289.

Fontana, M.F., Shin, S. and Vance, R.E. (2012). Activation of host mitogen-activated protein kinases by secreted *Legionella pneumophila* effectors that inhibit host protein translation. *Infect Immun* **80,** 3570-3575.

Guindon, S., Delsuc, F., Dufayard, J.F. and Gascuel, O. (2009). Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* **537,** 113-137.

Hartzog, G.A. and Fu, J. (2013). The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim Biophys Acta* **1829,** 105-115.

Huang da, 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.

Huang, L., Boyd, D., Amyot, W.M., Hempstead, A.D., Luo, Z.Q., O'Connor, T.J., *et al.* (2011). The E Block motif is associated with *Legionella pneumophila* translocated substrates. *Cell Microbiol* **13,** 227-245.

Hubber, A. and Roy, C.R. (2010). Modulation of host cell function by *Legionella pneumophila* type IV effectors. *Annu Rev Cell Dev Biol* **26,** 261-283.

Jepras, R.I., Fitzgeorge, R.B. and Baskerville, A. (1985). A comparison of virulence of two strains of *Legionella pneumophila* based on experimental aerosol infection of guinea-pigs. *J Hyg (Lond)* **95,** 29-38.

Jonkers, I. and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16,** 167-177.

Kay, S., Hahn, S., Marois, E., Hause, G. and Bonas, U. (2007). A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318,** 648-651.

Keilhauer, E.C., Hein, M.Y. and Mann, M. (2015). Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics* **14,** 120-135.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14,** R36.

King, N.P., Newton, P., Schuelein, R., Brown, D.L., Petru, M., Zarsky, V., *et al.* (2015). Soluble NSF attachment protein receptor molecular mimicry by a *Legionella pneumophila* Dot/Icm effector. *Cell Microbiol* **17,** 767-784.

Kwak, H. and Lis, J.T. (2013). Control of transcriptional elongation. *Annu Rev Genet* **47,** 483-508.

Li, H., Xu, H., Zhou, Y., Zhang, J., Long, C., Li, S., *et al.* (2007). The phosphothreonine lyase activity of a bacterial type III effector family. *Science* **315,** 1000-1003.

Li, T., Lu, Q., Wang, G., Xu, H., Huang, H., Cai, T., *et al.* (2013). SET-domain bacterial effectors target heterochromatin protein 1 to activate host rDNA transcription. *EMBO Rep* **14,** 733-740.

Malone, E.A., Fassler, J.S. and Winston, F. (1993). Molecular and genetic characterization of SPT4, a gene important for transcription initiation in *Saccharomyces cerevisiae*. *Mol Gen Genet* **237,** 449-459.

Merriam, J.J., Mathur, R., Maxfield-Boumil, R. and Isberg, R.R. (1997). Analysis of the *Legionella pneumophila fliI* gene: intracellular growth of a defined mutant defective for flagellum biosynthesis. *Infect Immun* **65,** 2497-2501.

Mi, H., Muruganujan, A., Casagrande, J.T. and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* **8,** 1551-1566.

Moujalled, D.M., Cook, W.D., Okamoto, T., Murphy, J., Lawlor, K.E., Vince, J.E. and Vaux, D.L. (2013). TNF can activate RIPK3 and cause programmed necrosis in the absence of RIPK1. *Cell Death Dis* **4,** e465.

Nguyen Ba, A.N., Pogoutse, A., Provart, N. and Moses, A.M. (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC bioinformatics* **10,** 202.

Ponting, C.P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic Acids Res* **30,** 3643-3652.

Qiu, J. and Luo, Z.Q. (2017). *Legionella* and *Coxiella* effectors: strength in diversity and activity. *Nat Rev Microbiol* **15,** 591-605.

Riedmaier, P., Sansom, F.M., Sofian, T., Beddoe, T., Schuelein, R., Newton, H.J. and Hartland, E.L. (2014). Multiple ecto-nucleoside triphosphate diphosphohydrolases facilitate intracellular replication of *Legionella pneumophila*. *Biochem J* **462,** 279-289.

Rolando, M. and Buchrieser, C. (2014). *Legionella pneumophila* type IV effectors hijack the transcription and translation machinery of the host cell. *Trends Cell Biol* **24,** 771-778.

Rolando, M., Sanulli, S., Rusniok, C., Gomez-Valero, L., Bertholet, C., Sahr, T., *et al.* (2013). *Legionella pneumophila* effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication. *Cell Host Microbe* **13,** 395-405.

Shao, W. and Zeitlinger, J. (2017). Paused RNA polymerase II inhibits new transcriptional initiation. *Nat Genet* **49,** 1045-1051.

Shen, X., Banga, S., Liu, Y., Xu, L., Gao, P., Shamovsky, I., *et al.* (2009). Targeting eEF1A by a *Legionella pneumophila* effector leads to inhibition of protein synthesis and induction of host stress response. *Cell Microbiol* **11,** 911-926.

Sinclair, S.H., Rennoll-Bankert, K.E. and Dumler, J.S. (2014). Effector bottleneck: microbial reprogramming of parasitized host cell transcription by epigenetic remodeling of chromatin structure. *Front Genet* **5,** 274.

Speir, M., Lawlor, K.E., Glaser, S.P., Abraham, G., Chow, S., Vogrin, A., *et al.* (2016). Eliminating *Legionella* by inhibiting BCL-XL to induce macrophage apoptosis. *Nat Microbiol* **1,** 15034.

Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., *et al.* (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9,** 493-501.

Tanaka, S., Nishimura, M., Ihara, F., Yamagishi, J., Suzuki, Y. and Nishikawa, Y. (2013). Transcriptome analysis of mouse brain infected with *Toxoplasma gondii*. *Infect Immun* **81,** 3609-3619.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., *et al.* (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes & development* **12,** 343-356.

Zhu, W., Banga, S., Tan, Y., Zheng, C., Stephenson, R., Gately, J. and Luo, Z.Q. (2011). Comprehensive identification of protein substrates of the Dot/Icm type IV transporter of *Legionella pneumophila*. *PLoS One* **6,** e17638.

Table 1. Strains and plasmids used in this study.

Strains/plasmids	Characteristics	Source
L. pneumophila		
130b (ATCC BAA-74)	O1; clinical isolate	(Edelstein *et al.*, 1986)
Δ*dotA*	130b Δ*dotA* in-frame markerless deletion	(Riedmaier *et al.* 2014)
Δ*snpL*	130b Δ*snpL* in-frame markerless deletion	This study
130b (pXDC61)	130b carrying pXDC61	(Riedmaier *et al.* 2014.)
130b (pTEM1-RalF)	130b carrying pTEM1-RalF	(Riedmaier *et al.* 2014)
130b (pTEM1-SnpL)	130b carrying pTEM1-SnpL	This study
130b (pTEM1-Lpw27451)	130b carrying pTEM1-Lpw27451	This study
130b (pTEM1-Lpw27441)	130b carrying pTEM1-Lpw27441	This study
Δ*dotA* (pTEM1-RalF)	Δ*dotA* carrying pTEM1-RalF	(Riedmaier *et al.* 2014)
Δ*dotA* (pTEM1-SnpL)	Δ*dotA* carrying pTEM1-SnpL	This study
Δ*dotA* (pTEM1-Lpw27451)	Δ*dotA* carrying pTEM1-Lpw27451	This study
Δ*dotA* (pTEM1-Lpw27441)	Δ*dotA* carrying pTEM1-Lpw27441	This study
130b (p4HA-SnpL)	130b carrying p4HA-SnpL	This study
Δ*dotA* (p4HA-SnpL)	Δ*dotA* carrying p4HA-SnpL	This study
Philadelphia JR32	Clinical isolate	(Berger *et al.* 1993)
Paris	Clinical isolate	(Cazalet *et al.* 2004)
Corby	Clinical isolate	(Jepras *et al.* 1985)
L. longbeachae		
NSW150	Clinical isolate	(Cazalet *et al.* 2010)
E. coli		
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^{R})]	Stratagene
BL21-C43	*E. coli* used for expression of proteins for affinity purification	Novagen
Plasmids		

pGEM®-T Easy	high copy cloning vector	Promega
pXDC61	Vector expressing TEM1 version of BlaM,	(de Felipe *et al.* 2008)
pTEM1-RalF	pXDC61 expressing TEM1-RalF fusion protein	(King *et al.*, 2015)
pTEM1-SnpL	pXDC61 expressing TEM1-Lpw27461 fusion protein	This study
pTEM1-Lpw27451	pXDC61 expressing TEM1-Lpw27451 fusion protein	This study
pTEM1-Lpw27441	pXDC61 expressing TEM1-Lpw27441 fusion protein	This study
pICC562	pMMB207c based expression vector for IPTG inducible 4xHA-tagged proteins	(Dolezal *et al.* 2012)
p4HA-SnpL	pICC562 expressing 4xHA-tagged SnpL	This study
pSR47s	*sacB* based suicide vector for mutagenesis	(Merriam *et al.* 1997)
pSR47s:Δ*snpL*	Vector carrying *snpL* deletion cassette	This study
p3xFLAG-*Myc*-CMV™-24	Dual tagged N-terminal Met-3xFLAG and C-terminal *c-myc* expression vector	Sigma
pFLAG-SnpL	Vector expressing 3xFLAG-tagged SnpL	This study
pFLAG-Lpw27451	Vector expressing 3xFLAG-tagged Lpw27451	This study
pFLAG-Lpp2587	Vector expressing 3xFLAG-tagged Lpp2587	This study
pFLAG-Lpg2519	Vector expressing 3xFLAG-tagged Lpg2519	This study
pFLAG-Lpc1951	Vector expressing 3xFLAG-tagged Lpc1951	This study
pFLAG-Llo2415	Vector expressing 3xFLAG-tagged Llo2415	This study
pVSV-G	Lentivirus packaging vector – envelope	(Stewart *et al.*, 2003)
PCMVδR8.2	Lentivirus packaging vector – transcriptional machinery	(Stewart *et al.*, 2003)
pTRE_H2b_GFP	Positive transfection control vector expressing GFP	(Moujalled *et al.* 2013)
pF_TRE3G_PGK_puro	Transduction vector carrying a tetracycline-responsive element	(Moujalled *et al.* 2013)
pF_TRE3G:3xFLAG-SnpL	Transduction vector expressing IPTG inducible SnpL	This study
pF_TRE3G:GFP	Transduction vector expressing IPTG inducible GFP (from pEGFP_N1)	This study
pEGFP_N1	Vector to create C-terminal EGFP fusions	Clontech
pCMV:4HA-SnpL	Vector expressing 4xHA-tagged SnpL	This study
pLX304:SUPT5H	Vector carrying SUPT5H	DNASU
pCMV:4xHA-SUPT5H$_{1-176}$	Vector expressing 4xHA-tagged SUPT5H$_{1-176}$	This study
pCMV:4xHA-SUPT5H$_{1-274}$	Vector expressing 4xHA-tagged SUPT5H$_{1-274}$	Amersham

pCMV:4xHA-SUPT5H$_{1-424}$	Vector expressing 4xHA-tagged SUPT5H$_{1-424}$	This study
pCMV:4xHA-SUPT5H$_{418-1087}$	Vector expressing 4xHA-tagged SUPT5H$_{418-1087}$	This study
pCMV:4xHA-SUPT5H$_{751-1087}$	Vector expressing 4xHA-tagged SUPT5H$_{751-1087}$	This study
pCMV:4xHA-SUPT5H$_{1-754}$	Vector expressing 4xHA-tagged SUPT5H$_{1-754}$	This study
pCMV:4xHA-SUPT5H$_{1-1087}$	Vector expressing 4xHA-tagged SUPT5H$_{1-1087}$	This study
pGEX-4T-1	Vector to create GST fusion proteins	GE Healthcare
pGEX-4T-1-SnpL	Vector expressing GST-tagged SnpL	This study

Table 2. List of primers used in this study

Oligo	Sequence (5'-3')
$SnpL_{(pTEM1)F}$	ACGTATGAATTCGGATCCATGCCGATTTTGATTCAAAAAG
$SnpL_{(pTEM1)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGTAAACTG
$Lpw27451_{(pTEM1)F}$	ACGTATGAATTCGGGATCCATGTCAATGACTCTCTCGTTAG
$Lpw27451_{(pTEM1)R}$	ATCCATGTCGACTCTAGATTAGAAAGAATAATGGTAATCTTC
$Lpw27441_{(pTEM1)F}$	ACGTATGAATTCGGGATCCATGATGCTCGATCGTACTGAT
$Lpw27441_{(pTEM1)R}$	ATCCATGTCGACTCTAGACTAATCCAAATGATCAAGGGG
$SnpL_{(pFLAG)F}$	ACGTATGAATTCGATGCCGATTTTGATTCAAAAAG
$SnpL_{(pFLAG)R}$	ATCCATGTCGACTTAGTTTGAACTCAATGTAAACTG
$Lpw27451_{(pFLAG)F}$	ACGTATGAATTCGGGATCCATGTCAATGACTCTCTCGTTAG
$Lpw27451_{(pFLAG)R}$	ATCCATGTCGACTCTAGATTAGAAAGAATAATGGTAATCTTC
$SnpL_{(p4HA)F}$	ACGTATGAATTCGGATCCATGCCGATTTTGATTCAAAAAG
$SnpL_{(p4HA)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGTAAACTG
$Lpg2519_{(pFLAG)F}$	ACGTATGAATTCGGGATCCATGCTGATTTTGATTCAAAAAG
$Lpg2519_{(pFLAG)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGTAAATTG
$Lpp2587_{(pFLAG)F}$	ACGTATGAATTCGGGATCCATGCTGATTTTGATTCAAAAAG
$Lpp2587_{(pFLAG)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGTAAACTG
$Lpc1951_{(pFLAG)F}$	ACGTATAAGCTTGGATCCATGTCTAAATTCCCGCATAAAAA
$Lpc1951_{(pFLAG)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGCAAAC
$Llo2415_{(pFLAG)F}$	ACGTATGAATTCGGGATCCATGTCAAGATTTCCAAATAAAAC
$Llo2415_{(pFLAG)R}$	ATCCATGTCGACTCTAGATTAATATCTGTGAACTACCTGG
$SnpL_{(pF\_TRE3G)F}$	CAGGATCCATGGACTACAAAGACCATGACGGTG
$SnpL_{(pF\_TRE3G)R}$	ATCCATGTCGACTCTAGATTAGTTTGAACTCAATGTAAACTG
$SUPT5H_{(1-176)F}$	ATGGTACCATGTCGGACAGCGAGGAC

SUPT5H$_{(1\text{-}176)R}$	ATGTTAACTCAATCCTTGACTCCTGGGAG
SUPT5H$_{(1\text{-}274)F}$	ATGGTACCATGTCGGACAGCGAGGAC
SUPT5H$_{(1\text{-}274)R}$	ATGTTAACTCATTTCAGGTTGGCCACCTC
SUPT5H$_{(1\text{-}424)F}$	ATGGTACCATGTCGGACAGCGAGGACAG
SUPT5H$_{(1\text{-}424)R}$	ATGTTAACTCAGTCCCCAGGTTGGAAGTTG
SUPT5H$_{(418\text{-}1087)F}$	ATGGTACCCACAACTTCCAACCTGGGG
SUPT5H$_{(418\text{-}1087)R}$	ATGTTAACTCAGGCTTCCAGGAGCTTC
SUPT5H$_{(751\text{-}1087)F}$	ATGGTACCACCACGGTGGGCTCACG
SUPT5H$_{(751\text{-}1087)R}$	ATGTTAACTCAGGCTTCCAGGAGCTTC
SUPT5H$_{(1\text{-}754)F}$	ATGGTACCATGTCGGACAGCGAGGAC
SUPT5H$_{(1\text{-}754)R}$	ATGTTAACTCAGCCCACCGTGGTGAGCC
SnpL $_{(\Delta 1)}$	AGCTAGGTCGACAGTAAAAAATGTCAATGACTCTC
SnpL $_{(\Delta 2)}$	CATATCTTATCATATGCTTTTTG
SnpL $_{(\Delta 3)}$	CAAAAAGCATATGATAAGATATGTAAAAATTGATAAGCCTAACCC
SnpL $_{(\Delta 4)}$	AGCTAGGTCGACCTGTAAAGATTATCCTGTTC

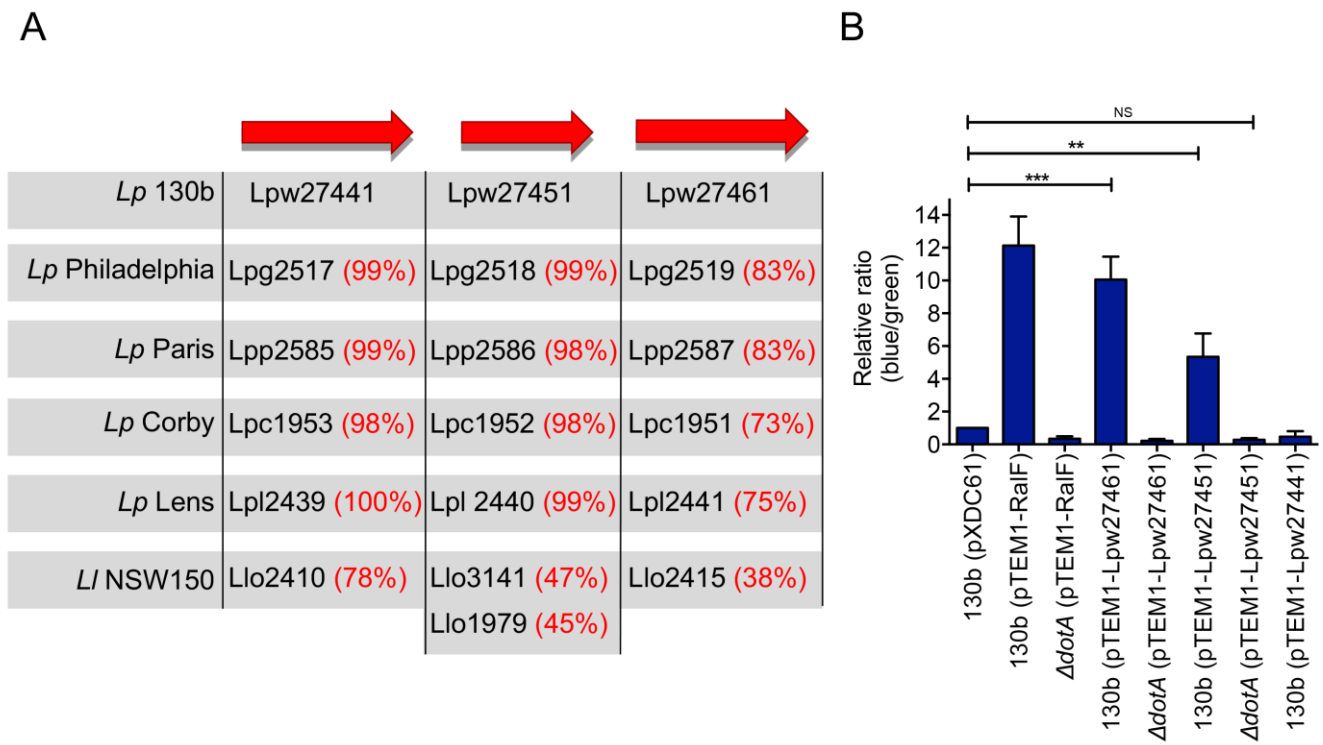

Fig. 1. Conservation and translocation of putative *L. pneumophila* effector proteins.

A. Identification of gene orthologues of *L. pneumophila* (130b) putative effector proteins through pair-wise sequence alignment against *L. pneumophila* Phil.-1, Paris, Corby, Lens and *L. longbeachae* strain NSW150 (BlastP 2.2.31+). Percentage amino acid sequence identities shown in red. **B.** TEM1-based translocation of putative *L. pneumophila* effector proteins by wild type *L. pneumophila* 130b compared to an isogenic Δ *dotA* derivative. Translocation expressed as a response ratio of blue/green fluorescence, with background fluorescence subtracted, and normalized against fluorescence seen in cells expressing the empty pXDC61 vector. TEM-RalF translocation was measured as a positive control. Error bars represent the standard error of the mean (** denotes $p < 0.005$, *** denotes $p < 0.001$, unpaired two tailed *t*-test) of three biological repeats. NS, not significant.

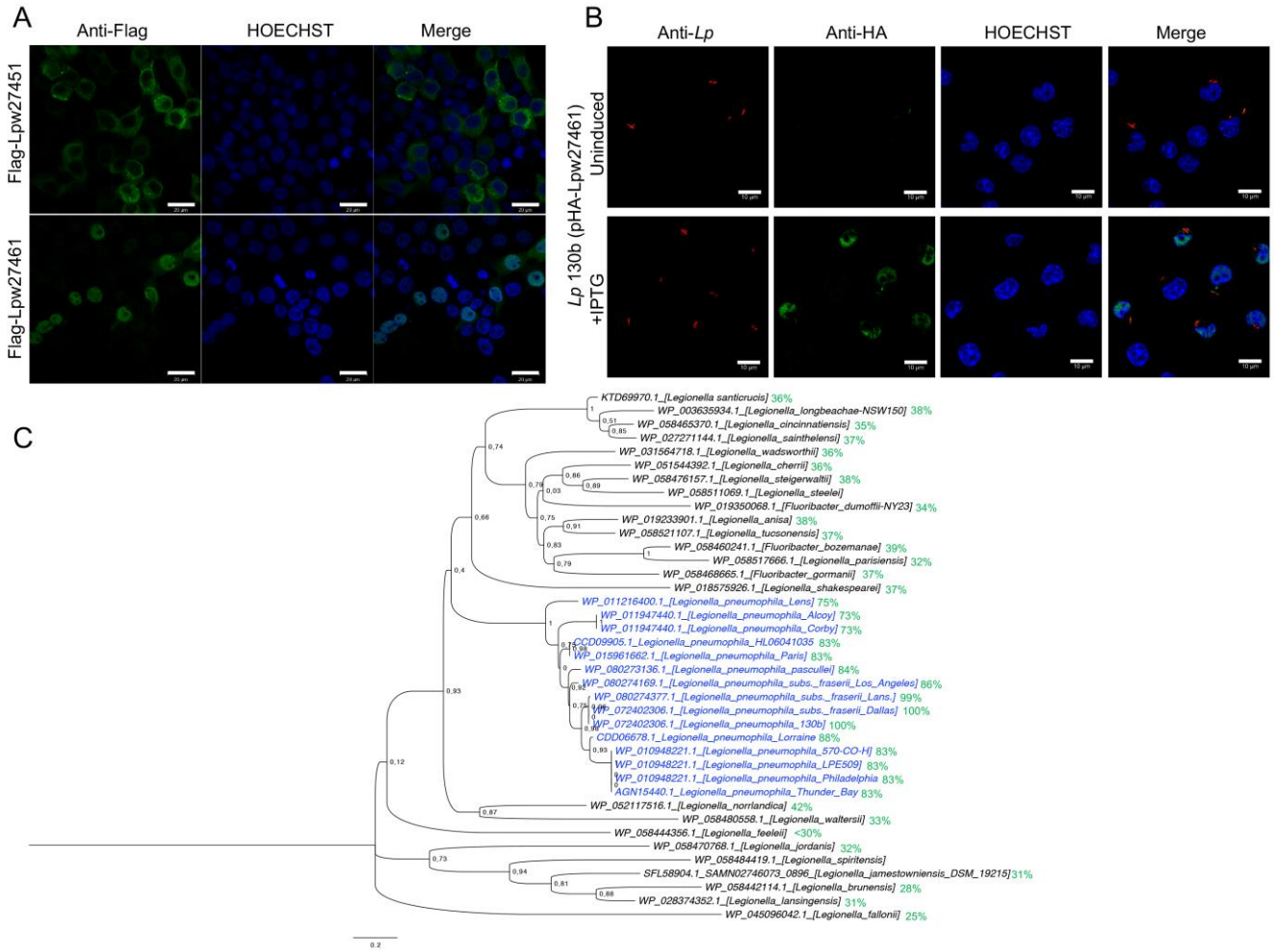

Fig. 2 Localisation of SnpL within mammalian cells, and the phylogenetic relatedness of SnpL homologues amongst *Legionella* species. **A**. Representative immunofluorescence fields of HEK293T cells transfected to express FLAG-Lpw27451 and FLAG-Lpw27461 (green). Cell nuclei were visualized with Hoechst stain (blue). **B.** Immortalized bone marrow derived macrophages (iBMDM) infected with *L. pneumophila* 130b (pHA-Lpw27461) at MOI of 10 for 8 h. 4HA-Lpw27461 expression was induced with the addition of IPTG prior to infection (green). **C.** Phylogenetic tree of Lpw27461 homologous proteins constructed by likelihood. *L. pneumophila* sequences are in blue. Numbers indicate branch support for nodes. Shown in green are the amino acid percentage identities of homologues with Lpw27461. Bar at the bottom represents the estimated evolutionary distance.

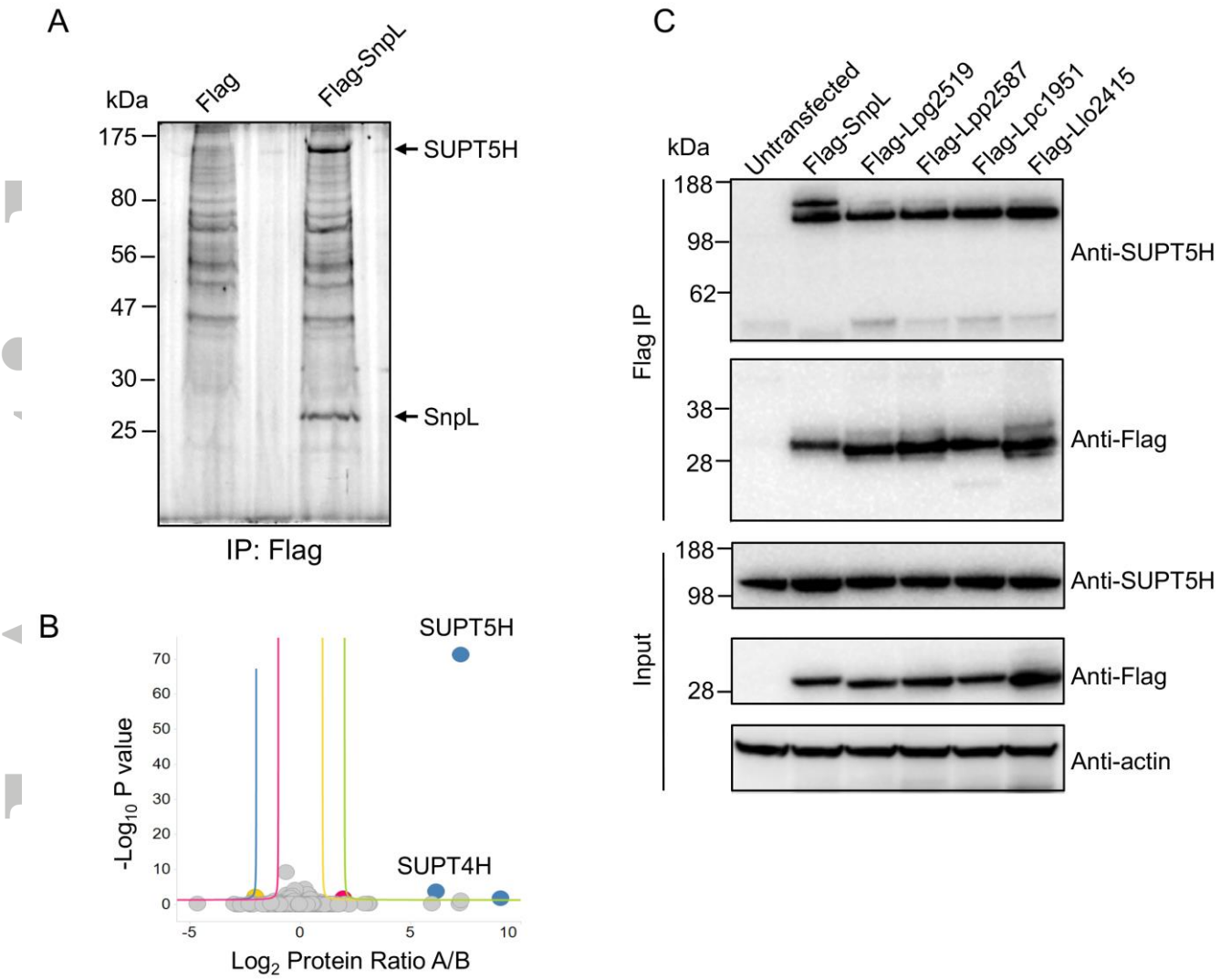

Fig. 3. Identification of SnpL host targets A. FLAG-immunoprecipitation (IP) of HEK293T cells expressing FLAG or FLAG-SnpL separated by SDS-PAGE and stained with SYPRO Ruby™. Representative SDS-PAGE gel of at least three independent experiments. **B.** Volcano plot illustrating the $\log_2$ protein ratios of proteins enriched by immunoprecipitation of FLAG-SnpL (A) and relative to the control (B), following quantitative pipeline analysis, including SUPT5H ($\log_2$ 7.3 fold) and SUPT4H ($\log_2$ 6.1 fold). Proteins were deemed differentially regulated if the $\log_2$ fold change in protein expression was greater than 2-fold (pink and yellow) or 4-fold (blue and green) and a $-\log_{10}$ p value $\geq$ 1.3, equivalent to a p value ≤ 0.05. **C.** Immunoblot showing co-immunoprecipitation of ectopically expressed SUPT5H and FLAG-tagged SnpL homologues Lpg2519, Lpp2587, Lpc1951 and Llo2415 in HEK293T cells. Cells were harvested for immunoprecipitation using anti-FLAG antibodies and subjected to SDS-PAGE followed by immunoblot using anti-FLAG and anti-SUPT5H antibodies. Antibodies to β-actin were used as a loading control. Representative immunoblot of at least three independent experiments.

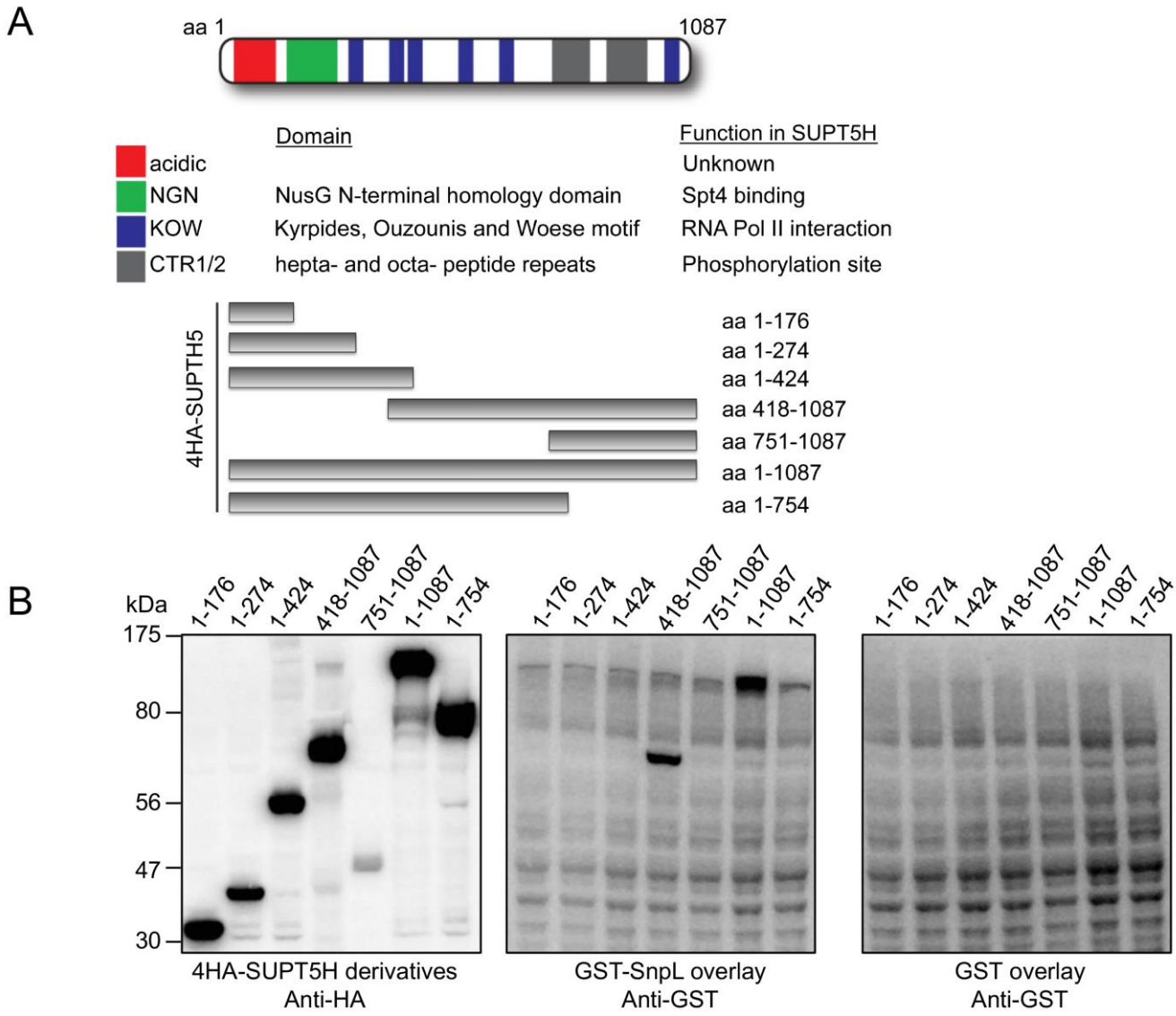

Fig. 4. Investigation of SnpL interaction with SUPT5H. A. Schematic representation of SUPT5H showing functional domains and SUPT5H fragments encompassing various domains. **B.** Far-western analysis of HEK293T cells transfected with pHA-SUPT5H domain fragments. HEK293T lysates were subjected to SDS-PAGE and transferred to PVDF membranes before the application of purified GST-SnpL. Protein expression was detected with anti-HA antibodies (left panel) and GST-SnpL interactions were detected using anti-GST antibodies (middle panel). GST alone was applied as a negative control (right panel). Representative immunoblot of at least three independent experiments.

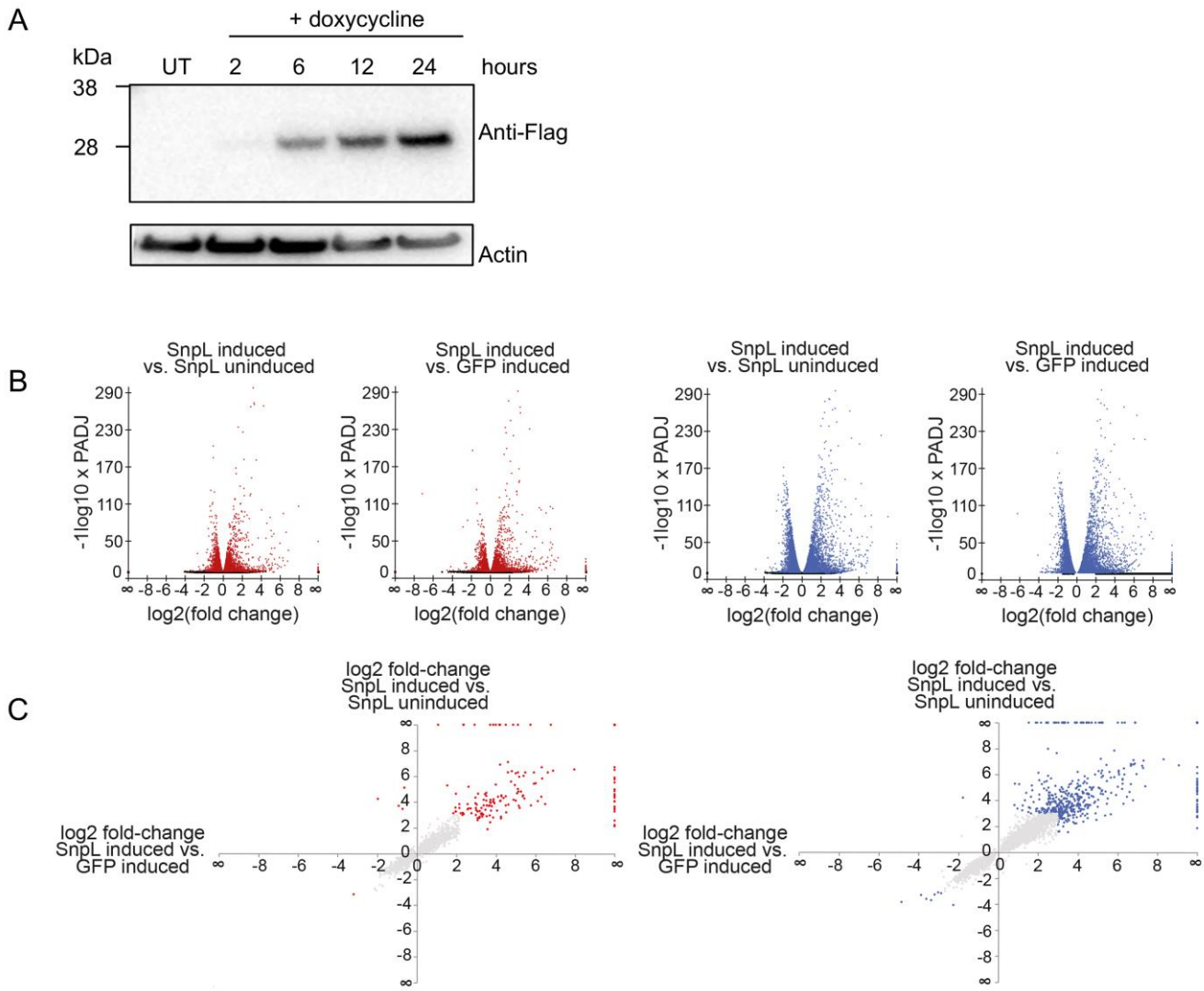

Fig. 5. Differential gene expression of iBMDM upon SnpL induction A. Immunoblot showing induction of FLAG-SnpL expression in transduced iBMDM over time with the addition of 50 ng/mL doxycycline. Cells were lysed and FLAG-SnpL expression detected using anti-FLAG antibody. Antibody to β-actin was used as a loading control. Representative immunoblot of at least three independent experiments. **B.** Global changes in iBMDM gene expression after SnpL induction at 7 h (red) and 12 h (blue) relative to uninduced controls (SnpL induced vs. SnpL uninduced), or GFP induced controls (SnpL induced vs. GFP induced). Volcano plots illustrate the distribution of change (log2(fold change)) detected against significance, *p*-value adjusted for multiple testing using Benjamini-Hochberg to estimate the false discovery rate (PADJ). **C.** Scatter plots of the log2(fold-change) in gene expression of SnpL induced vs. SnpL uninduced against SnpL induced vs GFP induced after 7 h (red) and 12 h (blue) of doxycycline (50ng/mL) treatment. Points shaded grey represent genes with a log2(fold change) difference of < 3 or > -3 comparing SnpL induction and controls. These genes were omitted from subsequent analysis. Points reaching infinity represent genes with mRNA transcripts detected only in the SnpL induced cell line or the control cell line and were also omitted from subsequent analysis.

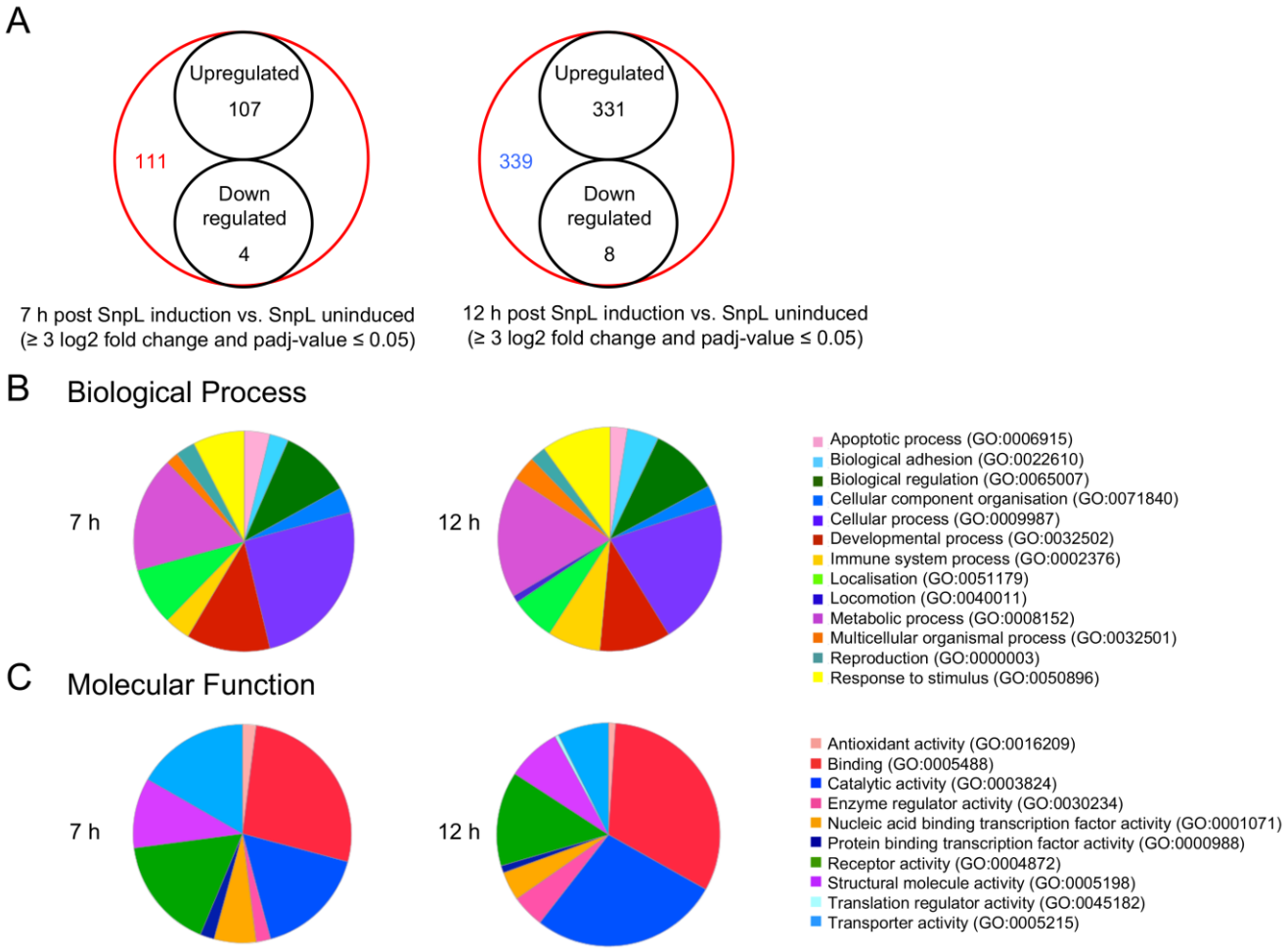

Fig. 6. Expression of SnpL results in global upregulation of the iBMDM transcriptome.

A. Diagram showing the number of differentially expressed iBMDM genes identified after 7 h or 12 h of SnpL induction compared to uninduced conditions. Differentially upregulated genes were deposited into the PANTHER classification system. **B** and **C**. Results illustrated as pie charts in which each gene has the potential to fall into more than one biological process (B) and/or molecular functional group (C).

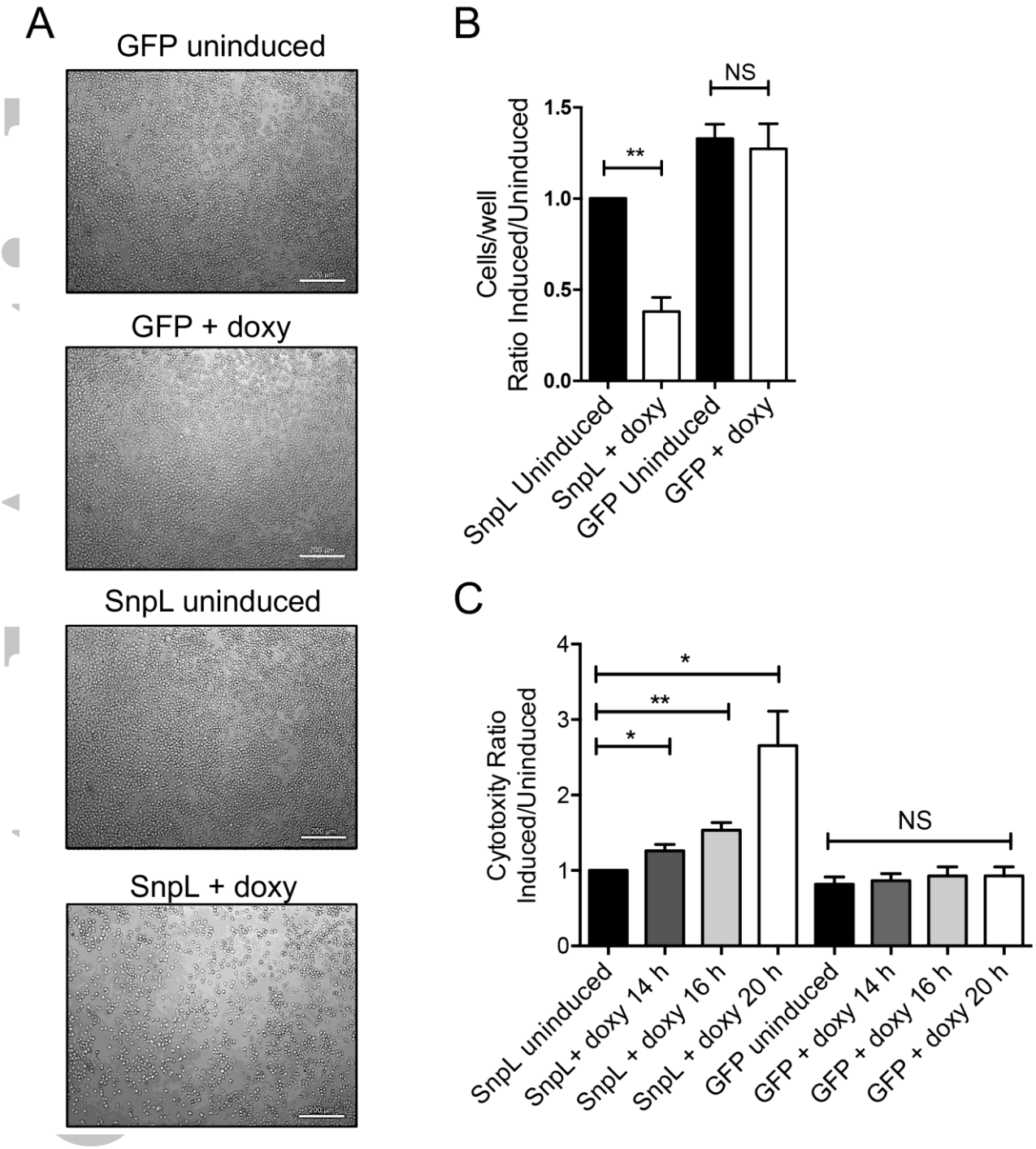

Fig. 7. Functional analysis of SnpL expression in iBMDMs. A. Representative phase contrast field comparing adhesion of iBMDM induced to express GFP or FLAG-SnpL, 16 h after induction with 50 ng/mL doxycycline. **B.** Number of iBMDM per well before and after 16 h of SnpL induction with 50 ng/mL doxycycline. **C.** LDH release assays measuring the effect of FLAG-SnpL expression on the cytotoxicity of iBMDM 14 h, 16 h and 20 h after induction with with 50 ng/mL doxycycline. iBMDM expressing GFP were used as negative controls. Experiment was performed in biological and technical triplicate. Values presented as a ratio of induced: non-induced cells, $p < 0.0001$, unpaired two-tailed *t*-test.

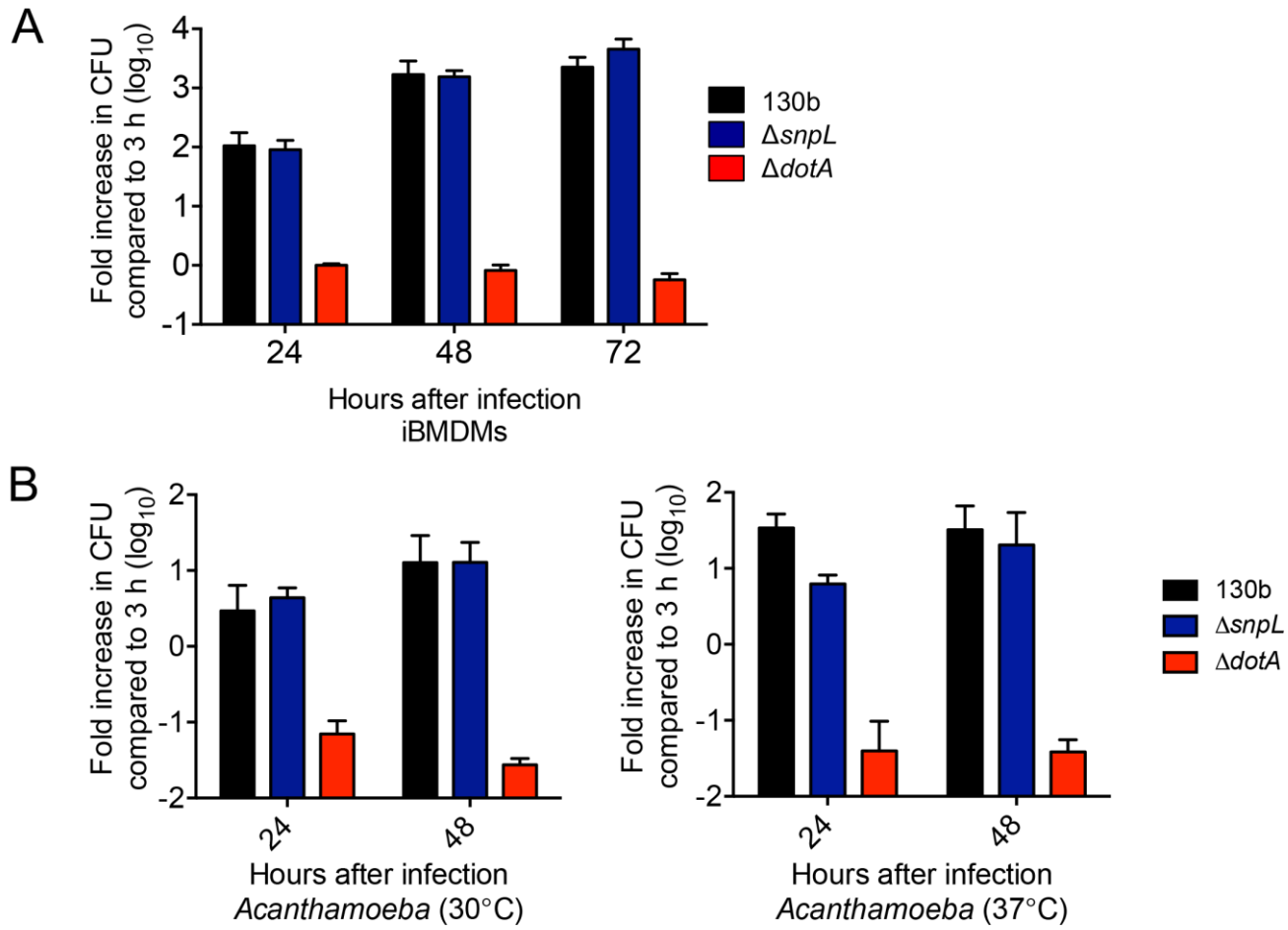

Fig. 8. Role of *snpL* in bacterial replication during infection. Fold increase in bacteria recovered from iBMDM (A) and *A. castellanii* (B) infected with wild type *L. pneumophila* 130b or isogenic Δ*snpL* and Δ*dotA* mutants. Colony-forming units (CFU) were enumerated and expressed as log_{10} fold-change normalized to CFU 3 h post infection. No significant differences were observed at any time points between wild type *L. pneumophila* 130b and Δ*snpL* (unpaired two-tailed *t*-test).