

RESEARCH ARTICLE

# Investigation of *Yersinia pestis* Laboratory Adaptation through a Combined Genomics and Proteomics Approach

Owen P. Leiser<sup>1☯✉1</sup>, Eric D. Merkley<sup>2☯</sup>, Brian H. Clowers<sup>3</sup>, Brooke L. Deatherage Kaiser<sup>2</sup>, Andy Lin<sup>2</sup>, Janine R. Hutchison<sup>2</sup>, Angela M. Melville<sup>2</sup>, David M. Wagner<sup>1</sup>, Paul S. Keim<sup>1</sup>, Jeffrey T. Foster<sup>1✉2</sup>, Helen W. Kreuzer<sup>2\*</sup>

**1** Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, AZ, 86001, United States of America, **2** Chemical and Biological Signature Sciences, Pacific Northwest National Laboratory, Richland, WA, 99352, United States of America, **3** Department of Chemistry, Washington State University, Pullman, WA, 99354, United States of America

☯ These authors contributed equally to this work.

✉1 Current address: Chemical and Biological Signature Sciences, Pacific Northwest National Laboratory, Richland, WA, 99352, United States of America.

✉2 Current address: Department of Molecular, Cellular, and Biomedical Science, University of New Hampshire, Durham, NH, 03824, United States of America.

\* [helen.kreuzer@pnnl.gov](mailto:helen.kreuzer@pnnl.gov)



OPEN ACCESS

**Citation:** Leiser OP, Merkley ED, Clowers BH, Deatherage Kaiser BL, Lin A, Hutchison JR, et al. (2015) Investigation of *Yersinia pestis* Laboratory Adaptation through a Combined Genomics and Proteomics Approach. PLoS ONE 10(11): e0142997. doi:10.1371/journal.pone.0142997

**Editor:** D. Ashley Robinson, University of Mississippi Medical Center, UNITED STATES

**Received:** December 12, 2014

**Accepted:** October 29, 2015

**Published:** November 24, 2015

**Copyright:** © 2015 Leiser et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Genome assemblies of ancestral strains Yp1945 and Yp2126 have been deposited as Whole Genome Shotgun projects in GenBank (accession numbers LIXX00000000 and LIXY00000000). Raw reads from evolved populations in FASTQ format have been deposited in the Sequence Read Archive (SRA; accession numbers SRR2183237-SRR2183258, and SRR2183308). All proteomic data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository (accession numbers PXD002955 and PXD002961).

## Abstract

The bacterial pathogen *Yersinia pestis*, the cause of plague in humans and animals, normally has a sylvatic lifestyle, cycling between fleas and mammals. In contrast, laboratory-grown *Y. pestis* experiences a more constant environment and conditions that it would not normally encounter. The transition from the natural environment to the laboratory results in a vastly different set of selective pressures, and represents what could be considered domestication. Understanding the kinds of adaptations *Y. pestis* undergoes as it becomes domesticated will contribute to understanding the basic biology of this important pathogen.

In this study, we performed a parallel serial passage experiment (PSPE) to explore the mechanisms by which *Y. pestis* adapts to laboratory conditions, hypothesizing that cells would undergo significant changes in virulence and nutrient acquisition systems. Two wild strains were serially passaged in 12 independent populations each for ~750 generations, after which each population was analyzed using whole-genome sequencing, LC-MS/MS proteomic analysis, and GC/MS metabolomics. We observed considerable parallel evolution in the endpoint populations, detecting multiple independent mutations in *ail*, *pepA*, and *zwf*, suggesting that specific selective pressures are shaping evolutionary responses. Complementary LC-MS/MS proteomic data provide physiological context to the observed mutations, and reveal regulatory changes not necessarily associated with specific mutations, including changes in amino acid metabolism and cell envelope biogenesis. Proteomic data support hypotheses generated by genomic data in addition to suggesting future mechanistic studies, indicating that future whole-genome sequencing studies be designed to leverage proteomics as a critical complement.

**Funding:** This work was supported by Defense Threat Reduction Agency Basic Research Award DTRA100271A-2129 and by the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. Battelle Memorial Institute operates Pacific Northwest National Laboratory for the U.S. DOE under Contract DE-AC06-76RLO. The funders authorized publication, but had no role in study design, data collection and analysis, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Microorganisms inhabiting the natural environment experience vastly different selective pressures from those growing in the laboratory. The natural environment is generally nutrient-limited for microbes, with spatial orientation playing an important role in nutrient acquisition and interaction with neighboring cells [1–3]. In contrast, laboratory conditions are often far removed from those in the environment, with a high likelihood that a cell will find itself in a nutrient-rich environment, growing in a monoculture without interspecies competition, and free from challenges from the host immune system. Little is known about the genomic adaptations arising during the transition from wild to laboratory conditions, largely because whole-genome sequencing has only recently become available as an economically viable means of addressing fine-scale evolutionary questions. Similarly, little is known regarding changes in protein expression arising during pathogen domestication.

*Yersinia pestis* is a recently emerged clone of *Yersinia pseudotuberculosis*, having arisen perhaps as recently as 2,600 years ago [4] and is the causative agent of plague. *Y. pestis* has been responsible for three historical pandemics, including the Justinian plague (first pandemic; 6<sup>th</sup>-18<sup>th</sup> centuries) [5], and the second pandemic from the 14<sup>th</sup>-17<sup>th</sup> centuries [6]. The third and most recent global plague pandemic has been attributed to a reemergence of *Y. pestis* in China in the mid-19<sup>th</sup> century, with subsequent distribution around the world via commercial shipping vessels [4, 7, 8].

*Y. pestis*' natural life cycle is primarily sylvatic; the organism occurs in specific flea vectors on rodent hosts [7], with humans acting as incidental hosts. It has become endemic in the western United States, with particular foci occurring in the Four Corners region of Arizona, Utah, Colorado, and New Mexico [9]. Endemic plague continues to be an issue for wildlife biologists and human public health worldwide, and a basic understanding of this organism's biology has important implications for disease control and prevention.

In this study, we explored genomic and proteomic changes arising during laboratory culturing conditions in recently isolated sylvatic strains of *Y. pestis*. We performed a parallel serial passage experiment (PSPE), in which multiple populations of two strains were passaged repeatedly under separate but identical growth conditions. This allowed each population to evolve on its own independent trajectory. Although PSPE has previously been carried out on several other *Y. pestis* strains [10], the purpose of those studies was to evaluate mutation rates in variable number tandem repeat loci and no whole-genome sequencing was performed. PSPE has been performed extensively with *Escherichia coli* in glucose- and maltose-limiting medium in order to elucidate evolutionary processes (known as the Lenski experiment, it has been running continuously since 1989 and has most recently been described in [11]). This long-term evolution experiment is ongoing, and has selected strongly for mutations that improve cells' ability to take up and utilize glucose/maltose. The *E. coli* experiment imposes a fairly stringent set of evolutionary pressures on a strain pre-adapted to laboratory growth. In our case selective pressure is relaxed due to the rich medium used, and cells are presented with multiple potential evolutionary pathways for adaptation from environmental to laboratory growth.

Mikkola and Kurland [12] have reported isolation and adaptation of wild *E. coli* cells, and Sjödin et al [13] have compared wild and laboratory strains of *Francisella tularensis*, but did not directly investigate mechanisms of domestication. Eydallin et al [14] investigated metabolic and phenotypic consequences of *E. coli* domestication, but did not explore the underlying genetic causes of observed phenotypes. Although Saxer et al [15] recently used proteomics and genomics to examine adaptation by two commensal bacteria to laboratory growth, to our knowledge this is the first reported experiment investigating specific genomic and proteomic adaptation(s) of a wild pathogen to laboratory conditions. *Y. pestis* has been the subject of

numerous proteomic studies investigating protein expression changes caused by temperature [16, 17] and iron availability [18] related to the transition between flea and mammalian hosts, and by intracellular growth in macrophages [19]. Proteomic approaches have also been used for the identification of virulence factors and mechanisms [20, 21] and for strain identification [22]. The purpose of this study was to conduct a thorough genomic and proteomic characterization of two wild *Y. pestis* strains passaged under laboratory conditions for ~750 generations through whole genome sequencing to discover single nucleotide polymorphisms (SNPs) and insertion/deletion mutations (indels), as well as LC-MS/MS proteomic analysis and GC-MS carbohydrate profiling to examine protein and monosaccharide abundance changes over time. The picture that emerges from the combined proteomic and genomic data is of a systems-level change in metabolism and physiology that results from a complex interplay between genetic and regulatory factors. Our results illustrate the value of systems-level phenotypic measurements in PSPE studies and suggest several promising avenues for future investigations.

## Materials and Methods

### Growth media, chemicals, and growth conditions

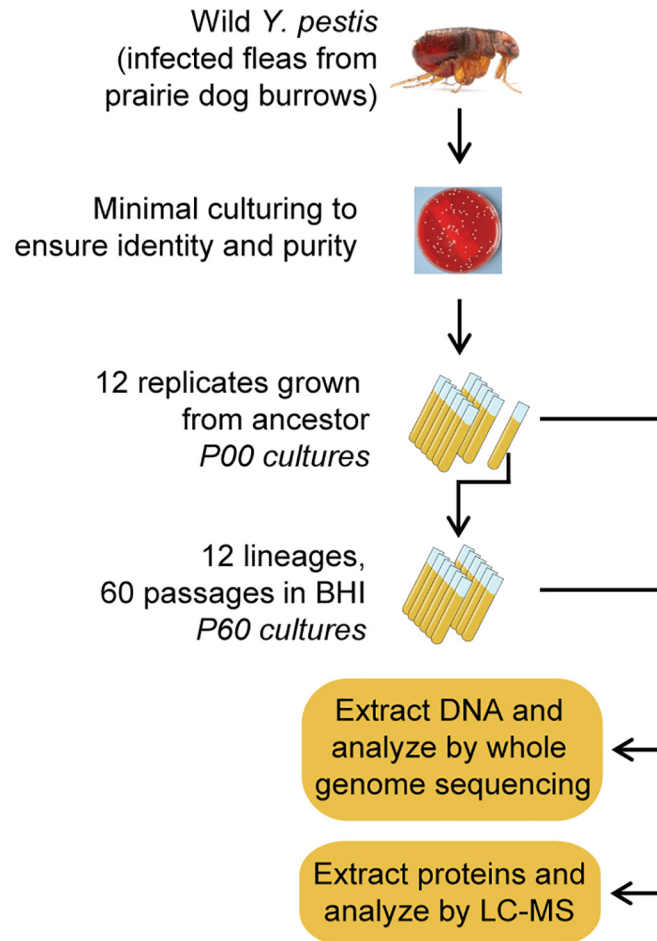
Unless otherwise noted, all chemicals and reagents used in this study were of analytical grade. All culturing of *Y. pestis* was carried out in a Centers for Disease Control and Prevention-approved Select Agent BSL3 facility at Northern Arizona University. *Y. pestis* strains were routinely cultured in Brain-Heart Infusion broth (BHI; BD Diagnostics), and, as required, on 5% sheep's blood agar (SBA; Hardy Diagnostics) or BHI supplemented with 1.5% w/v agar. Cells were grown at 28°C to minimize the likelihood of plasmid loss, with vigorous agitation under aerobic conditions.

### Isolation of *Yersinia pestis* strains used in this study

Two wild *Yersinia pestis* strains were isolated from fleas using methods developed in this laboratory [23, 24]. Briefly, fleas were collected from black-tailed prairie dog (*Cynomys ludovicianus*) colonies in Texas, USA in 2009 (Yp1945) and from Gunnison's prairie dog colonies (*C. gunnisoni*) in Arizona, USA in 2011 (Yp2126). Both colonies had shown signs of recent die-offs. The Arizona sample was collected from public land, and a research permit was provided by the State of Arizona for DMW. The Texas sample was collected from private land by a State of Texas veterinarian, who was given permission by the landowner to conduct the study at the site. Because fleas were collected from burrows after rodent hosts had already died, no prairie dogs were harmed or killed for the purpose of this study. Fleas were pooled by burrow and homogenized in BHI broth supplemented with 10% glycerol. The homogenized suspensions were plated onto cefsulodin, irgasan, and novobiocin (CIN) agar plates and incubated at 28°C for 48 h. Suspected *Y. pestis* colonies were purified onto SBA, and their identity confirmed by a real-time PCR-based assay targeting the plasmid-borne *pla* gene [25, 26]. Confirmed *Y. pestis* isolates were spread onto a fresh sheep blood agar plate. All colonies were scraped from this plate and used to make a freezer stock, which served as the inoculum for the starting culture for PSPE. Using this protocol, the *Y. pestis* strains used were passaged no more than three times during the isolation process.

### Parallel Serial Passage Experiment (PSPE)

The conceptual framework of this experiment is outlined in Fig 1. PSPE was carried out by repeatedly passaging the two *Y. pestis* strains in 10 ml BHI broth at 28°C for 48 h (stationary phase) in 50 ml conical tubes. Passages destined for proteomic and carbohydrate analyses at



**Fig 1. Simplified outline of PSPE.** Following isolation and minimal laboratory culturing ( $\leq 3$  passages on solid media), each isolate was grown in replicate broth cultures. A single culture was used to inoculate replicate 12 independent broth cultures carried through 60 passages ( $\sim 750$  generations). Evolved genomes and proteomes were compared to ancestral states.

doi:10.1371/journal.pone.0142997.g001

Pacific Northwest National Laboratory (PNNL) were grown in 15 ml BHI in order to achieve sufficient cellular biomass. In order to minimize the number of passages before the start of the experiment, Passage 0 (P00) cultures were inoculated directly from frozen glycerol stocks into two replicate 15 ml BHI broth cultures and agitated vigorously during growth. An additional ten replicate cultures were later inoculated and processed to provide sufficient statistical resolution (see below) for proteomic analysis of starting strains. Beginning with P01, twelve populations were inoculated, originating from one of the original cultures for each strain. Cell density was determined by measuring optical density (600 nm) prior to subculturing to ensure sufficient inoculum was used. To avoid a potential genetic bottleneck resulting in the loss or fixation of mutant genotype(s) by means other than selection, approximately  $10^6$  CFU were used as an inoculum for each passage. Passages 1–60 were carried out by diluting each stationary phase culture in two 10-fold dilutions, routinely 100  $\mu$ l into 900  $\mu$ l BHI broth, with 250  $\mu$ l of the final dilution inoculated into 10 ml fresh BHI broth (i.e.  $10^1 * 10^1 * 4 \times 10^1 = 4 \times 10^3$ ). As noted above, populations destined for proteomic analysis in addition to DNA sequencing were grown in 15 ml BHI broth. Under the conditions of this experiment, cell populations reached approximately  $5 \times 10^8$  CFU ml<sup>-1</sup> representing  $\sim 12$  generations per passage in 10 ml cultures

and ~17 generations in 15 ml cultures, and resulting in ~750 generations per population over the course of the experiment. Population 1 (L01) of Yp2126 became contaminated early (passage 3) in the experiment; this contaminant was later identified as *Bacillus licheniformis* (a common laboratory contaminant) by 16S sequencing and the culture was discarded, leaving a total of eleven independent populations for Yp2126 and twelve populations for Yp1945.

## DNA extraction and sequencing

Total cellular DNA from endpoint populations and starting isolates was extracted from 250  $\mu$ l stationary phase culture (approximately  $1.5 \times 10^8$  cells) using the DNeasy Blood and Tissue Kit (Qiagen, Valencia CA) according to the manufacturer's instructions. Cells were pelleted by centrifugation at 5,000  $\times$  g and supernatant removed prior to processing.

DNA extracts were prepared for sequencing according to the manufacturer's instructions, with the following modifications: DNA was fragmented using a SonicMan 96-well sonicator (Matrical Bioscience, Spokane WA) to yield a distribution of fragment sizes between 250–500 bp. After sonication, fragmented DNA was purified using a QIAquick PCR Purification kit (Qiagen, Valencia CA). All other purification steps were carried out using Agencourt AMPure XP beads (Beckman Coulter, Danvers MA) in a 96-well format. End repair, dA-tailing, and ligation reactions were carried out using NEBNext kit (New England Biolabs, Ipswich MA). Size selection was performed using E-Gel SizeSelect Gels (Life Technologies, Foster City CA). AIR DNA Barcodes– 48 (Bioo Scientific, Austin TX) were used during indexing PCR. Quantification of prepared libraries was carried out using Kapa SYBR Fast ABI Prism 2x qPCR Master Mix (Kapa Biosystems, Boston MA) on an Applied Biosystems ABI Prism 7900 (Life Technologies, Foster City CA). Flow cells were prepared for sequencing on a Cluster Station (Illumina, San Diego CA). Paired-end DNA sequencing was carried out on an Illumina Genome Analyzer IIx or MiSeq (Illumina, San Diego CA) using SBS version 3 reagents at the Translational Genomics Research Institute Division of Pathogen Genomics (TGen North, Flagstaff AZ), or using an Illumina MiSeq at the NAU Environmental Genetics and Genomics Laboratory (from 50 and 200 base reads). Read length varied depending on the library preparation chemistry and sequencer available at the time of preparation. Raw reads from evolved populations in FASTQ format have been deposited in the Sequence Read Archive (SRA; accession numbers SRR2183237-SRR2183258, and SRR2183308).

## Genome assembly and mutation analyses

Ancestral genomes from Yp1945 and Yp2126 were assembled from Illumina sequence reads (FASTQ) using ABySS v1.3.2 [27] in order to uncover large-scale genome rearrangements relative to YpCO92 reference. Because average fold-coverage for the chromosome and plasmids pCD1, pMT1, and pPCP1 was excessive for genome assembly when aligned to *Y. pestis* CO92 (accession numbers NC\_003143, NC\_003131, NC\_003134, and NC\_003132, respectively), and indeed was likely to generate misassemblies due to the accumulation of sequencing errors, we randomly down sampled to 7,000,000 read pairs for each strain before generating assemblies using CO92 as a reference sequence. Post-assembly correction using the PAGIT toolkit [28] yielded 12 contigs for Yp1945 and 19 contigs for Yp2126. Visualization of assembled contigs relative to YpCO92 reference was performed using MUMmer v3.23 [29]. Both assemblies have been deposited as Whole Genome Shotgun projects in GenBank (accession numbers LIXX00000000 and LIXY00000000)

Single nucleotide polymorphism (SNP), insertions/deletions (indels), and small-scale rearrangement (e.g. transposons) analyses for all strains were carried out using breseq v0.24-rc6 (<https://code.google.com/p/breseq/downloads/list>) using `—j2—p—c` arguments and a

customized YpCO92 annotated GenBank file as reference, generated by concatenating GenBank files of the chromosome and all three plasmids. Mutations were considered significant if they were present on at least 90% of reads for ancestral strains, or at least 75% of reads for evolved populations, and were verified by manually examining alignments using Tablet v1.14.04.10 (<http://bioinf.hutton.ac.uk/tablet>). Mutations called in all populations were discarded, as it is highly unlikely to observe the same mutation in all populations and these are most parsimoniously explained as sequencing errors.

### Inactivation of cells with ethanol and preparation of cells for proteomic analysis

Inactivation of *Y. pestis* cells prior to shipment was carried out as described in Lin et al [30], which has been demonstrated to be compatible with downstream proteomic analysis. Approximately  $10^9$  stationary phase cells were pelleted by centrifugation. A 500  $\mu$ l aliquot of culture supernatant was filter-sterilized using 0.22  $\mu$ m centrifugal filters (Corning); the remaining supernatant was discarded. Cells were inactivated by resuspending cell pellets in 40% ethanol and incubating for 20 min at room temperature. Inactivated cells were washed once with 1 ml sterile phosphate-buffered saline (PBS) and resuspended in 500  $\mu$ l PBS. Supernatant and cell mass were verified sterile prior to shipment on ice to PNNL. For the purposes of statistical analysis, 10 additional cultures of starting isolates were grown and processed in the same manner as the original cultures, bringing the total number of independent cultures to 12.

### Preparation of samples for proteomic analysis

For peptide preparation, 25–50  $\mu$ l of ethanol-inactivated *Y. pestis* cells were pelleted by centrifugation. Volumes varied because the initial blocking and randomization of run order required more analytical replicates of ancestor samples, and thus a higher starting biomass (see below). Pellets were resuspended in 50  $\mu$ l lysis buffer (6 M urea (Sigma) and 14.3 mM 2-mercaptoethanol (Sigma) in 100 mM triethylammonium bicarbonate (TEAB) pH 9 (Sigma)). Samples were then incubated for one hour at 60°C with shaking. Insoluble cell material was removed by a brief centrifugation. 400  $\mu$ l of 100 mM TEAB was added to the supernatant followed by 5  $\mu$ l 375 mM iodacetamide (Pierce) in 100 mM TEAB. Samples were incubated for 30 minutes at room temperature in the dark. Proteins were digested with 2.5  $\mu$ g of trypsin (Promega) for 14 hours at 37°C with gentle shaking. Solid phase extraction (SPE) was performed with a vacuum manifold using Strata C-18T columns according to the manufacturer's instructions. Briefly, 1 ml of 100% methanol (Sigma) was added to activate the resin, followed by a conditioning rinse of 0.1% v/v trifluoroacetic acid (TFA, Sigma), then addition of samples in TEAB. Samples were washed with 0.1% v/v TFA, and eluted with 80% v/v acetonitrile (Sigma) in 0.1% v/v TFA into clean low-protein binding 1.5 ml microfuge tubes (Fisher Scientific). Samples were dried to near completeness (~5  $\mu$ l) using an Eppendorf Vacufuge Plus. Peptides were resuspended in 0.1% v/v formic acid (Suprapure EMD) in water and the concentration was adjusted to 1 mg ml<sup>-1</sup> using the BCA assay (Pierce). Samples were transferred to high performance liquid chromatography (HPLC) vials with inert glass inserts and capped with screw caps, and stored at -20°C until analysis.

### Liquid chromatography-mass spectrometry measurements

Digested peptide samples were subjected to liquid chromatography on an Agilent Infinity 1260 HPLC system. The column was a fused silica capillary (40 cm x 150  $\mu$ m inner diameter) packed with 5  $\mu$ m particle size, 300 Å pore size Jupiter C18 resin (Phenomenex, Torrance CA). 1  $\mu$ l aliquots (total mass ~1  $\mu$ g) were injected and subjected to the following 160-minute gradient:

100% Solvent A for 10 minutes; 0%-7.5% Solvent B over 1 minute; 7.5%-45% Solvent B over 109 minutes; 45%-95% Solvent B over 2 minutes; 95% Solvent B for 10 minutes; 95%-0% Solvent B over 4 minutes; 100% Solvent A for 23 minutes. Solvent A was 5% v/v acetonitrile/0.1% v/v formic acid and Solvent B was 95% v/v acetonitrile/0.1% v/v formic acid. Blanks consisting of 5  $\mu$ l injections of 50% v/v isopropanol/50% v/v acetone/0.1% v/v formic acid were run with a shorter gradient between samples to minimize column carryover.

Each batch of samples, which comprised P00 and six P60 populations, was run in a block with a randomized run order. Although some MS drift was observed the effect was minimal and did not affect data analysis. Each block was repeated a total of three times with a different random run order each time. To monitor the quality of the chromatographic separation, standards were run before and after each block. The standard was a tryptic digest of ovalbumin, bovine serum albumin, bovine  $\alpha$ S1-casein, and bovine lactalbumin (all from Sigma) at equal mass concentrations.

The HPLC was coupled to a Thermo Scientific LTQ Orbitrap XL mass spectrometer via a custom electrospray emitter consisting of an etched fused silica capillary [31]. The MS was operated in data dependent “high-low” mode with a high-resolution ( $R = 30,000$ ) precursor scan collected in the Orbitrap followed by collision-induced dissociation (CID) fragment scans of the top seven most intense precursors collected in the ion trap. Data dependent acquisition parameters were: dynamic exclusion repeat count 2, repeat duration 30 seconds, exclusion list size 250, exclusion list duration 180 seconds.

## Proteomic data analysis

We used intensity-based label-free quantitation (LFQ) using MaxQuant v1.5.2.8 [32, 33], which provided detailed and quantitative data in addition to identifying a large number of proteins.

The protein sequence database used was built by annotating the assembled genomes of wild isolates (i.e. Yp1945 and Yp2126) using RAST [34, 35]. The experimentally determined genome sequences of the starting isolates were used in the database searches for the respective ancestor and evolved populations. Gene/locus names were mapped back to the published *Y. pestis* CO92 genome [36] using BLASTp [37]. LC-MS/MS data in the Thermo Scientific RAW file format and the protein sequence database in FASTA format were loaded into MaxQuant.

Peptide identification was accomplished in MaxQuant using the integrated Andromeda search engine [38] with the following parameters: precursor ion mass tolerance, 20 ppm for the first search and 4.5 ppm for the main search; tryptic enzyme specificity; maximum number of missed cleavages, 2; methionine oxidation, and protein N-terminal acetylation were included as variable modifications, and cysteine carbamidomethylation as a fixed modification; minimum number of peptides, 1 (additional analysis was performed using 2 peptide minimum; see below). The default 1% false discovery rate filter was used at both the peptide and protein level. The “match between runs” and “re-quantify” options were also used. Proteomic data have been deposited in the ProteomeXchange Consortium [39] via the PRIDE partner repository (accession numbers PXD002955 and PXD002961).

MaxQuant’s LFQ functionality was also used for quantification [40, 41]. Each dataset was compared to at least three and on average six other datasets for estimation of normalization coefficients. Statistical analysis of LFQ intensity values was carried out using Inferno (<http://omics.pnl.gov/software/infernordn>), a freely available version of DANTE [42]. Proteins were judged to have changed abundance significantly if the  $q$ -value from ANOVA (comparing all P00 replicates to all P60 replicates) was less than 0.05, and the fold change between two conditions was greater than 2; or if the protein was detected in only one condition, and in that case

more than half the LC-MS/MS analyses for that condition. Identified proteins were functionally classified using eggNOG v4.1beta [43] based on UniProt identification numbers.

## Carbohydrate analysis

Carbohydrate content of the biomass and medium samples was measured as the monosaccharide profile by the alditol acetate method [44–46]. Briefly, biomass samples were hydrolyzed with H<sub>2</sub>SO<sub>4</sub>, and the resulting monosaccharides were purified by SPE on a C18 stationary phase, converted to their volatile alditol acetate derivatives by reduction with sodium borodeuteride, acetylated with acetic anhydride, extracted with chloroform, purified again by reaction with ammonium hydroxide, followed by SPE on a hydrophilic stationary phase. The organic phase was evaporated to dryness and the sample dissolved in the appropriate amount of solvent for gas chromatography-mass spectrometry (GC-MS) analysis.

GC-MS analysis and quantitation of carbohydrates were carried out as described previously [44–46]. The polyamine putrescine (butane-1,4-diamine), although not a carbohydrate, was initially detected in these experiments as a prominent unknown peak. MS/MS and accurate mass data confirmed the identity of this peak, and putrescine was added to the external standard mixture in subsequent experiments to allow quantitation.

## Results

### Ancestral genome assembly and analysis

Assembled genomes of the ancestral strains, Yp1945 and Yp2126 had no apparent large-scale chromosomal rearrangements when compared to YpCO92 using MUMmer (data not shown). Yp1945 was assembled into 12 contigs with average contig length (N50) of 798,621 and total sequence of 4.82 Mb. Yp2126 was assembled to 19 contigs with N50 of 903,789 and total sequence of 4.81 Mb. A total of 42 SNPs in Yp1945 (5 synonymous, 19 nonsynonymous, 17 intergenic, and 1 noncoding RNA; S2 Table) and 40 SNPs in Yp2126 (5 synonymous, 16 nonsynonymous, and 19 intergenic; S3 Table) differentiated the ancestral strains from YpCO92. The high ratio of nonsynonymous to synonymous SNPs may imply positive selection as these two lineages have become established in their respective niches. In addition to SNPs, Yp1945 contained 38 indels relative to YpCO92 (S2 Table). Yp2126 differed from YpCO92 at 30 indels (S3 Table). All of these mutations were removed from SNP/indel calls for the respective evolved populations using the gdttools SUBTRACT function of breseq to remove spurious calls, i.e. mutations that were present in the genome at the beginning of the experiment but which differ from the YpCO92 sequence used as the reference.

### Mutational analysis of populations after laboratory evolution

We found that the loss of the *pgm* locus and the pCD1 plasmid were common, along with mutations in the genes *ail*, *pepA*, and *zwf*. We defined mutations as dominant in a given population if they were observed in >75% of sequence reads, in order to focus our analysis on mutations evidently well on the way to being fixed. A total of 75 mutations were identified across the 23 populations, with nearly half (n = 36) observed in intergenic regions. Mutations occurred in a wide variety of genes (S1 Spreadsheet), with the overwhelming majority of SNPs (24 of 30) resulting in nonsynonymous amino acid changes. Although intergenic mutations could have effects on gene expression, e.g. by affecting promoter or operator regions, any such effects were not investigated in this study. As our objective was to identify common pathways by which wild pathogens adapt to the laboratory, we looked for mutations arising in a given gene in populations derived from both starting strains, and have restricted further analysis



here to genes meeting this criterion. Mutations in *ail*, *pepA*, and *zwf* as well as loss of pCD1 were observed in populations derived from both ancestral strains, but loss of *pgm* was only observed in populations derived from Yp1945 (see below).

### Mutations in *ail* give rise to predicted truncated Ail protein

*Ail* (YPO2905) encodes an adhesin necessary for successful *Y. pestis* invasion of host cells [47]. It folds into a monomeric  $\beta$ -barrel protein in the outer membrane [48, 49], and is transcribed at a very high level in *Y. pestis* [50, 51]. *Ail* mutations were dominant in 10 of 12 Yp1945-derived populations and 6 of 11 Yp2126-derived populations (Table 1). *Ail* was mutated in two additional populations derived from Yp1945, although read frequency (29%, L06; 32%, L11)

**Table 1. Effect of *ail* alleles on Ail protein abundance.**

Population	Observed mutation	Ail protein abundance ratio <sup>1</sup>
Yp1945-L01	IS100 insertion <sup>2</sup> at nt 21	Ancestor only <sup>3</sup>
Yp1945-L02	E29*	0.021 <sup>5</sup>
Yp1945-L03	IS100 insertion at nt 19	0.007 <sup>5</sup>
Yp1945-L04	+20 bp <sup>4</sup> at nt 142; F69*	Ancestor only
Yp1945-L05	WT	0.011 <sup>5</sup>
Yp1945-L06	E62* (29% of reads)	0.902±0.587
Yp1945-L07	E29*	0.020±0.013
Yp1945-L08	+8 bp at nt 523; E182*	Ancestor only
Yp1945-L09	+8 bp at nt 523; E182*	Ancestor only
Yp1945-L10	+1 bp at nt 41; K18*	Ancestor only
Yp1945-L11	W67* (32% of reads)	Ancestor only
Yp1945-L12	Q132*	0.050±0.033
Yp2126-L02	WT	Ancestor only
Yp2126-L03	+4 bp at nt 43; F15*	Ancestor only
Yp2126-L04	E62*	Ancestor only
Yp2126-L05	+4 bp at nt 43; F15*	Ancestor only
Yp2126-L06	-34 bp at nt 122–155; T43*	Ancestor only
Yp2126-L07	WT	Ancestor only
Yp2126-L08	+8 bp at nt 531; E182*	Ancestor only
Yp2126-L09	IS100 insertion at nt 43 (40% of reads) IS100 insertion at nt 167 (60% of reads)	Ancestor only
Yp2126-L10	WT	Ancestor only
Yp2126-L11	WT	Ancestor only
Yp2126-L12	WT	0.026±0.015

<sup>1</sup>Ratio of protein abundance of evolved population relative to ancestor. Error reflects standard deviation of the mean using triplicate technical replicates.

<sup>2</sup>Insertion sequence mediated mutations are referred to by their nucleotide position (nt, nucleotide).

<sup>3</sup>Ancestor only, protein was observed in ancestor but not in evolved population.

<sup>4</sup>Indel mutations resulting in frame shift and premature stop codon are indicated by the relevant amino acid mutation

\* (bp, base pair).

<sup>5</sup>Protein was only identified in one of three technical replicates; therefore statistical analysis of variance was impossible.

doi:10.1371/journal.pone.0142997.t001

did not meet our cutoff for inclusion. All of the observed mutations resulted in either premature stop codons, or, in three populations, disruption by mobile elements.

Observed protein levels in evolved strains were consistent with disruptive mutations. Ail protein was not identified in most of the evolved populations, either due to complete absence or presence at levels below detectable limits of LC-MS/MS, and when identified was significantly reduced in abundance relative to ancestral strains. The notable exception to this trend is Yp1945-L06, with 29% of the population harboring an E62\* mutation but with Ail levels approximately the same as ancestral cells.

### PepA mutations result in downstream changes to CarAB expression

*PepA* (YPO3441) encodes a multifunctional leucyl aminopeptidase/DNA binding protein [52]. Aside from its eponymous function, it is an accessory protein required for proper plasmid resolution partitioning via the Xer system [53], is required for stable maintenance of P1 prophage [54], and is an accessory protein to Cer recombination [55]. At the end of 60 passages, mutant *pepA* alleles were dominant in seven populations for Yp1945 and eight populations for Yp2126, with an additional population, Yp1945-L06 harboring a mutation at moderate frequency (30% of reads) (Fig 2). When functioning as a DNA binding protein, PepA represses transcription of *carAB*, encoding the large and small subunits of carbamoylphosphate synthetase, by binding to sites upstream of *carA* [52, 56]. CarAB (YPO0481/0482) catalyzes the synthesis of carbamoylphosphate, an early step in arginine and pyrimidine metabolism [57–59]. As PepA represses *carAB* transcription, we hypothesized that the observed mutations serve to de-repress *carAB* transcription.

Proteomic data support this hypothesis. Yp1945-derived populations expressing mutant PepA have on average 6.0-fold higher protein levels of CarA and 5.4-fold higher levels of CarB than their wild progenitors (Fig 2A). Populations derived from Yp2126 expressing mutant PepA have on average 7.4-fold higher protein levels of CarA and 3.9-fold higher levels of CarB (Fig 2B). Yp1945-L06 was excluded from the above average because its mutant *pepA* allele is present in only 30% of reads. Interestingly, point mutations observed at the extreme N-terminus of the protein, G8C and A16T, appear to exert a smaller influence on *carAB* expression than any other observed mutations (Fig 2).

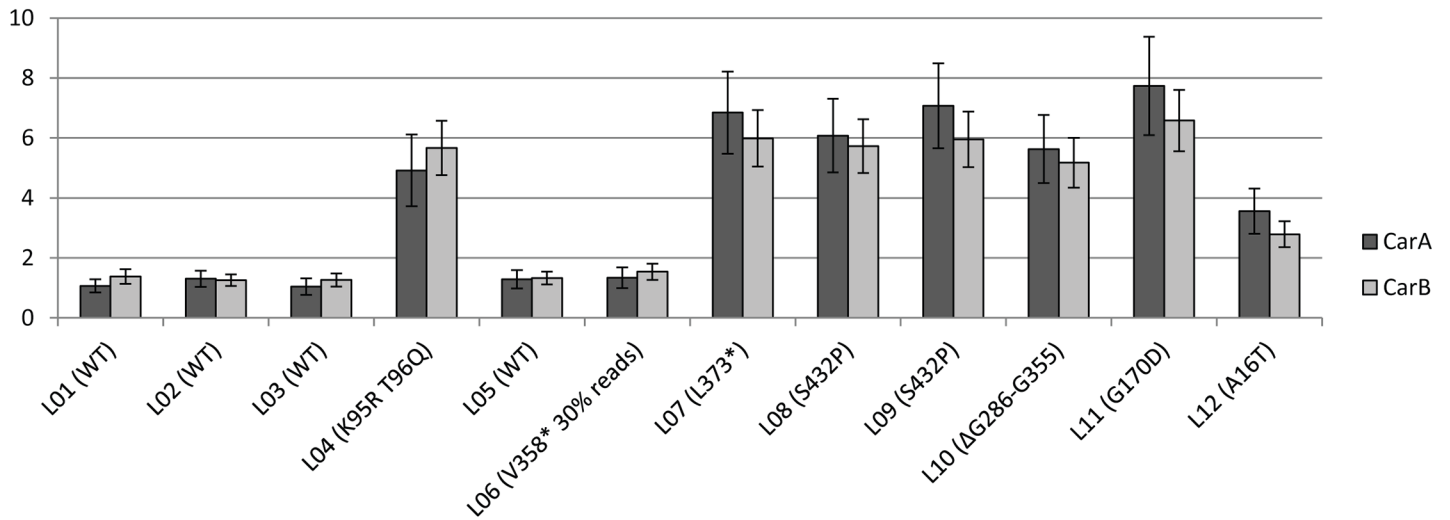
### Zwf mutations may restore protein function

*Zwf* (YPO2066) encodes glucose-6-phosphate dehydrogenase, which is responsible for the conversion of glucose-6-phosphate to 6-phosphogluconolactone as the entrance to the pentose-phosphate pathway (PPP), and which serves as a modulator of PPP [60–62]. *Zwf* additionally functions in anabolic metabolism and in maintaining a reducing environment within the cell cytoplasm [63]. *Zwf* is inactive in *Y. pestis* due to a S155P mutation relative to the *Y. pseudotuberculosis* sequence [64, 65]. After passaging, *zwf* mutations were discovered in 4 populations derived from Yp1945 and 1 population from Yp2126 (Table 2). Three of the mutations found in this study alter position 155, including a reversion to serine in Yp1945-L10. Two mutations distal to position 155, H397Y and I382L, were also observed. Potential effects of these mutations are discussed below.

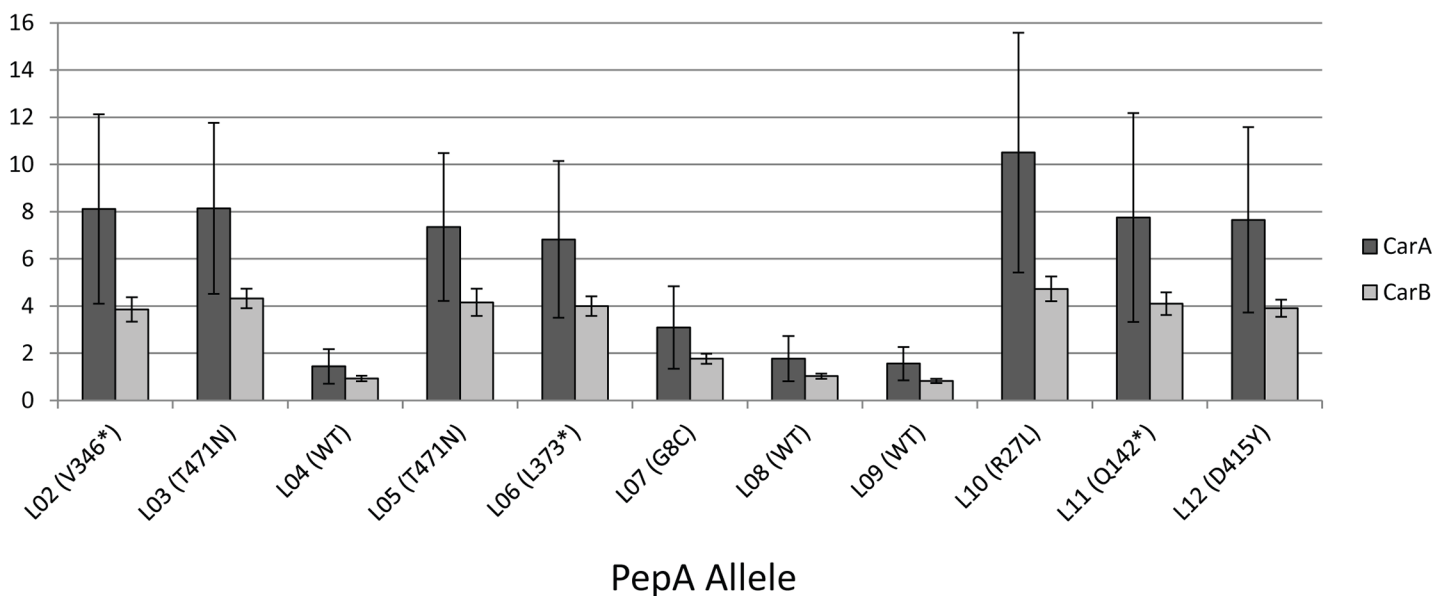
### pCD1 plasmid is a commonly lost genetic element

The pCD1 plasmid carries a number of genes encoding proteins required for virulence, including a number of outer membrane proteins [66] and supporting biosynthetic and secretory proteins [67], including those comprising a type 3 secretion system. Importantly, it also confers calcium-dependence during growth at 37°C [68]. After 60 passages at 28°C in BHI, only

**A. Yp1945**



**B. Yp2126**



**Fig 2. Effects of mutant PepA protein on CarAB protein levels.** Evolved populations expressing mutant PepA protein have increased CarA (dark grey bars) and CarB (light grey bars) relative to ancestral cells, consistent with derepression of the *carAB* operon. Each population was measured in triplicate. Specific *pepA* alleles are given in parentheses. Error bars reflect standard deviation of the mean. A, Yp1945-derived populations. B, Yp2126-derived populations.

doi:10.1371/journal.pone.0142997.g002

populations Yp1945-L10, Yp2126-L07, and Yp2126-L11 retained the pCD1 plasmid at high levels as judged by number of reads (185-, 308-, and 425-fold coverage as judged by analysis with MUMmer, respectively; data not shown). An additional two populations had low sequence coverage of pCD1: Yp1945-L04 and Yp2126-L09 had coverage of 12- and 17-fold coverage, respectively (data not shown). It is likely that there is strong selective pressure for loss of this plasmid, as low coverage levels were observed in multiple populations as early as

**Table 2. Identified mutations in Zwf.**

Population	Mutation
Yp1945-L01	P155T
Yp1945-L03	H397Y
Yp1945-L05	I382L
Yp1945-L10	P155S
Yp2126-L04	P155L

doi:10.1371/journal.pone.0142997.t002

passage 20 and complete loss was observed in most populations by passage 30. Plasmids pPCP1 and pMT1 were almost uniformly retained (data not shown).

### Large colony variants are a common occurrence in evolved populations

At the end of P60, cultures were serially diluted and plated onto BHI agar in order to directly measure cell density, which was not significantly different from that of the starting strains (data not shown). Although cell density was essentially unchanged at the end of the experiment, we were surprised to find colonies with markedly different morphologies than ancestral strains. Specifically, very large colonies were observed at high frequencies in almost all populations ([S1 Table](#)), with the notable exception of Yp1945-L06. We initially hypothesized that the large colony morphotype was due to loss of the *pgm* locus in these isolates, but close examination of sequence alignments revealed that lack of *pgm* is insufficient to explain this phenotype. Colony size variants after serial passage have been observed in *E. coli* and *Citrobacter freundii*. [[15](#)]. Large colony variants observed in this work are discussed in supporting information ([S1 Table](#)).

### Proteomic analysis of P60 cultures reveals common themes

Numerous proteins had significantly different abundances in evolved relative to ancestral strains (see [S2 Spreadsheet](#) for a complete list). Database searching as described above resulted in a total of 1074 identified proteins for Yp1945-derived populations and 1078 identified proteins for Yp2126-derived populations. Due to the stochastic nature of peptide/protein identification in shotgun proteomics (as opposed to mutation identification in genomics) we have not restricted our analysis to changes in protein abundance observed in individual lineages derived from both starting isolates. Instead we have identified proteins for which mean abundance ratios across all evolved lineages of a starting isolate have changed significantly from the mean abundance ratios across all replicate starting cultures, as described in Materials and Methods. Of the confidently identified proteins, 137 proteins in Yp1945-derived populations and 182 proteins in Yp2126-derived populations were present in significantly different abundances between P00 and P60 (a full listing of these proteins is found in the [S2 Spreadsheet](#)). We focus here on two of the broad functional categories of proteins classified using eggNOG [[43](#)]: 1) amino acid transport and metabolism, and 2) cell envelope and chaperones.

### Amino acid transport and metabolism

Numerous proteins identified as having involvement with amino acid transport and metabolism were expressed significantly differently in evolved vs. ancestral strains ([Table 3](#)). Of particular interest are two proteins responsible for the assimilation of amino nitrogen: GdhA and GlnA. *GdhA* (YPO3971) encodes glutamate dehydrogenase and is the key player in the glutamate dehydrogenase pathway of nitrogen metabolism. *GlnA* (YPO0024) encodes glutamine

**Table 3. Abundance ratios of proteins involved in amino acid metabolism.**

Name	Function	Uniprot ID	Abundance Ratio 1945 <sup>1</sup>	Abundance Ratio 2126 <sup>1</sup>
argG/YPO1570	Argininosuccinate synthase	Q8ZJV7	AO/AO	NI/NI
argH/YPO3924	Argininosuccinate lyase	Q8ZA88	NI/NI	EO/EO
aroF/YPO3286	Phospho-2-dehydro-3-deoxyheptonate aldolase	Q0WC04	AO/AO	NI/NI
aroH/YPO2411	Phospho-2-dehydro-3-deoxyheptonate aldolase	Q7CIT5	AO/NI	AO/NI
artI/YPO4111	Periplasmic solute-binding protein	Q7CFN1	4.91/5.55	10.23/17.83
asnB/YPO2623	Asparagine synthetase	Q7CJU5	0.37/NI	0.49/NI
dapD/YPO1041	Tetrahydropicolinate succinylase	Q8ZH69	2.87/2.76	2.04/2.04
dapE/YPO3053	Succinyl-diaminopimelate desuccinylase	Q7CJI9	AO/NI	AO/AO
gcvP/YPO0905	Glycine dehydrogenase	Q8ZHI8	2.15/2.08	NI/NI
gcvT/YPO0907	Aminomethyltransferase	Q8ZHI6	2.09/NI	NI/NI
gdhA/YPO3971	Glutamate dehydrogenase	Q0WA46	11.47/NI	104.63/94.65
glnA/YPO0024	Glutamine synthetase	Q74YC3	5.90/NI	20.87/47.80
ilvC/YPO3888	Ketol-acid reductoisomerase	Q8ZAC2	2.48/2.38	NI/NI
iscS/YPO2896	Cysteine desulfase	Q7CJN7	2.90/NI	NI/NI
pepA/YPO344	Leucyl aminopeptidase	Q8ZBH3	2.92/2.56	2.89/2.90
pepB/YPO2889	Aminopeptidase B	P58475	2.20/2.10	6.18/7.19
proV/YPO2647	Glycine betaine/L-proline transport ATP-binding protein	Q0WDN9	NI/NI	0.48/NI
selD/YPO2164	Selenophosphate synthase	Q8ZEK1	AO/AO	NI/NI
trpE/YPO2208	Anthranilate synthase component	Q7CIM3	NI/NI	AO/AO
ureA/YPO2665	Urease subunit alpha	P69994	2.38/2.32	4.63/4.56
ureB/YPO2666	Urease subunit beta	P69990	3.30/3.30	7.89/6.30
ureC/YPO2667	Urease subunit gamma	Q9ZFR9	2.772.70	4.17/4.11

<sup>1</sup>Abundance ratio reflects ratio of protein levels in evolved relative to ancestral populations.

NI, not identified (either protein was not identified during database search or abundance ratio did not meet significance threshold); EO, protein observed in evolved populations only; AO, protein observed in ancestral populations only.

Abundance ratios reflect “1 peptide”/“2 peptides” as described in discussion.

doi:10.1371/journal.pone.0142997.t003

synthetase, the first step of the glutamine synthetase/glutamate synthase pathway. GdhA and GlnA were present at higher abundance in evolved lineages of both Yp1945 and Yp2126. In addition to glutamate and glutamine metabolism, protein abundance changes were observed in lysine, glycine, serine, and threonine metabolism (generally increased abundance) as well as phenylalanine, tyrosine, and tryptophan metabolism (generally decreased abundance). Increased CarAB protein abundance has been described above.

Interestingly, we observed substantially increased abundance of the urease holoenzyme (UreABC; YPO2665/2666/2667), despite the fact that *Y. pestis* is phenotypically urease-negative due to a single base insertion in the cryptic *ureD* gene [69]. Implications of this finding are discussed below.

Finally, we also investigated carbohydrate content of these populations during the course of this experiment. Although not a carbohydrate, the polyamine putrescine was readily detected by GC-MS in these samples. Putrescine levels were elevated in Yp1945- and Yp2126-derived lineages (1.41-fold higher,  $p = 1.45e^{-5}$ ; 2.01-fold higher,  $p = 1.08e^{-7}$ , respectively). Polyamine production is related to amino acid production, and putrescine is a common intermediate in the metabolism of multiple amino acids.

## Envelope biogenesis and chaperones

Cell envelope biogenesis and chaperone activity are interconnected in bacteria, and so are reported here together (Table 4). In general, chaperone protein abundance is lower in evolved populations. In particular, DegP (YPO3382), IbpB (YPO4084), and RseP (YPO1051) are all involved in envelope stress response [70–72] and are present in lower abundance after laboratory evolution. In contrast, FkpA (YPO0195), GroL (YPO0351), Skp (YPO1053), and SlyD (YPO1093) are all primarily responsible for envelope biogenesis [73–76] and are present in higher abundance after laboratory evolution. SurA, thought to be the major periplasmic

**Table 4. Abundance ratios of proteins involved in cell wall/membrane biogenesis and chaperones/folding factors.**

Protein Name	Protein Function	UniProt ID	Expression Ratio 1945 <sup>1</sup>	Expression Ratio 2126 <sup>1</sup>
alr/YPO0321	Alanine racemase 1	Q8ZJ10	AO/AO	NI/NI
ampD/YPO1683	N-acetylmuramoyl-L-alanine amidase	Q7CIX7	NI/NI	0.47/0.48
clpB/YPO0506	ATP-dependent protease	Q7CG96	0.11/0.11	0.10/0.09
clpP/YPO3157	Endopeptidase Clp	Q8ZC65	NI/NI	EO/3.35
dacC/YPO1320	D-alanyl-D-alanine carboxypeptidase	Q7CHG0	2.05/2.92	2.32/2.43
degP/YPO33832	Periplasmic protease/chaperone	Q7CKD3	0.44/0.42	NI/NI
degQ/YPO3566	Protease	Q74PR7	NI/NI	0.35/0.37
fbaA/YPO0920	Fructose-bisphosphate aldolase class II	Q7CGS5	2.22/2.17	NI/NI
fkpA/YPO0195	Peptidyl-prolyl cis-trans isomerase	Q0WKB1	2.74/2.59	11.63/12.36
glnD/YPO1042	Protein-P11 uridylyltransferase	Q8ZH68	NI/NI	AO/NI
groL/YPO0351	60 kDa chaperonin	Q8ZIY3	2.91/2.90	4.25/4.15
grxA/YPO1327	Glutaredoxin 1	Q7CHG3	NI/NI	2.16/NI
ibpB/YPO4084	Small heat shock protein IbpB	Q8Z9V6	NI/NI	AO/AO
lgt/YPO0784	Prolipoprotein diacylglycerol transferase	Q8ZHV0	NI/NI	0.38/NI
lolB/YPO2015	Outer membrane lipoprotein LolB	Q8ZEY0	NI/NI	AO/AO
lpxD/YPO1054	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase	P58611	NI/NI	AO/NI
mpl	Putative ligase	Q7CKK2	NI/NI	AO/AO
ompF/YPO1411	Outer membrane porin protein F	Q0WH04	15.48/14.74	24.93/18.17
ompW/YPO2201	Putative exported protein	Q0WEW6	EO/EO	NI/NI
rafQ/YPO0186	Putative LPS biosynthetic enzyme	Q3V8F4	NI/NI	AO/NI
rffG/YPO2602	dTDP-glucose 4,6-dehydratase	Q0WAE8	NI/NI	AO/NI
rlpA/YPO2602	Rare lipoprotein A	Q8ZDG6	AO/NI	NI/NI
rseP/YPO1051	Protease RseP	Q8ZH59	NI/NI	AO/AO
skp/YPO1053	Chaperone protein Skp	P58607	NI/NI	2.30/2.03
slyD/YPO0193	Metallochaperone SlyD	Q7CFU4	2.14/2.08	NI/NI
surA/YPO0494	Chaperone SurA	Q7CG87	NI/NI	AO/NI
trxB/YPO1374	Thioredoxin reductase	Q7CHI6	3.69/3.28	NI/2.14
wecB/YPO3864	UDP-N-acetylglucosamine 2-epimerase	Q8ZAE3	NI/NI	AO/NI
YPO2072	Uncharacterized protein	Q0WF84	NI/NI	EO/NI
YPO2155	Putative exported protein	Q7CIH1	NI/NI	0.50/NI
YPO2855	Putative protease	Q7CJL2	2.12/2.05	NI/NI
YPO3524	Putative exported protein	Q7CKK5	NI/NI	AO/NI

<sup>1</sup>Abundance ratio reflects ratio of protein levels in evolved relative to ancestral populations.

NI, not identified (either protein was not identified during database search or abundance ratio did not meet significance threshold);

EO, protein observed in evolved populations only;

AO, protein observed in ancestral populations only.

Abundance ratios reflect “1 peptide”/“2 peptides” as described in discussion.

chaperone for outer membrane proteins (OMPs) in *E. coli* [75], does not follow this trend; it is only observed in ancestral populations of Yp2126.

While not themselves chaperones, LolB (YPO2015) and LpxD (YPO1054) are responsible for lipoprotein [77] and lipopolysaccharide (LPS) [78] localization in the OM and are only observed only in ancestral populations of Yp2126. OmpF (YPO1411), which is a major porin residing in the OM, is significantly more abundant after evolution in both Yp1945- and Yp2126-derived lineages. RseP, a critical component of the  $\sigma^E$  activation cascade, is only observed in ancestral Yp2126.

Proteins responsible for cell wall synthesis are also present in altered abundances in evolved populations: Alr (YPO0321) and AmpD (YPO1683) are significantly less abundant in evolved populations relative to ancestor strains (only observed in ancestral Yp1945 and expression ratio of 0.47 in Yp2126-derived populations, respectively). A notable exception to this pattern is DacC (YPO1320), an inner membrane penicillin-binding protein, which is present in significantly higher abundance after evolution.

## Discussion

### Genomic investigation of *Y. pestis* laboratory adaptation

Mutations in *ail* resulted universally in disruption or premature truncation of Ail protein (Table 1), and can be explained by invoking reduced metabolic load; since Ail is one of the most highly transcribed genes in *Yersinia* [50, 51] and is not required for laboratory growth, reduced synthesis of the protein in the mutants should concomitantly increase the ability of mutant cells to produce other cellular components necessary to outcompete other members of the population. Pieper et al [48, 79] showed that Ail is more abundant at 37°C than at 26°C, but it is important to note that Ail is abundant in the outer membrane even at 26°C (see Fig 4 of reference [48] and Fig 1 of reference [79]). With the exception of Yp1945-L06, all lineages containing mutant *ail* have either low or undetectable levels of Ail protein after evolution. Yp1945-L05, Yp2126-L02, Yp2126-L07, Yp2126-L10, Yp2126-L11, and Yp2126-L12 are all genetically wild-type with respect to *ail*, but do not express Ail at detectable levels after laboratory evolution. This suggests that there is a significant role played by Ail downregulation independent of direct mutation. Since Ail is required for invasion of host cells, it is likely that populations lacking the protein have attenuated virulence, although this hypothesis was not tested here.

*PepA* mutations were dispersed throughout the protein, occurring in both the DNA-binding and aminopeptidase domains (see [52] for a discussion of *E. coli* PepA structure and function). Similar mutations have been observed in other *pepA* mutants; many of them resulted in increased *carAB* expression [52, 53], suggesting that the observed PepA mutants in our study also result in increased *carAB* transcription. Proteomic data show that expression of CarAB was significantly elevated in the presence of all of these mutations (Fig 2A and 2B). These data suggest that three distinct but related mechanisms could be responsible for de-repression of the *carAB* operon. First, point mutations in the DNA-binding domain likely alter PepA function and prevent repression of *carAB* transcription. Second, indels leading to premature truncation throughout the protein likely result in loss of the protein entirely. The third possible mechanism of derepression of *carAB* transcription is to prevent interaction between PepA monomers by altering amino acids at the C-terminus of the protein. C-terminal amino acids interact with N-terminal amino acids of adjoining PepA monomers during Xer site-specific recombination [80]. It is likely that prevention of monomer interactions hinders proper protein assembly and function. The fact that a particular mutation at the N-terminus of the protein, G8T, only modestly increases CarAB expression suggests that a modest increase in the pool of cellular

carbamoylphosphate provides sufficient benefit to cells under these conditions. The data presented here suggest that increased levels of CarAB protein are of particular benefit to cells growing in rich media in the laboratory. This is supported by the observation that evolved populations expressing wild-type PepA also have slightly higher levels of CarA and CarB proteins relative to ancestral strains (Fig 2A and 2B), and also indicates that another mechanism or mechanisms may influence *carAB* transcription under these conditions.

Zwf is normally inactive in *Y. pestis* biovar Orientalis [65, 81] due to a mutation of S155P relative to *Y. pseudotuberculosis*. The observed mutations in this experiment, including a reversion to “wild-type” Zwf (i.e. P155S; Table 2) suggest that utilization of the oxidative steps of PPP is a selectable trait under these conditions. The ability of *Y. pestis* cells to effectively shunt the carbon contained in media glucose and cellular glucose-6-phosphate to biosynthetic pathways, i.e. PPP instead of TCA cycle, may confer an advantage during laboratory growth. We speculate that the observed P155S mutation enables use of PPP by restoring Zwf function, and it is possible that P155T and P155L mutations restore function as well, although probably to lower levels. The effects of the remaining mutations, H397Y and I382L, are not easy to predict. Cleavage of Zwf by ClpXP cytoplasmic protease produces extracellular death factor (EDF) [82], amplifying the activities of MazEF and ChpBK toxins [83]. It is unlikely that H397Y or I382L affect EDF production (positions 199–203; [82]), although they may alter Zwf protein folding in such a way as to overcome presumable misfolding caused by the presence of proline instead of serine at position 155.

Plasmid loss has been observed in other laboratories after passaging [84, 85]. Loss of the pCD1 plasmid may also be favored under the conditions of our experiment: pCD1 contains the genes for the low calcium response (LCR; [86]), which is not invoked at temperatures below 37°C. As our PSPE was performed at 28°C it is likely that due to a lack of selection for maintenance of the LCR, cells were cured of this plasmid and thus escaped the metabolic cost of maintenance. In contrast, plasmids pMT1 and pPCP1 were maintained during laboratory passage. Maintenance of pMT1 is likely related to the high expression of murine toxin (S2 Spreadsheet); the phospholipase activity of murine toxin may be advantageous in BHI and could therefore provide selective pressure to maintain this plasmid. Pesticin and the pesticin immunity protein are encoded by pPCP1. Toxin/antitoxin systems often consist of a stable toxin and labile antitoxin [87, 88], and although relative stability of the pesticin/pesticin immunity proteins has not been investigated it is likely that they follow the same pattern. Pesticin is expressed at detectable levels in evolved populations; it therefore follows that cells within the population must maintain pPCP1 in order to also express the immunity protein.

Pervasive loss of the *pgm* locus in populations derived from Yp1945 may be a completely neutral event not requiring selection. *Pgm* is bounded by IS100 elements, and is commonly lost during laboratory culture [89]. Nevertheless, a model in which *pgm* is lost due to neutral events is insufficient to explain the fact that Yp2126-derived populations uniformly maintained the *pgm* locus. Further investigation into this phenomenon is warranted.

## Proteomic investigation of *Y. pestis* laboratory adaptation

In contrast to the genomic analysis above, in which we focused on specific mutations present in lineages derived from both starting strains, proteomic analysis of evolved populations took a wider view and examined broad categories of proteins differing significantly in abundance from ancestral strains. We took this approach largely due to the stochastic nature of shotgun proteomics; inefficiencies in peptide detection and identification during LC-MS/MS can ultimately result in a less than complete biological picture for any one biological replicate, even with multiple technical replicates. In this experiment, shotgun proteomics provided a high-



level view of metabolic processes, and in some cases provided data to support or disprove hypotheses generated by genomic data. In addition to the initial analysis using a minimum of one peptide to identify a protein, we refined our analysis to require two peptides for identification. This increased stringency did not qualitatively affect our results. Most protein quantification was unaffected. The majority of protein quantification that was affected was due to subtle differences resulting in abundance changes that only just failed to meet our search criteria. On their own, proteomic data generated during this experiment have shown promising avenues for further hypothesis-driven research independent of observing specific mutations.

Alteration of cell envelope constituents is a key result of laboratory adaptation in this study (Table 4). In this study, we observed downregulation of key stress response proteins. DegP is the major periplasmic protease [90], and is thought to be responsible for rescuing OMPs that have fallen off the SurA-based assembly pathway, in concert with Skp [75]. DegP levels are reduced after laboratory adaptation, as are SurA levels, suggesting a low overall flux of OMP intermediates through the envelope. As Skp can function to rescue *degP*- cells under certain conditions [75], it is not surprising that Skp and DegP protein abundances follow opposite trends. Further support for low overall OMP flux through the cell envelope comes from low levels of RseP in evolved lineages. RseP functions to liberate the alternative sigma factor  $\sigma^E$  from the membrane protein RseA during envelope stress response [91], and from low RseP protein levels we infer that envelope stress is low. Lower abundance of additional envelope assembly factors after laboratory adaptation was also common. We observed lower levels of LolB and LpxD in adapted strains. These genes are responsible for proper localization of OM lipoproteins [77] and LPS [78]. LpxD is itself localized in the OM, again highlighting low overall OMP flux through the cell envelope. It is important to note here that both LolB and LpxD are essential in *E. coli*; therefore it is not possible to infer complete cellular absence from a lack of identification of these or other proteins in this discussion using proteomics. Overall, our observations of cell envelope proteins support the assertion that adapted cells experience far lower stress in the laboratory growth environment (in this case, BHI broth) than their wild ancestors.

We observed significantly elevated levels of the UreABC urease holoenzyme. Intriguingly, *Y. pestis* is phenotypically urease-negative due to a single base insertion in the cryptic *ureD* gene, which introduces a premature stop codon [69] in the metallochaperone protein required for inserting a nickel ion to form a mature holoenzyme. Since a reversion of *ureD* is possible, and since we observed elevated UreABC levels, we sequenced *ureD* using Sanger sequencing and tested laboratory-adapted strains using an agar-based urease assay (Beckton-Dickenson). Neither a reversion to full-length *ureD* nor urease-positive phenotypes were observed. However, our proteomic results unambiguously show that the *ureABC* genes are not only transcribed and translated, but also are upregulated during long-term laboratory cultivation on BHI.

The power of combining genomic with proteomic analyses was particularly highlighted when we examined the connection between mutant PepA and elevated putrescine levels. CarAB catalyzes the conversion of L-glutamine to L-glutamate and carbamoylphosphate. Carbamoylphosphate can enter the urea cycle by reacting with ornithine to form arginine. Because the urea cycle in *Y. pestis* is blocked by the lack of arginase, the cells convert arginine to agmatine via arginine decarboxylase. Agmatine is ultimately converted to putrescine by the successive action of agmatine deaminase and N-carbamoyl putrescine decarboxylase. We hypothesize that *Y. pestis* cells use putrescine and possibly other polyamine compounds as nitrogen sinks during growth on amino nitrogen-rich compounds, and that this use drives the fixation of mutations such as we observed in *pepA* (Fig 2). This fits generally with the differences between BHI and mammalian/flea host environments, although the lack of knowledge of

basic *Y. pestis* physiology leaves some of the details unclear. BHI is particularly rich in nitrogen, containing 0.31 moles nitrogen per 1 mole carbon, and after growth in BHI this ratio is further decreased in *Y. pestis* biomass (Kreuzer, unpublished data). Nitrogen in BHI is supplied mostly in the form of oligopeptides and amino acids. Since glucose is rapidly depleted during exponential phase growth, cells growing on BHI would presumably need to use amino acids as a carbon source by the time the culture reaches stationary phase. It is therefore reasonable to speculate that an increased ability to use amino acid carbon backbones for energy and biomass by stripping them of their amino groups would be advantageous in BHI. Free amines could then ultimately be stored as polyamine compounds such as putrescine.

This PSPE showed that alterations in global nitrogen metabolism, especially alterations in amino acid biosynthesis, are key adaptations to growth in the laboratory (Table 3). Mechanisms of global nitrogen metabolism in enterobacteria have been described, although they come primarily from studies of *E. coli*, *Klebsiella*, and *Salmonella* [92, 93], and are not well understood in *Y. pestis*. *Y. pestis* differs fundamentally from *E. coli* and *Klebsiella*. First *Y. pestis* does not possess the nitrogen assimilation control (NAC) gene, which encodes a secondary transcription factor implicated in global response to intracellular nitrogen [92]. A BLAST search of the entire Uniprot database using *K. pneumoniae* NAC failed to yield any hits to any *Yersinia* species. Second, *Y. pestis* lacks several key enzymes active in the downstream metabolism of amino acids and nitrogenous compounds: aspartase, arginase, and urease. Thus, models of nitrogen metabolism built using other enterobacteria may not be fully applicable when interpreting data from *Y. pestis*. Future work should be geared towards closing this knowledge gap.

## Supporting Information

**S1 Spreadsheet. Genomic mutations in laboratory-evolved *Y. pestis* populations derived from ancestor strains Yp1945 and Yp2126.** Provided as a fully searchable Microsoft Excel spreadsheet with multiple tabs. *Read evidence* tabs provide the fraction of reads in each evolved population exhibiting the indicated mutation. Green, yellow, and red highlighting indicate  $\geq 75\%$ , between 20% and 75%, 5%-20% of reads for a given population exhibit the indicated mutation. *Missing coverage evidence* tabs summarize genes for which evolved populations were missing read coverage, indicating genomic deletions. *New junction evidence* tabs detail the reads providing evidence for novel junctions as a complement to the missing coverage information.

(XLSX)

**S2 Spreadsheet. Proteomics results for ancestral Yp1945 and Yp2126 strains and their respective laboratory-evolved descendent populations.** Data are the normalized, log<sub>2</sub>-transformed protein abundance values derived from LC-MS/MS experiments as reported by the MaxQuant analysis software. Each column labeled "P00.x" (where x is an identifier, 1–10 or A, B) represents an independent biological replicate (average of three or more injection replicates) of the ancestral culture. Each column labeled "Ly" (where y identifies the parallel cultures from 1–12) represents an independently evolved (60 passages) population, (average of injection replicates). "q" is the q-value calculated by Inferno. "Unique P00" and "Unique P60" describe whether a given protein is observed exclusively in one or the other conditions, with 1 indicating logical "true" and 0 "false" (see text for details). "Log<sub>2</sub> Fold Change" is the overall expression ratio, log<sub>2</sub>-transformed. "Final Filter" indicates (1 = true, 0 = false) whether the given protein meets all of the criteria for consideration as significantly changing. See [Methods](#) section for details.

(XLSX)

**S1 Table. Large colony phenotype in evolved strains and relationship to loss of pCD1 and *pgm* loci.** Provided as a Microsoft Word document. The table records the frequency of the large colony phenotype and the loss or maintenance of the two loci. The image shows an example of the large colony morphotype together with wild-type colonies. Supplemental text describes the large colony phenotype.

(DOCX)

**S2 Table. SNPs and indels present in the ancestral Yp1945 strain relative to *Y. pestis* CO92.** The small number of differences shows that Yp1945 is closely related to CO92.

(DOCX)

**S3 Table. SNPs and indels present in the ancestral Yp2126 strain relative to *Y. pestis* CO92.** The small number of differences shows that Yp2126 is closely related to CO92.

(DOCX)

## Acknowledgments

We thank Adina Doyle and Nicolette Janke for preparing DNA libraries for sequencing, and Janine Hutchison for valuable assistance in preparing samples for proteomic analyses. This work was supported by Defense Threat Reduction Agency Basic Research Award DTRA100271A-2129 and by the Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. Battelle Memorial Institute operates Pacific Northwest National Laboratory for the U.S. DOE under Contract DE-AC06-76RLO. The funders authorized publication, but had no role in study design, data collection and analysis, or preparation of the manuscript.

## Author Contributions

Conceived and designed the experiments: HWK JTF BHC PSK. Performed the experiments: OPL EDM BHC JRH BLDK AMM. Analyzed the data: OPL EDM BHC AL. Contributed reagents/materials/analysis tools: DMW. Wrote the paper: OPL EDM JTF HWK BLDK JRH.

## References

1. Tolker-Nielsen T, Molin S. Spatial organization of microbial biofilm communities. *Microbial Ecology*. 2000; 40(2):75–84.
2. Wang YP, Law RM, Pak B. A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere. *Biogeosciences*. 2010; 7:2261–82.
3. Brookes P. The soil microbial biomass: Concept, measurement and applications in soil ecosystem research. *Microbes and Environments*. 2001; 16(3):131–40.
4. Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, et al. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nature Genetics*. 2010; 42(12):1140–3. doi: [10.1038/ng.705](https://doi.org/10.1038/ng.705) PMID: [21037571](https://pubmed.ncbi.nlm.nih.gov/21037571/)
5. Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, et al. *Yersinia pestis* and the Plague of Justinian 541–543 AD: A genomic analysis. *The Lancet Infectious Diseases*. 2014; 14(4):319–26. doi: [10.1016/S1473-3099\(13\)70323-2](https://doi.org/10.1016/S1473-3099(13)70323-2) PMID: [24480148](https://pubmed.ncbi.nlm.nih.gov/24480148/)
6. McGovern TW, Friedlander AM. Plague. In: Zajchuk R, Bellamy RF, editors. *Medical Aspects of Chemical and Biological Warfare* 1997. p. 479–502.
7. Gage KL, Kosoy MY. Natural history of plague: perspectives from more than a century of research. *Annual Review of Entomology*. 2005; 50:505–28. PMID: [15471529](https://pubmed.ncbi.nlm.nih.gov/15471529/)
8. Achtman M, Morelli G, Zhu P, Wirth T, Diehl I, Kusecek B, et al. Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(51):17837–42. doi: [10.1073/pnas.0408026101](https://doi.org/10.1073/pnas.0408026101) PMID: [15598742](https://pubmed.ncbi.nlm.nih.gov/15598742/)

9. Eisen RJ, Ensore RE, Biggerstaff BJ, Reynolds PJ, Ettestad P, Brown T, et al. Human plague in the southwestern United States, 1957–2004: Spatial models of elevated risk of human exposure to *Yersinia pestis*. *Journal of Medical Entomology*. 2007; 44(3):530–7. doi: [10.1603/0022-2585\(2007\)44\[530:hpitsuj2.0.co;2](https://doi.org/10.1603/0022-2585(2007)44[530:hpitsuj2.0.co;2) PMID: [17547242](https://pubmed.ncbi.nlm.nih.gov/17547242/)
10. Vogler AJ, Keys CE, Allender C, Bailey I, Girard J, Pearson T, et al. Mutations, mutation rates, and evolution at the hypervariable VNTR loci of *Yersinia pestis*. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 2007; 616(1–2):145–58. PMID: [17161849](https://pubmed.ncbi.nlm.nih.gov/17161849/)
11. Maddamsetti R, Lenski RE, Barrick JE. Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with *Escherichia coli*. *Genetics*. 2015;genetics. 115.176677.
12. Mikkola R, Kurland CG. Selection of laboratory wild-type phenotype from natural isolates of *Escherichia coli* in chemostats. *Molecular Biology and Evolution*. 1992; 9(3):394–402. PMID: [1584010](https://pubmed.ncbi.nlm.nih.gov/1584010/)
13. Sjodin A, Svensson K, Lindgren M, Forsman M, Larsson P. Whole-genome sequencing reveals distinct mutational patterns in closely related laboratory and naturally propagated *Francisella tularensis* strains. *PLoS One*. 2010; 5(7):e11556. doi: [10.1371/journal.pone.0011556](https://doi.org/10.1371/journal.pone.0011556) PMID: [20657845](https://pubmed.ncbi.nlm.nih.gov/20657845/)
14. Eydallin G, Ryall B, Maharjan R, Ferenci T. The nature of laboratory domestication changes in freshly isolated *Escherichia coli* strains. *Environmental Microbiology*. 2014; 16(3):813–28. doi: [10.1111/1462-2920.12208](https://doi.org/10.1111/1462-2920.12208) PMID: [23889812](https://pubmed.ncbi.nlm.nih.gov/23889812/)
15. Saxer G, Krepps MD, Merkley ED, Ansong C, Kaiser BLD, Valovska M-T, et al. Mutations in global regulators lead to metabolic selection during adaptation to complex environments. *PLoS genetics*. 2014; 10(12):e1004872. doi: [10.1371/journal.pgen.1004872](https://doi.org/10.1371/journal.pgen.1004872) PMID: [25501822](https://pubmed.ncbi.nlm.nih.gov/25501822/)
16. Pieper R, Huang ST, Clark DJ, Robinson JM, Parmar PP, Alami H, et al. Characterizing the dynamic nature of the *Yersinia pestis* periplasmic proteome in response to nutrient exhaustion and temperature change. *Proteomics*. 2008; 8(7):1442–58. doi: [10.1002/pmic.200700923](https://doi.org/10.1002/pmic.200700923) WOS:000254986200009. PMID: [18383009](https://pubmed.ncbi.nlm.nih.gov/18383009/)
17. Pieper R, Huang ST, Robinson JM, Clark DJ, Alami H, Parmar PP, et al. Temperature and growth phase influence the outer-membrane proteome and the expression of a type VI secretion system in *Yersinia pestis*. *Microbiology*. 2009; 155:498–512. doi: [10.1099/mic.0.022160-0](https://doi.org/10.1099/mic.0.022160-0) WOS:000263428900018. PMID: [19202098](https://pubmed.ncbi.nlm.nih.gov/19202098/)
18. Pieper R, Huang ST, Parmar PP, Clark DJ, Alami H, Fleischmann RD, et al. Proteomic analysis of iron acquisition, metabolic and regulatory responses of *Yersinia pestis* to iron starvation. *Bmc Microbiology*. 2010; 10. Art 30 doi: [10.1186/1471-2180-10-30](https://doi.org/10.1186/1471-2180-10-30) WOS:000275359600001.
19. Ponnusamy D, Hartson SD, Clinkenbeard KD. Intracellular *Yersinia pestis* expresses general stress response and tellurite resistance proteins in mouse macrophages. *Veterinary Microbiology*. 2011; 150(1–2):146–51. doi: [10.1016/j.vetmic.2010.12.025](https://doi.org/10.1016/j.vetmic.2010.12.025) WOS:000290696500019. PMID: [21295415](https://pubmed.ncbi.nlm.nih.gov/21295415/)
20. Hixson KK, Adkins JN, Baker SE, Moore RJ, Chromy BA, Smith RD, et al. Biomarker candidate identification in *Yersinia pestis* using organism-wide semiquantitative proteomics. *Journal of Proteome Research*. 2006; 5(11):3008–17. doi: [10.1021/pr060179y](https://doi.org/10.1021/pr060179y) WOS:000241755400014. PMID: [17081052](https://pubmed.ncbi.nlm.nih.gov/17081052/)
21. Ansong C, Schrimpe-Rutledge AC, Mitchell HD, Chauhan S, Jones MB, Kim YM, et al. A multi-omic systems approach to elucidating *Yersinia virulence* mechanisms. *Mol Biosyst*. 2013; 9(1):44–54. doi: [10.1039/c2mb25287b](https://doi.org/10.1039/c2mb25287b) WOS:000311822100006. PMID: [23147219](https://pubmed.ncbi.nlm.nih.gov/23147219/)
22. Jabbour RE, Wade MM, Deshpande SV, Stanford MF, Wick CH, Zulich AW, et al. Identification of *Yersinia pestis* and *Escherichia coli* strains by whole cell and outer membrane protein extracts with mass spectrometry-based proteomics. *Journal of Proteome Research*. 2010; 9(7):3647–55. doi: [10.1021/pr100402y](https://doi.org/10.1021/pr100402y) PMID: [20486690](https://pubmed.ncbi.nlm.nih.gov/20486690/)
23. Sarovich DS, Colman RE, Price EP, Chung WK, Lee J, Schupp JM, et al. Selective isolation of *Yersinia pestis* from plague-infected fleas. *Journal of Microbiological Methods*. 2010; 82(1):95–7. doi: [10.1016/j.mimet.2010.03.019](https://doi.org/10.1016/j.mimet.2010.03.019) PMID: [20385178](https://pubmed.ncbi.nlm.nih.gov/20385178/)
24. Girard JM, Wagner DM, Vogler AJ, Keys C, Allender CJ, Drickamer LC, et al. Differential plague-transmission dynamics determine *Yersinia pestis* population genetic structure on local, regional, and global scales. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(22):8408–13. doi: [10.1073/pnas.0401561101](https://doi.org/10.1073/pnas.0401561101) PMID: [15173603](https://pubmed.ncbi.nlm.nih.gov/15173603/)
25. Hinnebusch J, Schwan TG. New method for plague surveillance using polymerase chain reaction to detect *Yersinia pestis* in fleas. *Journal of Clinical Microbiology*. 1993; 31(6):1511–4. PMID: [8314993](https://pubmed.ncbi.nlm.nih.gov/8314993/)
26. Stevenson HL, Bai Y, Kosoy MY, Monteneri JA, Lowell JL, Chu MC, et al. Detection of novel *Bartonella* strains and *Yersinia pestis* in prairie dogs and their fleas (*Siphonaptera: Ceratophyllidae* and *Pulicidae*) using multiplex polymerase chain reaction. *Journal of Medical Entomology*. 2003; 40(3):329–37. doi: [10.1603/0022-2585-40.3.329](https://doi.org/10.1603/0022-2585-40.3.329) PMID: [12943112](https://pubmed.ncbi.nlm.nih.gov/12943112/)

27. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. ABySS: A parallel assembler for short read sequence data. *Genome Research*. 2009; 19:1117–23. doi: [10.1101/gr.089532.108](https://doi.org/10.1101/gr.089532.108) PMID: [19251739](https://pubmed.ncbi.nlm.nih.gov/19251739/)
28. Swain MT, Tsai IJ, Assefa SA, Newbold C, Berriman M, Otto TD. A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs. *Nature Protocols*. 2012; 7:1260–84. doi: [10.1038/nprot.2012.068](https://doi.org/10.1038/nprot.2012.068) PMID: [22678431](https://pubmed.ncbi.nlm.nih.gov/22678431/)
29. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. *Genome Biology*. 2004; 5:R12. PMID: [14759262](https://pubmed.ncbi.nlm.nih.gov/14759262/)
30. Lin A, Merkley ED, Clowers BH, Hutchison JR, Kreuzer HW. Effects of bacterial inactivation methods on downstream proteomic analysis. *Journal of microbiological methods*. 2015; 112:3–10. doi: [10.1016/j.mimet.2015.01.015](https://doi.org/10.1016/j.mimet.2015.01.015) PMID: [25620019](https://pubmed.ncbi.nlm.nih.gov/25620019/)
31. Kelly RT, Page JS, Luo Q, Moore RJ, Orton DJ, Tang K, et al. Chemically etched open tubular and monolithic emitters for nanoelectrospray ionization mass spectrometry. *Analytical Chemistry*. 2006; 78(22):7796–801. doi: [10.1021/ac061133r](https://doi.org/10.1021/ac061133r) PMID: [17105173](https://pubmed.ncbi.nlm.nih.gov/17105173/)
32. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008; 26(12):1367–72. [http://www.nature.com/nbt/journal/v26/n12/supinfo/nbt.1511\\_S1.html](http://www.nature.com/nbt/journal/v26/n12/supinfo/nbt.1511_S1.html). doi: [10.1038/nbt.1511](https://doi.org/10.1038/nbt.1511) PMID: [19029910](https://pubmed.ncbi.nlm.nih.gov/19029910/)
33. Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics: MCP*. 2014; 13(9):2513–26. doi: [10.1074/mcp.M113.031591](https://doi.org/10.1074/mcp.M113.031591) PMC4159666.
34. Aziz R, Bartels D, Best A, DeJongh M, Disz T, Edwards R, et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*. 2008; 9(1):75. doi: [10.1186/1471-2164-9-75](https://doi.org/10.1186/1471-2164-9-75)
35. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*. 2014; 42(D1):D206–D14. doi: [10.1093/nar/gkt1226](https://doi.org/10.1093/nar/gkt1226)
36. Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MTG, Prentice MB, et al. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature*. 2001; 413(6855):523–7. [http://www.nature.com/nature/journal/v413/n6855/supinfo/413523a0\\_S1.html](http://www.nature.com/nature/journal/v413/n6855/supinfo/413523a0_S1.html). PMID: [11586360](https://pubmed.ncbi.nlm.nih.gov/11586360/)
37. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990; 215(3):403–10. doi: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
38. Cox Jr, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. *Journal of Proteome Research*. 2011; 10(4):1794–805. doi: [10.1021/pr101065j](https://doi.org/10.1021/pr101065j) PMID: [21254760](https://pubmed.ncbi.nlm.nih.gov/21254760/)
39. Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol*. 2014; 32(3):223–6. doi: [10.1038/nbt.2839](https://doi.org/10.1038/nbt.2839) PMID: [24727771](https://pubmed.ncbi.nlm.nih.gov/24727771/)
40. Cox Jr, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Molecular and Cellular Proteomics*. 2014;M113.031591.
41. Lubner CA, Cox J, Lauterbach H, Fancke B, Selbach M, Tschopp J, et al. Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. *Immunity*. 2010; 32(2):279–89. doi: [10.1016/j.immuni.2010.01.013](https://doi.org/10.1016/j.immuni.2010.01.013) PMID: [20171123](https://pubmed.ncbi.nlm.nih.gov/20171123/)
42. Polpitiya AD, Qian W-J, Jaitly N, Petyuk VA, Adkins JN, Camp DG, et al. DANTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics*. 2008; 24(13):1556–8. doi: [10.1093/bioinformatics/btn217](https://doi.org/10.1093/bioinformatics/btn217) PMID: [18453552](https://pubmed.ncbi.nlm.nih.gov/18453552/)
43. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, et al. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Research*. 2014; 42(D1):D231–D9. doi: [10.1093/nar/gkt1253](https://doi.org/10.1093/nar/gkt1253)
44. Fox A, Morgan SL, Gilbert J. Preparation of alditol acetates and their analysis by gas chromatography and mass spectrometry. In: Biermann CJ, McGinnis GD, editors. *Analysis of Carbohydrates by GLC and MS*. Florida: CRC Press; 1988. p. 87–117.
45. Harley WM, Kozar MP, Fox A. Trace analysis of muramic acid in indoor air using an automated derivatization instrument and GC—MS2 or GC—MS3. *Journal of Microbiological Methods*. 2002; 51(1):95–104. doi: [10.1016/S0167-7012\(02\)00064-7](https://doi.org/10.1016/S0167-7012(02)00064-7) PMID: [12069894](https://pubmed.ncbi.nlm.nih.gov/12069894/)
46. Wunschel D, Fox KF, Black GE, Fox A. Discrimination among the *B. cereus* group, in comparison to *B. subtilis*, by structural carbohydrate profiles and ribosomal RNA spacer region PCR. *Systematic and Applied Microbiology*. 1994; 17:625–35.

47. Kolodziejek AM, Hovde CJ, Minnich SA. *Yersinia pestis* Ail: multiple roles of a single protein. *Frontiers in Cellular and Infection Microbiology*. 2012; 2(103). doi: [10.3389/fcimb.2012.00103](https://doi.org/10.3389/fcimb.2012.00103)
48. Pieper R, Huang S-T, Clark DJ, Robinson JM, Alami H, Parmar PP, et al. Integral and peripheral association of proteins and protein complexes with *Yersinia pestis* inner and outer membranes. *Proteome Science*. 2009; 7(5). doi: [10.1186/1477-5956-7-5](https://doi.org/10.1186/1477-5956-7-5)
49. Plesniak LA, Mahalakshmi R, Rypien C, Yang Y, Racic J, Marassi FM. Expression, refolding, and initial structural characterization of the *Y. pestis* Ail outer membrane protein in lipids. *Biochimica et Biophysica Acta—Biomembranes*. 2011; 1808(1):482–9.
50. Chauvaux S, Dillies M-A, Marceau M, Rosso M-L, Rousseau S, Moszer I, et al. In silico comparison of *Yersinia pestis* and *Yersinia pseudotuberculosis* transcriptomes reveals a higher expression level of crucial virulence determinants in the plague bacillus. *International Journal of Medical Microbiology*. 2011; 301(2):105–16. doi: [10.1016/j.ijmm.2010.08.013](https://doi.org/10.1016/j.ijmm.2010.08.013) PMID: [20951640](https://pubmed.ncbi.nlm.nih.gov/20951640/)
51. Chauvaux S, Rosso M-L, Frangeul L, Lacroix C, Labarre L, Schiavo A, et al. Transcriptome analysis of *Yersinia pestis* in human plasma: an approach for discovering bacterial genes involved in septicemic plague. *Microbiology*. 2007; 153(9):3112–24.
52. Charlier D, Kholti A, Huysveld N, Gigot DM, Thia-Toong T-L, Glansdorff N. Mutational analysis of *Escherichia coli* PepA, a multifunctional DNA-binding aminopeptidase. *Journal of Molecular Biology*. 2000; 302:411–26. PMID: [10970742](https://pubmed.ncbi.nlm.nih.gov/10970742/)
53. Reijns M, Lu Y, Leach S, Colloms SD. Mutagenesis of PepA suggests a new model for the Xer/ cer synaptic complex. *Molecular Microbiology*. 2005; 57(4):927–41. PMID: [16091035](https://pubmed.ncbi.nlm.nih.gov/16091035/)
54. Paul S, Summers D. ArgR and PepA, accessory proteins for XerCD-mediated resolutions of ColE1 dimers, are also required for stable maintenance of the P1 prophage. *Plasmid*. 2004; 52(1):63–8. PMID: [15212893](https://pubmed.ncbi.nlm.nih.gov/15212893/)
55. G S.C., Colloms SD. Control of Cre recombination by regulatory elements from Xer recombination systems. *Molecular Microbiology*. 2004; 52(1):53–65. PMID: [15049810](https://pubmed.ncbi.nlm.nih.gov/15049810/)
56. Devroede N, Huysveld N, Charlier D. Mutational analysis of intervening sequences connecting the binding sites for integration host factor, PepA, PurR, and RNA polymerase in the control region of the *Escherichia coli* carAB operon, encoding carbamoylphosphate synthase. *Journal of Bacteriology*. 2006; 188(9):3236–45. PMID: [16621816](https://pubmed.ncbi.nlm.nih.gov/16621816/)
57. Cunin R, Glansdorff N, Pierard A, Stalon V. Biosynthesis and metabolism of arginine in bacteria. *Microbiological Reviews*. 1986; 50(3):314–52. PMID: [3534538](https://pubmed.ncbi.nlm.nih.gov/3534538/)
58. Makoff A, Radford A. Genetics and biochemistry of carbamoyl phosphate biosynthesis and its utilization in the pyrimidine biosynthetic pathway. *Microbiological Reviews*. 1978; 42(2):307–28. PMID: [353478](https://pubmed.ncbi.nlm.nih.gov/353478/)
59. Charlier D, Hassanzadeh G, Kholti A, Gigot D, Pierard A, Glansdorff N. carP, involved in pyrimidine regulation of the *Escherichia coli* carbamoylphosphate synthetase operon encodes a sequence-specific DNA-binding protein identical to XerB and PepA, also required for resolution of ColEI multimers. *Journal of Molecular Biology*. 1995; 250(4):392–406. PMID: [7616564](https://pubmed.ncbi.nlm.nih.gov/7616564/)
60. Flores S, de Anda-Herrera R, Gosset G, Bolivar FG. Growth-rate recovery of *Escherichia coli* cultures carrying a multicopy plasmid, by engineering of the pentose-phosphate pathway. *Biotechnology and Bioengineering*. 2004; 87(4):485–94. PMID: [15286986](https://pubmed.ncbi.nlm.nih.gov/15286986/)
61. Zhao J, Baba T, Mori H, Shimizu K. Effect of zwf knockout on the metabolism of *Escherichia coli* grown on glucose or acetate. *Metabolic Engineering*. 2004; 6:164–74. PMID: [15113569](https://pubmed.ncbi.nlm.nih.gov/15113569/)
62. Nicolas C, Keifer P, Letisse F, Kromer J, Massou S, Soucaille P, et al. Response of the central metabolism of *Escherichia coli* to modified expression of the gene encoding the glucose-6-phosphate dehydrogenase. *FEBS Letters*. 2007; 581:3771–6. PMID: [17631881](https://pubmed.ncbi.nlm.nih.gov/17631881/)
63. Lu J, Holmgren A. The thioredoxin antioxidant system. *Free Radical Biology and Medicine*. 2014; 66(0):75–87.
64. Chain PSG, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, et al. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(38):13826–31. doi: [10.1073/pnas.0404012101](https://doi.org/10.1073/pnas.0404012101) PMID: [15358858](https://pubmed.ncbi.nlm.nih.gov/15358858/)
65. Charusanti P, Chauhan S, McAteer K, Lerman JA, Hyduke DR, Motin VL, et al. An experimentally-supported genome-scale metabolic network reconstruction for *Yersinia pestis* CO92. *BMC Systems Biology*. 2011; 5(163). doi: [10.1186/1752-0509-5-163](https://doi.org/10.1186/1752-0509-5-163)
66. Straley SC. The plasmid-encoded outer membrane proteins of *Yersinia pestis*. *Reviews of Infectious Diseases*. 1988; 10(Supplement 2):S323–S6.
67. Hu P, Elliott JM, McCready P, Skowronski E, Games J, Kobayashi A, et al. Structural organization of virulence-associated plasmids of *Yersinia pestis*. *Journal of Bacteriology*. 1998; 180(19):5192–202. PMID: [9748454](https://pubmed.ncbi.nlm.nih.gov/9748454/)

68. Straley SC, Bowmer WS. Virulence genes regulated at the transcriptional level by Ca<sup>2+</sup> in *Yersinia pestis* include structural genes for outer membrane proteins. *Infection and Immunity*. 1986; 51(2):445–54. PMID: [3002984](#)
69. Sebbane F, Devalckenaere A, Foulon J, Carniel E, Simonet M. Silencing and reactivation of urease in *Yersinia pestis* is determined by one G residue at a specific position in the *ureD* gene. *Infection and Immunity*. 2001; 69(1):170–6. doi: [10.1128/iai.69.1.170-176.2001](#) PMID: [11119503](#)
70. Strauch KL, Johnson K, Beckwith J. Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. *Journal of Bacteriology*. 1989; 171(5):2689–96. PMID: [2540154](#)
71. Veinger L, Diamant S, Buchner J, Goloubinoff P. The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *Journal of Biological Chemistry*. 1998; 273(18):11032–7. PMID: [9556585](#)
72. Ades SE. Regulation by destruction: design of the  $\sigma$  E envelope stress response. *Current opinion in microbiology*. 2008; 11(6):535–40. doi: [10.1016/j.mib.2008.10.004](#) PMID: [18983936](#)
73. Arié JP, Sassoon N, Betton JM. Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*. *Molecular Microbiology*. 2001; 39(1):199–210. PMID: [11123702](#)
74. Weissman JS, Hohl CM, Kovalenko O, Kashi Y, Chen S, Braig K, et al. Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. *Cell*. 1995; 83(4):577–87. PMID: [7585961](#)
75. Sklar JG, Wu T, Kahne D, Silhavy TJ. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes & Development*. 2007; 21(19):2473–84. doi: [10.1101/gad.1581007](#)
76. Scholz C, Eckert B, Hagn F, Schaarschmidt P, Balbach J, Schmid FX. SlyD proteins from different species exhibit high prolyl isomerase and chaperone activities. *Biochemistry*. 2006; 45(1):20–33. PMID: [16388577](#)
77. Tanaka K, Matsuyama S-I, Tokuda H. Deletion of *lofB*, encoding an outer membrane lipoprotein, is lethal for *Escherichia coli* and causes accumulation of lipoprotein localization intermediates in the periplasm. *Journal of Bacteriology*. 2001; 183(22):6538–42. PMID: [11673422](#)
78. Buetow L, Smith TK, Dawson A, Fyffe S, Hunter WN. Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis. *Proceedings of the National Academy of Sciences*. 2007; 104(11):4321–6.
79. Pieper R, Huang S-T, Robinson JM, Clark DJ, Alami H, Parmar PP, et al. Temperature and growth phase influence the outer-membrane proteome and the expression of a type VI secretion system in *Yersinia pestis*. *Microbiology*. 2009; 155:498–512. doi: [10.1099/mic.0.022160-0](#) PMID: [19202098](#)
80. Norbert S, Scherratt DJ, Colloms SD. X-ray structure of aminopeptidase A from *Escherichia coli* and a model for the nucleoprotein complex in Xer site-specific recombination. *The EMBO Journal*. 1999; 18(16):4513–22. PMID: [10449417](#)
81. Chain PSG, Hu P, Malfatti SA, Radnedge L, Larimer F, Vergez LM, et al. Complete genome sequence of *Yersinia pestis* strains Antiqua and Nepal516: Evidence of gene reduction in an emerging pathogen. *Journal of Bacteriology*. 2006; 188(12):4453–63. doi: [10.1128/jb.00124-06](#) PMID: [16740952](#)
82. Kolodkin-Gal I, Engelberg-Kulka H. The extracellular death factor: physiological and genetic factors influencing its production and response in *Escherichia coli*. *Journal of Bacteriology*. 2008; 190(9):3169–75. doi: [10.1128/JB.01918-07](#) PMID: [18310334](#)
83. Belitsky M, Avshalom H, Erental A, Yelin I, Kumar S, London N, et al. The *Escherichia coli* extracellular death factor EDF induces the endoribonucleolytic activities of the toxins MazF and ChpBK. *Molecular Cell*. 2011; 41(6):625–35. doi: [10.1016/j.molcel.2011.02.023](#) PMID: [21419338](#)
84. Perry RD, Fetherston JD. *Yersinia pestis*—etiologic agent of plague. *Clinical Microbiology Reviews*. 1997; 10(1):35–66. PMID: [8993858](#)
85. Leal-Balbino TC, Leal NC, Lopes CV, de Almeida AMP. Differences in stability of the plasmids of *Yersinia pestis* cultures in vitro: Impact on virulence. *Mem Inst Oswaldo Cruz*. 2004; 99(7):727–32. PMID: [15654429](#)
86. Perry RD, Straley SC, Fetherston JD, Rose DJ, Gregor J, Blattner FR. DNA sequencing and analysis of the low-Ca<sup>2+</sup>-response plasmid pCD1 of *Yersinia pestis* KIM. *Infection and Immunity*. 1998; 66(10):4611–23. PMID: [9746557](#)
87. Engelberg-Kulka H, Hazan R, Amitai S. mazEF: a chromosomal toxin-antitoxin module that triggers programmed cell death in bacteria. *Journal of cell science*. 2005; 118(19):4327–32.
88. Gerdes K, Christensen SK, Løbner-Olesen A. Prokaryotic toxin—antitoxin stress response loci. *Nature Reviews Microbiology*. 2005; 3(5):371–82. PMID: [15864262](#)
89. Fetherston JD, Schuetze P, Perry RD. Loss of the pigmentation phenotype in *Yersinia pestis* is due to the spontaneous deletion of 102 kb of chromosomal DNA which is flanked by a repetitive element. *Molecular Microbiology*. 1992; 6(18):2693–704. PMID: [1447977](#)

90. Sawa J, Heuck A, Ehrmann M, Clausen T. Molecular transformers in the cell: lessons learned from the DegP protease—chaperone. *Current opinion in structural biology*. 2010; 20(2):253–8. doi: [10.1016/j.sbi.2010.01.014](https://doi.org/10.1016/j.sbi.2010.01.014) PMID: [20188538](https://pubmed.ncbi.nlm.nih.gov/20188538/)
91. Akiyama Y, Kanehara K, Ito K. RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences. *The EMBO journal*. 2004; 23(22):4434–42. PMID: [15496982](https://pubmed.ncbi.nlm.nih.gov/15496982/)
92. Bender RA. A NAC for regulating metabolism: the Nitrogen Assimilation Control protein (NAC) from *Klebsiella pneumoniae*. *Journal of Bacteriology*. 2010; 192(19):4801–11. doi: [10.1128/jb.00266-10](https://doi.org/10.1128/jb.00266-10) PMID: [20675498](https://pubmed.ncbi.nlm.nih.gov/20675498/)
93. Reitzer L. Nitrogen assimilation and global regulation in *Escherichia coli*. *Annual Review of Microbiology*. 2003; 57(1):155–76. doi: [10.1146/annurev.micro.57.030502.090820](https://doi.org/10.1146/annurev.micro.57.030502.090820) PMID: [12730324](https://pubmed.ncbi.nlm.nih.gov/12730324/).